Journal of Medical Internet Research

Journal Impact Factor (JIF) (2023): 5.8
Volume 21 (2019), Issue 8 ISSN 1438-8871 Editor in Chief: Gunther Eysenbach, MD, MPH

Contents

Reviews

Viewpoints

XSL•FO
RenderX

Original Papers

XSL•FO
RenderX

XSL•FO
RenderX

XSL•FO
RenderX

Corrigenda and Addendas

XSL•FO
RenderX

XSL•FO
RenderX

Review

Current State of Digital Biomarker Technologies for Real-Life, Home-Based Monitoring of Cognitive Function for Mild Cognitive Impairment to Mild Alzheimer Disease and Implications for Clinical Care: Systematic Review

Antoine Piau[1,2], MD, PhD; Katherine Wild[2], PhD; Nora Mattek[2], MPH; Jeffrey Kaye[2], MD

[1]Gerontopole, University Hospital of Toulouse, Université Paul Sabatier, Toulouse, France

[2]Oregon Center for Aging and Technology, Oregon Health and Science University, Portland, OR, United States

Corresponding Author:
Antoine Piau, MD, PhD
Oregon Center for Aging and Technology
Oregon Health and Science University
3181 SW Sam Jackson Park Rd
Portland, OR, 97239
United States
Phone: 1 971 230 8664
Email: antoinepiau@hotmail.com

Abstract

Background: Among areas that have challenged the progress of dementia care has been the assessment of change in symptoms over time. Digital biomarkers are defined as objective, quantifiable, physiological, and behavioral data that are collected and measured by means of digital devices, such as embedded environmental sensors or wearables. Digital biomarkers provide an alternative assessment approach, as they allow objective, ecologically valid, and long-term follow-up with continuous assessment. Despite the promise of a multitude of sensors and devices that can be applied, there are no agreed-upon standards for digital biomarkers, nor are there comprehensive evidence-based results for which digital biomarkers may be demonstrated to be most effective.

Objective: In this review, we seek to answer the following questions: (1) What is the evidence for real-life, home-based use of technologies for early detection and follow-up of mild cognitive impairment (MCI) or dementia? And (2) What transformation might clinicians expect in their everyday practices?

Methods: A systematic search was conducted in PubMed, Cochrane, and Scopus databases for papers published from inception to July 2018. We searched for studies examining the implementation of digital biomarker technologies for mild cognitive impairment or mild Alzheimer disease follow-up and detection in nonclinic, home-based settings. All studies that included the following were examined: community-dwelling older adults (aged 65 years or older); cognitively healthy participants or those presenting with cognitive decline, from subjective cognitive complaints to early Alzheimer disease; a focus on home-based evaluation for noninterventional follow-up; and remote diagnosis of cognitive deterioration.

Results: An initial sample of 4811 English-language papers were retrieved. After screening and review, 26 studies were eligible for inclusion in the review. These studies ranged from 12 to 279 participants and lasted between 3 days to 3.6 years. Most common reasons for exclusion were as follows: inappropriate setting (eg, hospital setting), intervention (eg, drugs and rehabilitation), or population (eg, psychiatry and Parkinson disease). We summarized these studies into four groups, accounting for overlap and based on the proposed technological solutions, to extract relevant data: (1) data from dedicated embedded or passive sensors, (2) data from dedicated wearable sensors, (3) data from dedicated or purposive technological solutions (eg, games or surveys), and (4) data derived from use of nondedicated technological solutions (eg, computer mouse movements).

Conclusions: Few publications dealt with home-based, real-life evaluations. Most technologies were far removed from everyday life experiences and were not mature enough for use under nonoptimal or uncontrolled conditions. Evidence available from embedded passive sensors represents the most relatively mature research area, suggesting that some of these solutions could be proposed to larger populations in the coming decade. The clinical and research communities would benefit from increasing attention to these technologies going forward.

XSL•FO
RenderX

(J Med Internet Res 2019;21(8):e12785) doi:10.2196/12785

KEYWORDS

technology; Alzheimer disease; cognition disorders; dementia; older adults; digital biomarkers; digital phenotyping; digital health

Introduction

Dementia and New Technologies

Interest in technologies as solutions for the challenges of dementia is high. Despite a plethora of technologies ranging from companion robots to fully functional smart home assessment environments, development and adoption has been slow or inconsistent [1]. In general, there is a wide spectrum of opinion about the utility of these technologies; these range from convinced *technophiles*, who believe that new technologies, particularly information and communication technologies (ICT) and artificial intelligence (AI), will revolutionize medicine, to skeptics or those not interested at all or who are even fearful of potential outcomes.

Among the most important areas that have challenged the progress of dementia care and treatment has been the assessment of those affected, those who are either at risk or presymptomatic, as well as those with clear, manifest symptoms [2,3]. At the root of this challenge is the need to identify symptoms and, most importantly, identify change in symptoms over time [3]. The latter is the essence of the diagnosis of dementia (ie, that there is a change from a prior state of normal cognition to a point where function is disturbed) [4,5]. This fact drives the basic approach that every clinician involved in mild cognitive impairment (MCI) and dementia assessment and care follows in their practice. It results in the need to assess, through careful history taking and neuropsychological assessment, whether a patient is experiencing change that reflects underlying neuropathology. It is vital to directing appropriate therapies [4,5].

Digital Biomarkers Development

To aid in the more precise assessment of patients, clinicians increasingly use biological and imaging biomarkers (eg, cerebrospinal fluid and positron emission tomography) to determine the patient's particular risk for developing Alzheimer disease (AD) and other dementias, as well as to differentiate the dementia type [6-8]. Although these biomarkers are an advance to the current diagnostic schemas widely promoted [4,5], these now *conventional* biomarkers face several limitations: they are expensive, difficult to access, invasive or inconvenient, and they do not accommodate a high-frequency measurement strategy. In addition, clinical and neuropsychological assessments, although remaining the core gold standard, are time-consuming, require self-report, and are subject to interassessor variability. More importantly, they are performed at discrete points in time in contexts that can affect their sensitivity (eg, patient comorbid conditions, medications, motivation, etc).

To improve this current clinical paradigm, digital biomarkers provide an alternative and rapidly developing approach. Digital biomarkers are defined here as objective, quantifiable, physiological, and behavioral data that are collected and measured by means of digital devices, such as embedded environmental sensors, portables, wearables, implantables, or digestibles. Digital biomarkers allow objective, ecologically valid, long-term follow-up with frequent or continuous assessment that can be minimally obtrusive or function in the background of everyday activity. Further, these frequent measures can capture intraindividual variability in performance that may be the earliest indicator of change [9-12] and thus detect subtle health transitions (eg, healthy to MCI). Even more potentially transformative, this approach may also allow us to discover novel and innovative digital indicators, such as gait-speed variability over time [11,13] or computer use metadata [10,14].

The adoption of these methodologies has been hampered by a number of factors [15,16]. The approach requires an interdisciplinary team, there is a multitude of sensors and devices that can be used, and there are no agreed-upon standards for these digital biomarkers. Most importantly, there is not a large evidence base indicating which standards are most effective. Much of the literature focuses on a narrow perspective using a single device or technology (eg, a wearable or a cognitive testing app). Most research has been limited to small numbers of participants assessed in a smart apartment or bioengineering laboratory. However, there is a growing movement in this research area to bring the technologies out of the laboratory and to the larger community in so-called "living lab" or "life laboratory" settings. The focus in these settings is to develop and confirm the utility of these technologies in the everyday environment of older adults' homes. In this review, we take stock of this research to answer the following questions: (1) What is the evidence for real-life, home-based use of technologies for early detection and follow-up of MCI or dementia? And based on this current evidence, (2) What transformation might clinicians expect in their everyday practices?

Methods

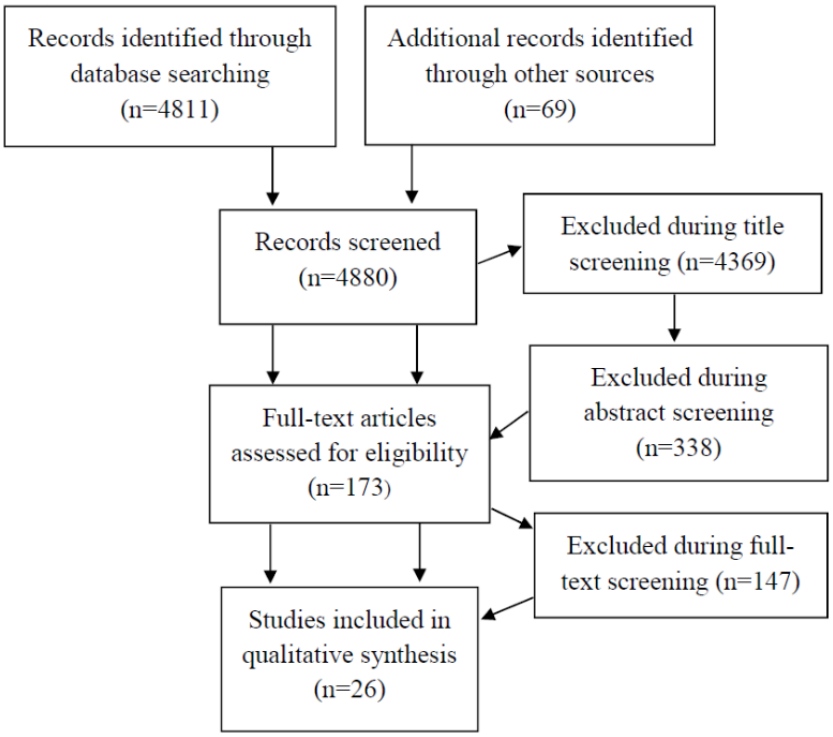
Information Sources and Study Selection

We followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement [17]. A systematic search was conducted of PubMed, including the Institute of Electrical and Electronics Engineers (IEEE); Cochrane; and Scopus databases. We searched for papers published from inception to July 14, 2018, for original research studies examining the implementation of ICT for MCI to mild AD follow-up and detection in real-life settings. We used the following Medical Subject Headings (MeSH) search terms and keywords: "clinical trial," "evaluation," "assessment," "Alzheimer*," "cognitive impairment," "MCI," "dementia," "cognition," "technology," "telehealth," "telemonitoring," "e-health," "internet," "sensors," "global positioning system," "phone," "smartphone*," "computer," "tablet," and "smart home*." We updated search terms after an initial review of our

XSL•FO
RenderX

search yield. We only considered English-language publications. Additional articles were obtained by scanning reference lists of literature collected on that basis. Two reviewers (AP and KW) conducted initial eligibility screening based on title and abstract, followed by assessment of full-text versions (see Figure 1 for more details). Any disagreements were resolved by consensus after a third opinion (JK).

Figure 1. Flow diagram of the study selection process.



Eligibility and Exclusion Criteria

Published studies that included the following were considered for inclusion: community-dwelling older adults (aged 65 years or more); healthy participants, if cognitive status was monitored, or those presenting cognitive decline (ie, from subjective cognitive complaints to early AD); a focus on home-based ICT evaluation for follow-up; and remote diagnosis of cognitive deterioration. Studies were excluded if they did or did not do the following: did not include data-generated results; included only moderate-to-severe AD; focused on caregiver support (eg, social support); focused on behavioral and psychological symptoms of dementia management (eg, global positioning system [GPS]-based wandering tracking); or took place in a controlled area (eg, smart, simulated, apartment laboratory or single-test home). Computerized cognitive tests, which mostly involve online evaluation at discrete points in time without longitudinal continuous assessment, have already been reviewed elsewhere [18,19].

Search Results

The initial search yielded a total of 4811 records. Articles were screened based on titles and abstracts, of which 173 full-text versions were assessed for inclusion. A total of 26 studies were finally eligible for inclusion in the review. There was an initial disagreement on eligibility between the two reviewers (AP and KW) concerning only 2 studies, which were finally excluded after consensus between three reviewers (JK, AP, and KW) because they evaluated sporadic, computerized, online testing rather than longitudinal follow-up. The most common reasons for exclusion were as follows: inappropriate setting (eg, hospital setting), intervention (eg, drugs and rehabilitation), or population (eg, psychiatry and Parkinson disease). Because of the great heterogeneity across selected studies in this developing research field, we did not perform a meta-analysis.

Results

An initial sample of 4811 English-language papers were retrieved from three electronic databases. After screening and review, 26 studies were eligible for inclusion in the review (see Tables 1-5) [10-14,20-40]. These 26 studies were observational studies taking place at home with community-dwelling older people, which is in line with the scope of the review. Sample size ranged from 12 to 279 participants. Mean age ranged from 64 to 89 years and percentage of female participants ranged from 49% to 92%. A total of 10 studies were considered comparative studies. Cognitive status was measured, with various methodological quality. There was a wide range of study duration, from 3 days to 3.6 years of follow-up.

We summarized and classified these 26 studies into four groups, although there was overlap, based on the proposed technological solutions to extract relevant data: (1) data from dedicated embedded or passive sensors, (2) data from dedicated wearable sensors, (3) data from dedicated or purposive technological solutions (eg, games or surveys), and (4) data derived from use of nondedicated technological solutions (eg, computer mouse movements). A fifth group includes solutions that fall into more than one category.

XSL•FO
RenderX

Table 1. Summary of 9 studies that included data from dedicated embedded or passive sensors in homes and cars (Group 1).

First author (year), country	Technology description	Study description: design; number and type of subjects (number living alone[a], if relevant) and setting; duration	Cognitive status and number of participants; age; number of male and/or female participants	Main results
Hayes (2008), United States [13]	Infrared motion sensors and magnetic contact door sensors	Comparative observational study; 14 elderly living alone in the community; 6-month follow-up (mean 315 days, SD 82)	Healthy group (n=7, CDR[b]=0, MMSE[c] ≥24), MCI[d] group (n=7, CDR=0.5, MMSE ≥24); mean age 89.3 years; 5 males, 9 females	Walking speed and activity of MCI group was more variable than that of the cognitively healthy controls.
Suzuki (2010), Japan [20]	Passive infrared sensors to record in-house movements	Observational study; 50 elderly living alone in the community; 1-year follow-up	MMSE ≥24; mean age 80.9 years; participant gender NC[e]	Association between lower numbers of outings with decrease of indoor movements and cognition declines.
Kaye (2012), United States [21]	Unobtrusively measures every instance of walking past a line of four passive infrared motion sensors fixed sequentially on the ceiling	Observational study; 76 persons living alone and independently; 4-week period	Mean MMSE=28.3; mean age 85.9 years; 86% women	Faster speeds were correlated with better cognitive test scores.
Dodge (2012), United States [11]	Passive infrared sensors fixed in series on the ceiling of the homes	Observational longitudinal study; 93 elderly living alone at home independently; mean follow-up of 2.6 years (SD 1.0)	54 cognitively intact, 8 with aMCI[f], 31 with naMCI[g]; mean age 84.9, 84.5, and 83.8 years, respectively; 88%, 84%, and 91% women, respectively	Daily walking speeds and their variability are associated with naMCI; naMCI presented a slowing of walking speed over 3 years. The highest and lowest variability were also found to be predominantly associated with naMCI.
Hayes (2014), United States [22]	Infrared motion sensors and magnetic contact door sensors	Comparative, observational, cross-sectional study; 45 elderly living independently and alone; 26 weeks	16 MCI, 29 cognitively intact; mean age 87 years; 89% female	aMCI volunteers had less disturbed sleep than both naMCI and cognitively intact volunteers, as measured by movement in bed, wake after sleep onset, and times up at night.
Petersen (2015), United States [23]	Total out-of-home daily time in hours assessed unobtrusively using an in-home activity sensor platform (eg, infrared sensors in each room and contact sensors on the doors to the home)	Observational study; 85 independent older adults who lived alone; 1 year	75 (CDR=0), 10 (CDR=0.5); mean age 86.4 years; 87% female	More hours spent outside the home was associated with better cognitive function.
Dawadi (2016), United States [24]	Smart homes: combination motion and light sensors on the ceilings and combination door and temperature sensors on cabinets and doors	Observational study; 18 community-dwelling seniors living alone; 2 years	7 cognitively healthy, 6 lowered performance, or cognitive difficulties (1 dementia, 4 MCI) (MMSE NC); age 84.7 years; 5 females, 13 males	Statistically significant correlation between sensor-based daily activity behaviors and clinician-provided cognitive assessment scores.
Urwyler (2017), Switzerland [25]	In-home, wireless, unobtrusive sensors network to detect activities of daily living	Comparative observational study; 20 participants living alone; 20 consecutive days	10 dementia, 10 healthy controls (MMSE=29.1 vs 23.0); age 76.7 vs 73.9 years; 70% female in both groups	Activity differed significantly between the healthy and diseased participants.
Seelye (2017), >United States [26]	Continuous routine driving-monitoring using an unobtrusive driving sensor: passive sensing device plugged into participants' vehicles data port	Observational study; 28 older adults living at home: 19 of 28 (68%) lived alone; average of 206 days	21 intact cognition, 7 MCI (average MMSE=28.6); mean age 82.0 years; 62% female	MCI participants drove fewer miles and spent less time on the highway per day than cognitively intact participants. MCI drivers showed less day-to-day fluctuations in their driving habits.

[a]The number of participants living alone is specified when the information is relevant; for example, for ambient sensors but not for wearables devices.

[b]CDR: Clinical Dementia Rating.

[c]MMSE: Mini Mental State Examination.

[d]MCI: mild cognitive impairment.

[e]NC: not communicated.

XSL•FO
RenderX

[f]aMCI: amnestic MCI.

[g]naMCI: nonamnestic MCI.

Table 2. Summary of 6 studies that included data from dedicated wearable sensors: accelerometers and GPS[a]-based solutions (Group 2).

First author (year), country	Technology description	Study description: design; number and type of subjects (number living alone[b], if relevant) and setting; duration	Cognitive status and number of participants; age; number of male and/or female participants	Main results
Westerberg (2010), United States [27]	Sleep monitoring with a wrist-worn activity sensor device	Comparative observational study; 20 volunteers; 2 weeks	10 aMCI[c] patients (MMSE[d]=27.8), 10 controls (MMSE=29.3); mean age 71.1 and 72.5 years, respectively; 8 and 7 females, respectively	Actigraphy parameters failed to reveal significant differences between groups.
Shoval (2011), Israel [28]	Tracking using a location kit: a GPS with radio frequency identification	Observational study; 41 community-dwelling participants; 28 days	13 healthy, 21 MCI[e], 7 mild dementia (MMSE and CDR[f] NC[g]); mean age 72.9, 78.3, and 81.9 years, respectively; 54% female	The spatial range of the mobility of elderly people with cognitive impairment is severely restricted, with most out-of-home time spent in close proximity.
Tung (2014), Canada [29]	GPS-enabled mobile phone	Observational comparative study; 52 older adults; 3 days	19 mild-to-moderate AD[h] (MMSE=23.1), 33 controls (MMSE NC); mean age 70.7 and 73.7 years, respectively; 40% and 64% female, respectively	GPS-derived area, perimeter, and mean distance from home were significantly smaller in the AD group compared to controls.
Wettstein (2015), Germany and Israel [30]	Mobility data: questionnaires and GPS receiver with a global system for mobile communications modem and a monitoring unit in the home	Observational comparative study; 257 older adults; 4 weeks	35 mild AD (mean MMSE=24.1), 76 MCI (mean MMSE=27.0), 146 healthy persons (mean MMSE=28.6); age 74.1, 72.9, and 72.5 years, respectively; 49% female	Questionnaire-based cognitively demanding activities showed a significant difference between MCI and cognitively healthy participants, and a significant difference between AD and cognitively healthy participants.
Takemoto (2015), United States [31]	GPS and accelerometer	Observational study; 279 older adults; 6 days	MMSE NC; mean age 83 years; 71% female	Number, distance, and minutes of pedestrian trips, as well as vehicle trips were not associated with cognitive functioning.
Mancini (2016), United States [32]	Quality and quantity of turning during normal daily activities by wearing three inertial sensors (one on their belt and two on shoes) during the day	Observational study; 35 elderly adults: 16 nonfallers, 12 one-time fallers, and 7 recurrent fallers; 7 days	Nonfallers (MMSE=28.3), one-time fallers (MMSE=28.9), recurrent fallers (MMSE=28.0); age 83.9, 86.0, and 88.4 years, respectively; 66% female	Visuospatial and memory function scores were associated with quality of turning.

[a]GPS: global positioning system.

[b]The number of participants living alone is specified when the information is relevant; for example, for ambient sensors but not for wearables devices.

[c]aMCI: amnestic MCI.

[d]MMSE: Mini Mental State Examination.

[e]MCI: mild cognitive impairment.

[f]CDR: Clinical Dementia Rating.

[g]NC: not communicated.

[h]AD: Alzheimer disease.

XSL•FO
RenderX

Table 3. Summary of 6 studies that included data from dedicated or purposive ICT[a]-monitoring solutions, such as phone-based automated interviews, Nintendo Wii, and virtual reality (Group 3).

First author (year), country	Technology description	Study description: design; number and type of subjects (number living alone[b], if relevant) and setting; duration	Cognitive status and number of participants; age; number of male and/or female participants	Main results
Mundt (2007), United States [33]	Use of IVR[c] technology (ie, pressing keys) to administer simple cognitive evaluations by phone during a 20-minute, computer-automated telephone call	Observational comparative study; 107 community-dwelling participants; 24 weeks: IVR administered at home at weeks 4, 12, and 20	36 cognitively normal, (MMSE[d]=28.1), 37 MCI[e] (MMSE=25.6), 34 mild dementia (MMSE=20.0); mean age 76.7 years; 42% female	The automated administration of IVR simple cognitive tests via phone calls reliably and validly discriminated cognitive functioning among normal, MCI, and mild dementia.
Allard (2014), France [34]	Monitoring of behavior, semantic memory performance, and daily life experiences using a personal digital assistant five times a day	Observational study; 60 older adults; 7 days	60 healthy participants (mean MMSE=27.0); mean age 75.1 years; 45% female	Magnetic resonance imagery markers were significantly associated with mobile assessments of semantic memory performance.
Brown (2016), United Kingdom [35]	Touch screen system to assess multiple domains of health and behavior; cognitive tasks scheduled once per day	Observational study; 40 community-dwelling adults; three periods of approximately 7 days	40 healthy participants (mean MMSE=28.63); mean age 72 years; 24 females, 16 males	Convergent validity with, and similar levels of, reliability to the standard cognitive battery.
Seelye (2016), United States [36]	Completion of a short 12-item weekly online questionnaire of health and life events, administered on desktop computers	Observational study; 83 independent, community-dwelling older adults; 1 year	59 healthy (MMSE=28.8), 24 MCI (MMSE=27.4); mean age 86.2 and 87.9 years, respectively; 88% and 75% female, respectively	Online questionnaire performance significantly correlated to cognitive test. MCI participants submitted their questionnaires progressively later in the day and they needed greater assistance from staff as compared with intact participants.
Zygouris (2017), Greece [37]	Tablet personal computer with software enabling the self-administration of a cognitive assessment through virtual reality	Comparative, two-arm, observational study; 12 elderly living at home; 1-month follow-up	6 healthy and 6 MCI; mean 64 years; 3 males, 9 females	Performances to complete the given exercise differed significantly between healthy and MCI groups, yielding a correct classification rate of 92% for MCI detection.
Leach (2018), United States [38]	A Nintendo Wii balance board used to quantify postural sway twice daily, under a single-task condition and under a dual-task condition, using a daily word-search task administered via a Nook tablet	Observational study; 20 healthy community-dwelling elderly; 30 days	Mean MMSE=28.6; mean age 87.0 years; 65% females	Linear relationships were observed between the day-to-day variability in postural sway and cognitive status.

[a]ICT: information and communication technologies.

[b]The number of participants living alone is specified when the information is relevant; for example, for ambient sensors but not for wearables devices.

[c]IVR: interactive voice response.

[d]MMSE: Mini Mental State Examination.

[e]MCI: mild cognitive impairment.

XSL•FO
RenderX

Table 4. Summary of 4 studies that included data derived from nondedicated ICT[a] solutions use, for example, secondary analysis of everyday computer use and pill box use (Group 4).

First author (year), country	Technology description	Study description: design; number and type of subjects (number living alone[b], if relevant) and setting; duration	Cognitive status and number of participants; age; number of male and female participants	Main results
Hayes (2009), United States [39]	Adherence to a twice-daily vitamin C regimen measured using an electronic 7-day pill box	Observational cross-sectional study; 38 participants living independently in the community; 5 weeks	A high cognitive function group (MMSE[c]=28.8) and a low cognitive function group (MMSE=28.0); mean age 82.8 years; 68% female	The low cognitive function group was significantly less adherent than the healthy elders. Very mild cognitive impairment had a detrimental and significant impact on medication adherence.
Kaye (2014), United States [10]	Remotely monitored computer use	Comparative observational study; 113 elderly living independently and alone or who were the only computer user; mean 36-month follow-up	38 MCI[d] and 75 cognitively intact; mean age 85 years; 92% female	Decrease in number of days with use, mean daily usage, and an increase in day-to-day use variability in MCI subjects.
Seelye (2015), United States [40]	Mouse pointer movement variables were computed during routine home computer use using algorithms that identified and characterized mouse movements within each computer use session	Observational comparative study; 62 older adults living at home alone or who were the only computer user in the household; 1 week	42 healthy (MMSE=28.8), 20 MCI (MMSE=27.3); mean age 87.9 and 87.5 years, respectively; 88% and 80% female, respectively	MCI was associated with making significantly fewer mouse moves and making mouse movements that were more variable, less efficient, and with longer pauses. Mouse movement significantly associated with several cognitive domains.
Austin (2017), United States [12]	Computer monitoring software used to track the terms people entered while conducting Internet searches as a measure of language and cognition	Observational study; 42 community-dwelling older adults living alone; 6 months	Cognitively intact, with the exception of 1 participant (CDR[e] score ≥0.5, suggesting MCI); average age 81.1 years; 83% female	Individuals with higher cognitive function used more unique terms per search and employed less-common terms in their searches.

[a]ICT: information and communication technologies.

[b]The number of participants living alone is specified when the information is relevant; for example, for ambient sensors but not for wearables devices.

[c]MMSE: Mini Mental State Examination.

[d]MCI: mild cognitive impairment.

[e]CDR: Clinical Dementia Rating.

Table 5. Summary of a solution that falls into more than one category (Group 5).

First author (year), country	Technology description	Study description: design; number and type of subjects (number living alone[a], if relevant) and setting; duration	Cognitive status and number of participants; age; number of male and/or female participants	Main results
Seelye (2018), United States [14]	Weekly online survey metadata metrics based on survey engagement patterns	Observational study; 110 healthy older adults; 3.6-year follow-up	110 with intact cognition at the beginning and 29 transitioned to MCI[b] during study follow-up (MMSE[c]=28.8); mean age 84.8 years; 77% female	At baseline, incident MCI participants completed surveys later in the day than cognitively intact participants. Longitudinally, incident MCI participants showed an increase in survey completion time compared with cognitively intact participants.

[a]The number of participants living alone is specified when the information is relevant; for example, for ambient sensors but not for wearables devices.

[b]MCI: mild cognitive impairment.

[c]MMSE: Mini Mental State Examination.

In the first group (ie, embedded dedicated sensors), we can principally cite *smart home* technologies [11,13,20-25] and *smart car* technologies [26]. Studies in the second group (ie, data from wearable dedicated technologies) mainly rely on accelerometers and GPS solutions [27-32]. The third group (ie, dedicated ICT solutions) imply ICT-supported monitoring solutions [33-38]. These mainly employ online surveys or touch-screen tests [34-36] as well as computer-automated telephone calls [33] or a Nintendo Wii-dedicated game [38]. The fourth group (ie, monitoring of nondedicated ICT solutions use) consists of secondary analyses of commonly used technologies, including everyday computer use [10,12,40] and

XSL•FO
RenderX

pill box use [39]. The fifth group included one study that dealt with monitoring of dedicated ICT solutions using survey metadata metrics analysis [14].

Discussion

Evidence for Real-Life, Home-Based Use of Technologies

The first aim of this paper was to provide an overview of technologies for real-life early detection and follow-up of cognitive function to practicing clinicians involved in management of AD and related disorders. A total of 26 studies were identified, with a variety of technologies and a wide range of study duration and sample size. The first key observation is that compared to the overall number of publications in the field, few papers dealt with home-based, real-world evaluations. Most excluded articles focused on technology functionality, tests of technical aspects in laboratory settings, and focused evaluations in single or a few *test-bed* homes or hospital settings. Most technologies were far removed from everyday life experiences or widely disseminated implementation in the community. Among the included study types, the first (ie, embedded dedicated sensors), the third (ie, dedicated ICT solutions), and the fourth groups (ie, monitoring of nondedicated ICT solutions use) have the common advantage of unobtrusiveness. They rely on everyday life observation without any, or very minimal, participant involvement. The fact that these are among the longest studies in this review, up to several years, likely speaks to the passive nature of the technologies. Several studies in the third group are partly similar to studies of computerized online tests, with the difference that a longitudinal follow-up and a self-administration of nonconventional cognitive tests at home is evaluated [34-36]. In contrast, wearable technology (eg, GPS and wrist-worn device) studies are generally short-term studies. This may be explained by the difficulties in implementing such solutions in real-world settings, as they demand more extensive end-user participation (eg, remembering to wear or charge the device) in this older adult population with various levels of cognitive impairment and technical capacity.

The exclusive use of ambient passive sensors in homes and cars does not guarantee good acceptability to end users. As an example, people may have an intrusive perception of a 3D camera or microphones. However, authors do not report any acceptability issues for experiments involving infrared, temperature, humidity, luminescence, and magnetic door contact sensors or driving sensors [11,13,20-26]. Studies on monitoring the use of personal computers have yielded comparable results [10,12,40]. However, this has only been validated in a select population thus far, as discussed in the Limitations section below, and acceptability outcomes are not always reported.

The completion rate of repeated remote assessments using ICT monitoring solutions was high, generally above 80% [33-35]. Over a longer period of time, Seelye et al [14,36] reported that online weekly health forms were submitted on schedule 75% of the time. Using a Nintendo Wii balance board, Leach and al [38] found an average of 3 days of missing posture and cognitive data over the 30-day observation period for each subject.

For the wearable devices, most studies do not report acceptability issues, such as refusal rate at inclusion and adherence data during follow-up. Mancini et al [32] reported that all 35 participants complied with the protocol (ie, wearing inertial sensors) for 7 consecutive days, while Shoval et al [28] found that participants actively wore the device for 88% of the days. Finally, Wettstein et al [30] reported that the major reasons given to refuse participation in their study were distrust and fear of being observed. In the case of wearables, samples are smaller and/or durations are shorter, which limits comparisons.

What Transformation Might Clinicians Expect?

With advances in monitoring technologies, we can anticipate what clinicians might expect in their practice in the coming years. A major limitation of all these studies is the selection of volunteers that are relatively homogeneous (ie, white, educated, receptive to technology, and living in urban areas). If, in the future, there is wider use of ICT and digital biomarkers in clinical practice, a perceived advantage of technologies over *traditional* biomarkers, we need to develop ICT in a way that ensures its acceptance and usability under nonoptimal conditions for long periods of time. Most of the solutions presented in these studies are not mature enough for this goal.

The evidence base from dedicated embedded sensors (Group 1) seems to be the most mature research area in this field. There is evidence that these sensor-based technologies are sensitive to detecting cognitive and functional change. Nevertheless, this is still an area of active research, not yet translated into wide clinical adoption. Several advances are needed. Most studies dealing with embedded sensors are limited to participants living alone. It remains technically complex to disambiguate activity in multi-person homes in a real environment. New sensing approaches and data fusion algorithms are in development, but this remains work in progress. Dedicated home systems require installation, which can be a barrier to wider dissemination. Several approaches to this need have been developed, including prepared field kits (ie, “sensors in a box”) for ready deployment [41] and online video installation guides [15]. With these and other advances, one can reasonably anticipate that some of these solutions could be proposed in everyday care to a large population in the coming decade. The Collaborative Aging Research Using Technology (CART) Initiative [42] in the United States is a multi-site, nationwide project that uses multiple embedded sensing technology and diverse data to facilitate research in the field of independence and health of older adults from diverse communities. Funded by the National Institutes of Health and the Department of Veterans Affairs, the CART Initiative has been designed to enable expansion of the home-based sensing platform to 10,000 homes across the United States in several years.

Dedicated wearable sensors (Group 2) have several barriers to overcome before readiness for large community-wide implementation. They have been perceived as obtrusive and stigmatizing. Although their adherence rate may be lower, their great advantages are that their use is not limited to people living alone and they can provide both indoor and outdoor information. We believe, however, that significant progress can be made by improving both their accuracy [43] and their development and

XSL•FO
RenderX

evaluation processes. Technical-centered development of these devices in laboratory settings disregards the barriers to successful implementation in real-life situations and, more particularly, acceptability for unselected (eg, low computer literacy and cognitively impaired) populations. This real-life implementation depends on technical improvements for deployment in suboptimal contexts (eg, uncertain Internet coverage) and on the end user's acceptability. In recent years, several research teams have tried to design creative multiphase studies, applying iterative modifications of a proposed digital technology-associated solution or using participatory design approaches [44,45]. Living laboratory, iterative design approaches, here defined as a "research method which brings together end users, developers, and health professionals in a cocreation and evaluation process," introduce an intermediate phase closer to *real life* before deployment. Participatory design methodology also emphasizes the involvement of users throughout the innovation process. Widespread diffusion of these concepts in medical research has not yet been achieved. However, health-related functions of dedicated wearable devices have become commonly offered in routine use devices (eg, GPS and accelerometers) and hopefully will be integrated in a more efficient way for health use in the near future [46,47].

Dedicated ICT solutions (Group 3), such as embedded assessment algorithms within home-based cognitive computer games, remains a promising area [48]. Surprisingly, few trials evaluated such solutions in a home environment. Trials involving mobile phone apps are also promising, notably in the psychiatry field [49,50], but they are not at an advanced stage in the dementia research area of interest, other than assessments taking place in a dedicated place with the presence of an operator outside of the home [51-53]. Finally, a promising field is in the monitoring of nondedicated ICT solutions (Group 4), as they have the advantage of unobtrusiveness in common with the embedded dedicated sensors (Group 1), and they do not require the installation of dedicated devices. This area is limited by the willingness of this population to adopt ICT solutions in their everyday lives. Nevertheless, generations follow one another, and future older adults may be more likely to be interested in new technologies. In that case, assessing those solutions in an ideal target population may be seen as less of a limitation if a large dissemination is expected within the next 5-10 years.

The wider adoption for all these technologies will require particular attention to specific use cases, ranging from detecting early cognitive decline to assessing change in function during treatment; ease of deployment; data provenance and analysis; and creating not just evidence for efficacy but evidence for effectiveness. The first drivers of this transformation will likely be the research community and the growing adoption of community-based studies. Clinical trials in particular are an area where this approach is very promising. Often, proof of efficacy in trials leads to practice adoption.

Limitations

As already stated by Pillai et al [54], keywords describing ICT and AI vary and they are not always specified in the papers, as this research field is not yet mature. Nevertheless, it is unlikely that we have missed many relevant studies, as we chose the broadest possible key search terms and followed up with an extensive hand search of full-text references and key terms. Promising results from studies related to screening and assessment rather than follow-up [55,56], or conducted in controlled settings, are not presented in this paper. All the studies reviewed here were conducted in real-life homes of older adults. Several issues were not highlighted in this paper, including health inequalities [57], ethical issues, data security, information overload for clinicians, and business models of technology implementations, among others. These key issues also need to be addressed early during the evaluation and development process of health technology research, before larger dissemination can occur [58].

Conclusions and Future Research Directions

The studies included in this review cover a diversity of designs and approaches representing many new avenues of research. There is no conclusive evidence at this stage on the superiority of one or many digital biomarker assessments over others. This is a new area of research. Even for studies based on the same cohorts and, therefore, on comparable populations and locations, the technologies used (eg, driving or computer tracking), sample sizes engaged, and statistical methods differ. For similar outcomes of interest (eg, activity), there is a wide range of digital biomarkers: walking speed, overall activity, outdoor time, etc. It is difficult to know which biomarkers will be most relevant for broader future applications. Further, it is not yet clear which outcomes are best correlated with cognitive decline or, more generally, with mental health. Choices will probably initially focus more on the ease of implementation of a technical solution in a given environment than on direct comparisons as to their accuracy, a comparison that is difficult to make in practice.

Nevertheless, monitoring cognitive and functional domains using ICT devices will grow rapidly and will likely involve AI [59,60] and innovative biomarkers derived from such methods as automated speech and language analysis, motor performance assessments, computer use abilities, and online questionnaire responses and their metadata. These advances will facilitate the transition to proactive, personalized, and participatory medicine. In achieving this goal, the gap between real-life clinical practice and clinical research will be narrowed with clinical trials reflecting patients' typical activities and outcomes [3,41]. Integration of heterogeneous data (eg, environmental data and multiple biomarkers) will improve our understanding and management of cognitive decline; accordingly, some of these solutions may become adopted into everyday care among the wider population in the coming decade.

XSL•FO
RenderX

Acknowledgments

This review was undertaken as part of the research project by AP at the Oregon Center for Aging and Technology (ORCATECH), which was funded by the National Institutes of Health and the Department of Veterans Affairs (grant numbers: R01AG024059, U2CAG054397, P30AG024978, and P30AG008017).

Conflicts of Interest

None declared.

References

1. LaMonica HM, English A, Hickie IB, Ip J, Ireland C, West S, et al. Examining Internet and eHealth practices and preferences: Survey study of Australian older adults with subjective memory complaints, mild cognitive impairment, or dementia. J Med Internet Res 2017 Oct 25;19(10):e358 [FREE Full text] [doi: 10.2196/jmir.7981] [Medline: 29070481]
2. Morley JE, Morris JC, Berg-Weger M, Borson S, Carpenter BD, Del Campo N, et al. Brain health: The importance of recognizing cognitive impairment: An IAGG consensus conference. J Am Med Dir Assoc 2015 Sep 01;16(9):731-739 [FREE Full text] [doi: 10.1016/j.jamda.2015.06.017] [Medline: 26315321]
3. Borson S, Frank L, Bayley PJ, Boustani M, Dean M, Lin P, et al. Improving dementia care: The role of screening and detection of cognitive impairment. Alzheimers Dement 2013 Mar;9(2):151-159 [FREE Full text] [doi: 10.1016/j.jalz.2012.08.008] [Medline: 23375564]
4. McKhann GM, Knopman DS, Chertkow H, Hyman BT, Jack CR, Kawas CH, et al. The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement 2011 May;7(3):263-269 [FREE Full text] [doi: 10.1016/j.jalz.2011.03.005] [Medline: 21514250]
5. Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, et al. The diagnosis of mild cognitive impairment due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. Alzheimers Dement 2011 May;7(3):270-279 [FREE Full text] [doi: 10.1016/j.jalz.2011.03.008] [Medline: 21514249]
6. Jack CR, Bennett DA, Blennow K, Carrillo MC, Feldman HH, Frisoni GB, et al. A/T/N: An unbiased descriptive classification scheme for Alzheimer disease biomarkers. Neurology 2016 Aug 02;87(5):539-547 [FREE Full text] [doi: 10.1212/WNL.0000000000002923] [Medline: 27371494]
7. Slavin MJ, Brodaty H, Sachdev PS. Challenges of diagnosing dementia in the oldest old population. J Gerontol A Biol Sci Med Sci 2013 Sep;68(9):1103-1111. [doi: 10.1093/gerona/glt051] [Medline: 23685769]
8. Blennow K, Zetterberg H. The application of cerebrospinal fluid biomarkers in early diagnosis of Alzheimer disease. Med Clin North Am 2013 May;97(3):369-376. [doi: 10.1016/j.mcna.2012.12.012] [Medline: 23642576]
9. Kaye J. Home-based technologies: A new paradigm for conducting dementia prevention trials. Alzheimers Dement 2008 Jan;4(1 Suppl 1):S60-S66. [doi: 10.1016/j.jalz.2007.10.003] [Medline: 18632003]
10. Kaye J, Mattek N, Dodge HH, Campbell I, Hayes T, Austin D, et al. Unobtrusive measurement of daily computer use to detect mild cognitive impairment. Alzheimers Dement 2014 Jan;10(1):10-17 [FREE Full text] [doi: 10.1016/j.jalz.2013.01.011] [Medline: 23688576]
11. Dodge HH, Mattek NC, Austin D, Hayes TL, Kaye JA. In-home walking speeds and variability trajectories associated with mild cognitive impairment. Neurology 2012 Jun 12;78(24):1946-1952 [FREE Full text] [doi: 10.1212/WNL.0b013e318259e1de] [Medline: 22689734]
12. Austin J, Hollingshead K, Kaye J. Internet searches and their relationship to cognitive function in older adults: Cross-sectional analysis. J Med Internet Res 2017 Sep 06;19(9):e307 [FREE Full text] [doi: 10.2196/jmir.7671] [Medline: 28877864]
13. Hayes TL, Abendroth F, Adami A, Pavel M, Zitzelberger TA, Kaye JA. Unobtrusive assessment of activity patterns associated with mild cognitive impairment. Alzheimers Dement 2008 Nov;4(6):395-405 [FREE Full text] [doi: 10.1016/j.jalz.2008.07.004] [Medline: 19012864]
14. Seelye A, Mattek N, Sharma N, Riley T, Austin J, Wild K, et al. Weekly observations of online survey metadata obtained through home computer use allow for detection of changes in everyday cognition before transition to mild cognitive impairment. Alzheimers Dement 2018 Feb;14(2):187-194 [FREE Full text] [doi: 10.1016/j.jalz.2017.07.756] [Medline: 29107052]
15. Kaye J, Reynolds C, Bowman M, Sharma N, Riley T, Golonka O, et al. Methodology for establishing a community-wide life laboratory for capturing unobtrusive and continuous remote activity and health data. J Vis Exp 2018 Jul 27(137):1-10 [FREE Full text] [doi: 10.3791/56942] [Medline: 30102277]
16. Arnerić SP, Cedarbaum JM, Khozin S, Papapetropoulos S, Hill DL, Ropacki M, et al. Biometric monitoring devices for assessing end points in clinical trials: Developing an ecosystem. Nat Rev Drug Discov 2017 Oct;16(10):736. [doi: 10.1038/nrd.2017.153] [Medline: 28935908]

XSL•FO
RenderX

17. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. PLoS Med 2009 Jul 21;6(7):e1000097 [FREE Full text] [doi: 10.1371/journal.pmed.1000097] [Medline: 19621072]
18. Wild K, Howieson D, Webbe F, Seelye A, Kaye J. Status of computerized cognitive testing in aging: A systematic review. Alzheimers Dement 2008 Nov;4(6):428-437 [FREE Full text] [doi: 10.1016/j.jalz.2008.07.003] [Medline: 19012868]
19. Zygouris S, Tsolaki M. Computerized cognitive testing for older adults: A review. Am J Alzheimers Dis Other Demen 2015 Feb;30(1):13-28. [doi: 10.1177/1533317514522852] [Medline: 24526761]
20. Suzuki T, Murase S. Influence of outdoor activity and indoor activity on cognition decline: Use of an infrared sensor to measure activity. Telemed J E Health 2010;16(6):686-690. [doi: 10.1089/tmj.2009.0175] [Medline: 20575611]
21. Kaye J, Mattek N, Dodge H, Buracchio T, Austin D, Hagler S, et al. One walk a year to 1000 within a year: Continuous in-home unobtrusive gait assessment of older adults. Gait Posture 2012 Feb;35(2):197-202 [FREE Full text] [doi: 10.1016/j.gaitpost.2011.09.006] [Medline: 22047773]
22. Hayes TL, Riley T, Mattek N, Pavel M, Kaye JA. Sleep habits in mild cognitive impairment. Alzheimer Dis Assoc Disord 2014;28(2):145-150 [FREE Full text] [doi: 10.1097/WAD.0000000000000010] [Medline: 24145694]
23. Petersen J, Austin D, Mattek N, Kaye J. Time out-of-home and cognitive, physical, and emotional wellbeing of older adults: A longitudinal mixed effects model. PLoS One 2015;10(10):e0139643 [FREE Full text] [doi: 10.1371/journal.pone.0139643] [Medline: 26437228]
24. Dawadi PN, Cook DJ, Schmitter-Edgecombe M. Automated cognitive health assessment from smart home-based behavior data. IEEE J Biomed Health Inform 2016 Jul;20(4):1188-1194 [FREE Full text] [doi: 10.1109/JBHI.2015.2445754] [Medline: 26292348]
25. Urwyler P, Stucki R, Rampa L, Müri R, Mosimann UP, Nef T. Cognitive impairment categorized in community-dwelling older adults with and without dementia using in-home sensors that recognise activities of daily living. Sci Rep 2017 Feb 08;7:42084 [FREE Full text] [doi: 10.1038/srep42084] [Medline: 28176828]
26. Seelye A, Mattek N, Sharma N, Witter P, Brenner A, Wild K, et al. Passive assessment of routine driving with unobtrusive sensors: A new approach for identifying and monitoring functional level in normal aging and mild cognitive impairment. J Alzheimers Dis 2017;59(4):1427-1437 [FREE Full text] [doi: 10.3233/JAD-170116] [Medline: 28731434]
27. Westerberg CE, Lundgren EM, Florczak SM, Mesulam M, Weintraub S, Zee PC, et al. Sleep influences the severity of memory disruption in amnestic mild cognitive impairment: Results from sleep self-assessment and continuous activity monitoring. Alzheimer Dis Assoc Disord 2010;24(4):325-333 [FREE Full text] [doi: 10.1097/WAD.0b013e3181e30846] [Medline: 20592579]
28. Shoval N, Wahl H, Auslander G, Isaacson M, Oswald F, Edry T, et al. Use of the global positioning system to measure the out-of-home mobility of older adults with differing cognitive functioning. Ageing Soc 2011 May 16;31(05):849-869. [doi: 10.1017/S0144686X10001455]
29. Tung JY, Rose RV, Gammada E, Lam I, Roy EA, Black SE, et al. Measuring life space in older adults with mild-to-moderate Alzheimer's disease using mobile phone GPS. Gerontology 2014;60(2):154-162. [doi: 10.1159/000355669] [Medline: 24356464]
30. Wettstein M, Wahl H, Shoval N, Oswald F, Voss E, Seidl U, et al. Out-of-home behavior and cognitive impairment in older adults: Findings of the SenTra Project. J Appl Gerontol 2015 Feb;34(1):3-25. [doi: 10.1177/0733464812459373] [Medline: 25548086]
31. Takemoto M, Carlson JA, Moran K, Godbole S, Crist K, Kerr J. Relationship between objectively measured transportation behaviors and health characteristics in older adults. Int J Environ Res Public Health 2015 Oct 30;12(11):13923-13937 [FREE Full text] [doi: 10.3390/ijerph121113923] [Medline: 26528999]
32. Mancini M, Schlueter H, El-Gohary M, Mattek N, Duncan C, Kaye J, et al. Continuous monitoring of turning mobility and its association to falls and cognitive function: A pilot study. J Gerontol A Biol Sci Med Sci 2016 Aug;71(8):1102-1108 [FREE Full text] [doi: 10.1093/gerona/glw019] [Medline: 26916339]
33. Mundt JC, Kinoshita LM, Hsu S, Yesavage JA, Greist JH. Telephonic Remote Evaluation of Neuropsychological Deficits (TREND): Longitudinal monitoring of elderly community-dwelling volunteers using touch-tone telephones. Alzheimer Dis Assoc Disord 2007;21(3):218-224. [doi: 10.1097/WAD.0b013e31811ff2c9] [Medline: 17804954]
34. Allard M, Husky M, Catheline G, Pelletier A, Dilharreguy B, Amieva H, et al. Mobile technologies in the early detection of cognitive decline. PLoS One 2014;9(12):e112197 [FREE Full text] [doi: 10.1371/journal.pone.0112197] [Medline: 25536290]
35. Brown LJE, Adlam T, Hwang F, Khadra H, Maclean LM, Rudd B, et al. Computer-based tools for assessing micro-longitudinal patterns of cognitive function in older adults. Age (Dordr) 2016 Aug;38(4):335-350 [FREE Full text] [doi: 10.1007/s11357-016-9934-x] [Medline: 27473748]
36. Seelye A, Mattek N, Howieson DB, Austin D, Wild K, Dodge HH, et al. Embedded online questionnaire measures are sensitive to identifying mild cognitive impairment. Alzheimer Dis Assoc Disord 2016;30(2):152-159 [FREE Full text] [doi: 10.1097/WAD.0000000000000100] [Medline: 26191967]

XSL•FO
RenderX

37. Zygouris S, Ntovas K, Giakoumis D, Votis K, Doumpoulakis S, Segkouli S, et al. A preliminary study on the feasibility of using a virtual reality cognitive training application for remote detection of mild cognitive impairment. J Alzheimers Dis 2017;56(2):619-627. [doi: 10.3233/JAD-160518] [Medline: 28035922]
38. Leach JM, Mancini M, Kaye JA, Hayes TL, Horak FB. Day-to-day variability of postural sway and its association with cognitive function in older adults: A pilot study. Front Aging Neurosci 2018;10:126 [FREE Full text] [doi: 10.3389/fnagi.2018.00126] [Medline: 29780319]
39. Hayes TL, Larimer N, Adami A, Kaye JA. Medication adherence in healthy elders: Small cognitive changes make a big difference. J Aging Health 2009 Jun;21(4):567-580 [FREE Full text] [doi: 10.1177/0898264309332836] [Medline: 19339680]
40. Seelye A, Hagler S, Mattek N, Howieson DB, Wild K, Dodge HH, et al. Computer mouse movement patterns: A potential marker of mild cognitive impairment. Alzheimers Dement (Amst) 2015 Dec 01;1(4):472-480 [FREE Full text] [doi: 10.1016/j.dadm.2015.09.006] [Medline: 26878035]
41. Hu Y, Tilke D, Adams T, Crandall AS, Cook DJ, Schmitter-Edgecombe M. Smart home in a box: Usability study for a large scale self-installation of smart home technologies. J Reliab Intell Environ 2016 Jul;2(2):93-106 [FREE Full text] [doi: 10.1007/s40860-016-0021-y] [Medline: 28936390]
42. Oregon Health & Science University. Collaborative Aging Research Using Technology (CART) URL: https://www.ohsu.edu/collaborative-aging-research-using-technology [accessed 2019-07-25]
43. Gomersall SR, Ng N, Burton NW, Pavey TG, Gilson ND, Brown WJ. Estimating physical activity and sedentary behavior in a free-living context: A pragmatic comparison of consumer-based activity trackers and ActiGraph accelerometry. J Med Internet Res 2016 Sep 07;18(9):e239 [FREE Full text] [doi: 10.2196/jmir.5531] [Medline: 27604226]
44. Swinkels ICS, Huygens MWJ, Schoenmakers TM, Oude Nijeweme-D'Hollosy W, van Velsen L, Vermeulen J, et al. Lessons learned from a living lab on the broad adoption of eHealth in primary health care. J Med Internet Res 2018 Mar 29;20(3):e83 [FREE Full text] [doi: 10.2196/jmir.9110] [Medline: 29599108]
45. DeSmet A, Thompson D, Baranowski T, Palmeira A, Verloigne M, De Bourdeaudhuij I. Is participatory design associated with the effectiveness of serious digital games for healthy lifestyle promotion? A meta-analysis. J Med Internet Res 2016 Apr 29;18(4):e94 [FREE Full text] [doi: 10.2196/jmir.4444] [Medline: 27129447]
46. Jongstra S, Wijsman LW, Cachucho R, Hoevenaar-Blom MP, Mooijaart SP, Richard E. Cognitive testing in people at increased risk of dementia using a smartphone app: The iVitality Proof-of-Principle Study. JMIR Mhealth Uhealth 2017 May 25;5(5):e68 [FREE Full text] [doi: 10.2196/mhealth.6939] [Medline: 28546139]
47. Goodspeed R, Yan X, Hardy J, Vydiswaran VGV, Berrocal VJ, Clarke P, et al. Comparing the data quality of global positioning system devices and mobile phones for assessing relationships between place, mobility, and health: Field study. JMIR Mhealth Uhealth 2018 Aug 13;6(8):e168 [FREE Full text] [doi: 10.2196/mhealth.9771] [Medline: 30104185]
48. Vallejo V, Wyss P, Rampa L, Mitache AV, Müri RM, Mosimann UP, et al. Evaluation of a novel serious game-based assessment tool for patients with Alzheimer's disease. PLoS One 2017;12(5):e0175999 [FREE Full text] [doi: 10.1371/journal.pone.0175999] [Medline: 28472049]
49. Zulueta J, Piscitello A, Rasic M, Easter R, Babu P, Langenecker SA, et al. Predicting mood disturbance severity with mobile phone keystroke metadata: A biaffect digital phenotyping study. J Med Internet Res 2018 Jul 20;20(7):e241 [FREE Full text] [doi: 10.2196/jmir.9775] [Medline: 30030209]
50. Place S, Blanch-Hartigan D, Rubin C, Gorrostieta C, Mead C, Kane J, et al. Behavioral indicators on a mobile sensing platform predict clinically validated psychiatric symptoms of mood and anxiety disorders. J Med Internet Res 2017 Mar 16;19(3):e75 [FREE Full text] [doi: 10.2196/jmir.6678] [Medline: 28302595]
51. Munro Cullum C, Hynan LS, Grosch M, Parikh M, Weiner MF. Teleneuropsychology: Evidence for video teleconference-based neuropsychological assessment. J Int Neuropsychol Soc 2014 Nov;20(10):1028-1033 [FREE Full text] [doi: 10.1017/S1355617714000873] [Medline: 25343269]
52. Brouillette RM, Foil H, Fontenot S, Correro A, Allen R, Martin CK, et al. Feasibility, reliability, and validity of a smartphone-based application for the assessment of cognitive function in the elderly. PLoS One 2013;8(6):e65925 [FREE Full text] [doi: 10.1371/journal.pone.0065925] [Medline: 23776570]
53. Konig A, Satt A, Sorin A, Hoory R, Derreumaux A, David R, et al. Use of speech analyses within a mobile application for the assessment of cognitive impairment in elderly people. Curr Alzheimer Res 2018;15(2):120-129. [doi: 10.2174/1567205014666170829111942] [Medline: 28847279]
54. Pillai JA, Bonner-Jackson A. Review of information and communication technology devices for monitoring functional and cognitive decline in Alzheimer's disease clinical trials. J Healthc Eng 2015;6(1):71-83 [FREE Full text] [doi: 10.1260/2040-2295.6.1.71] [Medline: 25708378]
55. Groppell S, Soto-Ruiz KM, Flores B, Dawkins W, Smith I, Eagleman DM, et al. A rapid, mobile neurocognitive screening test to aid in identifying cognitive impairment and dementia (BrainCheck): Cohort study. JMIR Aging 2019 Mar 21;2(1):e12615. [doi: 10.2196/12615]
56. Bott N, Madero EN, Glenn J, Lange A, Anderson J, Newton D, et al. Device-embedded cameras for eye tracking-based cognitive assessment: Validation with paper-pencil and computerized cognitive composites. J Med Internet Res 2018 Jul 24;20(7):e11143 [FREE Full text] [doi: 10.2196/11143] [Medline: 30042093]

XSL•FO
RenderX

57. Allen LN, Christie GP. The emergence of personalized health technology. J Med Internet Res 2016 May 10;18(5):e99 [FREE Full text] [doi: 10.2196/jmir.5357] [Medline: 27165944]
58. Mahoney DF, Purtilo RB, Webbe FM, Alwan M, Bharucha AJ, Adlam TD, Working Group on Technology of the Alzheimer's Association. In-home monitoring of persons with dementia: Ethical guidelines for technology research and development. Alzheimers Dement 2007 Jul;3(3):217-226. [doi: 10.1016/j.jalz.2007.04.388] [Medline: 19595941]
59. Cleret de Langavant L, Bayen E, Yaffe K. Unsupervised machine learning to identify high likelihood of dementia in population-based surveys: Development and validation study. J Med Internet Res 2018 Jul 09;20(7):e10493 [FREE Full text] [doi: 10.2196/10493] [Medline: 29986849]
60. Berrouiguet S, Ramírez D, Barrigón ML, Moreno-Muñoz P, Carmona Camacho R, Baca-García E, et al. Combining continuous smartphone native sensors data capture and unsupervised data mining techniques for behavioral changes detection: A case series of the Evidence-Based Behavior (eB2) study. JMIR Mhealth Uhealth 2018 Dec 10;6(12):e197 [FREE Full text] [doi: 10.2196/mhealth.9472] [Medline: 30530465]

Abbreviations

AD: Alzheimer disease
AI: artificial intelligence
aMCI: amnestic MCI
CART: Collaborative Aging Research Using Technology
CDR: Clinical Dementia Rating
GPS: global positioning system
ICT: information and communication technologies
IEEE: Institute of Electrical and Electronics Engineers
IVR: interactive voice response
MCI: mild cognitive impairment
MeSH: Medical Subject Headings
MMSE: Mini Mental State Examination
naMCI: nonamnestic MCI
NC: not communicated
ORCATECH: Oregon Center for Aging and Technology
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RFID: radio frequency identification

Edited by G Eysenbach; submitted 12.11.18; peer-reviewed by T Nef, A Billis, Q Zhang; comments to author 31.03.19; revised version received 22.05.19; accepted 29.06.19; published 30.08.19.

Please cite as:
Piau A, Wild K, Mattek N, Kaye J
Current State of Digital Biomarker Technologies for Real-Life, Home-Based Monitoring of Cognitive Function for Mild Cognitive Impairment to Mild Alzheimer Disease and Implications for Clinical Care: Systematic Review
J Med Internet Res 2019;21(8):e12785
URL: http://www.jmir.org/2019/8/e12785/
doi:10.2196/12785
PMID:31471958

©Antoine Piau, Katherine Wild, Nora Mattek, Jeffrey Kaye. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 30.08.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Viewpoint

Using Facebook for Qualitative Research: A Brief Primer

Daschel Franz[1], BS; Heather Elizabeth Marsh[1,2], MA; Jason I Chen[1,2], PhD; Alan R Teo[1,2,3], MD, MS

[1]Center to Improve Veteran Involvement in Care, Veterans Affairs Portland Health Care System, Department of Veterans Affairs, Portland, OR, United States

[2]Department of Psychiatry, Oregon Health & Science University, Portland, OR, United States

[3]School of Public Health, Oregon Health & Science University and Portland State University, Portland, OR, United States

Corresponding Author:
Alan R Teo, MD, MS
Center to Improve Veteran Involvement in Care
Veterans Affairs Portland Health Care System
Department of Veterans Affairs
3710 SW US Veterans Hospital Road
R&D 66
Portland, OR, 97239
United States
Phone: 1 5032208262 ext 52461
Email: teoa@ohsu.edu

Abstract

As Facebook continues to grow its number of active users, the potential to harness data generated by Facebook users also grows. As much of Facebook users' activity consists of creating (and commenting on) written posts, the potential use of text data for research is enormous. However, conducting a content analysis of text from Facebook users requires adaptation of research methods used for more traditional sources of qualitative data. Furthermore, best practice guidelines to assist researchers interested in conducting qualitative studies using data derived from Facebook are lacking. The purpose of this primer was to identify opportunities, as well as potential pitfalls, of conducting qualitative research with Facebook users and their activity on Facebook and provide potential options to address each of these issues. We begin with an overview of information obtained from a literature review of 23 studies published between 2011 and 2018 and our own research experience to summarize current approaches to conducting qualitative health research using data obtained from Facebook users. We then identify potential strategies to address limitations related to current approaches and propose 5 key considerations for the collection, organization, and analysis of text data from Facebook. Finally, we consider ethical issues around the use and protection of Facebook data obtained from research participants. In this primer, we have identified several key considerations that should aid health researchers in the planning and execution of qualitative studies involving content analysis of text data from Facebook users.

(J Med Internet Res 2019;21(8):e13544) doi:10.2196/13544

Introduction

Social media platforms provide an information-rich opportunity to reach diverse populations that would otherwise be difficult to identify. Facebook, in particular, is the most dominant player in the social media landscape. Over the past decade, the number of active Facebook users has grown from 145 million in 2008 to more than 1.2 billion in 2018 [1,2]. As of 2018, approximately two-thirds of US adults use Facebook [3]. In addition, about 75% of Facebook users visit the site at least once per day and spend upward of 50 min daily on Facebook [3,4], where they get entertainment, read news, communicate with friends and family, and exchange social support [5].

As a significant portion of individuals' social lives is conducted (and hence displayed and recorded) on Facebook, it is a potentially rich source of qualitative data for researchers [6]. Numerous studies ranging in topic from psychopathology [7,8] and chronic physical illnesses (eg, cancer or diabetes) [9,10] to substance use [11,12] have incorporated data from Facebook, recruited from and included Facebook users as study participants [13,14], or conducted behavioral interventions on the Facebook platform [12]. Despite the rising number of studies on Facebook, relatively little is understood about how qualitative data from Facebook users can best be captured and used for health research purposes. Individual and group interviewing, focus groups, individual and group ethnographic interviewing, and observational data are among the most common methods used to traditionally collect qualitative data [15-17]. These sources of qualitative data naturally allow researchers to unpack deep meaning within a select group of people [18], probe for underlying values, beliefs, and assumptions [19], and obtain

XSL•FO
RenderX

more nuanced or novel information than that derived from other methods such as close-ended survey questions [19]. However, because of the nature of Facebook data, qualitative research methods may require additional adaptation to best capture the visual, virtual, and textual interactions on social media with accuracy [20].

In this primer, we explore the opportunities, as well as potential pitfalls, of conducting qualitative research with Facebook users and their activity on Facebook. Our focus here is purposefully narrow. We limit our approach to content analysis and user-generated text related to health topics on Facebook. We begin with an overview of the forms of qualitative data and data analysis best suited to the Facebook environment, focusing on text data generated by Facebook users. Then, we consider gaps in current qualitative methods based on the existing published literature. Finally, we present 5 key issues that must be addressed in a successive manner when conducting qualitative content analyses of health-related topics involving Facebook data, and we offer potential options to address each of these issues.

Overview of Using Qualitative Data on Facebook

Data obtained from Facebook users offer substantial opportunities for qualitative researchers. As described in Table 1, user-generated videos, images, reactions, and text are a rich source of qualitative data on Facebook. For the purpose of this paper, we focused on user-generated textual data. There are 3 primary types of user-generated textual data on Facebook:

1. *Posts:* A post is written by a Facebook user, and that post then appears on another Facebook user's timeline. A *status update* is a common type of post in the Facebook environment, which will appear in the *news feed* of a user's Facebook friends. A news feed is a list of updates from a user's Facebook friends that is intended to provide the user a quick update on what their Facebook friends have been doing on Facebook.
2. *Comments*: A comment is a response to a Facebook post or a response to another comment itself.
3. *Messages:* A message is privately sent from one user to another Facebook user, typically a Facebook friend. A message does not appear on a user's Facebook timeline or in their news feed.

All 3 of these types of user-generated text on Facebook may be accompanied by image(s), video(s), and/or emoticon(s). An emoticon, or *emoji*, is a graphic facial expression that can appear embedded in text communication on Facebook and is primarily used to provide emotional information that would otherwise only be found in traditional face-to-face interactions (eg, tone of voice) [21].

Social media qualitative research methods can be described in 3 ways: active analysis, passive analysis, and research self-identification [22]. *Active analysis* on Facebook involves the participation of research members in communication with Facebook participants. For instance, Cheung et al [11] created a study Facebook group and invited participants to join. The research team member serving as the Facebook group moderator actively participated in generating content (ie, posts and comments) that aimed to stimulate engagement with study participants. *Passive analysis* on Facebook involves the study of information patterns observed on Facebook or the interactions between users in existing Facebook groups. For example, Kent et al [13] investigated public attitudes about obesity and cancer by performing a keyword search on Facebook to identify relational themes, grammatical elements, and valence of the sentiments contained in Facebook posts and associated comments. Finally, *research self-identification* is when researchers use Facebook as a research recruitment tool to gather participants for Web-based interviews, focus groups, or surveys. For example, Pedersen et al [14] designed 3 different sets of study advertisements that appeared on approximately 3.6 million targeted Facebook users' news feed. By clicking on the study advertisements, Facebook users were redirected to a study survey and were given the option to participate in the study.

To determine current approaches to the use of qualitative data on Facebook, we performed a literature search in April 2018 for papers that used qualitative methods to analyze user-generated Facebook text related to health topics (ie, any acute or chronic disease including substance abuse disorders). Our review identified 23 studies published between 2011 and 2018. The majority of these studies extracted data from public Facebook pages or groups [7-11,13,23-37]. Of 23 studies, 18 used passive analysis [7-10,13,23-28,30-36], 5 used active analysis [11,12,29,37,38], and none used research self-identification. Among the passive analysis studies, the number of posts, comments, and groups or pages analyzed ranged from 25 to 500, 233 to 15,972, and 1 to 840, respectively. In addition, among the active analysis studies, the number of posts analyzed and participants included ranged from 6 to 469 and 79 to 160 participants, respectively. Nearly all studies used a process of manual coding, although 1 study used machine learning techniques [37]. A wide range of health issues was examined from breast cancer to smoking cessation. Further descriptive characteristics can be found in Multimedia Appendix 1.

XSL•FO
RenderX

Table 1. Potential sources of data for qualitative data analysis within Facebook.

Filters	Data included
Timeline	User-generated and user-directed posts, comments, reactions, shares, photos, videos, tagged posts and photos, and when the participant added someone as a friend. Displays public data
Activity log	User-generated and user-directed posts, comments, reactions, shares, photos, videos, tagged posts and photos, pages liked, and when the participant added someone as a friend. Displays public and private data
Posts	Posts generated by the user
Posts tagged in	Posts where other users tag the user
Others' posts to your timeline	Posts that others generate on the user's timeline
Hidden from timeline	A privacy setting that limits who can see posts on a participant's timeline
Photos and videos	Photos and videos that the user posts, uploads, or is tagged in
Likes and reactions	Likes and reactions generated by the user
Comments	Comments generated by the user
Articles you have read	Articles read by the user
Notes	User-generated full-length posts without limited character length and can include tagging and pictures
Videos you have watched	Videos watched by the user
Following	A list of pages the user follows
Groups	A list of groups the user is a member of
Search history	Content the user searches on Facebook

Gaps in Current Qualitative Approaches

Our review identified a number of limitations within the existing literature. First, most studies did not provide detailed descriptions of their methods [39,40]. In particular, description of data extraction methods was frequently missing [7,11,13,23,25,27-31,33,34,37]. Furthermore, there are few existing resources that offer guidance for researchers seeking to use Facebook for health-related topics. Lack of methodological descriptions and advice in the literature pose as barriers to researchers trying to replicate study results or apply the same methods in pursuit of novel research questions in the health domain. Second, none of the studies analyzed bidirectional interactions among participants and other Facebook users. Bidirectional interactions are social exchanges of user-generated and received text between Facebook users. Received text is text directed to a Facebook user, such as a friend's comment to that Facebook user's post (hereafter, user-directed text). These interactions are commonly displayed as a chain of communication on a user's timeline or news feed that exemplify how individuals use and interact with others on Facebook. By collecting only user-generated text or user-directed text on Facebook, studies are only capturing one side of Facebook user's interactions with other Facebook members. However, collecting bidirectional interactions provides more context of social exchanges on Facebook, which can assist in more meaningful interpretations of the data. Therefore, it is important to establish methods for researchers seeking to capture this type of information. Third, most studies that included either manual or machine-coding techniques lacked familiarization methods before coding [8,11-13,24-27,29-38]. Familiarization methods include researchers immersing themselves with the data before coding by actively reading the data to understand the depth and context of the content [41]. To conduct rigorous and trustworthy thematic analyses, it is vital to read through the entire dataset at least once before coding [41,42].

Owing to these limitations, in this paper, we identify and discuss 5 key issues in the process of conducting qualitative research using data obtained from Facebook. These issues are summarized in Textbox 1 and described in detail below. In addition, we use our own experience from a recent research project to illustrate 1 potential approach to handle each of these issues. Our experience derives from a study in which we used Facebook advertisements to recruit a sample of military veterans [43]. Study participants completed a Web-based survey about their psychiatric symptoms and social support, and a subgroup was invited to participate in an additional in-person study visit in which they provided access to some of their Facebook data. For qualitative analysis in this project using Facebook data, we used content analysis, which, for our study, was a more directed approach that allowed us to begin by identifying key concepts and variables as initial coding categories.

XSL•FO
RenderX

Textbox 1. Key considerations for future studies using qualitative approaches for social media data.

Step 1. What kind of Facebook user will be included in the study?

- The method of recruitment of Facebook users will affect participants' characteristics and generalizability of results.
- The degree of activity on Facebook by a study subject will impact the amount of data available for analysis.

Step 2. What Facebook data will be analyzed?

- Facebook contains a combination of public and private information about individual users.
- Filters can be used to select desired variables and data about Facebook users.
- It is helpful to predetermine a period of Facebook use to be included in data analysis.

Step 3. How will the Facebook data be obtained?

- Options include partnering with Facebook, collecting publicly available data, creating a research study–specific Facebook page or group, or downloading participants' Facebook data.
- Each option has pros and cons related to the complexity of the process and comprehensiveness of data obtained.

Step 4. How will the Facebook data be analyzed?

- Depending on the size of the dataset, researchers may prefer a manual versus more automated approach to coding and data analysis.
- Qualitative data analysis and other software can assist with the data analysis.
- Consider the model of qualitative analysis used in the study.

Step 5. How will participant's Facebook data be protected?

- The Connected and Open Research Ethics is a Web-based resource [44] to help navigate ethical issues around social media research.
- Common ethical issues include the following: who will informed consent be obtained from, how will data of research subjects be kept secure, and how will the privacy of research subjects be maintained.

Step 1: What Kind of Facebook User Will Be Included in the Study?

In deciding what kind of Facebook user will be included in the study, it is important to consider how participants will be recruited.

For studies that involve delivery of an intervention through Facebook (ie, active analysis), the platform offers 2 main features that researchers can use to recruit and maintain participants: Facebook pages and Facebook groups. Facebook pages are public, whereas Facebook groups can be public, or private or secret. In public Facebook groups, only invited members can see content. However, in secret Facebook groups, only invited members can see content, and the group is hidden—it cannot be searched for, or found, using the Facebook search engine [45]. Facebook pages and all Facebook groups can be created to recruit and conduct an intervention. In addition, researchers can access existing public Facebook pages and groups comprising current members to collect data. However, these pages and groups cannot be tailored to a researcher's interventions. Furthermore, Facebook advertisements can be used to target a specific population by leveraging demographic profiles available on Facebook. Furthermore, Facebook advertisements can use additional information (eg, interests) added by a user to their profile. Some studies recruit both current Facebook users and other participants who are willing to open a Facebook account for the study [45]. In addition, it is important to consider the degree to which participants are regularly and actively using Facebook. Regular users will tend to have a richer record of their Facebook activity. That said, not all users of Facebook actively engage in behaviors that create a record of interaction on Facebook (eg, posting and commenting) [46]. Facebook users can be categorized into 2 types of users based on the frequency of engaging in these behaviors: active users and passive users. Active users contribute to Facebook interactions by posting and commenting frequently. Passive users tend to observe Facebook interactions and not actively contribute. For active analysis studies, both active and passive users can be considered for recruitment. Interventionists may consider designing posts to initiate interactions among participants, especially from passive users.

In addition, studies intending to observe Facebook user's interactions with other users (ie, passive analysis) can use 2 public group features available on Facebook: Facebook pages and public Facebook groups. As these pages and groups are public, researchers are able to openly view all Facebook data without restrictions. As a result, researchers can search for an existing public page or group related to a health topic of interest and then collect the data presented within the page or group. Data found in public Facebook pages and groups can be from both active and passive users. Typically, there is a direct relationship between the number of members part of a Facebook page or group and the amount of data available. One drawback about using public Facebook pages and groups is that the pages and groups about a health topic of interest must already exist. Alternatively, passive analysis studies can recruit Facebook participants individually through Facebook advertisements. An

XSL•FO
RenderX

advantage of this approach is the ability to continue an advertising campaign until enough participants and data are collected, whereas a disadvantage of it is the requirement for a nontrivial advertising budget.

Paid advertisements on Facebook are also useful for studies seeking to recruit participants from Facebook to participate in interviews, focus groups, surveys, or other research activities (ie, research self-identification). Facebook advertisements can be used to target particular users using the methods described above. Facebook users can be directed to a study website when they click on the advertisement, which then can further describe the study and include Web-based informed consent. Furthermore, Facebook advertisements can record user actions such as advertisement clicks (ie, number of times the advertisement was clicked on) and comments on the post containing the advertisement.

Finally, as with other Web-based studies in which in-person contact with a study participant does not occur, exclusion criteria should be carefully considered to reduce misrepresentation of participants and potentially counterfeit responders (ie, responders pretending to fit a certain demographic for study compensation).

An Applied Example

We used research self-identification methods to recruit participants through Facebook advertisements [43]. Advertisements contained a call to action to participate in a health research study. Study advertisements broadly targeted Facebook users in the United States of any age or gender who had interests relevant to military veterans. Advertisements were hosted by Facebook pages affiliated with our university. This allowed us to draw on the established base of Facebook users interested in and following our university on Facebook.

To reduce misrepresentation of participants, we excluded individuals who completed the survey in less than 5 min, had a duplicate or multiple survey responses, or incorrectly answered military-related *insider knowledge* questions [14,47]. To help ensure study subjects had enough Facebook data to analyze, we chose to collect qualitative data from participants who reported using Facebook at least once a day.

Step 2: What Facebook Data Will Be Analyzed?

In deciding what Facebook data will be analyzed, it is critical to determine the setting in which the data will be collected. For active or passive analysis studies collecting data from public, private, or secret Facebook groups or pages, it is important to consider downloading individual Facebook user's profile information in addition to the information exchanged in groups or pages. A Facebook user's profile information shows how the user interacts in multiple Facebook settings compared with a singular setting (ie, a Facebook page or group). Therefore, collecting and analyzing data from a user's Facebook profile provides more context to how they interact, whom they interact with, and in which environments (ie, public or private) they are more active. Understanding how research participants interact on Facebook can be used to supplement the context of the responses and inform future intervention processes.

In addition, given how expansive the amount of Facebook data can be, even just from a single Facebook user, it is vital to determine the scope of data that will be analyzed. As described in Table 1, Facebook features, such as *Filters*, allow data to be viewed in already separated Facebook variables such as user-generated data (ie, notes, posts tagged in, and timeline review). These filters can be manipulated to display specific data of interest. Although filters can help find user-generated and user-directed data, it is important to also capture these same data in the timeline. The timeline shows how Facebook users are interacting, which helps provide context when analyzing the data.

Furthermore, it is also important to determine how long it takes to collect the Facebook data. Data collection time is dependent on how active the Facebook user is and, for pages or groups, how many users are part of a page or group. These factors can impact additional study procedures (eg, interviews) at the time of the Facebook data collection period.

Our Experience and Applied Example

In our study, we sought to capture all our veteran participants' written social interactions on Facebook. We did this by collecting user-generated and user-directed comments, status updates, and posts from the activity log and the timeline. The timeline was also included as it contains data from both public and private settings on Facebook. By collecting both user-generated and user-directed data, we were able to capture bidirectional interactions between study participants and other Facebook users within their social network.

In addition, data were collected over a 4-week period around the time of the participants' survey completion. We decided to collect participant's Facebook data at the time of the in-person interview so that a research member could be physically present to assist a participant in the process of downloading his or her Facebook activity. After informed consent, the initial 10 min of the session were used to collect the participant's Facebook activity information, which was sufficient to collect users' Facebook data, ranging up to approximately 70 user-generated posts.

Step 3: How Will the Facebook Data Be Obtained?

Option 1: Partner With Facebook

Facebook data can be obtained through a research partnership with Facebook. Kramer et al [48], supported by Facebook resources, collected posts and manipulated news feeds of 689,003 Facebook users over a 20-year period. Burke and Kraut [49], led by a Facebook researcher, collected user-directed comments, private messages, timeline posts, likes, and pokes, as well as user information such as number of profiles viewed, news feed stories clicked on, and photos viewed from 10,557 Facebook users. Some advantages of partnering with Facebook are that studies can have access to massive amounts of data including Facebook variables that are not shared with users or

XSL•FO
RenderX

third parties [50]. In addition, one can leverage Facebook resources (ie, data processing systems) to track how much people are discussing specific topics of interest and the subsequent opinions of those topics expressed in everyday conversation. Such Facebook resources efficiently gather large-scale data in which data are retrieved almost instantaneously. However, a challenge of partnering with Facebook is meeting their *collaborative requirements*, such as finding a Facebook sponsor to lead the research effort, and the faculty principal investigator's institution paying up to 40% of overhead costs for a hosted researcher [51]. Therefore, this process can be resource intensive in terms of both time and financial investment by the partner researcher.

Option 2: Publicly Available Data

Active and passive analysis studies can obtain Facebook data through public Facebook pages and groups. There are several studies using extraction methods such as manual extraction (eg, copying and pasting data into a spreadsheet) or contracting through external models and third-party services for manual extraction. Abramson et al [9] copied and pasted each public timeline post from the Breast Cancer Organization page into a spreadsheet with the corresponding responses. Eghdam et al [8] used Netvizz version 1.25, a data collection software created by Facebook, to collect anonymous data from public Facebook groups. Kent et al [13] used a Web-crawling service that mined publicly available posts and comments from Facebook using keywords related to obesity. Furthermore, Kosinski et al [50] provide Pennebaker's Linguistic Inquiry and Word Count (LIWC), and the Apply Magic Sauce, a website developed by the University of Cambridge psychometrics center, [52] as an additional resource for data collection. An advantage of using public data is that there are a lot of data for a range of health topics, and informed consent by the participant is not required. However, the challenge of using data shared publicly could be biased because of social desirability influences and other censoring by a given participant. Studies suggest that both privacy concerns and the user's audience can impact self-disclosure on Facebook, especially when it comes to sharing health information [53-57]. Eysenbach and Till [22] recommend working with group moderators to develop an adequate plan for informing group members of the use of their data. Although they identify obtaining permission from the group moderator as insufficient on its own, group moderators have greater knowledge of their group members and may be able to provide important information on how to best obtain consent for use of data.

Option 3: Create and Monitor a Facebook Page or Group

In addition, for active analysis studies, Facebook data can be obtained by creating and monitoring a Facebook page or group. Beullens and Schepers [12] collected 2575 pictures and 92 status updates by creating a study Facebook profile and sending friend requests, including a study overview message, to 166 college students. Tower et al [38] collected post information by creating a Facebook group and inviting 198 nursing students to join the group through email. The invitation advised the group to post information related to their study. A faculty member initiated discussion in the Facebook group. The text and associated attributes were downloaded onto a spreadsheet. An advantage of creating and monitoring a Facebook page or group is that it allows a research team to customize a group specific to a particular health topic. Subsequently, targeted individuals can be invited to this page or group and be presented a set of specific questions/instructions to stimulate participant engagement. In addition, only group settings can be made private, which can create a more secure environment for participants to disclose personal information. However, a disadvantage of private groups is that there is a permanent setting that organizes user-directed posts such that the most recent interactions appear at the top of the group feed versus a chronological ordering of the post [45]. As a result, posts containing important content may be pushed to the bottom of the group feed because of frequent posting in the groups, thereby making it difficult for participants to find information posted by the groups interventionists [45]. In addition, although Facebook groups can be private or secret, they are still not the Facebook user's *natural environment*—that is, the social network comprising Facebook friends the user normally interacts with. Therefore, Facebook users recruited into an intervention conducted in a private or secret group may behave differently in groups created by researchers, especially when they know they are being observed by researchers [58].

Option 4: Private Messages

Furthermore, for active analysis studies, Facebook data can be obtained by asking participants to copy and paste user-generated Facebook text (eg, text from timeline posts or private messages) and provide it to a research team member through a Web-based portal or through private messaging to a Facebook account created by the research team. Bazarova et al [37] collected 474 most recent status updates, timeline posts, and private messages by inviting 79 participants to copy and paste their data into a Web survey. An advantage of having users provide their Facebook data through the private messaging feature or a Web-based portal is that it creates a secure environment in which participants' Facebook data can be kept confidential from other Facebook users or study participants. However, one disadvantage of this particular method is that researchers would neither be able to observe passive interactions among a particular group of Facebook users nor observe interactions as a result of a proposed set of questions/instructions regarding health-related topics.

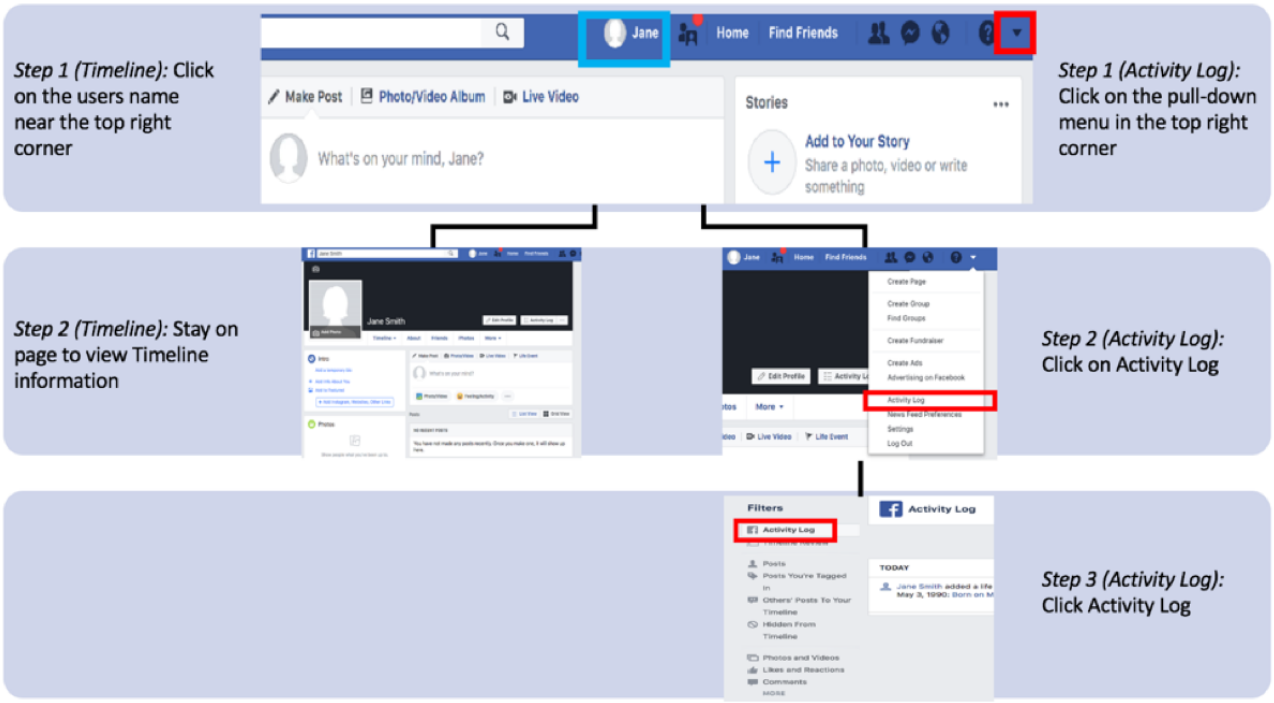
Our Experience and Applied Example

A fourth option, applicable to active and passive analyses and some research self-identification studies, is directly downloading participants' Facebook data during an in-person study visit. We chose this option because it was the only one that allowed us to download individual's Facebook profiles without establishing a partnership with Facebook. For instance, in our own study, we obtained Facebook data by downloading participants' Facebook activity information. During the in-person interview, users' Facebook activity log and timeline data were collected separately by study staff using the following steps: (1) ask participants to login to their Facebook account, (2) follow the steps described in Figure 1, (3) scroll *backward* on the selected page chronologically until 1-month period before the date of

XSL•FO
RenderX

the survey; (4) save as an HTML file on OHSU Box (a cloud-based data storage service that complies with local security and regulatory policies), (5) open saved file with Safari to view extracted data, (6) log participants out, and (7) ensure that no username or password information was retained by making sure user login information was not saved by the browser. We noted some advantages of downloading participants' Facebook profile information, such as a participants' Facebook profile can provide insight to how individuals interact, who they interact with, and what environment (ie, public and private) they are more active in. This helped us understand how study participants interacted on Facebook. However, a challenge of downloading participants' Facebook profile information is that it requires participant consent, and it can be more difficult to collect massive quantities of private data because of the length of the collection period.

Figure 1. Steps to access the timeline (eg, blue square) and activity log (eg, red squares) on Facebook.



Step 4: How Will the Facebook Data Be Analyzed?

Qualitative Facebook data are commonly analyzed using methods such as content analysis to assess a wide range of qualitative data or else constant comparison to identify themes [6]. In deciding how qualitative Facebook data will be analyzed, it is important to consider the quantity of the data as well as the qualitative approach being used. For active and passive analysis studies using larger datasets, it is preferable to analyze data using software programs. AlQarni et al [34] analyzed 1551 posts using predetermined themes, and further inductive codes were used to independently extract and analyze the Facebook posts to determine major content themes. Thematic analysis was performed using NVivo, a qualitative software used to code, store, and potentially exchange data with SPSS for further statistical analysis. Kramer et al [48] used LIWC (2007) software to analyze 689,003 posts to determine if the valence of the posts was positive or negative. Keller et al [32] used ATLAS.ti, a qualitative software used to code data, to code 1614 comments for major and minor themes. It is important to note that ATLAS.ti can be used to code HTML files of individual's Facebook downloads; however, this has not been done in social media qualitative research studies [59]. Instead, ATLAS.ti has been traditionally used to code Microsoft Word documents of transcribed interviews.

Our Experience and Applied Example

As our study contained a relatively small dataset (23 subjects with 201 posts and 424 comments), we opted to analyze data manually. User-generated text from status updates, posts, and comments and user-directed text from posts and comments from the HTML files were copied and pasted into an Excel spreadsheet and analyzed for markers of social support. Our codebook contained 3 different types of social support that have previously been described in the literature (emotional, instrumental, and informational) and a fourth category for other evidence of social support (eg, "Wow, that's a great joke"). In addition, we coded the valence of user-directed social support as positive, negative, or neutral. Before coding, each coder read over the entire dataset to familiarize themselves with the content of the data. The familiarization process helped lead to more meaningful interpretations of the data because we were able to easily provide context to each piece of text we coded. As is common in qualitative research, after an initial training period, 2 coders independently coded participants' data. Furthermore, each coder created a memo describing their experiences during the coding process. This highlighted the challenges and successes of the coding process, which guided conversations around any discrepancies. In addition, the memo process brought awareness to potential challenges of coding text on social media, which can be addressed early on for future social media

XSL•FO
RenderX

qualitative work. Furthermore, the memo process also identified general themes that were prevalent in the data.

Step 5: How Will Participants' Facebook Data Be Protected?

It is important to highlight that Facebook research raises several ethical questions. Owing to the nature of studying Facebook communities, researchers can potentially violate the privacy rights of Facebook users. Facebook users that are members of public Facebook pages or groups do not expect to become research subjects nor do the Facebook friends of study participants (ie, *nonparticipants*). The boundary between private and public Facebook data may sometimes be unclear. The majority of Facebook users are aware that their data may not be private [22], especially in a public setting on Facebook. However, the literature regarding social media users' comprehension of privacy literacy is limited [60]. As a result, researchers should ensure that informed consent language is clear regarding how a participant's Web-based data will be used. Pilot testing of informed consent language may help ensure that the information presented is easily comprehensible for a broad range of populations. Regardless, it is important to maintain the safety and anonymity of individuals' Facebook information whether or not they are a research participant.

In addition, it is important to note the potential ethical dilemmas associated with establishing a research partnership with Facebook. Facebook is a powerful company with a rich source of data; however, Facebook has received public scrutiny because of their misuse of their users' Facebook data. Therefore, the responsibly is placed on the research teams to ensure that Facebook users' data are obtained ethically and protected. Arigo et al [61] recommend including research team members who are well versed with Facebook's cooperate terms and conditions and privacy policies. It is strongly encouraged that research teams are knowledgeable of the peculiarities of Facebook before establishing a partnership to assist in the development of research methodological procedures regarding data collection and privacy.

As each institutional review board (IRB) will vary in its familiarity with social media research, we recommend closely consulting with professional and independent organizations (eg, Association of Internet Researchers Ethics Working Group Guidelines, The National Committee for Research Ethics, and The Humanities Research Ethics Guidelines for Internet Research) as well as Web-based resources such as the Connected and Open Research Ethics (CORE). CORE can provide assistance in how to address potential ethical issues for researchers and IRBs interested in social media research. Common ethical questions that have been raised on CORE include the following: (1) Who will informed consent be obtained from–is informed consent required for *nonparticipants* on a research subject's account?; (2) How will data from research subjects be kept secure on the social media platform?; and (3) How will the privacy of research subjects be maintained? CORE has created a collaborative platform where researchers can exchange expertise and questions pertaining to social media research. Features such as the *Resource Library*, *Q&A Forum*, and the *CORE Network* provide scientists access to IRB-approved research protocols and consent forms and allow researchers to discuss collaboratively ethical design or potential social media strategies [44].

Our Experience and Applied Example

In our study, participants interested in an optional, in-person interview provided contact information with which study staff used to arrange the study visit. For individuals who were unable to come in-person, we conducted interviews through phone but did not download their Facebook data. Overall, 2 separate informed consents were obtained, once online for those completing the survey and again in-person for those sharing their Facebook data. During the informed process for those sharing their Facebook data, participants were informed that their timeline and activity log would be collected to observe their online social interactions and Facebook usage. In addition, participants were informed that their Facebook data would be labeled with a unique code to protect their identity. All study procedures were approved by the IRB of Oregon Health & Science University.

Limitations

There are several limitations to this study. First, this study represents 1 proposed framework. Additional validation of this framework among other experts would be a helpful next step. Second, the scope of the study is limited. We primarily focused on content analysis of user-generated Facebook text related to health topics using a content analysis approach to qualitative analysis. Studies that intend to use other models of qualitative analysis may require somewhat different approaches to the use of data from Facebook. Nontext qualitative data from Facebook (eg, images, videos, and emoticons) also bear further examination. Third, because our key considerations are primarily directed toward health-related studies, it is unclear whether they are generalizable to other research topics that harness data from Facebook. Finally, our applied example did not address methods for collecting data from existing closed Facebook groups, although studies that did do so were identified in our literature review. Studies that involve interaction with Facebook group members require additional consideration, and future research could help elucidate this area by extending the work presented by Eysenbach and Till [22].

Conclusions

Although there are an increasing number of studies that are using qualitative data obtained from Facebook users, there has been little published to date, summarizing the current state of this research. Our review of the literature and own experience conducting this type of research have led us to identify several key considerations for health researchers interested in conducting qualitative studies involving Facebook data. Our hope is that future research continues to refine and develop approaches to conducting research in this exciting area.

XSL•FO
RenderX

Acknowledgments

This project was supported, in part, by a Career Development Award to ART from the Veterans Health Administration Health Service Research and Development (CDA 14-428) and the HSR&D Center to Improve Veteran Involvement in Care. This project was also supported by awards received by ART from the Collins Medical Trust Research Grant and the Medical Research Foundation of Oregon New Investigator Grant (1603). The US Department of Veterans Affairs, Collins Medical Trust Research Grant, and the Medical Research Foundation of Oregon New Investigator Grant 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; or decision to submit the manuscript for publication. The findings and conclusions in this document are those of the authors who are responsible for its contents; the findings and conclusions do not necessarily represent the views of the US Department of Veterans Affairs or the US government, the Collins Medical Trust Research Grant nor the Medical Research Foundation of Oregon New Investigator Grant.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Summary of qualitative peer-reviewed studies using Facebook as a data source.
[DOCX File, 24KB - jmir_v21i8e13544_app1.docx]

References

1. Desilver D. Pew Research Center. 2014. Overseas Users Power Facebook's Growth; More Going Mobile-Only URL: http://www.pewresearch.org/fact-tank/2014/02/04/overseas-users-power-facebooks-growth-more-going-mobile-only/ [WebCite Cache ID 75bOSGxDg]
2. 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/ [WebCite Cache ID 75bOgBZf9]
3. Gramlich J. Pew Research Center. 2019. 10 Facts About Americans and Facebook URL: http://www.pewresearch.org/fact-tank/2018/10/24/facts-about-americans-and-facebook/ [accessed 2019-01-21] [WebCite Cache ID 75bOp2CaI]
4. Stewart JB. The New York Times. 2016. Facebook Has 50 Minutes of Your Time Each Day. It Wants More URL: https://www.nytimes.com/2016/05/06/business/facebook-bends-the-rules-of-audience-engagement-to-its-advantage.html [accessed 2018-12-30] [WebCite Cache ID 75bPABJ7H]
5. Li X, Chen W, Popiel P. What happens on Facebook stays on Facebook? The implications of Facebook interaction for perceived, receiving, and giving social support. Comput Hum Behav 2015 Oct;51(Part A):106-113. [doi: 10.1016/j.chb.2015.04.066]
6. Snelson CL. Qualitative and mixed methods social media research. Int J Qual Methods 2016 Mar 4;15(1):160940691562457. [doi: 10.1177/1609406915624574]
7. Roffeei SH, Abdullah N, Basar SK. Seeking social support on Facebook for children with autism spectrum disorders (ASDs). Int J Med Inform 2015 May;84(5):375-385. [doi: 10.1016/j.ijmedinf.2015.01.015] [Medline: 25701266]
8. Eghdam A, Hamidi U, Bartfai A, Koch S. Facebook as communication support for persons with potential mild acquired cognitive impairment: a content and social network analysis study. PLoS One 2018;13(1):e0191878 [FREE Full text] [doi: 10.1371/journal.pone.0191878] [Medline: 29377930]
9. Abramson K, Keefe B, Chou WY. Communicating about cancer through Facebook: a qualitative analysis of a breast cancer awareness page. J Health Commun 2015;20(2):237-243. [doi: 10.1080/10810730.2014.927034] [Medline: 25495200]
10. Greene JA, Choudhry NK, Kilabuk E, Shrank WH. Online social networking by patients with diabetes: a qualitative evaluation of communication with Facebook. J Gen Intern Med 2011 Mar;26(3):287-292 [FREE Full text] [doi: 10.1007/s11606-010-1526-3] [Medline: 20945113]
11. Cheung YT, Chan CH, Wang MP, Li HC, Lam T. Online social support for the prevention of smoking relapse: a content analysis of the WhatsApp and Facebook social groups. Telemed J E Health 2017 Jun;23(6):507-516. [doi: 10.1089/tmj.2016.0176] [Medline: 27911654]
12. Beullens K, Schepers A. Display of alcohol use on Facebook: a content analysis. Cyberpsychol Behav Soc Netw 2013 Jul;16(7):497-503. [doi: 10.1089/cyber.2013.0044] [Medline: 23617225]
13. Kent EE, Prestin A, Gaysynsky A, Galica K, Rinker R, Graff K, et al. 'Obesity is the new major cause of cancer': connections between obesity and cancer on Facebook and Twitter. J Cancer Educ 2016 Sep;31(3):453-459. [doi: 10.1007/s13187-015-0824-1] [Medline: 25865399]
14. Pedersen ER, Helmuth ED, Marshall GN, Schell TL, PunKay M, Kurz J. Using Facebook to recruit young adult veterans: online mental health research. JMIR Res Protoc 2015 Jun 1;4(2):e63 [FREE Full text] [doi: 10.2196/resprot.3996] [Medline: 26033209]

XSL•FO
RenderX

15. Denzin NK, Lincoln YS. The Sage Handbook of Qualitative Research. Fourth Edition. Thousand Oaks, California: Sage Publications; 2011.
16. Barbour R. Doing Focus Groups. Thousand Oaks, California: Sage Publications; 2008.
17. Schutt RK. Investigating the Social World: The Process and Practice of Research. Seventh Edition. Thousand Oaks, California: Sage Publications; 2003.
18. Dudwick N, Keuhnast K, Jones VN, Woolcock M. Documents & Reports - All Documents | The World Bank. 2006. Analyzing Social Capital in Context : A Guide to Using Qualitative Methods and Data URL: http://documents.worldbank.org/curated/en/601831468338476652/Analyzing-social-capital-in-context-a-guide-to-using-qualitative-methods-and-data [WebCite Cache ID 75bOB7wNJ]
19. Yauch CA, Steudel HJ. Complementary use of qualitative and quantitative cultural assessment methods. Organ Res Methods 2016 Jun 29;6(4):465-481. [doi: 10.1177/1094428103257362]
20. Atieno OP. An analysis of the strengths and limitation of qualitative and quantitative research paradigms. Probl Educ 21st Century 2009;13(1):13-38 [FREE Full text]
21. Kaye LK, Wall HJ, Malone SA. 'Turn that frown upside-down': a contextual account of emoticon usage on different virtual platforms. Comput Hum Behav 2016 Jul;60:463-467. [doi: 10.1016/j.chb.2016.02.088]
22. Eysenbach G, Till JE. Ethical issues in qualitative research on internet communities. Br Med J 2001 Nov 10;323(7321):1103-1105 [FREE Full text] [doi: 10.1136/bmj.323.7321.1103] [Medline: 11701577]
23. Thoren EM, Metze B, Bührer C, Garten L. Online support for parents of preterm infants: a qualitative and content analysis of Facebook 'preemie' groups. Arch Dis Child Fetal Neonatal Ed 2013 Nov;98(6):F534-F538. [doi: 10.1136/archdischild-2012-303572] [Medline: 23759520]
24. Cole-Lewis H, Perotte A, Galica K, Dreyer L, Griffith C, Schwarz M, et al. Social network behavior and engagement within a smoking cessation Facebook page. J Med Internet Res 2016 Aug 2;18(8):e205 [FREE Full text] [doi: 10.2196/jmir.5574] [Medline: 27485315]
25. de la Torre-Díez I, Díaz-Pernas FJ, Antón-Rodríguez M. A content analysis of chronic diseases social groups on Facebook and Twitter. Telemed J E Health 2012;18(6):404-408. [doi: 10.1089/tmj.2011.0227] [Medline: 22650380]
26. Seidel RL, Jalilvand A, Kunjummen J, Gilliland L, Duszak R. Radiologists and social media: do not forget about Facebook. J Am Coll Radiol 2018 Jan;15(1 Pt B):224-228. [doi: 10.1016/j.jacr.2017.09.013] [Medline: 29132999]
27. Struck JP, Siegel F, Kramer MW, Tsaur I, Heidenreich A, Haferkamp A, et al. Substantial utilization of Facebook, Twitter, YouTube, and Instagram in the prostate cancer community. World J Urol 2018 Aug;36(8):1241-1246. [doi: 10.1007/s00345-018-2254-2] [Medline: 29523948]
28. Meng Y, Elkaim L, Wang J, Liu J, Alotaibi NM, Ibrahim GM, et al. Social media in epilepsy: a quantitative and qualitative analysis. Epilepsy Behav 2017 Jun;71(Pt A):79-84. [doi: 10.1016/j.yebeh.2017.04.033] [Medline: 28554148]
29. Dhar VK, Kim Y, Graff JT, Jung AD, Garrett J, Dick LE, et al. Benefit of social media on patient engagement and satisfaction: results of a 9-month, qualitative pilot study using Facebook. Surgery 2018 Mar;163(3):565-570. [doi: 10.1016/j.surg.2017.09.056] [Medline: 29195735]
30. Ahmed S, Haines-Saah RJ, Afzal AR, Tam-Tham H, Al Mamun M, Hemmelgarn BR, et al. User-driven conversations about dialysis through Facebook: a qualitative thematic analysis. Nephrology (Carlton) 2017 Apr;22(4):301-307. [doi: 10.1111/nep.12780] [Medline: 27003709]
31. Gonsalves CA, McGannon KR, Schinke RJ. Social media for health promotion: what messages are women receiving about cardiovascular disease risk by the Canadian Heart and Stroke Foundation? J Health Psychol 2017 Nov 1:1359105317743802. [doi: 10.1177/1359105317743802] [Medline: 29226735]
32. Keller MS, Park HJ, Cunningham ME, Fouladian JE, Chen M, Spiegel BM. Public perceptions regarding use of virtual reality in health care: a social media content analysis using Facebook. J Med Internet Res 2017 Dec 19;19(12):e419 [FREE Full text] [doi: 10.2196/jmir.7467] [Medline: 29258975]
33. Struik LL, Baskerville NB. The role of Facebook in Crush the Crave, a mobile- and social media-based smoking cessation intervention: qualitative framework analysis of posts. J Med Internet Res 2014 Jul 11;16(7):e170 [FREE Full text] [doi: 10.2196/jmir.3189] [Medline: 25016998]
34. AlQarni ZA, Yunus F, Househ MS. Health information sharing on Facebook: an exploratory study on diabetes mellitus. J Infect Public Health 2016;9(6):708-712 [FREE Full text] [doi: 10.1016/j.jiph.2016.08.015] [Medline: 27618634]
35. Hale TM, Pathipati AS, Zan S, Jethwani K. Representation of health conditions on Facebook: content analysis and evaluation of user engagement. J Med Internet Res 2014 Aug 4;16(8):e182 [FREE Full text] [doi: 10.2196/jmir.3275] [Medline: 25092386]
36. Bender JL, Jimenez-Marroquin M, Jadad AR. Seeking support on Facebook: a content analysis of breast cancer groups. J Med Internet Res 2011 Feb 4;13(1):e16 [FREE Full text] [doi: 10.2196/jmir.1560] [Medline: 21371990]
37. Bazarova NN, Taft JG, Choi YH, Cosley D. Managing impressions and relationships on Facebook: self-presentational and relational concerns revealed through the analysis of language style. J Lang Soc Psychol 2012 Aug 30;32(2):121-141. [doi: 10.1177/0261927X12456384]

XSL•FO
RenderX

38. Tower M, Blacklock E, Watson B, Heffernan C, Tronoff G. Using social media as a strategy to address 'sophomore slump' in second year nursing students: a qualitative study. Nurse Educ Today 2015 Nov;35(11):1130-1134. [doi: 10.1016/j.nedt.2015.06.011] [Medline: 26220639]
39. Nelson CG. Power of the people: a case study using Facebook as a data source in qualitative research. Issues Inf Syst 2015;16(4):232-238 [FREE Full text]
40. Società Italiana Marketing. 2017. Emerging Trends in Qualitative Research: A Focus on Social Media Research URL: http://www.simktg.it/MTF/Content/convegno/CfPEmergingtrendsinqualitativeresearch__5914239856948768046.pdf [accessed 2019-01-21] [WebCite Cache ID 75bPFZNRm]
41. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
42. Nowell L, Norris J, White D, Moules N. Thematic analysis: striving to meet the trustworthiness criteria. Int J Qual Methods 2017 Oct 2;16(1):160940691773384. [doi: 10.1177/1609406917733847]
43. Teo AR, Liebow SB, Chan B, Dobscha SK, Graham AL. Reaching those at risk for psychiatric disorders and suicidal ideation: Facebook advertisements to recruit military veterans. JMIR Ment Health 2018 Jul 5;5(3):e10078 [FREE Full text] [doi: 10.2196/10078] [Medline: 29980498]
44. Connected and Open Research Ethics (CORE). 2018. URL: https://thecore.ucsd.edu/ [accessed 2019-01-19] [WebCite Cache ID 75bPuVixM]
45. Pagoto S, Waring ME, May CN, Ding EY, Kunz WH, Hayes R, et al. Adapting behavioral interventions for social media delivery. J Med Internet Res 2016 Jan 29;18(1):e24 [FREE Full text] [doi: 10.2196/jmir.5086] [Medline: 26825969]
46. Verduyn P, Lee DS, Park J, Shablack H, Orvell A, Bayer J, et al. Passive Facebook usage undermines affective well-being: experimental and longitudinal evidence. J Exp Psychol Gen 2015 Apr;144(2):480-488. [doi: 10.1037/xge0000057] [Medline: 25706656]
47. Reger MA, Smolenski DJ, Skopp NA, Metzger-Abamukang MJ, Kang HK, Bullman TA, et al. Risk of suicide among US military service members following operation enduring freedom or operation Iraqi freedom deployment and separation from the US military. JAMA Psychiatry 2015 Jun;72(6):561-569. [doi: 10.1001/jamapsychiatry.2014.3195] [Medline: 25830941]
48. Kramer AD, Guillory JE, Hancock JT. Experimental evidence of massive-scale emotional contagion through social networks. Proc Natl Acad Sci U S A 2014 Jun 17;111(24):8788-8790 [FREE Full text] [doi: 10.1073/pnas.1320040111] [Medline: 24889601]
49. Burke M, Kraut RE. The relationship between Facebook use and well-being depends on communication type and tie strength. J Comput-Mediat Comm 2016 Jul 26;21(4):265-281. [doi: 10.1111/jcc4.12162]
50. Kosinski M, Matz SC, Gosling SD, Popov V, Stillwell D. Facebook as a research tool for the social sciences: opportunities, challenges, ethical considerations, and practical guidelines. Am Psychol 2015 Sep;70(6):543-556. [doi: 10.1037/a0039210] [Medline: 26348336]
51. Facebook Research. 2018. Research Collaborations URL: https://research.fb.com/programs/research-collaborations/ [accessed 2019-01-20] [WebCite Cache ID 75bPZCWUS]
52. University of Cambridge. 2019. Apply Magic Sauce URL: https://applymagicsauce.com/demo
53. Vitak J, Ellison NB. 'There's a network out there you might as well tap': exploring the benefits of and barriers to exchanging informational and support-based resources on Facebook. New Media Soc 2012 Jul 23;15(2):243-259. [doi: 10.1177/1461444812451566]
54. Bansal G, Zahedi FM, Gefen D. The impact of personal dispositions on information sensitivity, privacy concern and trust in disclosing health information online. Decis Support Syst 2010 May;49(2):138-150. [doi: 10.1016/j.dss.2010.01.010]
55. Syn SY, Kim SU. College students' health information activities on Facebook: investigating the impacts of health topic sensitivity, information sources, and demographics. J Health Commun 2016 Jul;21(7):743-754. [doi: 10.1080/10810730.2016.1157652] [Medline: 27220029]
56. Leary MR. Self-Presentation: Impression Management and Interpersonal Behavior. Madison, WI: Brown & Benchmark Publishers; 1996.
57. Stutzman F, Capra R, Thompson J. Factors mediating disclosure in social network sites. Comput Hum Behav 2011 Jan;27(1):590-598. [doi: 10.1016/j.chb.2010.10.017]
58. Orne M, Whitehouse W. Demand characteristics. In: Kazdin AE, editor. Encyclopedia of Psychology. Washington, DC: American Psychological Association and Oxford Univeristy Press; 2000:469-470.
59. ATLAS.ti: The Qualitative Data Analysis & Research Software. 2018. Knowledge Base URL: https://atlasti.com/support/faq-overview/ [accessed 2019-01-21] [WebCite Cache ID 75bPgfQ5z]
60. Baruh L, Secinti E, Cemalcilar Z. Online privacy concerns and privacy management: a meta-analytical review. J Commun 2017 Jan 17;67(1):26-53. [doi: 10.1111/jcom.12276]
61. Arigo D, Pagoto S, Carter-Harris L, Lillie SE, Nebeker C. Using social media for health research: methodological and ethical considerations for recruitment and intervention delivery. Digit Health 2018;4:2055207618771757 [FREE Full text] [doi: 10.1177/2055207618771757] [Medline: 29942634]

XSL•FO
RenderX

Abbreviations

CORE: Connected and Open Research Ethics
IRB: institutional review board
LIWC: Linguistic Inquiry and Word Count

Edited by G Eysenbach; submitted 11.02.19; peer-reviewed by B Chaudhry, E Zibrowski, GE Iyawa, R Alkoudmani; comments to author 31.03.19; revised version received 21.05.19; accepted 05.07.19; published 13.08.19.

Please cite as:
Franz D, Marsh HE, Chen JI, Teo AR
Using Facebook for Qualitative Research: A Brief Primer
J Med Internet Res 2019;21(8):e13544
URL: http://www.jmir.org/2019/8/e13544/
doi:10.2196/13544
PMID:31411143

©Daschel Franz, Heather Elizabeth Marsh, Jason I Chen, Alan R Teo. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 13.08.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Viewpoint

Multimodal Human and Environmental Sensing for Longitudinal Behavioral Studies in Naturalistic Settings: Framework for Sensor Selection, Deployment, and Management

Brandon M Booth[1*], MSc; Karel Mundnich[1*], Eng; Tiantian Feng[1*], MSc; Amrutha Nadarajan[1], MSc; Tiago H Falk[2], PhD; Jennifer L Villatte[3], PhD; Emilio Ferrara[4], PhD; Shrikanth Narayanan[1], PhD

[1]Signal Analysis and Interpretation Laboratory, University of Southern California, Los Angeles, CA, United States

[2]Multimedia/Multimodal Signal Analysis and Enhancement Lab, Institut National de la Recherche Scientifique, University of Québec, Montréal, QC, Canada

[3]Department of Psychiatry and Behavioral Sciences, University of Washington School of Medicine, Seattle, WA, United States

[4]Information Sciences Institute, University of Southern California, Marina del Rey, CA, United States

[*]these authors contributed equally

Corresponding Author:
Brandon M Booth, MSc
Signal Analysis and Interpretation Laboratory
University of Southern California
3740 McClintock Ave, EEB 400
Los Angeles, CA, 90089
United States
Phone: 1 7204704284
Email: brandon.m.booth@gmail.com

Abstract

Background: Recent advances in mobile technologies for sensing human biosignals are empowering researchers to collect real-world data outside of the laboratory, in natural settings where participants can perform their daily activities with minimal disruption. These new sensing opportunities usher a host of challenges and constraints for both researchers and participants.

Objective: This viewpoint paper aims to provide a comprehensive guide to aid research teams in the selection and management of sensors before beginning and while conducting human behavior studies in the wild. The guide aims to help researchers achieve satisfactory participant compliance and minimize the number of unexpected procedural outcomes.

Methods: This paper presents a collection of challenges, consideration criteria, and potential solutions for enabling researchers to select and manage appropriate sensors for their research studies. It explains a general data collection framework suitable for use with modern consumer sensors, enabling researchers to address many of the described challenges. In addition, it provides a description of the criteria affecting sensor selection, management, and integration that researchers should consider before beginning human behavior studies involving sensors. On the basis of a survey conducted in mid-2018, this paper further illustrates an organized snapshot of consumer-grade human sensing technologies that can be used for human behavior research in natural settings.

Results: The research team applied the collection of methods and criteria to a case study aimed at predicting the well-being of nurses and other staff in a hospital. Average daily compliance for sensor usage measured by the presence of data exceeding half the total possible hours each day was about 65%, yielding over 355,000 hours of usable sensor data across 212 participants. A total of 6 notable unexpected events occurred during the data collection period, all of which had minimal impact on the research project.

Conclusions: The satisfactory compliance rates and minimal impact of unexpected events during the case study suggest that the challenges, criteria, methods, and mitigation strategies presented as a guide for researchers are helpful for sensor selection and management in longitudinal human behavior studies in the wild.

(J Med Internet Res 2019;21(8):e12832) doi:10.2196/12832

XSL•FO
RenderX

KEYWORDS

research design; human activities; behavioral research; longitudinal studies; wearable electronic devices; organizational case studies; in situ research

Introduction

Overview

Recent advances in portable consumer technologies have led to a surge in the development of electronic devices [1] for monitoring and tracking human activity, wellness, and behavior. Aided by the ubiquity of personal smartphones, Bluetooth, and Wi-Fi, many devices currently on the market can discreetly collect physiologic and behavioral signals and upload the information to remote servers. Because of the growing support for distributed and personalized sensing, diverse research communities are taking a keen interest in this field, empowering the coordination of research studies of populations outside the laboratory and in natural home or work environments (also known as studies *in the wild*) [2]. For research into everyday human behavior, such as daily routines, studies conducted in natural settings can yield more relevant and insightful data than those performed in the laboratory [3-8].

Several factors need to be considered for the collection of data in natural human settings using sensing devices. Different sensors have different sampling rates, power restrictions, and communication capabilities. Participants also have their own habits and daily routines into which the sensors and the data collection procedures need to be embedded. A data collection framework designed to operate in the wild should therefore be flexible enough to accommodate different data communication channels and be capable of capturing information from different people with different needs at different times. These factors and a host of other challenges mentioned in this work complicate the data collection process and ultimately affect the quality of data available for analysis.

Background

Several previous studies have described some of these challenges [9-11] and offered strategies for mitigating them [12-15]. Other works offer data collection plans for particular fields of study that address the unique concerns of their research areas [16,17]. This paper subsumes many of the challenges and suggestions from these other works and aims to provide a comprehensive collection of methods and suggestions that help researchers address the challenges related to sensor selection and management in research studies. It specifically focuses on longitudinal studies aiming to unobtrusively capture and assess aspects of human experience and natural behavior; thus, it assumes a participatory study framework instead of a provocative approach [2]. Some examples of unobtrusive human behavior studies are *StudentLife* [6], *AffectiveROAD* [18], and a dataset on emotion recognition from wearable physiological sensing [19].

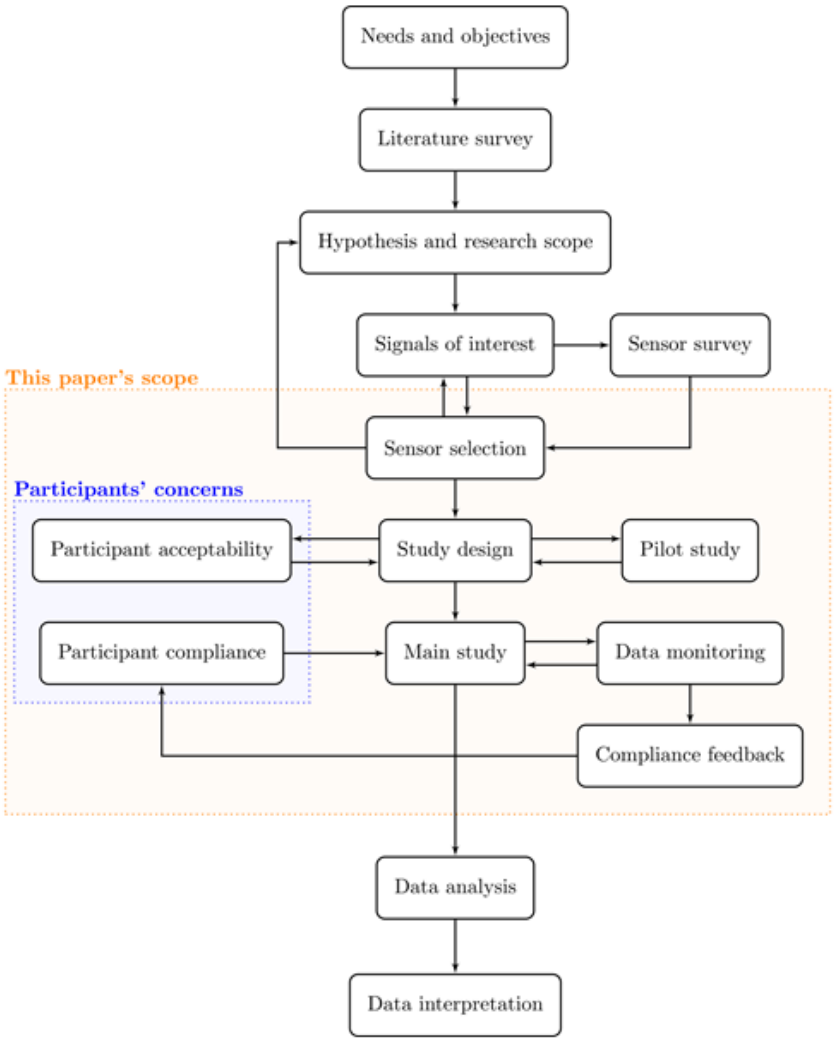
Objectives

Figure 1 illustrates a sequence of research program states at various stages for these types of studies. The scope of this paper covers the preplanning stages pertaining to sensor selection and the stages during a study related to sensor and data management. The key assumptions were as follows: (1) researchers already have a clear research objective in mind and have researched previous literature to develop a sense of the types of physiological and behavioral signals that may be helpful in achieving the goals and (2) researchers have surveyed the landscape of sensing technologies and are beginning to design a study protocol and select the appropriate sensors.

To the best of the authors' knowledge, this paper represents a first attempt to present a comprehensive guide for selecting and managing sensors for *in situ* research studies. The guide is based on a survey of related work [9-15] and the authors' experiences in designing a multiweek research study. It describes the main challenges that differentiate longitudinal and unobtrusive [20] studies in the wild from studies conducted in controlled laboratory settings. It also provides an overview of modern portable sensing capabilities and information workflows and outlines a general data collection framework that leverages an internet-enabled infrastructure for real-time data collection and feedback. It enumerates several criteria (or dimensions) that researchers should consider when designing a data collection protocol using portable sensors for a known participant population, and discusses the manner in which these dimensions can affect human subjects' concerns and data quality. Furthermore, it illustrates a snapshot of some of the many consumer technologies and products available for sensing in the wild as of mid-2018. This case study employs all the criteria and methods discussed, evaluating these with respect to participant compliance and the number of notable unplanned events occurring during data collection.

XSL•FO
RenderX

Figure 1. An overview of the general scientific process for human research studies involving sensing.



Methods

Overview

This section discusses the challenges involved in designing protocols and using sensors to collect data about human behavior in the wild, presenting a general-purpose framework that researchers can use for envisioning and orchestrating sensor data flow. This framework subsumes the most common information flows provided by modern sensing technologies. This section also presents an exposition of the various criteria and dimensions for which all sensors should be evaluated before the beginning of the data collection period, as well as a snapshot summary of many modern consumer technologies and products available for each type of sensing. The checklist form in Multimedia Appendix 1 provides a concise checklist of the challenges in this section. The authors' hope is that researchers will use this checklist in their discussions and planning about protocol design to help account for the numerous sensing challenges.

Challenges and Risk Mitigation Strategies

Studies conducted outside of controlled laboratory settings are of interest to researchers, as participants can be examined in their day-to-day environments where natural behaviors occur. Nevertheless, in the wild, many potentially confounding variables cannot be fully controlled, yielding unpredictable sources of variability alongside logistical difficulties. Some challenges in this kind of data collection are mitigated through careful planning and effective communication before the study begins. Other challenges are predictable, but they occur spontaneously, and they must be managed reactively with the aid of semiautomated systems. This section highlights the primary difficulties that are unique to studies in the wild and suggests strategies to help overcome them.

Sensor Logistics, Deployment, and Maintenance

One of the foremost difficulties is the logistical burden of deploying and maintaining sensors. As research teams have limited direct control over the environment for *in situ* studies, they should be aware of the different potential sensor failure modes and have a plan for quickly detecting and recovering from them.

Sensor failure is often inevitable, especially for studies conducted at scale, and an effective solution is to simply replace the devices by preplanning to streamline this process. For example, arranging to have trained personnel available to meet with participants in their environment to swap defective devices can minimize data losses because of downtime. For sensors

XSL•FO
RenderX

deployed in the environment itself (as opposed to wearable sensors), devising a mounting scheme that will allow for easy replacement may also help.

To aid the tracking of the status of all sensors in a large study, planning in an upfront manner to create semiautomated tools that monitor the state of each sensor as often as possible can help identify failures quickly, report them to personnel for maintenance, and further decrease data collection downtime [21]. Moreover, the use of automated tools may become a necessity if the number of participants, sensors, or hours of recording becomes large. Data-driven approaches for detecting and identifying anomalous sensor data streams have been recently proposed in the literature [22,23]. Implementing a strategy for automated ongoing maintenance of the deployed sensors is much easier once the research team has direct access to recent sensor data. A data flow framework (presented in a later section; *Information Flow Layers*) outlines and describes the communication channels that carry sensor data to the data servers (the collection of systems where the data are collected and securely stored for later processing). Researchers can use this framework to plan communication paths for each sensor and then set up a script to run on the research server, which monitors these data streams and notifies assistants when sensors malfunction. For example, automatic programs can be used to assess the quality of electrocardiography (ECG) signals and give feedback to the research support staff about potential fitting and usage problems [24].

Specific logistics and deployment strategies will be unique to each research study, and they will largely be influenced by the restrictions and constraints imposed by the research environment. For example, some hospitals require all equipment to be powered using 3-prong plugs; therefore, all sensor chargers are required to be used through 3-prong adapters. Other restrictions may include Wi-Fi availability, permission to mount sensors on the walls, availability of charging ports for sensors, and space for sensor storage, to name a few. Permissions for the research personnel to access all areas in which the study takes place should also be considered.

Data Loss

Data loss may occur for several reasons, including sensor or data pipeline malfunctions, poor participant compliance, and attrition, among other reasons. For example, sensors may fail to deliver data, as they run out of battery power or break, or they may fail to deliver when network outages interfere with data transfers. Subjects may also neglect the data collection protocol (including forgetting to wear the sensor or wearing the sensors without following instructions), forget to recharge a worn device, or fail to upload data at the end of each session, for example, the *Hexoskin* garment requires manual data upload via Universal Serial Bus (USB). In more extreme cases, subjects may become frustrated with the study and elect to drop out, thereby reducing the total amount of available data.

The key to mitigating these various sources of data loss is being aware of where in the data stream pipeline the losses occur. The *Data Acquisition and Flow Framework* section of this paper enumerates the communication paths that help carry sensor data to their destination on a research server. Once researchers have decided on a sensor suite, and once they know which paths are required, small scripts or monitoring systems can be instrumented to test or infer status of each communication channel and report failures to the research team. For cases where data loss occurs at the source (ie, the participants), this section describes a mechanism for sending feedback to the participants to notify them of the data loss and encourage them to remedy it.

Data Signal Quality and Unintentional Variability

Related to data loss, the signal quality of sensor data is a concern that presents a substantial challenge for research in the wild. The term *signal quality* used here refers to the ability of each sensor to measure its signal(s) of interest. Poor data quality may occur when sensors are not properly worn or maintained, such as when a wristband photoplethysmography (PPG) sensor to measure heart rate is worn too loosely or when a microphone is obscured. Instances of improper or inconsistent sensor usage are inevitable in large studies in the wild, and they can lead to an unintentionally higher degree of variability in the data captured across all participants, which may consequently skew the resulting statistical analyses.

Early steps should be taken to ensure that participants receive proper training for using the adopted sensors before the study begins and that clear and accessible instructions are made available to serve as a reminder. Making plans to monitor the quality of sensor data streams so that appropriate actions can be taken to rectify problems is also highly beneficial, especially for long-term studies. Once a process is in place to determine the quality rating of recent data, different intervening actions may be appropriate, depending on the participant population, study environment, and the goals of the research project. Some example interventions for improving data quality include the following: retraining participants in sensor usage, adjusting sensor fit, improving the network infrastructure to reduce downtime, or simply sending reminders to participants (eg, smartphone push notifications) to remind them to wear their devices and upload the data. Quickly responding to rectify data quality drops can help preserve the value of the data and minimize data loss. If low-quality data persist despite these measures, automated signal enhancement methods may still be employed to algorithmically improve data quality. The *Data Acquisition and Flow Framework* section illustrates how data from the sensors can be aggregated on a research server.

Privacy and Security

Among the opportunities to generate scientific knowledge are significant challenges to the ethical conduct of research on human subjects [25]. Threats to privacy and data security constitute the greatest risk to participants of behavioral research in the wild. As sensing technologies become ubiquitous and data science advances, it is possible to use passively collected digital data to identify and predict a surprising range of human behaviors with increasing accuracy [26]. Participants are often unaware that when they consent to share data from their fitness tracker, they may be allowing researchers to infer information about their alcohol consumption, sexual activity, and mental health symptoms. The accidental or malicious release of this information could cause significant social, occupational, and

XSL•FO
RenderX

psychological consequences to participants. The informed consent process must provide clear and transparent communication about what data are collected and how the data will be used by researchers, how data are either anonymized or kept confidential, and how data are securely transferred, stored, and destroyed. Researchers must stay up to date on evolving privacy and security concerns and best practices for mitigating risk.

Another significant challenge when conducting studies in the wild is respecting and protecting the privacy of nonparticipants coincidentally present in the research environment. For research scenarios in which raw audiovisual data are collected, extra steps must be taken to ensure that either no personally identifiable information (PII) is recorded about nonparticipants or that they are informed that they may be recorded, where appropriate and depending on municipal or state regulations and institutional review board (IRB) approval. A tactic for avoiding the collection of PII, even accidently, is to immediately transform the collected raw data streams, such as audio or video, into anonymized features—intonation, mel-frequency cepstral coefficients, gestures, and posture—and record these instead [27].

Another important step toward maintaining privacy is to ensure secure transmission of sensor data to the research server with as few transfers to intermediate nodes as possible. The *Criteria Related to Protection of Human Subjects* section in this paper discusses methods for securely transmitting information across a network, and the *Data Acquisition and Flow Framework* section can help researchers plan secure communication paths.

Data Acquisition and Flow Framework

State-of-the-art electronics and sensing technologies offer a wide variety of communication protocols for sending information among devices. Selection of the appropriate sensors for a research project depends on many factors, which are discussed in more detail in the *Considerations and Criteria for Sensor Selection* section. A crucial step toward evaluating each sensor is to understand the ways in which its data can be transmitted through different communication channels and how its data flow may be affected by the choice of other sensors and data hubs.

The proposed general sensing framework, deemed suitable for studies in the wild, depicts common transmission paths for data flowing from multiple sensor streams through disparate network paths and arriving on a secure server that is accessible only by the research team. The framework aggregates data in a single place, allowing for simpler implementations of automatic stream monitoring and participant feedback systems.

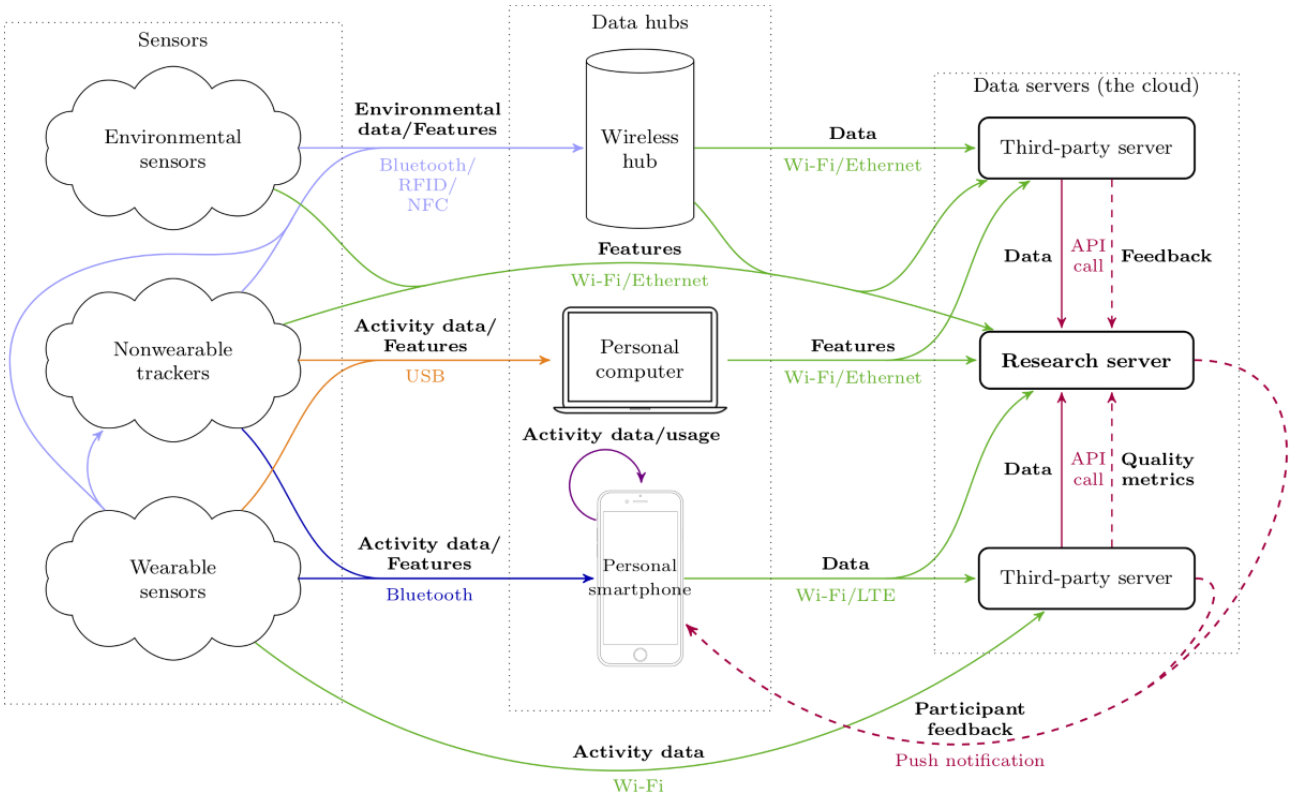
Information Flow Layers

Figure 2 depicts potential information pathways through different communication channels for passing data obtained from sensors (in the left column) to data servers (right column), where all the information passes through an intermediate data hub layer (middle column). These intermediate hubs are any devices that act as bridges to facilitate the aggregation and delivery of transient sensor data into long-term storage. Most of the available sensors in the market support a data flow matching some combination of paths in this figure.

The primary aim in this framework is the aggregation of all sensor data onto a single research server where additional processing, monitoring, and feedback can be performed. The following subsections describe each of these layers (columns) in detail.

XSL•FO
RenderX

Figure 2. A framework for studies of human behavior in the wild, showing common potential information pathways for data produced by sensors (eg, physiologic and activity), destined to be stored on a single research server. This type of data flow paradigm enables centralized data monitoring and facilitates immediate automatic participant feedback regarding data quality and compliance via the participant's smartphones. RFID: radio-frequency identification; NFC: near-field communication; USB: universal serial bus; API: application programming interface.



Sensors

Sensors for studies in the wild can broadly be grouped into 3 categories: environmental sensors, nonwearable (human) trackers, and wearable sensors.

Environmental Sensors

These devices passively capture information about their surroundings. Some examples of data captured by these types of sensors include the following: temperature/humidity/CO_2 levels, inertial measurements (eg, from accelerometers, gyroscopes, or magnetometers), and acoustics. These devices often perpetually broadcast information about their surroundings, using low-energy Bluetooth or radio-frequency identification (RFID) signals. Sampling rates below 1 Hz or event-based sampling techniques are typical, as environmental data usually change slowly (at least compared with physiological signals).

Nonwearable Trackers

These devices are placed in the environment, and they capture information about subjects and their behaviors indirectly or in a passive way. A few instances of these types of devices include the following: RFID scanners, Doppler effect and under-the-mattress sleep trackers, infrared gaze trackers, and video cameras. These sensors often operate on wall power and may include network capabilities for simplifying data transmission to long-term storage on the data servers. They also often include companion websites or smartphone apps for visualizing metrics extracted from the sensor data.

Wearable Sensors

These types of sensors encompass the set of custom-built or consumer products that are worn or carried on the subject's body to collect physiologic or contextual data or features extracted from data, for example, heart rate from electrocardiogram data, for behavior and activity tracking. Devices such as smart watches/wristbands, smart undergarments (underwear, T-shirts, and bras that collect data), smart rings, voice activity detectors, and smart shoe soles are some examples. Many of these devices can be recharged for long-term use over multiple sessions, and they generally either communicate via Bluetooth with companion apps installed on users' smartphones or via USB connections with personal computers. The companion apps tend to provide visualizations of the received data and upload functionality for long-term storage on third-party servers. Interestingly, when running certain tracking apps in the background, smartphones themselves can also serve as wearable sensors, collecting information about user movement and smartphone usage patterns.

Data Hubs

Data hubs are devices dedicated to collecting, aggregating, and transmitting sensor information to data servers. Transient data sources, such as many environmental sensors, have little memory, and they need to have their data collected continuously and retransmitted to a data server for long-term storage. Wi-Fi–enabled data hub devices with Bluetooth capabilities can serve as conduits for these types of data streams, whereas personal computers can act as data hubs for USB-only sensors. Battery-powered sensors collecting data at a high rate usually

XSL•FO
RenderX

communicate via USB, as the bandwidth and transmission speed are higher, and as wireless data transmission drains more power. Battery-powered sensors can afford to send a smaller amount of data through Bluetooth, whereas smartphones can often serve as both data hubs and data visualizers.

Data Servers

The term *data servers* refers to the collection of machines in which all the sensor data are stored. A typical consumer off-the-shelf sensing device will provide some pipeline for getting data off the sensor and into a data store in the cloud, usually owned by the sensor product's company. These companies often provide an application programming interface (API) for accessing the data, using automated tools that transmit the data securely to protect subject privacy. Eventually, all the sensor information needs to be aggregated in a single place, the research team's own server, so that the team has permanent and easy access to it. The aggregation of all sensor data on this server continuously throughout the data collection process enables monitoring and feedback systems to help manage some of the challenges mentioned in the *Studies in the Wild: Challenges and Risk Mitigation Strategies* section.

Considerations and Criteria for Sensor Selection

Minimizing participant risk and burden while maximizing the amount and quality of data is of primary importance. The set of sensors used plays a major role in a study's outcome, as data quality is inherently constrained by the sensors' characteristics and the participants' interactions with those sensors. Selecting the appropriate sensors to employ in a research study can be complicated, as the market provides many options, and each device has unique qualities and capabilities.

This section establishes a comprehensive list of the different criteria that should be considered before data collection begins. In practice, researchers must strike a balance between meeting their research objectives and ensuring a smooth participant experience to maximize attrition and minimize data loss. Both needs are constrained by the properties of the sensors that are available or can be produced. The criteria are partitioned according to whether each criterion is a characteristic of the sensor or whether it primarily concerns either the researchers or the participants. It is important for researchers to carefully review each one of these criteria, as they are highly connected, and each choice affects the outcome and experiences for both researchers and participants.

Table 1 lists different sensor criteria grouped according to whether they primarily concern research objectives and logistics, sensor characteristics, participant engagement, or human subject protection during the study. The categorization is not perfect, as some of the criteria pertain to more than 1 group, but it helps emphasize the different perspectives researchers should examine when selecting sensors. Key criteria are included in this table, which are expected to remain relevant as technology changes; however, there may be other factors worth considering, depending on the specific needs of a research project. The following subsections describe each criterion in detail.

Table 1. Considerations and criteria for sensor selection.

Research objectives and logistics	Sensor characteristics	Participant engagement	Human subject protection
Signals of interest	Sensor customizability	Cohort and individual suitability	Access and usability
Data properties and quality	Cost	Burden to participants	Privacy
Data access logistics	Battery life	—[a]	Data security
Sensor synergy	Operating system support	—	—
Additional experiment setup costs	Robustness	—	—
Sensor acceptance among target population	Provider support	—	—
On-site infrastructure requirements	—	—	—

[a]Empty cells are filled with a dash for visual clarity.

Criteria Related to Research Objectives and Logistics

The criteria in this section pertain to the logistical implications of the selected sensors and ways in which the selection affects the final outcomes and goals of a research project.

Signals of Interest

These criteria relate to the signals and how they are measured.

Target Signals

Before data collection begins, researchers need to consider what type of signals they want to measure from the participants or the environment. Varying amounts of potentially relevant information can be obtained from signals collected from different sources, such as physiology, for example, heart rate, breathing rate, electrodermal activity (EDA), behavior (eg, time spent speaking, sleeping duration and stage progression, number of steps per time interval, social interactions, and surveys), and the environment (eg, temperature, humidity, and CO_2 levels). The utility and overall quality of the chosen signals depend on the sensors' measurement mechanisms.

Measurement Mechanism

The physical mechanism through which a signal is acquired affects its quality and overall utility for future analysis. As an example, human location and kinematic data can either be reconstructed from a series of Global Positioning System (GPS) coordinates or inferred from an inertial measurement unit (IMU), such as an accelerometer and gyroscope. The GPS data tend to produce more accurate location measurements and less accurate

XSL•FO
RenderX

kinematic ones; the IMU location accuracy drifts over time but yields better kinematic figures, whereas GPS data can be used in the aid of calibrating step count from an IMU [28]. Another example is heart rate data, which can be obtained through PPG or ECG, each of which yields significantly different signal qualities and properties. The measurement mechanism may be constrained by a sensor's form factor requirements (wristband vs garment), which may limit the quality of data that can be obtained.

Data Properties and Quality

These criteria are important for assessing the quality and potentially undesirable aspects of gathered data.

Sampling Rate

For most consumer sensing technologies, the sampling rate is fixed by hardware design and power constraints, and it cannot be altered. It is always possible to decrease the number of samples considered for analysis purposes by downsampling data originally collected at a higher rate. However, upsampling data collected at a lower rate introduces distortions into the signal [29], and that may impact its utility for later analysis. The sampling rate of any sensing device should be at least twice that of the desired underlying signal's maximum frequency for the recording to provide reasonable fidelity (per the Nyquist-Shannon sampling theorem). The human voice, for example, can be characterized by pitch and formants (among many other features), which require sampling rates at least twice the maximum vocal frequency (typically greater than 8 kHz) for adequate analysis. However, tracking the position of a person inside of a building can be sampled around once per second, with meter-level accuracy based on average indoor walking speeds [30]. Researchers should be aware of the analytical power of the target signals and choose sensors capable of capturing data at a frequency where meaningful information can be extracted.

Signal-to-Noise Ratio

The data will only be useful if the signal-to-noise ratio (SNR) of the measurements is higher than a certain threshold. Noise in this case refers to any unwanted alterations to a signal during the measurement process, and it can appear for many reasons. If the noise is too high, it might not be possible to extract the relevant information from the measurements. For example, ECG-based heart pulse measurements may be subjected to noise when a participant moves or when the electrodes attached to the skin briefly detach during physical activity. Audio recordings of people socializing may also include unwanted background sounds. As unexpected sources of noise can occur in a research study, test runs with a small cohort should be conducted for sensors under consideration and then inspected to determine whether the SNR is adequate to extract meaningful information. Researchers may be able to improve a sensor's SNR by understanding where noise is introduced into the measurements and taking steps to reduce it.

Accuracy and Precision

Accuracy refers to the bias of the measurements, and precision is a representation of the variance of the measurements over time. High-accuracy (low bias) and high-precision (low variance) sensors are the most desirable. Published scientific validation studies pertaining to the accuracy and precision of measurements are available for some commercial and research-grade sensors. In situations where no previous validation work exists for a device, researchers should consider performing their own validation tests, using state-of-the-art, gold-standard sensors as the basis for comparison. As an example, in a study examining the measurement accuracy and precision of wrist-worn PPG devices (eg, Fitbit) among a diverse group of participants performing various physical activities, heart rate measurements were accurate to within 5% of clinical-grade devices, and the measured number of step counts varied within 15% of the actual number [31].

Drift

Measurement drift is a natural phenomenon that can occur in any sensor, caused by unintentional modifications to the device or object being measured [32]. When all other factors are held constant, measurements of a signal may drift up or down because of, for example, temperature or humidity shifts, changes in electrode impedance, or physical movement of the body. In many cases, drift is caused by physiological or environmental factors that cannot be controlled in the wild, but there are many common techniques for removing drift effects, including high-pass filters [33], adaptive filters [34], and time-variant filters [35]. In other cases, drift can be caused by sensor wear or material corrosion; therefore, it is important for research teams to consider the impact that normal usage and time will have on the sensors, and it is important to consider how this may cause a drift in the measurements.

Data Access at Various Stages of Processing

In some applications, it is important to be able to access the raw (unprocessed and unfiltered) signals. This is most relevant for research involving the denoising of signals, artifact removal, feature extraction, or even the estimation of other data streams from correlated signals [36]. Many consumer sensor devices provide preprocessed signals with artifacts already removed and which have been transformed into higher-level features, such as step count, heart rate, sleep quality, or physical readiness. Some sensor product companies elect to keep their preprocessing techniques unpublished; therefore, it can be difficult for researchers to understand exactly what each feature represents. These ready-made features can be useful for analysis, but researchers should be cautious when using features with no published methodology unless the features have been previously validated in scientific experiments. In cases where a provided feature cannot be trusted or is proven unhelpful in analysis, having access to the raw data to extract more meaningful features may be beneficial.

Data Access Logistics

These criteria concern the ease with which data are stored and accessed by researchers.

Data Upload Procedure

How and when data are transferred from sensor devices through the network to a data server can have a profound impact on a research project. As far as data upload procedures are concerned, there are 2 types of sensors: the ones that require manual

XSL•FO
RenderX

interaction and the ones that automatically and transparently upload data once configured. Manual interaction is often required for devices that collect a large amount of data and need to transfer it in bulk (eg, a Hexoskin ECG sensor uploads to a personal computer via USB). Automatic uploading is typically available for sensors that can stream information transparently to a data hub or smartphone over either Wi-Fi or Bluetooth (eg, an OMsignal ECG sensor uploads data wirelessly to a smartphone app). Both researchers and participants usually benefit from the automated paradigm, as there is less work involved for both parties, and data becomes available sooner, but the researchers need to consider its impact on smartphone battery drain and network bandwidth contention.

Ease of Data Access

Once the data have been successfully transferred from sensors to the data servers, data need to be stored on a research server that is easily accessible to the research team. Some sensors may be configured to upload information to the research server directly. For example, some companies supply a website where researchers can log in, visualize, and download participant data. Some companies track uploaded sensor data separately per user, in which case the research team would be responsible for creating and managing the participant accounts. Companies may provide tools to facilitate the download of data, such as Web-based (eg, REST) interfaces or APIs. The existence of well-documented guidebooks or a responsive technical support staff for these tools should be considered when selecting sensors.

Sensor Synergy

These criteria concern potential symbioses among sensors and signals.

Redundancy of Signals

There are situations in which measuring the same underlying signals using different measurement devices might be advantageous to a research effort. One such circumstance is when a sensor's accuracy and precision are unknown, but it is otherwise an appropriate pick for research. For example, if this device is a PPG-based wrist-worn sensor for heart rate tracking, then collecting heart information in parallel (perhaps on a subset of the participants), using an ECG sensor that has been validated against a gold standard, can enable researchers to infer the measurement quality of the PPG sensor. In a different scenario, researchers may decide that a certain signal is so important to capture in its full fidelity that using a single sensor that may occasionally fail or experience higher noise levels is not adequate. Using multiple sensors to capture the same signal adds fault tolerance to the measurement of that target signal, and this may also help reduce systemic measurement errors (eg, by averaging).

Sensor Versatility

Using a sensor that can adequately serve multiple purposes may be preferable to using multiple sensors instead. There are many reasons why this may be beneficial, such as cost, reductions in participant and research staff burdens, and simplicity. For example, it is possible to program a smartphone to gather human-produced audio and record participant proximity to known locations within a building by exploiting its Bluetooth or Wi-Fi connectivity. This approach uses a single sensor to achieve both goals instead of using 2 separate devices to capture each signal.

Additional Setup Costs

These criteria describe the (perhaps hidden) extra time and financial costs associated with setting sensors up for experiments.

Installation and Maintenance Costs

Once purchased, sensors require installation and maintenance throughout a research study to ensure measurement consistency and minimize data loss. Some sensors, such as Bluetooth beacons, may come packaged with installation tools that interfere with maintenance objectives (eg, double-stick tape for wall mounting). Using alternative installation devices (eg, adhesive Velcro strips) in anticipation of device malfunctions or required battery replacements can help expedite repairing or replacing these devices when necessary. This may add a small additional per-unit cost to some of the chosen sensors, but this can save time and may help save money in other ways.

Participant Training and Support

Participants who will wear sensors throughout a study should be trained to use these devices according to study rules and objectives. Generally, support staffing may be required, as the complexity and number of sensors increases or the sensors' robustness decreases.

Service Costs

Some companies, such as those producing sensors targeted for research rather than consumer use, may offer additional services for some cost. These services may include data aggregation and storage, data visualization, more convenient data access, or real-time monitoring and quality tracking for incoming data. Researchers should identify which services, if any, may be necessary.

Sensor Acceptance Among Target Population

Regardless of every desirable quality a sensor may possess for the research team and objectives, it cannot be beneficial if participants recruited from the target population will not accept or use it. There are many reasons why participants may reject any specific sensor, such as discomfort, obtrusiveness, complexity, or fashionability. These objections cannot be anticipated fully; thus, researchers should assess beforehand whether the target population would be generally willing to engage with the potential sensor set selected for use.

On-site Infrastructure Requirements

Studies conducted in the wild, which use sensors, depend on the study site infrastructure. As researchers converge on a set of desired sensors for a specific study, the infrastructural resources necessary at the study site(s), which can satisfy the sensor requirements, will emerge. In some cases, the existing infrastructure may not provide the resources required, but it can sometimes be augmented (eg, with additional wireless data hubs or power extension cables) to suit the needs of the research project. Supplementing the infrastructure may not be possible in other situations because of costs or prohibitive regulations, and researchers may have to settle for less desirable sensors

XSL•FO
RenderX

with fewer requirements. Some other examples of the infrastructural considerations that should be accounted for include the following: the total network bandwidth usage for all participants, the availability of power and network outlets, and access to a secured network for sensitive or private data transfer.

Criteria Related to the Evaluation of Sensor Characteristics

The criteria in this section describe various ways to evaluate sensors compared with other potential sensor choices. Each choice poses a certain set of constraints on the study, which can affect the research team, the study objectives, and the participants; thus, this merits vigilant consideration.

Sensor Customizability

These criteria address the alterability of sensor functionality.

Hardware Design

Presently, most commercial sensors are not designed with extensibility or hardware-level customization in mind. Therefore, it is difficult to alter the sampling rate, storage capacity, or battery life to suit the needs of a research study. There exist customizable do-it-yourself (DIY) hardware platforms (eg, Arduinos or Raspberry Pis) that researchers may want to consider in cases where no existing ready-made option is sufficient.

Software Customization

Many sensors on the market, which stream data to a smartphone, have a companion phone app, typically providing data visualization, high-level data summaries, or some types of behavioral interventions (eg, *stand up and stretch*, or *get extra sleep tonight*). Some devices, such as smartwatches, contain their own displays for visualizing data and haptic feedback for alerts and interventions. These features can be useful to participants, but they may misalign or interfere with the goals of a research study; therefore, customized versions may be desired.

Certain sensors offer software development kits, enabling researchers to build their own software for collecting, visualizing, and storing sensor data. Other devices, such as the Apple Watch or Wear OS–enabled gear, support software extensions installed on the device, giving researchers more control over the visual and haptic feedback to suit the needs of a study.

Cost

The total monetary cost of a sensor device itself is an important factor for researchers to consider, and it may impact the total number of participants who can be recruited and supported throughout a study. Sensor prices can sometimes be negotiated with their providers, depending on the number of devices desired.

Battery Life

Sensor battery lives vary greatly and depend on the device types and their functionality. On the basis of the survey of devices available today, wearable sensor battery life spans range from several hours to nearly a week. Most devices are rechargeable in just a few hours, but researchers should offer suggestions to participants about when to recharge to maximize the analytical utility of the data. Some strategies for minimizing the impact of data loss caused by recharging are as follows: staggering the recharge periods for different participants (so at least some data are always present) and choosing recharge times that coincide with periods where the devices could not normally be worn anyway (eg, while sleeping or taking a shower). It is inevitable that participants will at times forget to recharge their sensors, and researchers should have a plan for handling these situations as well. Other devices, such as many tiny and portable environmental sensors, consume a small amount of power, and they can operate continuously for over a year. These devices are often not designed for recharging, and they may need to be replaced throughout the research study.

Operating System Support

Some wearable sensors designed to stream data to a smartphone companion app may only support phones running on a particular operating system (eg, iOS or Android), which can create difficulties for the research team. Researchers could elect to recruit only those participants with compatible smartphones, but this will introduce a selection bias that may impact the generalizability of the research findings or may reduce the number of potential participants. If researchers determine that a sensor with partial smartphone support is necessary, these negative effects could be mitigated by providing the interested participants using incompatible smartphones with a temporary and inexpensive compatible smartphone for use during the study.

Robustness

These criteria concern the ability of sensors to endure repeated use and proneness to failure.

Physical Design

Different sensors have distinct physical characteristics that make them more or less suitable for reliable operations over an extended period of time. Some properties worth evaluating are as follows: whether a device is sturdy and can handle mild physical wear, how easily a worn device may fall off, whether its buttons and other inputs function well after prolonged use, how well it stays in place without shifting, and how quickly it resumes operation after being reattached. Researchers should consider performing a pilot study to fully understand and evaluate the sensors beforehand.

Firmware

The reliability of a sensor's firmware is important, as any failure may lead to loss of data. A few probing questions worth answering are as follows: is the firmware code stable or does it crash? Can it handle a barrage of unexpected inputs and continue to function? If the device sleeps, does it resume data collection once awakened? Researchers should stress-test sensor firmware before committing to any device to ensure they understand the possible failure modes and recovery procedures.

Companion Software

Some sensors require a companion app running on a second device, such as a smartphone or computer, to facilitate data collection and long-term storage. This software needs to be

XSL•FO
RenderX

resilient to minimize data loss. Ideally, it functions consistently, showing no signs of glitches or crashing. Its ability to receive data from the sensor and either cache or upload data to a data server should be seamless and fault tolerant. Research teams should stress-test this software to understand when and how it fails, so that the support staff will be prepared to help participants. A few tests worth performing are as follows: disconnecting the sensor from the app or removing the network uplink during a data transfer to see if sensor data are lost and switching foreground apps or providing random inputs to see if the app crashes. Once the failure modes of the companion software are understood, steps can be taken to remedy them or at least to alert participants.

Provider Support

Some companies are interested in building a scientific reputation for their sensor products; therefore, they are concerned with supporting research studies. This support comes in a few forms, and the criteria below pertain to the beneficial impact this support can have during and on the outcome of data collection.

Prestudy Support

Before data collection, it is essential for researchers to fully understand the properties and unique characteristics of each sensor under consideration to make the most informed choices. Virtually all sensor providers offer documentation and a communication channel for answering specific technical questions. Some of these providers may offer additional services for research teams, including direct communication to key technical or support personnel and free samples for testing.

Logistics

Research teams should seek any available logistical aid, offered by the sensor product companies, that may help the study function more smoothly. Teams should ensure that sensors can be provided on time and that there is a backup plan for any sensors that need replacement. It is advisable to seek help from the product companies to train the research staff for proper fitting of the sensors, especially for those requiring specialized knowledge. Other kinds of logistical help may include preconfiguration of sensors (eg, to specific Wi-Fi networks), custom delivery options (packaging, rush shipping), tailored fittings, or an emergency contact. Moreover, some sensor providers offer ongoing assistance, ranging from providing quality metrics and statistical reports of the study participants to ensuring APIs support the types of data monitoring and quality assessment metrics researchers desire.

Criteria Related to Participant Engagement

These criteria pertain to how sensors affect the participants' perception of a research experiment and willingness to engage with a study throughout its duration.

Cohort and Individual Suitability

These criteria relate to the ability of sensors to meet the needs of members of a cohort.

Sizing and Fit

Garments and sensors that match each participant's unique physical characteristics are best equipped to provide usable data. Devices that are too large or too small can cause discomfort, possibly leading to side effects, such as blistering or reductions in data quality.

Technological Literacy

Each sensor provides a unique interface for operating with its hardware and companion app software. Researchers should ensure interfaces are simple enough for all potential participants in the target population. In cases where the interface is unfamiliar, researchers will need to provide instructions, describing not only how to operate and interact with the devices but also how to check that they are in a proper state and performing the desired function at any time.

Fashionability

The selected suit of sensors should comply with dress codes of the environment in which they will be worn. Moreover, the design and appeal to wear sensors should be considered by the research team to ensure that all participants are comfortable wearing the sensors from an esthetic perspective.

Burden to Participants

These criteria address the physical and mental burdens sensors impose on research participants.

Physical Interference

Obtrusive sensors may physically interfere with normal activities, causing frustration or eventually leading participants to avoid wearing these sensors or drop out of the study. For example, undergarment or chest strap sensors may become uncomfortable after a few hours or produce skin rashes, preventing participants from using them further. Another example is desk-mounted, infrared eye-tracking devices that require participants to keep their heads in view, which may incidentally encourage poor posture. Other job-specific scenarios should be considered, such as the use of smart rings in hospital settings, where they can interfere with minimal hygiene requirements. Sensors that can adequately collect the intended signals without interfering or causing discomfort will improve the participants' acceptance of the devices, potentially minimizing attrition.

Time Investment

Studies conducted in the wild, which ask participants to wear or interact with sensors over an extended period, inherently push more responsibility onto the participants to manage and operate the sensors. Daily upkeep, such as cleaning and charging the devices and verifying that they are functioning as intended, requires a time investment that burdens participants and can cause frustration if the demands are too high. Choosing sensors with low upkeep and training costs will reduce this burden and can improve compliance and overall data quality [37].

Cognitive Load

An implicit stipulation in any study is that the participants understand they are responsible for adhering to the study protocol. This requires that participants remain mindful of the study throughout its duration. Researchers should aim to choose sensors and an overall study design that requires a small or occasional investment of the participants' time and mental energy. For example, helping participants with reminders to

XSL•FO
RenderX

charge their sensors every night, and supporting them with a charging hub may increase sensor usage.

Criteria Related to Protection of Human Subjects

Research investigating human behavior, using sensing technology, is subject to review by IRBs, which evaluate the risks and benefits to human participants and ensure that the study adheres to ethical principles detailed in the Belmont Report [38]. Researchers must consider how the passive collection of behavioral data will respect participants' autonomy and privacy, how it will maximize the benefits of the research while minimizing risks to participants, and how it will ensure that benefits and risks are equitably distributed. Some of the most relevant themes are reviewed here, but it is important to be aware of ethical guidelines that apply to specific populations or data types. Connected and Open Research Ethics CORE provides a checklist to guide researchers in deciding which technologies are appropriate for a study, with respect to protecting human subjects [39].

Access and Usability

Researchers are responsible for ensuring that potential benefits of a study are likely to apply to all members of the population under investigation. This means that sensor selection must not inadvertently exclude members of the study population from participation or result in poorer data quality because of individual differences. For example, wearable sensors may be affected by factors related to body shape and size, skin tone, body hair, or tattoos. It would violate the ethical principle of justice to exclude individuals as study participants on the basis of these factors, solely as the sensors selected did not perform well on them. Researchers should aim to select sensors that have demonstrated validity across diverse participants (eg, a heart rate monitor that relies on ECG instead of optical technologies), can be adapted to individual differences (eg, a respiration monitor that can be worn on a bra or belt), and employ inclusive design features (eg, accessibility settings to accommodate those with visual impairments) to ensure equitable representation and data quality.

Privacy

Privacy refers to the persons' right to control what information about them is shared, with whom it is shared, and how these data are used. The most common privacy protection is to separate information that could identify the participant from the data collected about the participant, but some passively collected behavioral data are inherently identifiable and sensitive. For example, GPS features can predict depression symptom severity [40], and 95% of individuals can be identified with as few as 4 GPS data points [41]. Participants electing to engage in a study that requires the collection of sensitive and personal data need assurances that researchers will take steps to mitigate the risk that their behaviors can be linked to their identities.

Given the array of data types available through passive sensing technologies and the low cost of collecting data unobtrusively, it is tempting to collect as much data as possible. However, researchers are ethically obligated to only collect data that are pertinent to specific research questions. When possible, researchers should disable sensors that are irrelevant and securely dispose of data that are not specifically related to study aims. In addition, participants should be able to select which data they are willing to share, with whom, for what duration, and for what purpose. Ideally, sensors should allow participants to deny or revoke access to particular data types. If these user controls are not permitted by third-party providers, researchers should consider providing additional data management tools that help participants exercise their right to privacy.

Many sensors on the market today require participants to register their own accounts, using their own personal information, which creates a link between potentially sensitive data and each identifiable participant. Studies needing to access these data while guaranteeing participant privacy have a few options. Researchers could register dummy accounts, allowing the participants to remain anonymous, or they may alternatively acquire data directly from each participant's personal profile (eg, by using an API) and then immediately remove PII. In the latter case, researchers should also check that both the network channels from the sensors to data servers and the network channel for researchers to access the data are encrypted and secured to avoid any privacy breaches.

Data Security

Proper protection of the PII sensor data gathered from participants requires all communication channels for the data streams to be secured (refer to Figure 2), and it requires protected long-term data storage with limited accessibility. Information sent over a Bluetooth link is naturally secure, as only paired Bluetooth devices can communicate. Similarly, USB transfers are secured between the 2 devices at either end of the USB cable. Data sent over Ethernet or Wi-Fi require an extra encryption layer (eg, https, secure file transfer protocol) to ensure the information cannot be intercepted. RFID and near-field communications are generally not considered safe, but sensors are typically using these channels to infer information (eg, about the movement of people indoors) rather than transferring PII directly. Stored data are typically secured by limiting physical access to the storage device itself, but encryption of the data is also possible. Access to the sensitive stored data should be limited to select members of the research team, and it is usually controlled through credential-based authentication (eg, usernames and passwords). Unfortunately, today, there are many other ways for hackers to obtain PII data (eg, malware, spyware, and cyberattacks), and research teams and participants may wish for every precaution to be implemented. Readers are referred to the study by Filkins et al [42] for more information about protecting private data in a mobile sensing landscape.

Current Sensing Technology

Much research has remarked on the variety of options and capabilities of sensors for research purposes [43-47], including several tests [48-50] and validation experiments [31,51-54]. These studies overlook the qualities of each sensor, which make them more or less suitable for different research applications.

This section provides a snapshot of some of the more recent and prominent unobtrusive sensing technologies worth considering for human behavior studies in noncontrolled

XSL•FO
RenderX

environments. Figure 3 provides a visual overview of this landscape. The following subsections give basic descriptions of these sensors, along with comments about the research trade-offs among different technologies. Compiled in mid-2018 while preparing a research effort examining the relationship between human behavior patterns at home and work, as well as mental states and job performance, the snapshot provided here covers a wide range of sensors that can capture many different signals. The current pace of innovation in the sensing market is too rapid to summarize in its entirety. Therefore, although this list is not exhaustive, it provides an approximate overview of the currently available products. The authors hope these suggestions will be beneficial to researchers formulating large-scale studies.

To make this compilation easier to understand, the sensing technologies are categorized with respect to the type of information they capture. At the highest level, the sensors are grouped into 3 categories: environmental sensors, nonwearable trackers, and wearable sensors, which are also reflected in Figure 3. *Environmental sensors* encompass the devices that capture information about the surrounding environment, which are not intended to directly measure information about people. *Nonwearable trackers* describe devices that are placed in an environment and capture information about people and their behaviors. *Wearable sensors* are the portable devices worn or carried by people, which capture physiological or behavioral data.

Environmental Sensors

As the *Internet of Things* movement continues to push more technology into portable devices, environmental sensors that capture multiple kinds of data from their ambient surroundings have become common. Many of the devices in this category are small, battery powered, and can easily be stuck onto walls or placed out of the way. Although their primary purpose is to collect and report environmental data, these sensors can also be used to capture other kinds of information. For example, protocols have been developed for proximity awareness and location-based services, including iBeacon, Eddystone, and Quuppa.

This section focuses on the environmental sensing capabilities of these multipurpose sensors, particularly on measurements of light, sound, and atmosphere. Figure 3 shows a breakdown of these properties and lists several products available on the market, which can be used for measuring each type of data. Environmental sensors are typically designed to remain turned on and collect information from their surroundings at a fixed rate. This information is usually made available to surrounding data hubs or smartphone devices so it can be stored or monitored by people (see Figure 2).

Atmosphere and Light

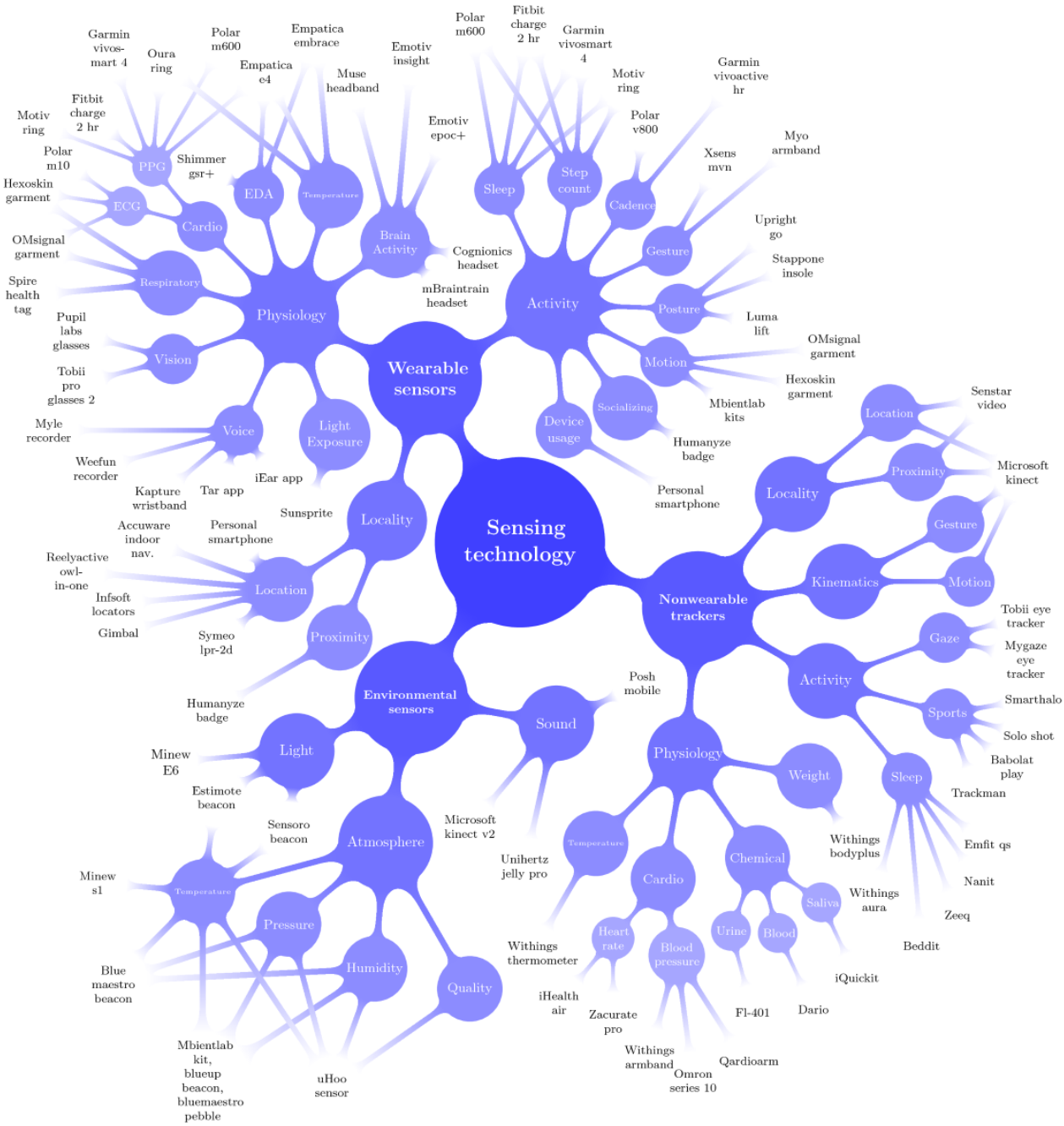
Many of the environmental sensors listed in Figure 3 use microelectromechanical systems devices to measure various properties of the environment. This technology has enabled sensors to be miniaturized and deployed with year-long battery lives, but this comes with a small cost in measurement accuracy. The underlying chips used to assess light levels [55], air pressure [56], and air quality [57] have very low measurement errors; therefore, any inaccuracies will likely be imperceptible at a human scale. The chips often used to measure temperature and humidity have standard errors that may be more significant: approximately 0.5°C and 3% relative humidity, respectively [58]. Near the boundaries of acceptable temperature ranges for controlled indoor environments [59], an error of 1°C may represent the difference between someone feeling comfortable or not. Having this kind of measurement error means researchers would not be able to distinguish between these 2 states. Deploying duplicate sensors in the same environment provides some redundancy in measurement, which allows these errors to be averaged out and may offer more discriminative power during analysis.

Sound

Portable sensor products for measuring or classifying ambient sounds are not widely available on the market. The Microsoft Kinect camera is a device capable of capturing audio (and video), but it requires a connected computer, operates on wall power, and is not easily portable. Researchers interested in capturing audio with portable devices in an unobtrusive manner have a few options. The Electronically Activated Recorder [60] periodically samples ambient audio from a smartphone in short bursts throughout the day. Several DIY platforms are also purchasable, such as the Raspberry Pi and Arduino, which can be customized to record microphone audio and transmit acoustic information via Wi-Fi or Bluetooth. Audio signal quality is highly dependent on the setup of the microphone [61] and the types of sounds that researchers want to capture or filter out. The DIY platforms typically offer small omnidirectional microphones with midrange quality, but higher-grade microphones can be integrated. If human-produced audio is of primary interest, researchers may consider a similar solution, such as TILES Audio Recorder (TAR) [27], which is described further in the *Wearable Sensors* section.

XSL•FO
RenderX

Figure 3. A snapshot of current consumer and research sensing technologies for human behavior studies in natural environments. This is not an exhaustive diagram of sensors on the market, but it provides an overview of the kinds of data can be captured using readily available technology. PPG: photoplethysmography; ECG : electrocardiography; EDA: electrodermal activity.



Nonwearable Trackers

Nonwearable trackers live in the environment and monitor and capture information about human behavior. These sensors can be categorized according to the types of signals they are primarily designed to capture: activity, kinematics, locality, or physiology. Some of these devices are self-contained and offer both a data collection mechanism and a means to transmit the information to a data server without a data hub. As these types of devices are not constrained by portability, they can sample data at a higher rate and may also offer more functionality. This section concentrates on nonwearable trackers designed for use at home or in an office.

Activity

Mentioned below are some of the activities that consumer nonwearable devices can track.

Sleep

Sleep monitoring sensors are installed in the bedroom, and these can capture information about the ambient environment and personal physiological information of someone while resting. These sensors may be installed under the mattress, on a nearby nightstand, or on a wall near the bed. They can collect rich information about sleep patterns, tossing and turning, sleep cycles and duration, snoring, temperature, humidity, breathing,

XSL•FO
RenderX

and heart rate, and they can ultimately provide details that can be used to assess the quality of sleep.

Sport Activities

Many unique kinds of nonwearable sensors used to track sport activities are available, including radars, cameras, tablet apps, and sensors built into sport equipment itself. Some sensors come integrated into sports gear, such as bikes, balls, helmets, and even tennis rackets, each of which blurs the line between wearables and nonwearable trackers. These devices record data, such as posture, gesture type, gesture accuracy (eg, golf swing), and practice duration, and these devices offer approximations for exercise intensity and calories burned.

Kinematics

Mentioned below are some kinematic features of human movement that can be captured by consumer sensors.

Motion

Human body movement dynamics have been linked to affect and cognition [62]. Camera-based computer vision systems allow noninvasive, scalable, and inexpensive motion tracking. A critical issue with camera-based tracking systems is the big brother effect (see *The Big Brother Effect* section). To mitigate participant concerns, researchers may extract relevant kinematic features from the videos in real time and store only relevant features, instead of the video recordings themselves. If physical space is available and the budget permits, using multiple cameras can improve tracking accuracy. Among the available off-the-shelf devices, the Microsoft Kinect [63], particularly with its depth camera technology, is a common choice for real-time full-body motion capture and gesture recognition. Motion tracking of multiple people in a fixed space is possible using cameras, with solutions such as CrowdVision.

Gesture

The camera-based systems described in the *Motion* section can also be used to track gestures, especially the Microsoft Kinect [63]. There are freely available tools, such as the Gesture Recognition Toolkit [64], that enable gesture recognition from any video source. Gesture recognition for multiple bodies in real time from single-camera sources is an active area of research, and there are no consumer products available.

Locality

Descriptions of sensors that capture rough estimations of position (proximity) and more precise location measurements are provided below.

Proximity

Modern personal smartphones support Bluetooth and can be programmed to broadcast Bluetooth packets at fixed rates for proximity sensing [65]. Proximity is measured using Bluetooth hubs that receive these broadcasts and use the received signal strength indicator values to determine a smartphone's approximate distance from each hub. This information can be used, for example, to monitor the amount of time participants spend in front of their workstations or in break rooms.

Location

Participant location data are a valuable information channel for human behavior monitoring, offering a means to track the movement patterns of individuals. Like proximity sampling and kinematics tracking, camera-based systems have been widely used for localization in indoor settings [66]. These camera-based solutions require the positions and orientation of each camera to be known. Usually, the cameras are placed out of the way and remain stationary, but they may require maintenance if perturbed. RFID systems are also available on the market, and they use distributed hubs to localize individuals wearing RFID badges. Usually, the location of individuals can be inferred more accurately when more cameras or hubs are added to the environment. It is recommended that research teams test localization systems in the target environment before committing to a solution.

Physiology

As shown in Figure 3, nonwearable sensors can measure a wide variety of physiological signals, including weight, body temperature, chemical indicators, and cardiovascular information.

Weight and Body Mass Index

Some smart scales can measure weight (mass) and assess a subject's body mass index (BMI) by using bioelectrical impedance through the bare feet [67]. Devices, such as the Withings Body+, capture both of these measurements [68], and they can also assess a person's total body water (TBW) [69]. Weight and BMI values are linked to obesity risk and heart disease [70], and TBW is linked to subjects' hydration levels [71]. These signals may be of interest to certain research studies, and presently, these signals are difficult to measure using wearable sensors.

Body Temperature

Body temperature is a known proxy measure for health and arousal [72]. Infrared thermography devices [73] measure skin surface temperature without direct body contact. These types of sensors are fast and easy to operate, but the accuracy and precision may be worse for some individuals [74]. Researchers desiring to use these types of no-contact body thermal sensors are encouraged to average several measurements to improve data quality.

Chemical Indicators

Nonwearable smart sensors are available, and they can sample blood, urine, or saliva to measure blood glucose level, blood oxygen level, and pH level. Measurements from these body fluids provide rich information about health and well-being, and they are considered the gold-standard source for some signals. These devices may require additional consumables for measurement (eg, electrochemical strips).

Cardiovascular

A variety of available nonwearable sensors provide cardiovascular information, such as blood pressure and heart rate. For example, pulse oximeters collect heart rate and blood oxygen saturation from pulse waves measured via PPG at the

XSL•FO
RenderX

fingertips. Pulse oximeters are widely used, given their ease of use and low cost. Smart blood pressure sensors are also available, such as the Withings or Omron 10 Series blood pressure monitors. Although these sensors must be attached to the body to capture cardiovascular data, they are not intended to be worn for any length of time. Recently, noncontact, video-based methods of inferring heart rate have been proposed and shown to obtain accurate and reliable measurements as well [75].

Wearable Sensors

Wearable sensors are separated into 2 primary groups, 1 for devices that directly capture physiological measurements and 1 for devices that infer behavioral or activity states. A third group encompasses sensors that infer information about the positioning of individuals.

Physiology

This subsection describes the (relatively unobtrusive) commercial sensors that can capture some of the wide variety of human physiological signals.

Electrodermal Activity

EDA, also known as galvanic skin response, can be used to track states of emotional arousal through the skin conductance level and responses. Sensors placed around the fingertips are among the most accurate, but they are more intrusive and likely to interfere with the participants' daily activities. Wristband sensors, such as the Empatica E4, measure EDA, but they can capture a large amount of noise in the signal when wearers move or flex their arm muscles. In practice, it can be difficult to capture this signal reliably and unobtrusively in the wild, without substantial noise. Denoising these signals to obtain more meaningful measurements is an active area of research [76].

Speech

Different wearable devices have been proposed by researchers for understanding emotions and other aspects of speech in social situations, such as the Sociometer [77], the EAR [60] and subsequent iEAR app, and the TAR [27] app. Privacy is a major concern when audio recordings are collected in public settings; thus, some apps, such as TAR, are designed to only collect and record anonymized acoustic features from human-produced audio. Commercial wristband devices for collecting raw audio in the wild are available, such as the Kapture audio wristband or the Weefun voice recorder, but these are designed to capture audio on demand and at the request of the wearer. The EAR and iEAR apps autonomously and periodically record ambient audio, but they may pose a privacy concern, especially for nonparticipants.

Vision

Eye trackers allow researchers to study human gaze patterns and points of interest that attract visual attention. Wearable consumer products for tracking gaze typically look like glasses and use cameras to track eye movement relative to each participant's forward head direction. Some of these products require calibration, where users are asked to look at a fiducial marker to realign the calculated gaze direction. Researchers should be aware that participants may need to calibrate periodically to maximize gaze tracking accuracy.

Cardio

Heart rate and heart rate variability measurements have been linked to activity levels, emotional arousal, stress, restfulness, and general fitness [78]. Wearable sensors are well suited to track the heart's behavior. It is currently possible to obtain unfiltered 200 Hz electrocardiograms throughout the day with chest straps, smart shirts, and undergarments. Wristband sensors offer PPG technology that collects volumetric measurements of blood flow. PPG-based wristbands provide heart rate and, sometimes, blood volume pulse metrics, but researchers often cannot access the unfiltered PPG data, as these are processed and transformed by the hardware in the device (usually to save power). Presently, heart rate information, as provided by PPG sensors, may not be accurate (see *Data Properties and Quality*), as evidenced by the study by Benedetto et al [48]; therefore, researchers should exercise caution when using this technology.

Respiration

Many chest strap [79] devices that capture ECG data are also capable of measuring respiratory information. These devices use stretch sensors in the strap wrapped around the chest to capture inhalation and exhalation and produce breathing frequency and volume per breath measurement. There are also other accelerometry-based devices that attach to the waist and extract the same measures by ignoring all motions, except those caused by breathing. Respiration sensors offer insight into physical activity intensity, recovery, and calmness when participants are at rest [80].

Temperature

Skin temperature provides information about participants' comfort levels, exercise efficiency, and physical well-being (eg, because of a fever). Some wearable garments that already measure ECG and respiration rate (eg, QardioCore) can capture skin temperature conveniently. Some wristband sensors can measure skin temperature as well (see Figure 3).

Brain Activity

Devices measuring brain activity using electroencephalography (EEG) have become more abundant in recent years. Numerous portable and wearable EEG headsets exist, with varying numbers of electrodes for capturing voltage levels at the scalp, and each one offers different sampling rates and monitors activity in different regions of the brain (Brodmann areas). Many of these portable headsets transmit data via Bluetooth, enabling smartphone apps to receive, process, and upload the data. Some devices require the data to be received and processed by a companion app running on a personal computer. Although the underlying technology is very similar for most devices and a standard exists for electrode placement (the 10-20 system), the captured signals may vary from device to device because of several factors: dry versus wet electrodes, sampling rates, number of channels, and degree of sensitivity to ambient (noisy) electromagnetic radiation.

XSL•FO
RenderX

Light Exposure

Sunlight exposure has been linked to sleep behavior and overall mental health. Special-purpose ultraviolet and visible light devices with smartphone integration, such as the Sun-Sprite light tracker, are available and can help researchers monitor participants' exposure to sunlight during the day. Many of the wearable wristbands that capture heart rate and other physiologic information can also track ambient light levels (see Figure 3).

Activity Tracking

Unique sensors have been developed to track various physical and contextual human activities, which are outlined here.

Socializing

Some wearable technologies can capture information about person-to-person social interactions. Active RFID or Bluetooth devices, such as the Humanyze badge, detect when 2 or more people are standing next to and facing each other. These devices and other wearable voice detectors (previously mentioned in the *Physiology* section) can also help determine when vocal exchanges occur between people nearby. Together, these types of sensors can aid researchers in assessing when group socialization occurs, how long it lasts, and who is involved. In practice, it is often difficult to determine who is speaking, and when anonymized audio features are collected instead of raw audio [81]; deciding whether a vocal utterance is intended as part of a conversation is an open research question. Nevertheless, even noisy inference of group social activity may be beneficial for a research endeavor.

Gestures

Human gestures contain valuable information in social contexts and provide insight into kinesiologic activity while at work or at home. Wearable sensors for gesture tracking commonly use IMUs to record motion of the arms and legs, and they may use electromyographic muscle sensors to detect certain kinds of hand gestures. The data can be streamed in real time via Bluetooth to smartphones or data hubs.

Motion

Many research efforts have previously observed a strong correlation between physical activity and both physical and mental health. With the explosion of fitness trackers in the last few years, it is now possible to track body movement and exercise patterns with relatively inexpensive wearable devices. These devices come in a variety of form factors with very different qualities and captured signal characteristics. One of the most common forms is the wristband, with other options including chest straps, shirts, and undergarments. Many of these devices contain embedded accelerometers and gyroscopes to record translations and rotations of the body over time. In addition, some devices may directly provide time series IMU data, whereas others may digest this to produce higher-level motion features, such as step count or distance traveled.

Posture

Some smart shirts and chest straps (mentioned in the previous section) can also be used to track the posture of participants. These devices contain IMU data that can be processed to obtain information about whether the wearer is sitting or standing, what his or her angle is with respect to the ground, and about sleeping posture (eg, resting on one's back, front, or side). Other sensors for gesture tracking attach to the arms and legs and can help provide a more holistic view of the entire body's posture over time.

Cadence

Cadence measures capture the consistency of repetition of motion over time and offer information about physical fitness, activity intensity, and physical exertion. Some devices provide cadence measures directly (usually for physical activity), but cadence can also be inferred and analyzed from other available motion modalities, such as IMU data.

Sleep

The *Nonwearable Trackers* section mentions nonwearable devices for tracking sleep, but many unobtrusive wearable devices can do so as well. Some wristbands and smart garments that track heart rate and motion can detect when a person is sleeping, and they can infer information about sleep stages and sleep quality from these data streams. Accurate sleep stage tracking from these types of data streams is a continuing area of research, and although many devices offer sleep metrics, they have not been validated thoroughly by the scientific community. These reported values may still be useful, but researchers should be cautious when using these sleep metrics.

Step Count

Step count can be inferred from wearables using IMUs that track motion and posture; therefore, it can be measured using wristband and garment devices. Bluetooth-enabled pedometers worn at the waist or tied to shoes are other options. Newer iPhones can also easily measure step count, which is calibrated using GPS information. The number of steps counted for any given activity will vary among devices and across participants because of differences in motion and how the motion is interpreted by the sensors. Particularly, wristband sensors may misinterpret activities involving repetitive arm motion (eg, washing dishes) as steps and introduce measurement error.

Personal Device Use

Apps for tracking smartphone usage have been developed to help people monitor and manage their own time, and they can also be used for research purposes. These apps primarily track how frequently users pick up their smartphones and how long they spend using different apps each day. Some of these tools can also track social media and internet use.

Locality

Some wearable sensors are able to provide coarse approximations of location (proximity) and finer location estimates (localization), which are described below.

Proximity

Proximity-based locality measurements yield rough estimations of location by proximity to other *a priori* known locations, usually measures through time-of-flight or received signal strength. The systems available for tracking proximity use a deployment of Bluetooth, Wi-Fi, or RFID hubs, with known

XSL•FO
RenderX

locations, to track the presence of wearable devices. Wearable badges and cards can be purchased for each participant, which are detectable by these hubs, but it is also possible in some cases to track devices that participants might already have on their person, such as smartphones. These proximity tracking systems are useful for detecting potential social interactions among nearby people or detecting when people are present in a known area.

Localization

Localization is the process of measuring or inferring a precise approximation of a person or object's location. Most smartphones today provide GPS-based location services, which provide accurate location measurements that are useful for tracking human movement at city scales. Other systems, such as dense Bluetooth or RFID hub network, can be used in conjunction with Bluetooth beacons or RFID cards carried by participants, and they provide precise estimations of their position in indoor environments. Location data are highly sensitive, and extra steps may need to be taken to securely collect, deidentify, and transmit this type of information.

Results

Overview

In early 2018, a research team (including the authors) began preparations for an *in situ* study at the University of Southern California's Keck Hospital, per the MOSAIC program [82], using sensors to track nurse and hospital staff behavior in the workplace and at home. The project aimed to understand how physiological dynamics and behavior both at work and at home are associated with personality, well-being, and work performance.

This section shares results from the application of the methods previously described. The team's experiences and rationale for selecting sensors to help achieve the research objectives are discussed, as well as how compliance was monitored and encouraged during the study. Metrics for attrition and compliance rates are provided. For a more detailed overview of the data collection itself, including IRB information, readers are referred to the retrospective study by Hasan et al [83]. A full description of the dataset and collection methodology will appear in a future publication; this section focuses on aspects related to sensing and data flow.

Study Goals and Constraints

The primary goal of MOSAIC was to use information gathered through commercially available sensors to study the predictive power of these types of sensors for assessing personality traits, as well as work-related behaviors and mental states throughout time. Owing to the complex trade space encompassing consumer sensors, creating a data collection protocol that met the project goals and was satisfactory to the participants and hospital environment required many iterations and challenging decisions. These deliberations and the data collection protocol that resulted led to a study, including over 200 hospital staff participants over a 10-week period and with an attrition (dropout) rate of 4% (primarily because of vacation conflicts).

Signals and Sensors

The signals of interest and sensor selection rationale are described below. The sensors employed in this case study, on the basis of the various study constraints, are also described.

Signals of Interest

Previous literature, related studies, and experience all revealed many physiologic signals of interest for capturing data likely related to work behaviors and mental states. Some of these signals, such as EDA and brain waves, were not possible to capture accurately in the wild over extended periods using consumer sensors. The research team initially reduced the list of potential signals of interest down to the ones that could be captured with unobtrusive sensors, based upon a survey of existing technologies (see Figure 3). Table 2 shows these signals and a short explanation of the expected utility for each in meeting the research objectives.

Table 2. Signals of interest in the case study that were measurable using consumer sensors.

Signal	Reason for interest
Cardiac	Connection to exercise, fitness level, and stress levels [84,85]
Physical activity	Linked with stress [86]
Sleep	Health (physical and emotional) [87,88]
Speech	Contains information about emotional expressions [89] and information related to social interaction
Breath	Calmness, stress, anxiety, and speech activity detection [80,90]
Environment and distractions	Connection with workplace performance, anxiety, and stress [91]
Locality	Captures workplace behavior and job role dynamics [92] and context for the job types of interest

Sensor Selection Rationale

As the study continuously required the collection of data over several months, one of the top priorities was to minimize the burden on participants to achieve a high compliance rate and capture representations of behavior in the wild. As previously described in the *Burden to Participants* section, the study took a holistic approach to assessing the participants' responsibilities and duties, including their time invested in compliance, physical disruption, cognitive load, and interference with their daily activities. While keeping these burdens in mind, each paragraph below describes how sensors were chosen to capture each signal of interest.

XSL•FO
RenderX

Cardiac and Physical Activity

Several form-fitting garments with ECG sensors were tested, and many provided the data quality desired (see Figure 3 for the list). Chest strap sensors were found to be uncomfortable for daylong use (as they are designed for exercise sessions), but the existence of different form factors of ECG garments (eg, shirts, bras) made it possible to gather high-quality data across genders. Some of these garments continuously collected high-quality data throughout the day, but they required that the physical box recording the data and hidden inside was hooked up to a computer via USB on a daily basis for data transfer. This step seemed cumbersome for participants; therefore, another similar garment that could stream the data to the subjects' personal smartphones was selected. The caveat with this second device was its companion app, which required a manual start and stop of the data recording process. The research team elected to have subjects wear these garments only during work hours to avoid potential discomfort associated with wearing them all day. Participants were also assisted in setting location-based reminders on their personal phones to start and stop the recordings. Heart-related information and other physical activities outside of work were also tracked by asking participants to continuously wear a wristband.

Sleep

Many unobtrusive sensors were capable of capturing information about sleep duration and sleep stages. Some sensors required a one-time installation on or near the bed, and then they would automatically detect and monitor participants when the participants were sleeping. Nurse focus groups had privacy concerns; therefore, wearable sensors were deemed more appropriate. To minimize cost and the burden of wearing multiple sensors, a wristband sensor was chosen, which was capable of capturing sleep and the cardiac and physical activity signals mentioned previously. Participants were asked to wear the band every day, including during sleep.

Speech

At the time of the study, no portable consumer devices were available for automatically sampling only human-produced audio. The research team programmed a smartphone app to automatically start, run in the background, and collect audio samples of ambient human utterances [27]. To address Health Insurance Portability and Accountability Act concerns about hospital patient and nonparticipant privacy, relevant information about the emotional content of the voice signal was computed by the device, and the raw audio signal was immediately discarded. Moreover, participants could disable the recording process for intervals of half an hour, by pressing a button in the app, after which the recording was resumed. Collecting low-noise audio required the smartphone's microphone to be placed near the mouth, and the research team wished to avoid using external microphones to avoid further participant burdens. Research staff met with representatives from the potential participant pool to discuss unobtrusive solutions and discovered that hospital personnel were already accustomed to wearing hospital badges on their lapels. Credit card–sized smartphones were acquired to run the custom software, and then these were attached to the participants' shirts, with a clip to get the microphone closer to the mouth [27], as shown in Figure 4. Although this solution may have been unacceptable for some subject populations, it was appropriate for the hospital workers in this study.

Figure 4. Setup of the TILES audio recorder [27].

Breath

Commercially available portable breath sensors measured the expansions and contractions of the chest. Some of these sensors were stand-alone devices attached to the waist or chest, and some were integrated into other multipurpose sensing garments. Once the research team decided on a comfortable device for capturing ECG, they found that breathing rate information was already available; therefore, the same device was used.

Environment and Distractions

Environmental sensors for capturing temperature, humidity, and door motion were used. Statistics about social media and general phone usage were acquired with the participants' permission and with the help of smartphone apps running in the background on their personal phones, requiring little power and no interaction after the initial setup.

Locality

Precise localization of subjects inside the hospital was deemed prohibitively expensive and would have required several months of installation time; therefore, approximate measurements of location by proximity to known locations were used instead. As described previously in the *Wearable Sensors* section, using a dense hub network and wearable consumer sensors, there were 2 general ways to achieve this: tracking participants' smartphones or tracking other worn wireless communication devices. The latter option was chosen using the audio recording phones for tracking to avoid any power draw from participants' personal phones.

XSL•FO
RenderX

Selected Sensors and Expected Use

Table 3 shows the selected sensors and the intended usage period for participants, per the study protocol.

Table 3. Selected sensors and their expected use.

Sensor	Measurements	Intended usage period
Fitbit Charge 2	Photoplethysmography-based heart rate, step count, and sleep	24 hours per day
OMsignal garments	Electrocardiography-based heartbeat, breath, motion	At work (12-hour shifts)
Unihertz Jelly Pro	Audio features, Bluetooth-based localization	At work (12-hour shifts)
reelyActive's Owl-in-One	Bluetooth-based localization, data hub for environmental sensors	Installed at the University of Southern California's Keck Hospital, 24 hours per day
Minew E6, E8, S1	Light, motion, temperature, and humidity	Installed at the University of Southern California's Keck Hospital, 24 hours per day

Data Flow

Figure 2 depicts a general flow of information for measurements obtained through sensors. In the study, all 3 kinds of sensors (in the left column) were used: environmental sensors, nonwearable, and wearable. All of the sampled data flowed through 2 different intermediate types of data hubs: Bluetooth data hubs connected to Wi-Fi and personal smartphones. Personal computers were not used to retrieve any data in an effort to reduce the time spent by participants uploading data to different servers.

Wireless passive sensors capturing information about light levels, temperature, and humidity were used, which transmitted information over Bluetooth. In addition, the participants wore Jelly Pro phones that were programmed to send Bluetooth pings with unique identifiers. The Owl-in-Ones received the data. They were connected to the public Wi-Fi network of the hospital and transmitted the data over this network to reelyActive's servers, from which the data were retrieved in real time, using a provided API.

Audio data recorded by Jelly Pro phones were directly uploaded to the research server, using hospital or home Wi-Fi networks. Wi-Fi was necessary because of the size of the files, approximately 8 GB per day.

Data transfer took place from Fitbit Charge 2 devices to participants' smartphones over Bluetooth, followed by data upload to Fitbit's servers through the smartphones' internet connections. The research server then retrieved these data using Fitbit's API. The same flow was employed by the OMsignal garments, using OMsignal's API.

Feedback for participants happened through a custom app (the TILES app) via push notifications. This app sent surveys to participants and gave them notifications about sensor usage and the quality of their previously received data when necessary.

Monitoring and Encouraging Compliance

Minimizing participant frustration in a study can help improve compliance and overall data quality [14]. This was one of the top priorities in this case study, and this was achieved by reducing cognitive burdens on participants, offering monetary incentives and consistent feedback to participants for compliance and providing convenient help whenever the participants encountered difficulties.

A custom smartphone app for the participants was developed, and it served as the primary resource for all aspects of the study. This app provided progress and monetary reward tracking, information about the study and protocol, and direct contact links for requesting help, and it also distributed questionnaires and reminders. Participants were rewarded for uploading their data daily, per the study protocol, which allowed the research team to monitor compliance and data quality every night. Each morning, the app provided feedback to the participants by letting them know whether their previous day's data had been received and whether the quality was sufficient. If the data were missing or quality was poor, the app reminded participants to double-check their sensors or seek help from the research team.

On-site assistants were always available during work hours to help participants who encountered difficulties during the study. Participants were able to drop in for help, or they could request for assistants to visit them and provide in-person support. These assistants actively engaged with participants who had recently uploaded poor-quality data to help make sure their devices were worn and functioning properly.

Metrics

Table 4 shows the average data compliance rates across different 10-week waves of this study for different sensors. The attrition rate was under 4% across all participants, and most of the participants dropped out because of vacation time conflicting with the study's participant inclusion criteria. More details about the study, including information about poststudy surveys on user experience, are available in the study by Hasan et al [83].

Figure 5 shows a histogram of the number of hours each sensor was used per day across all participants, where days with no logged data are not shown. This figure illustrates that on average, Fitbit was used about twice the amount of time as other sensors, which was in line with expectations. Moreover, although both the Jelly Pro and OM garments were designed to be used by participants at work, there is a noticeable difference in usage. This is partly explained by participants starting the recording of their OM garments at home rather than at work. It can also be explained by the fact that the Jelly Pro recording is activated only when participants or nearby persons are speaking.

XSL•FO
RenderX

The Fitbit usage of the subject cohort is in line with other studies [93], which claim 70% to 90% of compliance using wristband sensors. Compliance rates for the OM garments and the Jelly Pro are and were expected to be lower, as these devices required more attention from participants.

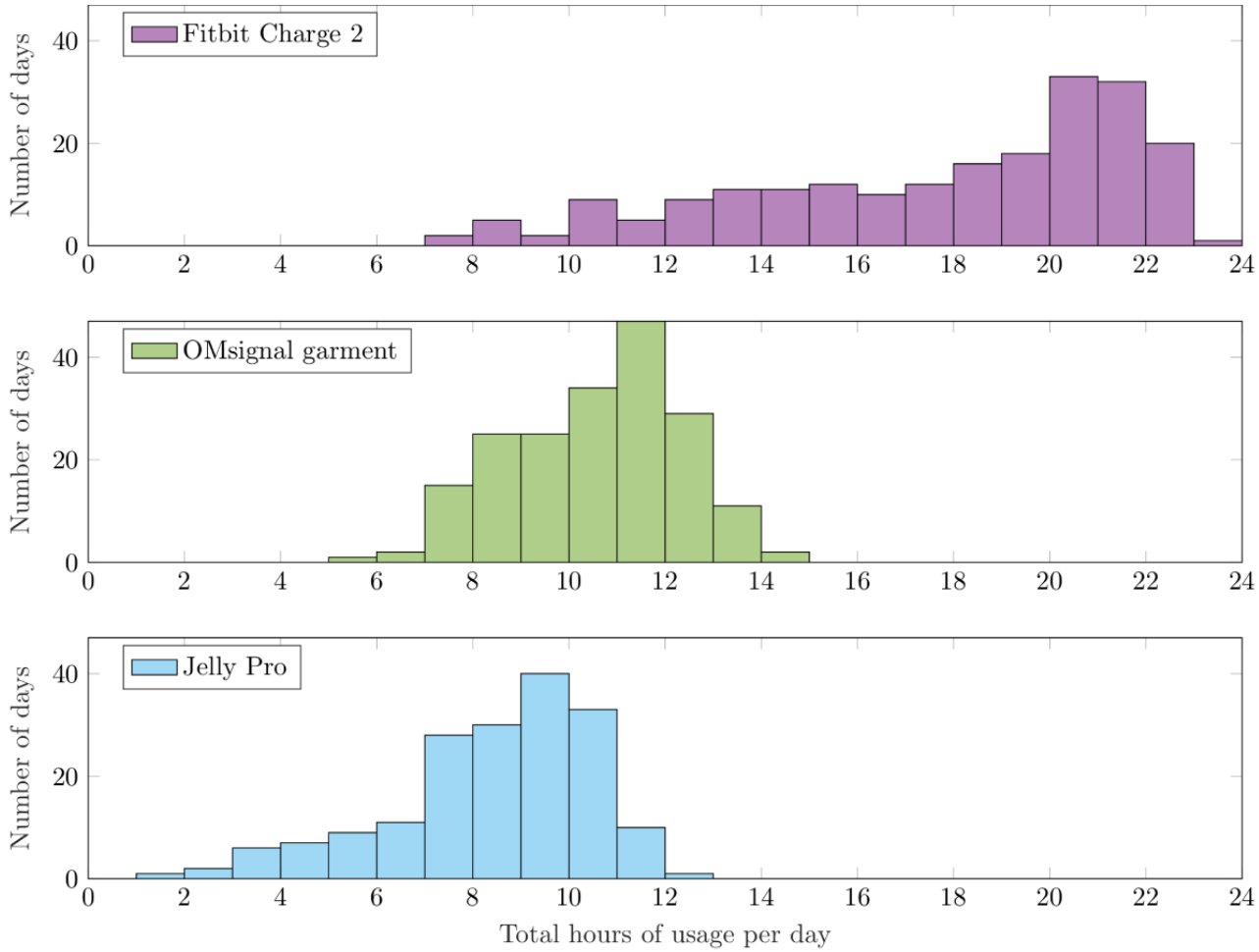
For Fitbit, the mean usage among all days with logged data is 17.8 hours, with an SD of 4.0 hours. For the OM garments, the mean is 10.6 hours, with an SD of 1.8, and for the Jelly Pro audio recorder and localizer, the mean is 8.4, with an SD of 2.1 hours.

Table 4. Compliance rates for participant-tracking sensors (n=212) and environment sensors (n=244) in the case study.

Sensor type and signals	Sensors	Participant who opted, n (%)	Total hours	Compliance rate[a], n (%)	Definition of compliance
Participant-tracking					
Cardio, sleep, and steps	Fitbit	208 (98.1)	236,725	152 (73.1)	Average fraction of days per participant with >12 hours of data
Cardio, breath, and motion	OMsignal	208 (98.1)	44,240	125 (60.1)	Average fraction of work days per participant with >6 hours of data
Audio	Jelly	184 (86.8)	37,065	131 (61.8)	Average fraction of work days per participant with >6 hours of data
Locality	Jelly+Owl-in-one	184 (86.8)	37,065	131 (61.8)	Average fraction of work days per participant with >6 hours of data
Environment					
Temperature, humidity, and motion	Minews	—	—	239 (98.0)	Uptime of the sensor network

[a]Compliance is computed as the presence of data exceeding half of the measurement period per day among the participants who opted in for each sensor.

Figure 5. Histograms of the total number of hours of recorded sensor data per day, across all participants. These plots only show data from days where data was logged.



XSL•FO
RenderX

Discussion

Principal Findings

The methods previously presented are evaluated with respect to the case study outcomes in 2 ways: participant compliance and the number of emergent unexpected challenges during the data collection period. More details are provided in the following subsections, but they can be summed up in the following manner: (1) participant compliance was satisfactory, yielding a large average number of hours of data per participant per day; (2) the unexpected challenges were manageable and had either very short-lived or isolated impact on the study.

These metrics suggest that the methods and mitigation strategies presented in this paper as a guide for researchers are helpful for sensor selection and management during longitudinal human behavior studies in the wild.

Participant Compliance

Participant compliance rates in the case study fall within an expected range when compared with observed compliance ratios in similar study conditions from the study by Merilahti et al [93]. The work from the study by Lima et al [94] observes that participant compliance decreases over time in long-term studies, which is also observed in this study. Overall, a sufficient number of hours of data per day per participant are collected for statistical analysis; therefore, the compliance levels are satisfactory.

Unexpected Challenges During the Case Study

This section recounts the unanticipated challenges encountered during the study despite efforts to avoid them during study planning. Unexpected challenges are defined as the events that were deemed unlikely to happen or that were not considered *a priori*, and these negatively affected the project budget, schedule, participants, or data. Each of the occurrences below were either isolated incidents, affecting only a narrow piece of the research project, or were short-lived, as the research staff was able to address them quickly. The following subsections present potential strategies for mitigating each of these events in future studies.

Shipping Dependencies and Customs

Some bundled sensor shipments were delayed because of product dependencies on secondary companies with limited shipping capacities. Urgent sensor package shipments from other countries were sometimes held up by the customs authority. In future studies, it would be best to be aware of the shipping capabilities of each product company and any potential shipping delays when preparing a study schedule.

Installation Time

The research staff underestimated the time required to install on-site sensors at the hospital. Although floor plans were used heavily for placement planning, they did not include locations of the electrical outlets. Several iterations and supplemental cabling were needed to install sensors across 16 different nursing units with similar layouts but different electrical circuit restrictions. Moreover, as most sensors were installed in patients' rooms, more trips to the data collection site were needed than expected to accommodate patient needs. Starting the installation process early can help researchers identify this problem in advance and budget time accordingly.

Battery Life

The Jelly Pro devices running the custom TAR app ran out of power for some of the participants early on during data collection. The parameters of the TAR app were tuned on the basis of the data collected during a pilot study from a subset of the final participant pool, but this subset did not reflect the worst-case scenario for power consumption. The battery life in this case depended on how many times vocal audio recording was triggered by the automatic voice activity detector, and the hospital staff in highly social environments triggered it more often than the worst case in the pilot study. The research team responded by recollecting the Jelly devices and modifying the parameters overnight. A possible strategy for mitigating this issue would be to design a pilot study that includes more participants at the expected extremes of the measurement spectra, but this may negatively affect the expected average-case findings. Perhaps a better strategy would be to implement tools to remotely or more easily update the parameters for all participants in anticipation of this type of issue.

Sensor Synergy

As the Jelly Pro devices served 2 functions in this study (collecting vocalized audio and proximity detection), when the power consumption exceeded expectations, 2 data streams were affected instead of 1. For sensors serving multiple purposes, there is greater risk to the data quality when they fail; therefore, proper stress testing and tooling (as mentioned in the previous paragraph) should be prepared before the main study.

Sensor Discomfort

Some participants acquired rashes caused by skin friction with the wrist-worn or undergarment sensors. This occurred because the sensors these participants used were improperly fitted or sized, and the discomfort they produced led to a short-term loss of data while the participants recovered. The pilot study helped the research staff identify and mitigate some fitting concerns, but it was not enough to handle all the cases during the main study. The team reached out to the product companies for these sensors to get help with proper fitting procedures, and with their guidance, they were able to find proper fits for each affected participant. Better approaches for mitigating the risk of data loss here would be to solicit help with fitting and sizing from the product companies earlier and then incorporate that wisdom into the study (as mentioned in the *Provider Support* section), as well as consider different options for materials that are in contact with the skin (eg, Fitbit offers wristbands of different materials).

Data Pipeline Failure

Months into the main data collection, 2 site-wide disconnections of the environmental and proximity sensors occurred. These devices were all connected to the existing hospital Wi-Fi network, and the research server's data monitoring processes identified this event immediately. Within 24 hours, research staff was dispatched to manually power cycle the devices and

XSL•FO
RenderX

ensure they reported gathered data upstream. Although these sensors were stress-tested during the pilot study and determined to be robust to power and network outages, they did not all recover automatically in these 2 instances. Having a separate backup system in place (eg, an extra firmware layer to perform a soft reboot) may help improve robustness in these unexpected situations, but the data monitoring processes enabled researchers to respond quickly in this instance.

Conclusions

This viewpoint highlights and enumerates many of the research challenges faced during studies conducted in the wild, when using sensors for unobtrusively capturing human activity and behavior; presents a diagram illustrating information flow and an explanation of the roles of different computerized devices for data collection, transmission, and storage; and provides as a comprehensive list of criteria that researchers should carefully consider when conducting their own studies in natural settings, including explanations of trade-offs among them. The paper offers an overview of the state of current consumer technology for unobtrusive sensing in the wild, and it provides a snapshot of many of the products available for measuring different types of environmental, physiological, and behavioral data. The information presented is based on previous work and the team's experiences in executing a large-scale 10-week study for assessing human behavior, well-being, and performance in a hospital environment using a variety of sensors. The collection of methods and criteria for sensor selection and management were evaluated, using this study, with respect to compliance rates and the impact of unexpected emergent challenges that arose during data collection.

Acknowledgments

The research is based on work supported by the Office of the Director of National Intelligence (ODNI), Intelligence Advanced Research Projects Activity (IARPA), via IARPA Contract No 2017-17042800005. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the ODNI, IARPA, or the US government. The US government is authorized to reproduce and distribute reprints for governmental purposes, notwithstanding any copyright annotation thereon.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Sensing checklist for studies in the wild.
[PDF File (Adobe PDF File), 836KB - jmir_v21i8e12832_app1.pdf]

References

1. Lomas N. TechCrunch. 2017. Global Wearables Market to Grow 17% in 2017, 310M Devices Sold, $30.5BN Revenue: Gartner URL: https://techcrunch.com/2017/08/24/global-wearables-market-to-grow-17-in-2017-310m-devices-sold-30-5bn-revenue-gartner [accessed 2019-06-20] [WebCite Cache ID 73rwTZndm]
2. Rogers Y, Marshall P. In: Carroll JM, editor. Research in the Wild: Synthesis Lectures on Human-Centered Informatics. San Rafael, California: Morgan & Claypool Publishers; Apr 04, 2017.
3. Sun FT, Yeh YT, Cheng HT, Kuo C, Griss M. Nonparametric Discovery of Human Routines From Sensor Data. In: Proceedings of the International Conference on Pervasive Computing and Communications. 2014 Presented at: PerCom'14; March 24-28, 2014; Budapest, Hungary. [doi: 10.1109/percom.2014.6813938]
4. Banovic N, Buzali T, Chevalier F, Mankoff J, Dey AK. Modeling and Understanding Human Routine Behavior. In: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems. 2016 Presented at: CHI'16; May 7-12, 2016; San Jose, California, USA p. 248-260. [doi: 10.1145/2858036.2858557]
5. Pierson E, Althoff T, Leskovec J. Modeling Individual Cyclic Variation in Human Behavior. In: Proceedings of the 2018 World Wide Web Conference. 2018 Presented at: WWW'18; April 23-27, 2018; Lyon, France p. 107-116. [doi: 10.1145/3178876.3186052]
6. Wang R, Chen F, Chen Z, Li T, Harari G, Tignor S, et al. Studentlife: Assessing Mental Health, Academic Performance and Behavioral Trends of College Students Using Smartphones. In: Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing. 2014 Presented at: UbiComp'14; September 13-17, 2014; Seattle, Washington p. 3-14. [doi: 10.1145/2632048.2632054]
7. Rogers Y, Connelly K, Tedesco L, Hazlewood W, Kurtz A, Hall RE, et al. Why It's Worth the Hassle: The Value of In-Situ Studies When Designing Ubicomp. In: Proceedings of the International Conference on Ubiquitous Computing. 2007 Presented at: UbiComp'07; September 16-19, 2007; Innsbruck, Austria p. 336-353. [doi: 10.1007/978-3-540-74853-3_20]
8. Bonsignore E, Quinn AJ, Druin A, Bederson BB. Sharing stories 'in the wild': a mobile storytelling case study using StoryKit. ACM T Comput-Hum Int 2013 Jul 1;20(3):1-38. [doi: 10.1145/2491500.2491506]

XSL•FO
RenderX

9. Benford S, Adams M, Tandavanitj N, Farr JR, Greenhalgh C, Crabtree A, et al. Performance-led research in the wild. ACM T Comput-Hum Int 2013 Jul 1;20(3):1-22. [doi: 10.1145/2491500.2491502]
10. Adams A, Fitzgerald E, Priestnall G. Of catwalk technologies and boundary creatures. ACM T Comput-Hum Int 2013 Jul 1;20(3):1-34. [doi: 10.1145/2491500.2491503]
11. Izmailova ES, Wagner JA, Perakslis ED. Wearable devices in clinical trials: hype and hypothesis. Clin Pharmacol Ther 2018 Dec;104(1):42-52 [FREE Full text] [doi: 10.1002/cpt.966] [Medline: 29205294]
12. Kanjo E, Al-Husain L, Chamberlain A. Emotions in context: examining pervasive affective sensing systems, applications, and analyses. Pers Ubiquit Comput 2015 Apr 25;19(7):1197-1212. [doi: 10.1007/s00779-015-0842-3]
13. Rennick-Egglestone S, Knowles S, Toms G, Bee P, Lovell K, Bower P. Health Technologies 'In the Wild': Experiences of Engagement With Computerised CBT. In: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems. 2016 Presented at: CHI'16; May 7-12, 2016; San Jose, California, USA p. 2124-2135. [doi: 10.1145/2858036.2858128]
14. Carroll JM, Rosson MB. Wild at home: the neighborhood as a living laboratory for HCI. ACM T Comput-Hum Int 2013 Jul 1;20(3):1-28. [doi: 10.1145/2491500.2491504]
15. Memarovic N, Langheinrich M, Cheverst K, Taylor N, Alt F. P-LAYERS -- a layered framework addressing the multifaceted issues facing community-supporting public display deployments. ACM T Comput-Hum Int 2013 Jul 1;20(3):1-34. [doi: 10.1145/2491500.2491505]
16. Sung YH, Kim HY, Son HH, Chang J. How to design in situ studies: an evaluation of experimental protocols. Restor Dent Endod 2014 Aug;39(3):164-171 [FREE Full text] [doi: 10.5395/rde.2014.39.3.164] [Medline: 25110639]
17. Whittemore R, Melkus GD. Designing a research study. Diabetes Educ 2008;34(2):201-216. [doi: 10.1177/0145721708315678] [Medline: 18375773]
18. Haouij NE, Poggi JM, Sevestre-Ghalila S, Ghozi R, Jaïdane M. AffectiveROAD System and Database to Assess Driver's Attention. In: Proceedings of the 33rd Annual ACM Symposium on Applied Computing. 2018 Presented at: SAC'18; April 9-13, 2018; Pau, France p. 800-803. [doi: 10.1145/3167132.3167395]
19. Healey JA. DSpace@MIT. 2000. Wearable and Automotive Systems for Affect Recognition From Physiology URL: https://dspace.mit.edu/handle/1721.1/9067
20. Zheng YL, Ding XR, Poon CC, Lo BP, Zhang H, Zhou XL, et al. Unobtrusive sensing and wearable devices for health informatics. IEEE Trans Biomed Eng 2014 May;61(5):1538-1554. [doi: 10.1109/TBME.2014.2309951] [Medline: 24759283]
21. Haghi M, Thurow K, Stoll R. Wearable devices in medical internet of things: scientific research and commercially available devices. Healthc Inform Res 2017 Jan;23(1):4-15 [FREE Full text] [doi: 10.4258/hir.2017.23.1.4] [Medline: 28261526]
22. Sivaraks H, Ratanamahatana CA. Robust and accurate anomaly detection in ECG artifacts using time series motif discovery. Comput Math Methods Med 2015;2015:453214 [FREE Full text] [doi: 10.1155/2015/453214] [Medline: 25688284]
23. Zou S, Liang Y, Poor HV, Shi X. Data-driven approaches for detecting and identifying anomalous data streams. In: Sejdic E, Falk TH, editors. Signal Processing and Machine Learning for Biomedical Big Data. Abingdon, England: Taylor & Francis Group; 2018:57-72.
24. Satija U, Ramkumar B, Manikandan MS. An automated ECG signal quality assessment method for unsupervised diagnostic systems. Biocybern Biomed Eng 2018;38(1):54-70. [doi: 10.1016/j.bbe.2017.10.002]
25. Nebeker C, Harlow J, Giacinto RE, Orozco-Linares R, Bloss CS, Weibel N. Ethical and regulatory challenges of research using pervasive sensing and other emerging technologies: IRB perspectives. AJOB Empir Bioeth 2017;8(4):266-276. [doi: 10.1080/23294515.2017.1403980] [Medline: 29125425]
26. Althoff T, Horvitz E, White RW, Zeitzer J. Harnessing the Web for Population-Scale Physiological Sensing: A Case Study of Sleep and Performance. In: Proceedings of the 26th International Conference on World Wide Web. 2017 Presented at: WWW'17; April 3-7, 2017; Perth, Australia p. 113-122. [doi: 10.1145/3038912.3052637]
27. Feng T, Nadarajan A, Vaz C, Booth B, Narayanan S. TILES Audio Recorder: an Unobtrusive Wearable Solution to Track Audio Activity. In: Proceedings of the 4th ACM Workshop on Wearable Systems and Applications. 2018 Presented at: WearSys'18; June 10, 2018; Munich, Germany p. 33-38. [doi: 10.1145/3211960.3211975]
28. Aggarwal P, Syed Z, Niu X, El-Sheimy N. A standard testing and calibration procedure for low cost MEMS inertial sensors and units. J Navig 2008 Mar 25;61(2):323-336. [doi: 10.1017/S0373463307004560]
29. Jerri A. The Shannon sampling theorem—its various extensions and applications: a tutorial review. Proc IEEE 1977 Nov;65(11):1565-1596. [doi: 10.1109/PROC.1977.10771]
30. Ekman F, Keränen A, Karvo J, Ott J. Working Day Movement Model. In: Proceedings of the 1st ACM SIGMOBILE Workshop on Mobility Models. 2008 Presented at: MobilityModels'08; May 26, 2008; Hong Kong, China p. 33-40. [doi: 10.1145/1374688.1374695]
31. Shcherbina A, Mattsson CM, Waggott D, Salisbury H, Christle JW, Hastie T, et al. Accuracy in wrist-worn, sensor-based measurements of heart rate and energy expenditure in a diverse cohort. J Pers Med 2017 May 24;7(2):pii: E3 [FREE Full text] [doi: 10.3390/jpm7020003] [Medline: 28538708]
32. Enderle JD, Bronzino JD, Mendelson Y. Sensitivity drift. In: Introduction to Biomedical Engineering. Third Edition. Burlington: Academic Press; 2012.

XSL•FO
RenderX

33. van Alsté JA, van Eck W, Herrmann O. ECG baseline wander reduction using linear phase filters. Comput Biomed Res 1986 Oct;19(5):417-427. [doi: 10.1016/0010-4809(86)90037-6] [Medline: 3769470]
34. Jane R, Laguna P, Thakor N, Caminal P. Adaptive Baseline Wander Removal in the ECG: Comparative Analysis With Cubic Spline Technique. In: Proceedings of Computers in Cardiology. 1992 Presented at: CinC'92; October 11-14, 1992; Durham, NC, USA p. 143-146. [doi: 10.1109/cic.1992.269426]
35. Sörnmo L. Time-varying digital filtering of ECG baseline wander. Med Biol Eng Comput 1993 Sep;31(5):503-508. [doi: 10.1007/BF02441986] [Medline: 8295440]
36. Cassani R, Narayanan S, Falk TH. Respiration Rate Estimation From Noisy Electrocardiograms Based on Modulation Spectral Analysis. In: Proceedings of the 2018 CMBEC41-ACCES23 Joint Conference. 2018 Presented at: CMBES'18; May 8-11, 2018; Delta Prince Edward Island, Charlottetown p. 1-6.
37. Harari GM, Wang W, Müller SR, Wang R, Campbell AT. Participants Compliance and Experiences With Self-Tracking Using a Smartphone Sensing App. In: Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers. 2017 Presented at: UbiComp'17; September 11-15, 2017; Maui, Hawaii p. 57-60. [doi: 10.1145/3123024.3123164]
38. Eckstein S, editor. The Belmont report: ethical principles and guidelines for the protection of human subjects of research. In: Manual for Research Ethics Committees. Cambridge: Cambridge University Press; 1978:126-132.
39. Nebeker C, Ellis RJ, Torous J. Development of a decision-making checklist tool to support technology selection in digital health research. Transl Behav Med 2019 May 23:pii: ibz074. [doi: 10.1093/tbm/ibz074] [Medline: 31120511]
40. Saeb S, Lattie EG, Schueller SM, Kording KP, Mohr DC. The relationship between mobile phone location sensor data and depressive symptom severity. PeerJ 2016;4:e2537 [FREE Full text] [doi: 10.7717/peerj.2537] [Medline: 28344895]
41. de Montjoye YA, Hidalgo CA, Verleysen M, Blondel VD. Unique in the crowd: the privacy bounds of human mobility. Sci Rep 2013;3:1376 [FREE Full text] [doi: 10.1038/srep01376] [Medline: 23524645]
42. Filkins BL, Kim JY, Roberts B, Armstrong W, Miller MA, Hultner ML, et al. Privacy and security in the era of digital health: what should translational researchers know and do about it? Am J Transl Res 2016;8(3):1560-1580 [FREE Full text] [doi: 10.1109/BigData.2016.7841101] [Medline: 27186282]
43. Patel S, Park H, Bonato P, Chan L, Rodgers M. A review of wearable sensors and systems with application in rehabilitation. J Neuroeng Rehabil 2012 Apr 20;9:21 [FREE Full text] [doi: 10.1186/1743-0003-9-21] [Medline: 22520559]
44. Acampora G, Cook DJ, Rashidi P, Vasilakos AV. A survey on ambient intelligence in health care. Proc IEEE Inst Electr Electron Eng 2013 Dec 1;101(12):2470-2494 [FREE Full text] [doi: 10.1109/JPROC.2013.2262913] [Medline: 24431472]
45. Custodio V, Herrera FJ, López G, Moreno JI. A review on architectures and communications technologies for wearable health-monitoring systems. Sensors (Basel) 2012 Oct 16;12(10):13907-13946 [FREE Full text] [doi: 10.3390/s121013907] [Medline: 23202028]
46. Giordano S, Puccinelli D. When sensing goes pervasive. Pervasive Mob Comput 2015 Feb;17:175-183. [doi: 10.1016/j.pmcj.2014.09.008]
47. Majumder S, Mondal T, Deen M. Wearable sensors for remote health monitoring. Sensors (Basel) 2017 Jan 12;17(1):pii: E130 [FREE Full text] [doi: 10.3390/s17010130] [Medline: 28085085]
48. Benedetto S, Caldato C, Bazzan E, Greenwood DC, Pensabene V, Actis P. Assessment of the FitBit charge 2 for monitoring heart rate. PLoS One 2018;13(2):e0192691 [FREE Full text] [doi: 10.1371/journal.pone.0192691] [Medline: 29489850]
49. OM Signal. 2017. Validation of Breathing Rate Algorithm During Running URL: https://omsignal.com/wp-content/uploads/2018/04/OMsignal-Note-2017-002.pdf [accessed 2018-11-12] [WebCite Cache ID 73s59n0Ew]
50. Paalasmaa J, Sarkanen T, Partinen M. Long-term sleep measurement with a smartphone-connected flexible bed sensor strip. Sleep Med 2013 Dec;14:e258. [doi: 10.1016/j.sleep.2013.11.627]
51. de Zambotti M, Goldstone A, Claudatos S, Colrain IM, Baker FC. A validation study of Fitbit Charge 2™ compared with polysomnography in adults. Chronobiol Int 2018 Dec;35(4):465-476. [doi: 10.1080/07420528.2017.1413578] [Medline: 29235907]
52. Hargens TA, Deyarmin KN, Snyder KM, Mihalik AG, Sharpe LE. Comparison of wrist-worn and hip-worn activity monitors under free living conditions. J Med Eng Technol 2017 Apr;41(3):200-207. [doi: 10.1080/03091902.2016.1271046] [Medline: 28078908]
53. Horton JF, Stergiou P, Fung TS, Katz L. Comparison of Polar M600 optical heart rate and ECG heart rate during Exercise. Med Sci Sports Exerc 2017 Dec;49(12):2600-2607. [doi: 10.1249/MSS.0000000000001388] [Medline: 29135785]
54. McCarthy C, Pradhan N, Redpath C, Adler A. Validation of the Empatica E4 Wristband. In: Proceeings of the IEEE EMBS International Student Conference. 2016 Presented at: ISC'16; May 29-31, 2016; Ottawa, ON, Canada p. 1-4. [doi: 10.1109/embsisc.2016.7508621]
55. Texas Instruments. OPT3001: Digital Ambient Light Sensor (ALS) With High-Precision Human-Eye Response URL: http://www.ti.com/product/OPT3001 [accessed 2019-05-31] [WebCite Cache ID 78mnoLrpt]
56. Servoflo. MS5540C: Miniature Barometer Module URL: https://www.servoflo.com/download-archive/data-sheets/download/438/607/17 [accessed 2019-05-31] [WebCite Cache ID 78mnvbWrI]

XSL•FO
RenderX

57. Mouser Electronics: United Kingdom. Air Quality Sensor B5W-LD0101-1/2: Highly Sensitive and Compact Air Quality Sensor URL: https://www.mouser.co.uk/ds/2/307/en_b5w-ld0101-1_2-1274538.pdf [accessed 2019-05-31] [WebCite Cache ID 78moFtDrq]
58. Shenzhen Minew Technologies Co. Ltd. Bluetooth 5.0 Wireless Temperature Sensor Beacon URL: https://minew.en.alibaba.com/product/60723619989-805195002/Bluetooth_5_0_wireless_temperature_sensor_beacon.html [accessed 2019-05-31] [WebCite Cache ID 78mopfacD]
59. Zhang H, Arens E, Pasut W. Air temperature thresholds for indoor comfort and perceived air quality. Build Res Inf 2011 Apr;39(2):134-144. [doi: 10.1080/09613218.2011.552703]
60. Mehl MR. The Electronically Activated Recorder (EAR): a method for the naturalistic observation of daily social behavior. Curr Dir Psychol Sci 2017 Apr;26(2):184-190 [FREE Full text] [doi: 10.1177/0963721416680611] [Medline: 28529411]
61. Crocco M, Cristani M, Trucco A, Murino V. Audio surveillance: a systematic review. ACM Comput Surv 2016 May 2;48(4):1-46. [doi: 10.1145/2871183]
62. Castellano G, Villalba SD, Camurri A. Recognising Human Emotions from Body Movement and Gesture Dynamics. In: Proceedings of the 2nd International Conference on Affective Computing and Intelligent Interaction. 2007 Presented at: ACII'07; September 12-14, 2007; Lisbon, Portugal p. 71-82. [doi: 10.1007/978-3-540-74889-2_7]
63. Microsoft Developer. Kinect for Windows URL: https://developer.microsoft.com/en-us/windows/kinect [accessed 2018-10-01] [WebCite Cache ID 78p3Eai0e]
64. Gillian N, Paradiso JA. The gesture recognition toolkit. In: Escalera S, Guyon I, Athitsos V, editors. Gesture Recognition. New York: Springer; 2017:497-502.
65. Zhuang Y, Yang J, Li Y, Qi L, El-Sheimy N. Smartphone-based indoor localization with Bluetooth low energy beacons. Sensors (Basel) 2016 Dec 26;16(5):pii: E596 [FREE Full text] [doi: 10.3390/s16050596] [Medline: 27128917]
66. Domingo JD, Cerrada C, Valero E, Cerrada JA. An improved indoor positioning system using RGB-D cameras and wireless networks for use in complex environments. Sensors (Basel) 2017 Oct 20;17(10):pii: E2391 [FREE Full text] [doi: 10.3390/s17102391] [Medline: 29053629]
67. Kyle UG, Bosaeus I, de Lorenzo AD, Deurenberg P, Elia M, Gómez JM, Composition of the ESPEN Working Group. Bioelectrical impedance analysis--part I: review of principles and methods. Clin Nutr 2004 Oct;23(5):1226-1243. [doi: 10.1016/j.clnu.2004.06.004] [Medline: 15380917]
68. Gallagher D, Visser M, Sepúlveda D, Pierson RN, Harris T, Heymsfield SB. How useful is body mass index for comparison of body fatness across age, sex, and ethnic groups? Am J Epidemiol 1996 Feb 1;143(3):228-239. [doi: 10.1093/oxfordjournals.aje.a008733] [Medline: 8561156]
69. Watson PE, Watson ID, Batt RD. Total body water volumes for adult males and females estimated from simple anthropometric measurements. Am J Clin Nutr 1980 Jan;33(1):27-39. [doi: 10.1093/ajcn/33.1.27] [Medline: 6986753]
70. Pi-Sunyer X. The medical risks of obesity. Postgrad Med 2009 Nov;121(6):21-33 [FREE Full text] [doi: 10.3810/pgm.2009.11.2074] [Medline: 19940414]
71. Powers JS, Choi L, Bitting R, Gupta N, Buchowski M. Rapid measurement of total body water to facilitate clinical decision making in hospitalized elderly patients. J Gerontol A Biol Sci Med Sci 2009 Jun;64(6):664-669 [FREE Full text] [doi: 10.1093/gerona/glp018] [Medline: 19228780]
72. Herborn KA, Graves JL, Jerem P, Evans NP, Nager R, McCafferty DJ, et al. Skin temperature reveals the intensity of acute stress. Physiol Behav 2015 Dec 1;152(Pt A):225-230 [FREE Full text] [doi: 10.1016/j.physbeh.2015.09.032] [Medline: 26434785]
73. Lahiri BB, Bagavathiappan S, Jayakumar T, Philip J. Medical applications of infrared thermography: a review. Infrared Phys Technol 2012 Jul;55(4):221-235. [doi: 10.1016/j.infrared.2012.03.007]
74. Cheung BM, Chan LS, Lauder IJ, Kumana CR. Detection of body temperature with infrared thermography: accuracy in detection of fever. Hong Kong Med J 2012 Aug;18(Suppl 3):31-34 [FREE Full text] [Medline: 22865221]
75. Poh MZ, McDuff DJ, Picard RW. Non-contact, automated cardiac pulse measurements using video imaging and blind source separation. Opt Express 2010 May 10;18(10):10762-10774. [doi: 10.1364/OE.18.010762] [Medline: 20588929]
76. Taylor S, Jaques N, Chen W, Fedor S, Sano A, Picard R. Automatic identification of artifacts in electrodermal activity data. Conf Proc IEEE Eng Med Biol Soc 2015;2015:1934-1937 [FREE Full text] [doi: 10.1109/EMBC.2015.7318762] [Medline: 26736662]
77. Choudhury T, Pentland A. Sensing and Modeling Human Networks Using the Sociometer. In: Proceedings of the Seventh IEEE International Symposium on Wearable Computers. 2003 Presented at: ISWC'03; October 21-23, 2003; White Plains, NY, USA. [doi: 10.1109/iswc.2003.1241414]
78. Shaffer F, Ginsberg JP. An overview of heart rate variability metrics and norms. Front Public Health 2017;5:258 [FREE Full text] [doi: 10.3389/fpubh.2017.00258] [Medline: 29034226]
79. Villar R, Beltrame T, Hughson RL. Validation of the Hexoskin wearable vest during lying, sitting, standing, and walking activities. Appl Physiol Nutr Metab 2015 Oct;40(10):1019-1024. [doi: 10.1139/apnm-2015-0140] [Medline: 26360814]
80. Loudon RG, Lee L, Holcomb BJ. Volumes and breathing patterns during speech in healthy and asthmatic subjects. J Speech Hear Res 1988 Jun;31(2):219-227. [doi: 10.1044/jshr.3102.219] [Medline: 3294505]

XSL•FO
RenderX

81. Nadarajan A, Somandepalli K, Narayanan SS. Speaker Agnostic Foreground Speech Detection From Audio Recordings in Workplace Settings From Wearable Recorders. In: Proceedings of the International Conference on Acoustics, Speech and Signal Processing. 2019 Presented at: ICASSP'19; May 12-17, 2019; Brighton, United Kingdom. [doi: 10.1109/icassp.2019.8683244]
82. The Intelligence Advanced Research Projects Activity. Multimodal Objective Sensing to Assess Individuals with Context (MOSAIC) URL: https://www.iarpa.gov/index.php/research-programs/mosaic [accessed 2019-05-31] [WebCite Cache ID 78mpye3w7]
83. Hasan M, L'Hommedieu J, Begay C, Schenone A, Dimitropoulou L, Margolin G, et al. Lessons learned: recommendations for implementing a longitudinal research study using sensors in an organizational setting. J Med Internet Res 2019:1-46 (forthcoming). [doi: 10.2196/preprints.13305]
84. Kim HG, Cheon EJ, Bai DS, Lee YH, Koo BH. Stress and heart rate variability: a meta-analysis and review of the literature. Psychiatry Investig 2018 Mar;15(3):235-245 [FREE Full text] [doi: 10.30773/pi.2017.08.17] [Medline: 29486547]
85. Lee JM, Kim HC, Kang JI, Suh I. Association between stressful life events and resting heart rate. BMC Psychol 2014;2:29 [FREE Full text] [doi: 10.1186/s40359-014-0029-0] [Medline: 25628890]
86. Anderson E, Shivakumar G. Effects of exercise and physical activity on anxiety. Front Psychiatry 2013;4:27 [FREE Full text] [doi: 10.3389/fpsyt.2013.00027] [Medline: 23630504]
87. Goldstein AN, Walker MP. The role of sleep in emotional brain function. Annu Rev Clin Psychol 2014;10:679-708 [FREE Full text] [doi: 10.1146/annurev-clinpsy-032813-153716] [Medline: 24499013]
88. Strine TW, Chapman DP. Associations of frequent sleep insufficiency with health-related quality of life and health behaviors. Sleep Med 2005 Jan;6(1):23-27. [doi: 10.1016/j.sleep.2004.06.003] [Medline: 15680291]
89. Schuller BW. Speech emotion recognition: two decades in a nutshell, benchmarks, and ongoing trends. Commun ACM 2018 Apr 24;61(5):90-99. [doi: 10.1145/3129340]
90. Conrad A, Müller A, Doberenz S, Kim S, Meuret AE, Wollburg E, et al. Psychophysiological effects of breathing instructions for stress management. Appl Psychophysiol Biofeedback 2007 Jun;32(2):89-98. [doi: 10.1007/s10484-007-9034-x] [Medline: 17520360]
91. Roper KO, Juneja P. Distractions in the workplace revisited. J Facil Manag 2008 May 2;6(2):91-109. [doi: 10.1108/14725960810872622]
92. Saravakos P, Sirakoulis GC. Modeling employees behavior in workplace dynamics. J Comput Sci 2014 Sep;5(5):821-833. [doi: 10.1016/j.jocs.2014.05.001]
93. Merilahti J, Pärkkä J, Antila K, Paavilainen P, Mattila E, Malm E, et al. Compliance and technical feasibility of long-term health monitoring with wearable and ambient technologies. J Telemed Telecare 2009;15(6):302-309. [doi: 10.1258/jtt.2009.081106] [Medline: 19720768]
94. de Lima AL, Hahn T, Evers LJ, de Vries NM, Cohen E, Afek M, et al. Feasibility of large-scale deployment of multiple wearable sensors in Parkinson's disease. PLoS One 2017;12(12):e0189161 [FREE Full text] [doi: 10.1371/journal.pone.0189161] [Medline: 29261709]

Abbreviations

API: application programming interface
BMI: body mass index
DIY: do it yourself
ECG: electrocardiography
EDA: electrodermal activity
EEG: electroencephalography
GPS: Global Positioning System
IARPA: Intelligence Advanced Research Projects Activity
IMU: inertial measurement unit
IRB: institutional review board
ODNI: Office of the Director of National Intelligence
PII: personally identifiable information
PPG: photoplethysmography
RFID: radio-frequency identification
SNR: signal-to-noise ratio
TAR: TILES Audio Recorder
TBW: total body water
USB: Universal Serial Bus

XSL•FO
RenderX

Edited by G Eysenbach; submitted 20.11.18; peer-reviewed by A Sano, Y Wang, J Goris, B Chaudhry, K Rosen, J Bennett; comments to author 21.03.19; revised version received 02.06.19; accepted 19.06.19; published 20.08.19.

Please cite as:
Booth BM, Mundnich K, Feng T, Nadarajan A, Falk TH, Villatte JL, Ferrara E, Narayanan S
Multimodal Human and Environmental Sensing for Longitudinal Behavioral Studies in Naturalistic Settings: Framework for Sensor Selection, Deployment, and Management
J Med Internet Res 2019;21(8):e12832
URL: http://www.jmir.org/2019/8/e12832/
doi:10.2196/12832
PMID:31432781

©Brandon M Booth, Karel Mundnich, Tiantian Feng, Amrutha Nadarajan, Tiago H Falk, Jennifer L Villatte, Emilio Ferrara, Shrikanth Narayanan. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 20.08.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

The Effect of the Interactive Mobile Health and Rehabilitation System on Health and Psychosocial Outcomes in Spinal Cord Injury: Randomized Controlled Trial

Michael Alan Kryger[1,2], MSc, MD; Theresa M Crytzer[3], PT, DPT; Andrea Fairman[4], PhD OTR/L, CRPR; Eleanor J Quinby[1], BSc; Meredith Karavolis[1], MS OTR/L; Gede Pramana[5], PhD; I Made Agus Setiawan[5,6], MSc; Gina Pugliano McKernan[1,3], PhD; Bambang Parmanto[3,5], PhD; Brad E Dicianno[1,3], MD

[1]Department of Physical Medicine and Rehabilitation, University of Pittsburgh, Pittsburgh, PA, United States

[2]Department of Physical Medicine and Rehabilitation, Penn State University Milton Hershey Medical Center, Hershey, PA, United States

[3]Human Engineering Research Laboratories, Department of Veterans Affairs, VA Pittsburgh Healthcare System, Pittsburgh, PA, United States

[4]Department of Occupational Therapy, MGH Institute of Health Professions, Boston, MA, United States

[5]Department of Health Information Management, School of Health and Rehabilitation, University of Pittsburgh, Pittsburgh, PA, United States

[6]Department of Computer Science, Udayana University, Badung, Indonesia

Corresponding Author:
Brad E Dicianno, MD
Department of Physical Medicine and Rehabilitation
University of Pittsburgh
6425 Penn Ave, Suite 400, Office 4115
Pittsburgh, PA, 15206
United States
Phone: 1 412 822 3700
Email: dicianno@pitt.edu

Abstract

Background: Individuals with spinal cord injury (SCI) are at risk for secondary medical complications, such as urinary tract infections (UTIs) and pressure injuries, that could potentially be mitigated through improved self-management techniques. The Interactive Mobile Health and Rehabilitation (iMHere) mobile health (mHealth) system was developed to support self-management for individuals with disabilities.

Objective: The main objective of this study was to determine if the use of iMHere would be associated with improved health outcomes over a 9-month period. A secondary objective was to determine if the use of iMHere would be associated with improved psychosocial outcomes. Phone usage, app usage, and training time data were also collected to analyze trends in iMHere use.

Methods: Overall, 38 participants with SCI were randomized into either the intervention group who used the iMHere system and received standard care or the control group who received standard care without any technology intervention. Health outcomes were recorded for the year before entry into the study and during the 9 months of the study. Participants completed surveys at baseline and every 3 months to measure psychosocial outcomes.

Results: The intervention group had a statistically significant reduction in UTIs (0.47 events per person; *P*=.03; number needed to treat=2.11). Although no psychosocial outcomes changed significantly, there was a nonsignificant trend toward a reduction in mood symptoms in the intervention group compared with the control group meeting the threshold for clinical significance. Approximately 34 min per participant per month were needed on average to manage the system and provide technical support through this mHealth system.

Conclusions: The use of the iMHere mHealth system may be a valuable tool in the prevention of UTIs or reductions in depressive symptoms. Given these findings, iMHere has potential scalability for larger populations.

Trial Registration: ClinicalTrials.gov NCT02592291; https://clinicaltrials.gov/ct2/show/NCT02592291.

(J Med Internet Res 2019;21(8):e14305) doi:10.2196/14305

XSL•FO
RenderX

KEYWORDS

cellular phone; emergency departments; hospitalization; mobile applications; pressure ulcer; rehabilitation; self-care; spinal cord injury; telemedicine; urinary tract infections

Introduction

Spinal cord injury (SCI), an insult to the spinal cord that is most commonly traumatic, can be a life-changing diagnosis. In the United States, approximately 17,000 new injuries occur each year, resulting in a prevalence of 285,000 [1]. Beyond the acute injury, the impact to these individuals and the health care system is lifelong and costly. Incomplete tetraplegia, for example, accounts for 47.7% of SCI cases, entailing an average annual cost of care of US $1,102,403 for the first year and US $191,436 for subsequent years because of the multiple potential chronic complications of SCI [1]. While the primary characteristics of SCI include strength loss and sensory loss, these chronic complications can result in increased mortality, health care costs, treatments, and hospitalizations [1,2].

There are multiple complications that can occur after SCI. Two of the most common chronic complications are urinary tract infections (UTIs) and skin pressure injuries. UTIs occur because of neurogenic bladder [3] and bacteria entering the bladder during catheterization or catheterization not occurring on a consistent schedule, leading to retention of urine and growth of bacteria [4]. Diseases of the genitourinary system were the most common cause of death in SCI populations 40 years ago; however, the introduction of clean intermittent catheterization, treatment of bladder spasticity, and appropriate antibiotic treatment have resulted in a decrease in UTIs and related complications [2]. Skin pressure injury is a loss of oxygenation to the tissues caused by inadequate pressure relief that is triggered by poor sensation and impaired mobility [5]. Additional complications frequently associated with SCI include neurogenic bowel, pulmonary compromise, spasticity, and depression [2]. Depression is prevalent in the United States in 1 out of every 5 individuals with SCI as compared with 1 out of 20 people without disabilities [6].

Given the complex nature of SCI, people with SCIs and their families require extensive training and constant vigilance to prevent secondary complications [7]. Frequent communication is required between the patient and their medical team to prevent or treat complications [8]. As a result, the potential exists for using mobile health (mHealth) platforms to allow patients with SCI to proactively monitor their health and gain self-management skills to prevent complications.

Smartphones have become ubiquitous in American society, with over 77% of Americans owning a device in January 2018, compared with 55% in 2014 and 35% in 2011 [9]. Over 98.7% of individuals in developed countries and 70.4% in lesser developed nations have mobile broadband subscriptions [10]. However, the prevalence of mobile phone and mobile internet usage in the SCI population is not well studied. From 2010 to 2014, 46% of participants in the SCI Model Systems Centers reported using the internet on their phone [11]. It can be expected that smartphone usage among individuals with SCI will continue to increase in the future, as those who use smartphones preinjury will continue to use them post injury.

The use of mHealth platforms is likewise gaining popularity. New systems are being studied for many different types of patient populations and conditions, including older adults [12], chronic obstructive pulmonary disease [13], diabetes [14], and bipolar disorder [15]. The Apple App store and Google Play store each contain over 100,000 health and wellness smartphone apps, and developers are beginning to target chronically ill individuals, particularly those with diabetes, obesity, and hypertension [16]. Some apps used within rehabilitation populations have had positive impacts on mobility and self-management [17]. However, few randomized controlled trials using robust mHealth self-management interventions have been conducted.

The Interactive Mobile Health and Rehabilitation (iMHere) system (Figure 1) was developed to promote self-management for persons with disabilities and to facilitate communication between patients and their medical teams [18].

The first clinical trial of the iMHere system was conducted in the spina bifida (SB) population. It was found that higher usage of iMHere was associated with improved self-management skills, less caregiver assistance, and a decreasing trend in UTIs and emergency department (ED) visits [19]. Cost savings from the use of the system were estimated to be over US $15,000 per user per year. In a separate study, about 80% of individuals with SB and their caregivers felt the app would be easy to use and make a positive impact [19].

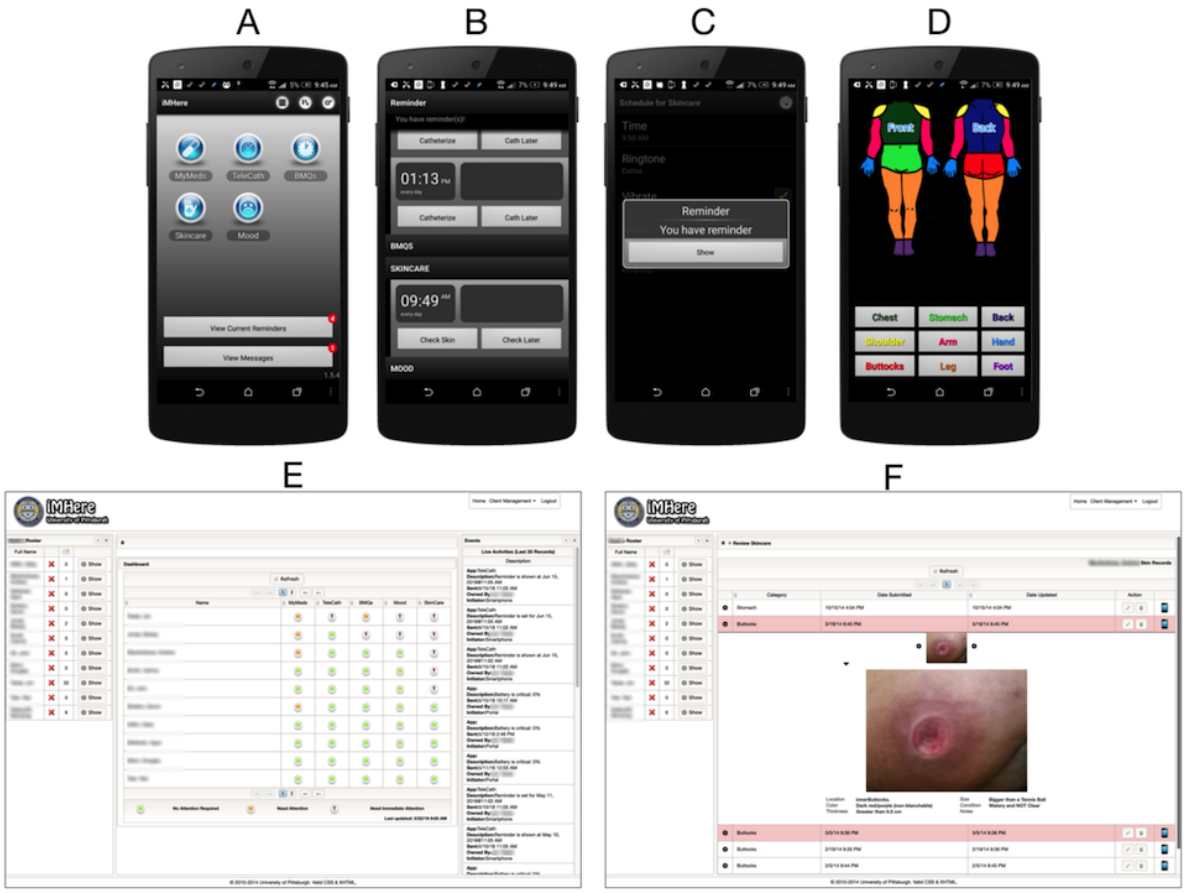
Many of the medical challenges in those with SB, including impaired mobility, neurogenic bowel, neurogenic bladder, insensate skin, polypharmacy, and depressed mood, are also present for those with SCI. Therefore, it was a natural extension to apply the iMHere system to the SCI population.

The goal of this study was to determine whether the use of the iMHere system would be associated with better health and psychosocial outcomes in patients with SCI. We hypothesized that the use of the app in addition to standard care would result in a larger magnitude of improvement in health outcomes (the primary outcome measures) and psychosocial outcomes (the secondary outcome measures) compared with a control group receiving standard care. Health outcomes were defined as the number of UTIs, number of pressure injuries, and number of ED visits and hospitalizations. Several specific survey outcomes were used as psychosocial metrics that described functional independence, quality of life, and mood.

An ancillary aim of the study was to determine whether any metrics related to phone use or app compliance impacted health and psychosocial outcomes. The time associated with managing the system and providing technical support was also evaluated.

XSL•FO
RenderX

Figure 1. The iMHere interface. (a) Home screen with modules; (b) Screen for skin check reminder; (c) Reminder example; (d) Screen for charting a wound location; (e) Web-based portal used by coordinator to track iMHere users; (f) Example of a wound photo uploaded through the iMHere app. iMHere: Interactive Mobile Health and Rehabilitation.



Methods

Study Design

This study was a randomized control study. Baseline data were collected before randomization. Participants were randomized using a random number generator in Microsoft Excel (Microsoft). The control group received the standard of care in an outpatient physiatry SCI clinic and no technologic intervention. *Standard care* involves a patient being seen by an SCI-trained physician in an outpatient clinic on an intermittent basis, with follow-up as needed based on current health status, as determined by the physician. As part of standard care, the patient is able to call in to speak to a nurse, who can triage cases, offer recommendations, and pass concerns onto the physician. The physician can then decide whether any further evaluation is needed, which may include a clinic visit, recommendation to go to the ED, diagnostic testing, etc. The intervention group was given a Samsung Galaxy S5 smartphone with the iMHere app and received standard care in the same outpatient physiatry SCI clinic. Owing to the nature of the intervention, the study participants could not be blinded. However, the investigators who reviewed medical records and collected retrospective data and the individuals who conducted interviews were blinded to the participant group. It should be noted that the investigators who reviewed the electronic medical record were a physiatrist and physical therapist. Surveys were conducted by occupational therapy, medical, and nursing students. Consolidated Standards of Reporting Trials guidelines were used in the development of this study and in reporting the results. This study was registered in ClinicalTrials.gov, under registration number NCT02592291.

Recruitment and Participants

This study was approved by the institutional review board of the University of Pittsburgh; all participants provided written informed consent. Participants were recruited from local physiatry-based SCI and assistive technology clinics. The inclusion criteria were (1) age 18 years and older; (2) diagnosis of SCI; (3) attends an outpatient physiatry clinic for SCI; and (4) lives in a community setting, rather than in a residential facility that provides care. The exclusion criteria were (1) diagnosis of severe intellectual disability or severe and persistent psychiatric illness and (2) actively participating in a concurrent wellness pilot program.

Interactive Mobile Health and Rehabilitation System

The iMHere system (Figure 1) consists of an app used by the participant in the community and a Web-based portal for the clinician. The app includes several modules: (1) medication management, including medication administration reminders, the ability to upload photos of the medications, and customizable descriptions of the purpose for taking them; (2) urinary and bowel program reminders, with a system for reporting concerning symptoms; (3) skincare tracking with photo capabilities to monitor for pressure injuries and skin breakdown; (4) mood tracking with validated surveys; and (5) messaging,

XSL•FO
RenderX

to communicate with a clinician [20]. The system has undergone multiple patient-centered design iterations to optimize the app for use by individuals with disabilities and their caregivers [19-25]. Intervention participants received 30 min of training to use the app. This involved app navigation, how to set up reminders, and how to record information in each module. The participants were told to use only the modules relevant to their recommended care regimen. After setting up the modules, the app would send them reminders in conjunction with their personal self-management routine. Participants were asked to respond to all reminders when they appeared on their device. If during their skin check reminder, they found a pressure injury, they were instructed to upload the location and a photo to the system. A physical therapist acted as the *wellness coordinator*, monitoring participant data using a Web portal and communicating with them electronically via their app.

Health Outcomes

Health outcomes were collected by retrospective chart review for the 9 months before the study as well as for the 9 months during which each participant was enrolled in the study. Individual phone interviews with patients were used to verify or clarify information in the medical record. The number of UTIs and pressure injuries were both used because of the high incidence of such events in individuals with SCI [1]. Number of ED visits and hospitalizations were included because ED visits and hospitalizations both result in increased health care costs. The following health outcome measures were collected:

- Number of UTIs: Number of symptomatic UTIs with positive urine cultures that were subsequently treated with antibiotics.
- Number of pressure injuries: Number of unique episodes of skin breakdown, at least stage 2 or above, based on the National Pressure Ulcer Advisory Panel guidelines [26]. A unique pressure injury was defined either as a wound in a different area of the skin or in the same area with documented complete healing before reinjury.
- Number of ED visits: Number of encounters in the ED for any reason.
- Number of ED visits because of UTIs or pressure injury: Number of encounters in the ED specifically for UTI or pressure injury diagnosis, evaluation, or treatment.
- Number of hospitalizations: Number of admissions to the hospital for any reason.
- Number of hospitalizations because of UTIs or pressure injury: Number of admissions to the hospital specifically for UTI or pressure injury diagnosis, evaluation, or treatment.

Psychosocial Outcomes

All participants were individually interviewed over the phone at baseline and every 3 months for 9 months, for a total of 4 interviews, using several psychosocial outcomes that are widely employed and validated to assess independence, mood, and quality of life in individuals with disabilities. The following questionnaires were used:

1. Canadian Occupational Performance Measure (COPM), which is a self-reported measure of self-care, productivity, and leisure [27].
2. Adolescent Self-Management and Independence Scale, which measures independence and self-management skills. The scale contains 10 items that measure independent living and 7 items that measure self-management skills and is valid for use in adults [28].
3. Beck Depression Inventory-II (BDI-II), which is a screening questionnaire that evaluates for symptoms of clinical depression, including guilt, self-blame, disappointment, satisfaction, and suicidal ideation [29].
4. Patient Assessment of Chronic Illness Care, which measures experience and satisfaction of chronic care [30].
5. World Health Organization Quality of Life Brief Instrument, which is a validated measure of perceived quality of life based on individual culture, values, and goals [31].
6. The physical independence domain of the Craig Handicap Assessment and Reporting Technique Short Form, which is a measure of perceived disability and independence [32]. This domain measures paid and unpaid caregiver hours on a 0 to 100 scale.

Phone Usage, Interactive Mobile Health and Rehabilitation Usage, and Support Time

Usage statistics were recorded to gain a better understanding of how participants used their smartphone and the iMHere system to provide a potential explanation for differences in study results or rule out any potential confounding factors. Phone use habits were recorded using cellular phone bill data. The number of calls sent and received, text messages sent and received, and data used in megabytes were calculated for each participant. An iMHere compliance rate was also determined for each module and each participant by calculating the number of times the participant input data into each module, divided by the number of times the participant was prompted to input data into the module. If the participants input information more often than they were prompted, they were given a compliance rate of 1.

Toggle software (Tallinn) was used to record the amount of time that support was provided to individual participants. *Wellness Time* was defined as the time that the wellness coordinator spent triaging issues for participants or communicating with them about concerns. *Tech Support Time* was defined as the amount of time that each participant required for help with setting up the app, training, and any minor technical issues that arose during the study.

Statistical Analysis

Sample size calculation was based on a previous study in which iMHere was used by participants with SB [25]. A moderate effect size of 0.30 was used and was based upon the primary outcome measures used in this study. A repeated-measures analysis of variance (ANOVA) yielded a sample size of 18 participants in each group for a power of 80%. Alpha values were set to .05 a priori.

The demographic information collected included gender, race, ethnicity, marital status, education, type of SCI, smoking status,

XSL•FO
RenderX

assistance at home, and technology experience. The demographics of the intervention and control groups were compared to confirm that the randomization was effective using the Student *t* test, chi-square test, Fisher exact test, or Mann-Whitney test (Table 1). Baseline psychosocial outcomes were compared using the Mann-Whitney test.

Table 1. Participant demographics (N=19).

Demographic details	Intervention group	Control group
Age (years), mean (SD)	37.9 (13.4)	44.1 (15.3)
Gender, n (%)		
Male	13 (68)	12 (63)
Female	6 (32)	7 (37)
Race, n (%)		
White	13 (68)	15 (79)
Black	6 (32)	4 (21)
Ethnicity, n (%)		
Hispanic	1 (5)	0 (0)
Non-Hispanic	18 (95)	19 (100)
Marital status, n (%)		
Single	12 (63)	11 (58)
Not single	7 (37)	8 (42)
Highest level of education, n (%)		
High school	11 (58)	10 (53)
Higher education	6 (32)	8 (42)
Completeness of injury, n (%)		
Complete	9 (47)	12 (63)
Incomplete	10 (53)	7 (37)
Functional status, n (%)		
Tetraplegia	8 (42)	9 (47)
Paraplegia	11 (58)	10 (53)
Time since injury, mean (SD)	9.9 (8)	13.5 (11)
Living status, n (%)		
Alone	2 (11)	2 (11)
With others	17 (89)	17 (89)
Student status, n (%)		
Student	3 (16)	1 (5)
Not a student	16 (84)	18 (95)
Smoking history, n (%)		
Smoker	8 (42)	6 (32)
Nonsmoker	11 (58)	13 (68)
Previous experience with smartphones, n (%)		
Yes	9 (47)	13 (68)
No	10 (53)	6 (32)
Previous experience with apps, n (%)		
Yes	9 (47)	11 (58)
No	10 (53)	8 (42)

XSL•FO
RenderX

Primary health outcomes were tallied for the periods before study enrollment and during study enrollment and were compared pre- and postintervention using the Wilcoxon signed-rank test. A number needed to treat (NNT) analysis was performed for statistically significant and trending health outcomes. Generalized linear models with both fixed and random effects were used to evaluate changes in the secondary psychosocial outcomes over time. As this was an intention-to-treat analysis, participants with missing interview data were still included in the analysis.

Participants were split into high-usage phone users and low-usage phone users based on their average monthly general phone use habits. A high-usage phone user was defined as a participant who sent or received over 500 calls, sent or received over 1000 texts, or used over 3000 MB data using cellular connectivity. All other participants were classified as low-usage phone users.

The overall iMHere compliance rate for each participant was defined as the average compliance rate of all modules used by that participant. It should be noted that not all modules were used by all participants. A Student *t* test was performed to evaluate for an association between phone usage and overall compliance rate.

Intervention participants were divided into 2 groups: high overall compliance users (n=10) and low overall compliance users (n=9). A repeated-measures ANOVA was then used to evaluate whether there were any between-group differences in psychosocial measures with respect to overall compliance.

Statistical analyses of primary outcomes were performed using IBM SPSS Statistics for Windows (IBM Corp), and secondary outcomes analyses were performed using SAS version 9.4 (SAS Institute).

Results

A total of 41 participants were recruited to participate, and of those, 38 completed informed consent and baseline interviews.

Figure 2 is a flow diagram demonstrating participant selection, randomization, and dropout. Tables 1 and 2 display participant demographics and baseline psychosocial outcome measures. No significant differences were detected at baseline between control and intervention groups.

Figure 3 displays the incidence of health outcomes before and during the study period for the intervention and control groups. Participants in the intervention group experienced about half as many UTIs during the study period, when compared with the period before the intervention (*P*=.03). There was a reduction of 0.47 UTIs per person in the study group during the intervention compared with before the intervention. Such a reduction was not seen in the control group. No other primary outcome measures were found to change significantly in the intervention or control groups.

Table 3 presents the changes in psychosocial outcome measures in both groups during the study period. No statistically significant trends were seen between the intervention and control groups over time.

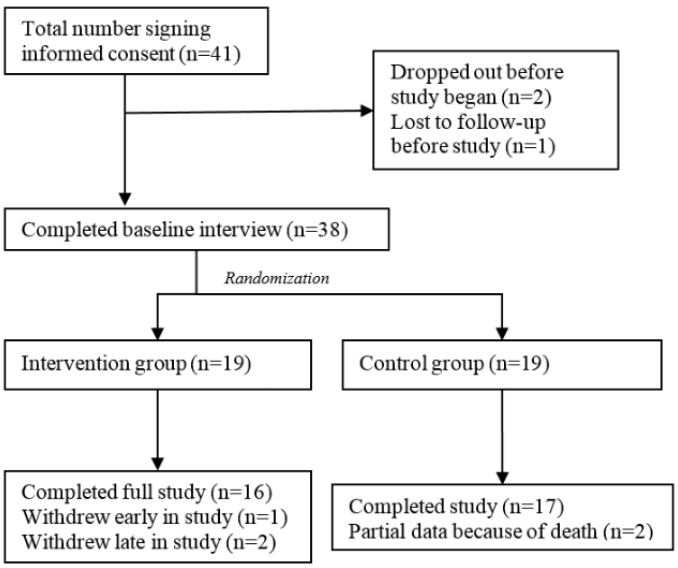
Figure 4 demonstrates some of the general trends seen in Table 3 from baseline to 9 months for certain secondary outcomes.

Table 4 shows high-usage and low-usage phone users with corresponding overall iMHere compliance rates. There was no statistically significant difference in overall iMHere compliance rates between the high-usage and low-usage phone users (*P*=.41).

No statistically significant differences were seen in the 2 overall iMHere compliance rate groups with respect to changes in psychosocial outcomes (all *P* values were .45 or higher).

As shown in Table 5, approximately 34 min per month per participant was spent on providing wellness coordinator and technical support.

Figure 2. Flow diagram of patient enrollment and randomization.

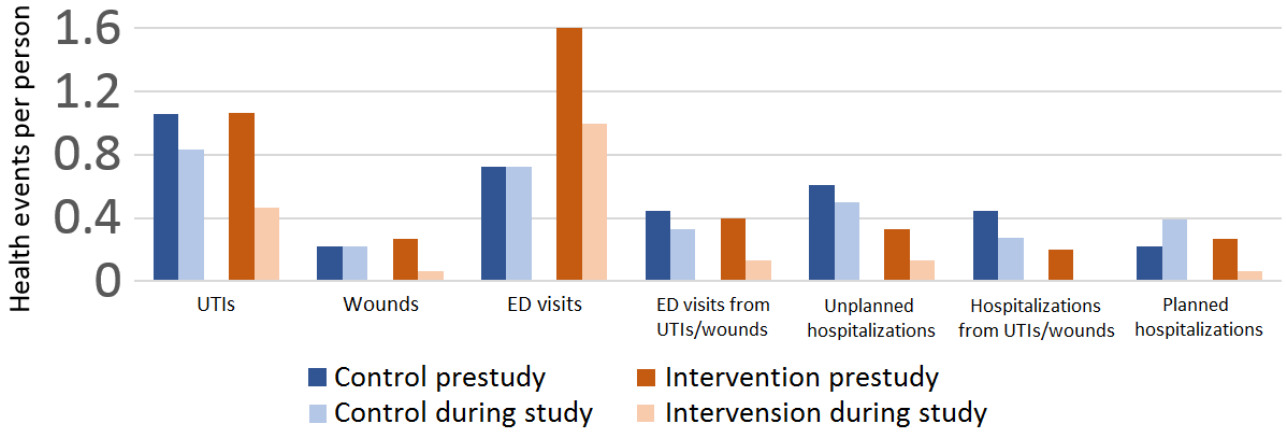


XSL•FO
RenderX

Table 2. Baseline comparison of psychosocial outcome measurements (N=19).

Outcome measure	Intervention, mean (SD)	Control, mean (SD)
Canadian occupational performance measure	8.59 (6.43)	8.58 (7.47)
Adolescent Self-Management and Independence Scale-II		
Independence subscale	5.68 (1.08)	5.21 (0.98)
Self-management subscale	5.86 (1.03)	5.70 (1.08)
Total	91.06 (15.16)	86.21 (13.89)
Beck Depression Inventory-II	11.18 (8.65)	12.05 (10.79)
Patient Assessment of Chronic Illness Care	3.39 (0.83)	3.02 (0.54)
World Health Organization Quality of Life		
Physical subscale	57.06 (12.21)	56.47 (12.01)
Psychological subscale	64.06 (15.95)	62.05 (15.84)
Social subscale	65.76 (24.96)	69.50 (24.27)
Environment subscale	69.53 (18.71)	73.53 (13.44)
Craig Handicap Assessment and Reporting Technique Short Form	66.59 (37.36)	62.53 (29.00)

Figure 3. Health outcomes. ED: emergency department; UTI: urinary tract infection.



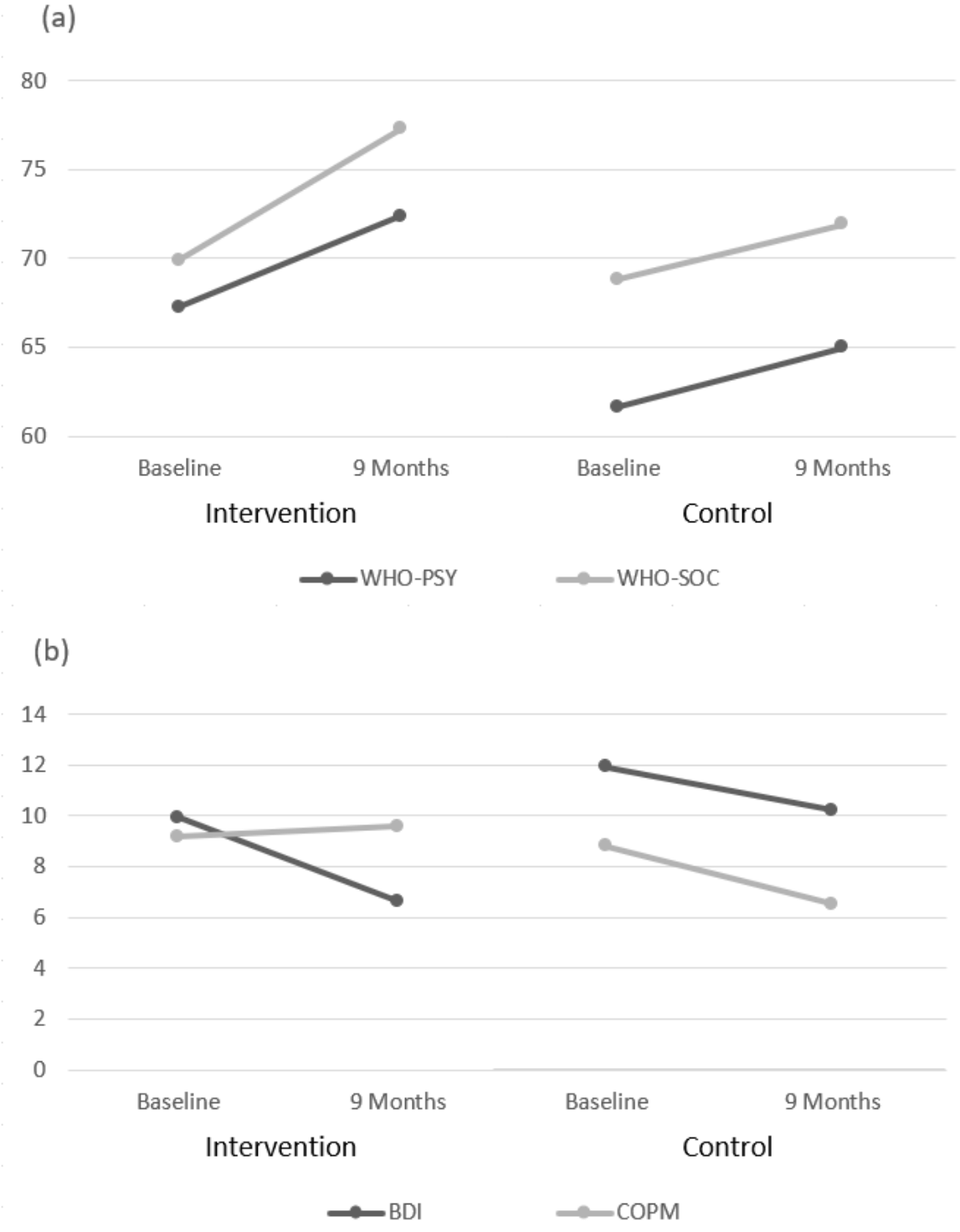
XSL•FO
RenderX

Table 3. Survey outcomes.

Outcome measure	Baseline, mean (SD)	3 months, mean (SD)	6 months, mean (SD)	9 months, mean (SD)
Canadian Occupational Performance Measure				
Intervention	9.19 (7.91)	5.57 (9.10)	7.36 (12.12)	9.57 (14.07)
Control	8.78 (7.64)	6.65 (5.48)	7.94 (8.03)	6.50 (6.45)
Adolescent Self-Management and Independence Scale-II				
Intervention	90.56 (15.40)	91.29 (14.67)	89.29 (14.86)	90.79 (14.22)
Control	87.39 (13.28)	87.18 (17.10)	89.31 (13.91)	90.06 (10.90)
Adolescent Self-Management and Independence Scale-II				
Independence subscale				
Intervention	5.62 (1.09)	5.62 (1.18)	5.50 (1.21)	5.54 (1.19)
Control	5.33 (0.84)	5.37 (1.01)	5.46 (0.86)	5.43 (0.75)
Self-management subscale				
Intervention	5.88 (1.03)	6.06 (1.02)	5.54 (1.76)	6.62 (0.72)
Control	5.69 (1.11)	5.62 (1.40)	5.82 (1.02)	5.99 (0.92)
Beck Depression Inventory -II				
Intervention	9.94 (6.74)	8.07 (5.65)	4.86 (5.87)	6.64 (4.53)
Control	11.89 (11.08)	12.35 (13.30)	11.38 (9.92)	10.19 (9.61)
Patient Assessment of Chronic Illness Care				
Intervention	3.54 (0.68)	3.34 (0.85)	3.24 (0.97)	3.44 (0.78)
Control	3.04 (0.56)	3.31 (0.70)	3.09 (0.68)	3.09 (0.74)
World Health Organization Quality of Life				
Physical subscale				
Intervention	59.06 (12.84)	63.07 (13.12)	61.29 (11.42)	60.43 (12.33)
Control	56.11 (12.26)	54.94 (13.36)	53.31 (14.20)	59.56 (7.99)
Psychological subscale				
Intervention	67.25 (14.38)	70.71 (10.30)	72.79 (13.13)	72.36 (9.21)
Control	61.67 (16.21)	60.47 (16.42)	60.75 (15.07)	65.00 (13.66)
Social subscale				
Intervention	69.88 (25.37)	77.64 (12.97)	78.14 (13.69)	77.29 (16.82)
Control	68.82 (24.84)	68.00 (22.27)	71.81 (22.32)	71.94 (23.22)
Environment subscale				
Intervention	72.69 (19.98)	75.14 (15.60)	73.21 (17.24)	77.43 (15.36)
Control	73.44 (13.83)	75.47 (13.70)	77.44 (14.61)	81.81 (12.16)
Craig Handicap Assessment and Reporting Technique Short Form				
Intervention	74.00 (34.78)	75.86 (34.95)	72.57 (32.83)	74.86 (27.60)
Control	62.22 (29.81)	66.59 (34.32)	70.00 (28.13)	66.00 (25.17)

XSL•FO
RenderX

Figure 4. (a) Difference in World Health Organization Quality of Life subscores over time for intervention and control participants. (b) Difference in Beck Depression Inventory-II and Canadian Occupational Performance Measure scores over time for intervention and control participants. WHO-PSY: World Health Organization Psychiatric subscore; WHO-SOC: Social subscore; BDI: Beck Depression Inventory; COPM: Canadian Occupational Performance Measure.



XSL•FO
RenderX

Table 4. Phone and Interactive Mobile Health and Rehabilitation (iMHere) app usage (N=19).

ID	Phone usage group	Calls sent and received, n	Text messages sent and received, n	Data used (MB)	Overall iMHere app compliance (%)
1	High	992	1955	3022	98
3	Low	1	3	2817	18
7	High	0	11	3234	46
8	High	34	2963	1565	45
9	High	4105	47,689	11,853	39
11	Low	14	58	1676	57
12	High	234	2445	25,563	42
14	Low	15	9	1079	30
16	Low	5	34	1097	98
18	Low	83	20	1097	51
19	Low	57	38	902	23
20	High	4463	3429	63,074	49
25	High	2120	8173	99,191	88
30	Low	81	62	317	18
31	Low	313	68	1464	74
32	Low	179	161	2641	42
40	High	3495	1589	21,268	85
47	High	2663	18,596	21,364	47
49	High	1362	8640	28,556	20

Table 5. Support time (N=19).

Contact time descriptions	Wellness time	Tech support time	Total time
Total contact time for 19 participants in 9 months, hours:minutes	49:00	56:23	105:23
Average contact time per participant in 9 months, hours:minutes (SD)	2:08 (2:08)	2:27 (2:44)	4:35 (3:35)
Average contact time per participant per month, hours:minutes (SD)	0:14 (0:14)	0:16 (0:18)	0:34 (0:24)

Discussion

Principal Findings

This study contributes to the literature by demonstrating successful use of an mHealth platform in individuals with SCI. This study demonstrated a significant reduction in UTIs (one of the primary outcomes) in those who used iMHere over time when compared with the control group. Given the reduction in UTIs during the intervention, 0.47 fewer UTIs per person, the NNT to prevent 1 UTI is 2.11. As not all intervention participants used the catheterization module, the study was not powered to determine whether the use of this specific module was associated with the reduction in UTIs. One explanation for reduction in UTIs aside from the use of the catheterization module is increased general health awareness and improved self-management that resulted from using the app in general. Unfortunately, it was challenging to determine the financial implications of these findings as there is a lack of literature examining health care costs associated with the treatment of UTIs in an outpatient setting in individuals with chronic SCI. However, in similarly aged adults with SB, the cost to treat a UTI was found to be approximately US $511 per event [33]. Although no other significant changes were seen in health outcomes over the study period, the other 6 primary outcome measures also decreased in the intervention group, which was a trend also observed in a similar study in the SB population [25]. More studies are warranted to determine whether larger patient cohorts might result in significant changes in these variables.

Of the secondary outcomes, the reduction in depressive symptoms based on BDI-II was the closest to approach significance between groups. The decrease over time in BDI-II in the intervention group, an average of 3.3 points or 33% (3.3/9.94), was twice that of the control group. Previous research has suggested that a decrease in BDI-II score of 17.5% may be clinically significant because it correlates with an individual *feeling better*. Therefore, this change could be considered clinically significant. It should be noted that the clinical

XSL•FO
RenderX

significance of the magnitude of change depends upon an individual's initial score [34]. Lower initial scores, meaning a person has fewer depressive symptoms, may require smaller changes in BDI-II score for an individual to *feel better*. This further supports the suggestion of clinical significance in the intervention group as the BDI-II was lower at baseline. As not all intervention participants used the mood module, the study was not powered to determine whether the use of this specific module was associated with the reduction in BDI-II score. Many of the control group participants used smartphones in everyday life, and those in the intervention group who used their phones more did not have better mood outcomes than those who used their phones less. These 2 findings suggest that the communication afforded by the phone itself was not solely responsible for this change. It is possible that iMHere's ability to facilitate communication with the health care team, participants' increased awareness of their own mood symptoms, or other improvements in health may have impacted mood positively. Notably, a review by Thota et al confirmed that collaborative care models that use case managers to connect patients, primary care physicians, and mental health professionals provide a supportive care network that empowers people with depression to take a self-management role in their own care [35].

As our study was underpowered for the secondary outcomes, it is not unexpected that the other psychosocial outcomes were not statistically significant. A descriptive analysis reveals that there are some trends (as shown in Figure 4). There is a larger general improvement in World Health Organization Quality of Life psychological subscore at 9 months, which makes sense in the context of the improvement in BDI-II. There is also a markedly higher social subscore. It should be noted that Figure 4 also shows that COPM improved slightly in the intervention group and decreased substantially in the control group, suggesting overall that participants in the intervention group perceived their self-care, productivity, and leisure to be maintained after 9 months, whereas those in the control group had declined in this perception. We postulate that a larger study cohort may have allowed us to detect significant changes.

Overall, iMHere compliance rates were not related to psychosocial outcomes or the amount of phone usage. This contrasts with findings from a previous study in SB in which more frequent users of iMHere had positive changes in self-management skill and amount of caregiver assistance needed [25]. One possible explanation for the contrasting findings between studies is that those with tetraplegia in this study may not have been able to reduce the need for hands-on care even if they did gain small improvements in knowledge about self-management. Although significant accessibility features have been implemented for individuals with tetraplegia and other impairments [22-24], it is also possible that users with paraplegia were able to use the system more proficiently. In addition, because we used billing data to calculate usage data, we may have underestimated the usage of individuals who primarily used Wi-Fi.

The integration of mHealth support into outpatient care depends in part on the requirements for staff effort [23]. This study demonstrated that wellness and technical support requires on average approximately 34 min per user per month. This information may be useful when scaling mHealth interventions to larger populations.

Study Limitations

Some limitations of this study warrant discussion. First, this study was powered to detect statistical differences in the primary outcomes related to health and not the secondary psychosocial outcomes. As a result, even though there was a trend toward improved psychosocial outcomes in the intervention group, the study was unable to find statistical significance in these trends. Future studies with larger sample sizes are planned to help with this issue. One potential confounder was that additional contact with study staff for wellness coordination or technical support may have had an impact on outcomes in the intervention group. Although a small amount of support was provided in person or via phone, the majority of contact was virtual, through the mHealth system. To evaluate whether such contact may have offset other types of contact, we conducted a post hoc analysis. However, no statistically significant changes were seen within or between groups with respect to the number of outpatient visits, phone calls to the clinic, and hospital health portal messages. This was likely because the number of instances of these occurrences was low on average. It is also possible that the use of iMHere shifted use for some nonurgent issues from the ED to the outpatient setting. It is important to note that the control group may also have had more contact with clinicians through ED visits and hospitalizations. More work will be needed to understand which aspects of an mHealth delivery system are most beneficial to outcomes and to provide more insight into how a self-management app might affect health care utilization and service delivery. A second limitation is the small sample size, which may have reduced our ability to detect changes in outcome measures with lower effect sizes. A larger population with greater usage levels of individual modules may have enabled us to do a subgroup analysis to determine if there was a correlation between individual module usage and health outcomes. Third, the inclusion of individuals with tetraplegia may have resulted in a ceiling effect of how much improvement can occur in some outcomes such as self-management given that they will still likely rely on caregiver assistance. Future studies will be aimed at evaluating outcomes using mHealth support for caregivers [8,36]. We are also investigating the addition of more accessibility features to support users with tetraplegia such as voice control. Fourth, the iMHere system has multiple features, but not all features were relevant to all users. Larger studies will be needed to evaluate the individual effects of different aspects of the system. Finally, fully functional smartphones were provided for study purposes, but some individuals also used their personal phone, which may have reduced the usage of study phones. The new version of iMHere (2.0) operates cross-platform and can now be used on personal devices.

Future Directions

Concurrent work on iMHere has produced a subsequent version (iMHere 2.0) with additional features. A new smartphone app will support family or formal caregivers and interface with the client app. A personal health record and additional modules

XSL•FO
RenderX

have also been built to support community integration, physical activity, nutrition, goal setting, and education [21,37]. Future work will be conducted to evaluate the implementation of these features into clinical workflows, translation to larger and different disability populations and clinical settings, and interfacing with other electronic health systems.

Conclusions

Overall, the use of the iMHere mHealth platform resulted in a statistically significant reduction in UTIs over time compared with the control group. On the basis of an NNT analysis, 2.11 users were needed to prevent 1 UTI. There was also a decrease in several other outcome measures (eg, symptoms of depression), which trended toward, but did not reach, statistical significance. Approximately 34 min per participant per month is needed to provide education, care coordination, and technical support through this mHealth system, thus suggesting scalability.

Acknowledgments

The authors would like to thank Carly Sullivan, Taya Irizarry, Yongbin *Matthew* Kwon, and Elizabeth Mueller, for assistance with data collection and Zara Ambadar, PhD, and Daniel Rusnak for assistance with clinical coordination.

The contents of this publication were developed under grants from the Craig H Neilsen Foundation and the National Institute on Disability, Independent Living, and Rehabilitation Research (NIDILRR; grant numbers 90DP0064-01-00, 90DPGE0002-01-00, and 90DP5004-01-00). NIDILRR is a center within the Administration for Community Living (ACL), Department of Health and Human Services (HHS). The contents of this publication do not necessarily represent the views of the Department of Veterans Affairs or the United States government nor do they necessarily represent the policy of NIDILRR, ACL, or HHS.

Conflicts of Interest

BED, AF, BP, GP, and IS are inventors of the iMHere system with no other financial interests in this technology.

Multimedia Appendix 1
CONSORT-EHEALTH Checklist (V 1.6.1).
[PDF File (Adobe PDF File), 2MB - jmir_v21i8e14305_app1.pdf]

References

1. The National Spinal Cord Injury Statistical Center. 2018. Spinal Cord Injury: Facts and Figures at a Glance URL: https://www.nscisc.uab.edu/Public/Facts%20and%20Figures%20-%202018.pdf [WebCite Cache ID 77gNlf7Se]
2. Campagnolo DI, Kirshblum S, Nash MS, Heary RF, Gorman PH, editors. Spinal Cord Medicine. Philadelphia, PA: Lippincott Williams & Wilkins; 2011.
3. de Ruz AE, Leoni EG, Cabrera RH. Epidemiology and risk factors for urinary tract infection in patients with spinal cord injury. J Urol 2000 Oct;164(4):1285-1289. [doi: 10.1097/00005392-200010000-00032] [Medline: 10992382]
4. Klebine P, Kirksey K. Model Systems Knowledge Translation Center (MSKTC). 2017. Urinary Tract Infection and Spinal Cord Injury URL: https://msktc.org/sites/default/files/SCI-UTI-508.pdf [accessed 2019-04-16] [WebCite Cache ID 77gO4Vwbo]
5. Chen Y, Devivo MJ, Jackson AB. Pressure ulcer prevalence in people with spinal cord injury: age-period-duration effects. Arch Phys Med Rehabil 2005 Jun;86(6):1208-1213. [doi: 10.1016/j.apmr.2004.12.023] [Medline: 15954061]
6. Bombardier CH. Model Systems Knowledge Translation Center (MSKTC). 2010. Depression and Spinal Cord Injury URL: https://msktc.org/lib/docs/Factsheets/SCI_Depression_and_SCI.pdf [accessed 2019-04-16] [WebCite Cache ID 77gO6w7u8]
7. Guilcher SJ, Craven BC, Lemieux-Charles L, Casciaro T, McColl MA, Jaglal SB. Secondary health conditions and spinal cord injury: an uphill battle in the journey of care. Disabil Rehabil 2013 Jun;35(11):894-906 [FREE Full text] [doi: 10.3109/09638288.2012.721048] [Medline: 23020250]
8. Munce SE, Allin S, Wolfe DL, Anzai K, Linassi G, Noonan VK, et al. Using the theoretical domains framework to guide the development of a self-management program for individuals with spinal cord injury: results from a national stakeholder advisory group. J Spinal Cord Med 2017 Nov;40(6):687-695 [FREE Full text] [doi: 10.1080/10790268.2017.1356437] [Medline: 28758540]
9. Pew Research Center. 2019. Mobile Fact Sheet URL: http://www.pewinternet.org/fact-sheet/mobile/ [WebCite Cache ID 77gOKlm7O]
10. ITU. 2017. ICT: Facts and Figures URL: https://www.itu.int/en/ITU-D/Statistics/Documents/facts/ICTFactsFigures2017.pdf [accessed 2019-04-16] [WebCite Cache ID 77gOOvlhg]
11. The National Spinal Cord Injury Statistical Center. 2017. National Spinal Cord Injury Model Systems Database URL: https://www.nscisc.uab.edu/Public_Pages/Database_files/What_Is_The_National_SCIMS_Database.pdf
12. Bourouis A, Feham M, Bouchachia A. Ubiquitous mobile health monitoring system for elderly (UMHMSE). Int J Comput Sci Inf Technol 2011 Jun 18;3(3):74-82. [doi: 10.5121/ijcsit.2011.3306]

XSL•FO
RenderX

13. Ding H, Moodley Y, Kanagasingam Y, Karunanithi M. A mobile-health system to manage chronic obstructive pulmonary disease patients at home. Conf Proc IEEE Eng Med Biol Soc 2012;2012:2178-2181. [doi: 10.1109/EMBC.2012.6346393] [Medline: 23366354]
14. Istepanian RS, Zitouni K, Harry D, Moutosammy N, Sungoor A, Tang B, et al. Evaluation of a mobile phone telemonitoring system for glycaemic control in patients with diabetes. J Telemed Telecare 2009;15(3):125-128. [doi: 10.1258/jtt.2009.003006] [Medline: 19364893]
15. Bardram JE, Frost M, Szántó K, Faurholt-Jepsen M, Vinberg M, Kessing LV. Designing Mobile Health Technology for Bipolar Disorder: A Field Trial of the Monarca System. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2013 Presented at: CHI'13; April 27-May 2, 2013; Paris, France p. 2627-2636. [doi: 10.1145/2470654.2481364]
16. Research2Guidance. 2016. mHealth Economics 2016 – Current Status and Trends of the mHealth App Market URL: https://research2guidance.com/product/mhealth-app-developer-economics-2016/ [accessed 2019-04-16] [WebCite Cache ID 77gOipr6V]
17. Nussbaum R, Kelly C, Quinby E, Mac A, Parmanto B, Dicianno BE. Systematic review of mobile health applications in rehabilitation. Arch Phys Med Rehabil 2019 Jan;100(1):115-127. [doi: 10.1016/j.apmr.2018.07.439] [Medline: 30171827]
18. Parmanto B, Pramana G, Yu DX, Fairman AD, Dicianno BE, McCue MP. iMHere: a novel mhealth system for supporting self-care in management of complex and chronic conditions. JMIR Mhealth Uhealth 2013 Jul 11;1(2):e10 [FREE Full text] [doi: 10.2196/mhealth.2391] [Medline: 25100682]
19. Fairman AD, Dicianno BE, Datt N, Garver A, Parmanto B, McCue M. Outcomes of clinicians, caregivers, family members and adults with Spina Bifida regarding receptivity to use of the iMHere mhealth solution to promote wellness. Int J Telerehabil 2013;5(1):3-16 [FREE Full text] [doi: 10.5195/ijt.2013.6116] [Medline: 25945209]
20. Fairman AD, Yih ET, McCoy DF, Lopresti EF, McCue MP, Parmanto B, et al. Iterative design and usability testing of the iMHere system for managing chronic conditions and disability. Int J Telerehabil 2016;8(1):11-20 [FREE Full text] [doi: 10.5195/ijt.2016.6194] [Medline: 27563387]
21. 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]
22. Yu DX, Parmanto B, Dicianno BE, Pramana G. Accessibility of mhealth self-care apps for individuals with Spina Bifida. Perspect Health Inf Manag 2015;12:1h [FREE Full text] [Medline: 26755902]
23. Yu DX, Parmanto B, Dicianno BE, Watzlaf VJ, Seelman KD. Accessibility needs and challenges of a mhealth system for patients with dexterity impairments. Disabil Rehabil Assist Technol 2017 Jan;12(1):56-64. [doi: 10.3109/17483107.2015.1063171] [Medline: 26153097]
24. Yu D, Parmanto B, Dicianno BE. An mhealth app for users with dexterity impairments: accessibility study. JMIR Mhealth Uhealth 2019 Jan 8;7(1):e202 [FREE Full text] [doi: 10.2196/mhealth.9931] [Medline: 30622096]
25. Dicianno BE, Fairman AD, McCue M, Parmanto B, Yih E, McCoy A, et al. Feasibility of using mobile health to promote self-management in Spina Bifida. Am J Phys Med Rehabil 2016 Jun;95(6):425-437. [doi: 10.1097/PHM.0000000000000400] [Medline: 26488144]
26. Edsberg LE, Black JM, Goldberg M, McNichol L, Moore L, Sieggreen M. Revised national pressure ulcer advisory panel pressure injury staging system: revised pressure injury staging system. J Wound Ostomy Continence Nurs 2016;43(6):585-597 [FREE Full text] [doi: 10.1097/WON.0000000000000281] [Medline: 27749790]
27. Law M, Baptiste S, McColl M, Opzoomer A, Polatajko H, Pollock N. The Canadian occupational performance measure: an outcome measure for occupational therapy. Can J Occup Ther 1990 Apr;57(2):82-87. [doi: 10.1177/000841749005700207] [Medline: 10104738]
28. Buran CF, Brei TJ, Sawin KJ, Stevens S, Neufeld J. Further development of the adolescent self management and independence scale: AMIS II. Cerebrospinal Fluid Res 2006 Dec 21;3(S1):S37. [doi: 10.1186/1743-8454-3-S1-S37]
29. Beck AT, Steer RA, Brown GK. BDI-II, Beck Depression Inventory: Manual. San Antonio, Texas: Psychological Corporation; 1996.
30. Glasgow RE, Wagner EH, Schaefer J, Mahoney LD, Reid RJ, Greene SM. Development and validation of the Patient Assessment of Chronic Illness Care (PACIC). Med Care 2005 May;43(5):436-444. [doi: 10.1097/01.mlr.0000160375.47920.8c] [Medline: 15838407]
31. Skevington SM, Lotfy M, O'Connell KA, WHOQOL Group. The World Health Organization's WHOQOL-BREF quality of life assessment: psychometric properties and results of the international field trial. A report from the WHOQOL group. Qual Life Res 2004 Mar;13(2):299-310. [doi: 10.1023/B:QURE.0000018486.91360.00] [Medline: 15085902]
32. Mellick D. The Center For Outcome Measurement in Brain Injury. 2000. Introduction to the Craig Handicap Assessment and Reporting Technique - Short Form URL: http://www.tbims.org/combi/chartsf/ [accessed 2019-04-16] [WebCite Cache ID 77gNtUv69]
33. Ouyang L, Grosse SD, Thibadeau J, Swanson M, Campbell VA. Outpatient medical conditions among children and adults with Spina Bifida in the United States: frequency and expenditures. J Pediatr Rehabil Med 2010;3(3):177-185. [doi: 10.3233/PRM-2010-0127] [Medline: 21791849]

XSL•FO
RenderX

34. Button KS, Kounali D, Thomas L, Wiles NJ, Peters TJ, Welton NJ, et al. Minimal clinically important difference on the Beck Depression Inventory--II according to the patient's perspective. Psychol Med 2015 Nov;45(15):3269-3279 [FREE Full text] [doi: 10.1017/S0033291715001270] [Medline: 26165748]
35. Thota AB, Sipe TA, Byard GJ, Zometa CS, Hahn RA, McKnight-Eily LR, Community Preventive Services Task Force. Collaborative care to improve the management of depressive disorders: a community guide systematic review and meta-analysis. Am J Prev Med 2012 May;42(5):525-538. [doi: 10.1016/j.amepre.2012.01.019] [Medline: 22516495]
36. Munce SE, Webster F, Fehlings MG, Straus SE, Jang E, Jaglal SB. Meaning of self-management from the perspective of individuals with traumatic spinal cord injury, their caregivers, and acute care and rehabilitation managers: an opportunity for improved care delivery. BMC Neurol 2016 Jan 23;16:11 [FREE Full text] [doi: 10.1186/s12883-016-0534-2] [Medline: 26801243]
37. Setiawan IM, Zhou L, Alfikri Z, Saptono A, Fairman AD, Dicianno BE, et al. An adaptive mobile health system to support self-management for persons with chronic conditions and disabilities: usability and feasibility studies. JMIR Form Res 2019 Apr 25;3(2):e12982 [FREE Full text] [doi: 10.2196/12982] [Medline: 31021324]

Abbreviations

ACL: Administration for Community Living
ANOVA: analysis of variance
BDI-II: Beck Depression Inventory-II
COPM: Canadian Occupational Performance Measure
ED: emergency department
HHS: Health and Human Services
iMHere: Interactive Mobile Health and Rehabilitation
mHealth: mobile health
NIDILRR: National Institute on Disability, Independent Living, and Rehabilitation Research
NNT: number needed to treat
SB: spina bifida
SCI: spinal cord injury
UTI: urinary tract infection

Edited by G Eysenbach; submitted 24.04.19; peer-reviewed by K Best, B Mortenson; comments to author 28.05.19; revised version received 21.07.19; accepted 21.07.19; published 28.08.19.

Please cite as:
Kryger MA, Crytzer TM, Fairman A, Quinby EJ, Karavolis M, Pramana G, Setiawan IMA, McKernan GP, Parmanto B, Dicianno BE
The Effect of the Interactive Mobile Health and Rehabilitation System on Health and Psychosocial Outcomes in Spinal Cord Injury: Randomized Controlled Trial
J Med Internet Res 2019;21(8):e14305
URL: http://www.jmir.org/2019/8/e14305/
doi:10.2196/14305
PMID:31464189

©Michael Alan Kryger, Theresa M Crytzer, Andrea Fairman, Eleanor J Quinby, Meredith Karavolis, Gede Pramana, I Made Agus Setiawan, Gina Pugliano McKernan, Bambang Parmanto, Brad E Dicianno. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 28.08.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Exploring the Relationship Between Usage and Outcomes of an Internet-Based Intervention for Individuals With Depressive Symptoms: Secondary Analysis of Data From a Randomized Controlled Trial

Angel Enrique[1,2], PhD; Jorge E Palacios[1,2], MD, PhD; Holly Ryan[2], MSc; Derek Richards[1,2], PhD

[1]E-mental Health Research Group, School of Psychology, Dublin, Ireland

[2]Clinical Research & Innovation, Silvercloud Health Ltd, Dublin, Ireland

Corresponding Author:
Angel Enrique, PhD
E-mental Health Research Group
School of Psychology
University of Dublin Trinity College
Dublin, 2
Ireland
Phone: 353 1 554 9771
Email: enriquea@tcd.ie

Abstract

Background: Internet interventions can easily generate objective data about program usage. Increasingly, more studies explore the relationship between usage and outcomes, but they often report different metrics of use, and the findings are mixed. Thus, current evaluations fail to demonstrate which metrics should be considered and how these metrics are related to clinically meaningful change.

Objective: This study aimed to explore the relationship between several usage metrics and outcomes of an internet-based intervention for depression.

Methods: This is a secondary analysis of data from a randomized controlled trial that examined the efficacy of an internet-based cognitive behavioral therapy for depression (Space from Depression) in an adult community sample. All participants who enrolled in the intervention, regardless of meeting the inclusion criteria, were included in this study. Space from Depression is a 7-module supported intervention, delivered over a period of 8 weeks. Different usage metrics (ie, time spent, modules and activities completed, and percentage of program completion) were automatically collected by the platform, and composite variables from these (eg, activities per session) were computed. A breakdown of the usage metrics was obtained by weeks. For the analysis, the sample was divided into those who obtained a reliable change (RC)—and those who did not.

Results: Data from 216 users who completed pre- and posttreatment outcomes were included in the analyses. A total of 89 participants obtained an RC, and 127 participants did not obtain an RC. Those in the RC group significantly spent more time, had more log-ins, used more tools, viewed a higher percentage of the program, and got more reviews from their supporter compared with those who did not obtain an RC. Differences between groups in usage were observed from the first week in advance across the different metrics, although they vanished over time. In the RC group, the usage was higher during the first 4 weeks, and then a significant decrease was observed. Our results showed that specific levels of platform usage, 7 hours total time spent, 15 sessions, 30 tools used, and 50% of program completion, were associated with RC.

Conclusions: Overall, the results showed that those individuals who obtained an RC after the intervention had higher levels of exposure to the platform. The usage during the first half of the intervention was higher, and differences between groups were observed from the first week. This study also showed specific usage levels associated with outcomes that could be tested in controlled studies to inform the minimal usage to establish adherence. These results will help to better understand how to use internet-based interventions and what optimal level of engagement can most affect outcomes.

Trial Registration: ISRCTN Registry ISRCTN03704676; http://www.isrctn.com/ISRCTN03704676

International Registered Report Identifier (IRRID): RR2-10.1186/1471-244X-14-147

XSL•FO
RenderX

(J Med Internet Res 2019;21(8):e12775) doi:10.2196/12775

KEYWORDS

Web-based intervention; depression; adherence; engagement; eHealth; internet

Introduction

Background

Internet- and computer-based interventions for depression have been shown to be effective in several meta-analytic reviews, reporting comparable effects to face-to-face treatments [1-5]. These findings have led to the use of these interventions within stepped care and collaborative care models of mental health provision, which operate on the premise that not everybody requires a high-intensity treatment in the form of face-to-face therapy provided by a trained psychologist [6]. The inclusion of these interventions into mental health services may lead to benefits, such as reduced costs associated with treatment, reduced waiting-list burden, and increased accessibility to services [7,8]; however, results are mixed, and more research is needed to draw firmer conclusions [9].

Despite the success of internet-based interventions, many questions about how these interventions work and for whom they are most suited remain unanswered. Most of the studies about internet-based cognitive behavioral therapy (iCBT) only analyze outcomes at fixed time points, treating the intervention as a singular entity, without considering the way the program is actually adopted and used by the users [10]. In this sense, some research has shown that the users' uptake and long-term use of these technologies are lower than might be expected, and a median of 56% of the participants complete the whole program [11]. However, other research has shown that the users do not necessarily need to go through the entire program to benefit clinically [12]. Although treatment dropout and spontaneous remission might explain some of these trends, as they could in other psychological interventions, understanding more completely how internet-delivered interventions are used and whether their usage differs depending on user characteristics is an important area of investigation [13].

One of the benefits of internet-delivered interventions at a research level is that they facilitate the collection of objective data on usage and engagement. In recent years, several studies have turned their attention to the usage of internet-delivered intervention and how usage is related to outcomes [10,14]. Previous research has revealed a relationship between how much users were actually exposed to the program and outcomes; nevertheless, as different metrics of usage have been used across studies, and results are somewhat mixed, the specific contributions of these variables to outcomes remain unclear [15,16]. A review conducted by Donkin et al [15] compiled studies that looked at usage metrics and their relation to outcomes. Their findings showed that the number of log-ins and the number of modules completed were the most commonly reported metrics among different trials. Module completion was therefore found to be the most related metric to outcomes. However, these 2 metrics do not necessarily account for the depth of involvement of the users with the platform and its content. In this regard, it is possible to distinguish between *active* and *passive* engagement in internet-based interventions, where the former involves users interacting with the program and completing the activities, and the latter speaks about users who go through the program only superficially and therefore barely interacting with it [17,18]. Thus, to get the most from the intervention users' exposure to the platform, one not only has to consider module completion but also completing prescribed activities and homework [19,20]. For this reason, exploring the relationship among the various available usage metrics through the development of composite metrics could shed more light on the actual effects of usage and adherence on outcomes [15]. In fact, Donkin et al [21] found that a composite measure (average number of activities completed per log-in) was the only predictor of clinically significant change, in contrast to other metrics, such as the time spent on the platform, the number of modules completed, or the number of log-ins. More studies are needed to explore composite metrics and their relation to outcomes, as this may inform the minimal dose of iCBT needed to achieve significant clinical benefit.

Furthermore, there is a lack of agreement in how adherence to treatment is actually defined and measured [16]. Adherence is often reported in terms of attrition or dropout from a trial [13], that is, the number of users who cease to use the intervention and therefore do not complete the per protocol treatment. However, several authors highlight that this information provides limited insight into users' interaction with Web-based interventions and how their use might influence outcome measures [21]. On the other hand, recent studies have conceptualized adherence as the intended use of the platform or "the extent to which individuals should experience the content to derive maximum benefit as implied by its creators" [22], which is also known as the therapeutic dose [16]. Following Sieverink's argument [16], intended use should be considered the minimum use to establish adherence; however, current electronic health evaluations fail to demonstrate the optimal dose-response relationship. In this sense, some authors suggest that designers need to find the balance between the theoretically efficacious dose and the effective or actual dose that can be only determined upon application [23]. Therefore, finding the most relevant metrics and determining thresholds of usage for these metrics are key to determining the optimal dosage, which in turn may be different depending on the target population, setting, or the type of intervention that is evaluated [23].

Another important factor that can have an impact on the efficacy of and adherence to internet-based interventions is the role of support [24]. Support may take different forms, but generally speaking, it involves someone checking in with the patient, encouraging the user to continue going through the platform, and providing feedback on the basis of the patient's progress. Different meta-analytic reviews have shown that supported interventions have higher rates of adherence and better outcomes than self-guided ones [1,3,4,25]. However, the amount and frequency of the support needed to produce clinically significant

XSL•FO
RenderX

improvements are barely understood and results so far do not show differences among different doses of support [24,26,27]. In this sense, no studies have explored the role of metrics related to the support, such as the number of reviews and the number of user replies to these reviews, which could shed more light on the role of support in the usage of the program and their outcomes [28].

In summary, the literature shows that there is an association between higher usage leading to better outcomes, but it remains unclear which of these usage metrics are more strongly related to outcomes. Furthermore, few studies have attempted to determine an optimal dose-response relationship that could inform a threshold to establish adherence.

Objectives

This study was aimed at exploring the relationship between several usage metrics and outcomes from a sample of individuals with depressive symptoms and who were involved in a trial that evaluated the efficacy of a Web-based supported intervention for depression [29,30]. The specific goals of the study were as follows: (1) to explore the differences in usage between those who significantly improved and those who did not, (2) to analyze differences in usage across different sociodemographic and clinical variables of both groups, (3) to explore which of the usage metrics were more important in predicting clinically significant changes, and (4) to explore whether specific usage levels are associated with a clinically meaningful change.

Methods

Study Design

This study is a secondary analysis of data from a randomized controlled trial (RCT) that examined the efficacy of an iCBT intervention for depression in a sample of adults from a community setting. The protocol and the main outcome paper have been published elsewhere [29,30]. In the main study, participants were randomized to the internet-delivered intervention with support or the waiting-list control group. Assessment took place at baseline and at posttreatment, 8 weeks after randomization. The study protocol, information on the study, informed consent, and related materials were approved by the ethics committee at the School of Psychology, Trinity College Dublin (November 22, 2013). Participants who were excluded from the main RCT also completed consent, agreeing to have their data included for analysis. The trial is registered as a controlled trial with ISRCTN (ISRCTN03704676).

Sample and Recruitment

For details on the participant flow and characteristics, see Richards et al [29]. In summary, 641 users from the Aware charity expressed interest and applied to participate in the research. From them, 379 users were excluded for different reasons, such as Beck Depression Inventory 2nd edition (BDI-II) scores <14 (n=114); BDI-II>28 (n=211); suicidal intent/ideation (n=16); if they were currently receiving psychological treatment for depression (n=104); organic mental health condition (n=82); on medication for less than 1 month (n=106); alcohol or drug misuse (n=50); and reported depressive symptoms that preceded or coincided with a diagnosed medical condition (n=138). Even though these excluded participants were not included in the trial, they were offered the intervention with support, and they were also administered the primary outcome measures. The only difference between those individuals who were included and excluded from the trial is that the latter were not actively followed up to complete the posttreatment measures. For the purposes of this study and given that those participants excluded in the trial received the same intervention, all participants who logged in to the platform and completed the outcome measures upon completion of their treatment were included in the secondary analyses. Similarly, all those participants who were assigned to the waiting-list group and received the intervention after the waiting-list period were included in the secondary analyses, taking as their posttreatment scores the scores they provided upon completion of the intervention. With regard to those participants who were not part of the main RCT, we only selected those who completed the posttreatment outcomes within a period of 85 days after the first log-in. This time period was computed by calculating the average number of days that participants included in the trial took to fill the posttreatment measures (mean 66.86, SD 9.15), and 2 SDs were added. Thus, we excluded users who completed posttreatment measures beyond the intended assessment period.

Procedure

The study was advertised through the Aware website, and those individuals who expressed an interest to participate were emailed about the intervention study and directed to a website to access further information on the study and what would be involved in participating. Informed consent and baseline screening questionnaires were completed on the Web. Thereafter, those participants who met the inclusion criteria were randomly assigned to the intervention group or the waiting-list group, and those who did not meet the inclusion criteria but were interested in participating were also given access to the intervention. Once a participant was assigned to the active treatment at the first log-in, the participant received a message from the participant's supporter, and this support was then offered once a week for a period of 8 weeks. At the end of the 8-week period, participants were automatically asked to complete the outcome measures, and those in the intervention group of the trial, who did not complete the measures, were followed up by the research team to achieve the completion.

Intervention

Computerized Cognitive Behavioral Therapy Program

Space from Depression is a 7-module Web-based, cognitive behavioral therapy–based program for depression. This program was developed by SilverCloud Health, which is a company that develops Web-based interventions for mental health conditions. The intervention is delivered on a Web 2.0 platform, using media-rich interactive content. The treatment comprises cognitive and behavioral components, including self-monitoring and thought recording, behavioral activation, cognitive restructuring, and challenging core beliefs. These components are included across the 7 modules, although the program follows a nonlinear fashion, which means that the user can go directly to the module that is of interest to him or her. Each module follows a structured format that includes introductory quizzes,

XSL•FO
RenderX

audios, videos, informational content, personal stories, interactive activities, and homework suggestions. Space from Depression has been described in detail elsewhere [29,30].

Support

Participants were assigned a trained supporter who monitored their progress throughout the trial. These supporters were trained volunteers of the charity who received training in the SilverCloud platform and on how to deliver feedback. A dashboard interface provided supporters with an overview of their participants' level of engagement with the program content. The role of the supporter mainly comprised encouraging, supporting, and providing feedback to the users, and this feedback used to take between 10 and 15 min per participant. Support was offered once a week during the period of 8 weeks.

Measures

Primary Outcome

The primary outcome of the main RCT was the BDI-II [31]. The 21-item measure is a widely used questionnaire that assesses severity of depressive symptoms using a Likert scale ranging from 0 to 3 on the basis of the Diagnostic and Statistical Manual of Mental Disorders, 4th edition, diagnostic criteria. The scale designates levels of severity, minimal (0-13), mild (14-19), moderate (20-28), and severe (29-63) [31]. This instrument has shown good psychometric properties in several studies.

Usage Metrics

Total Time on the Platform

This metric corresponds to the combination of the time spent in each session (in min) from the first to the last log-in. Interactions lasting longer than 30 min are automatically counted as 1 min, to avoid counting long idle periods when the program is open toward the total count.

Number of Sessions

This metrics relates to the number of times (log-ins) the user accessed the program. If a specific session has inactivity periods longer than 3 hours, the next moment of activity will count as a new session.

Average Time (in min) Per Session

This composite measure results from dividing the total time on the platform by the number of sessions.

Number of Activities

This metric is calculated by counting all the times users interacted actively with the platform, that is, every time that they completed a journal entry, used an interactive tool, or downloaded or played relaxation audios. The program has a total of 17 interactive activities distributed across the 8 modules. Participants were able to use these activities as many times as they wished.

Activities Per Session

This is a composite measure resulting from dividing the number of activities completed by the number of sessions.

Percentage of the Program Viewed

This metric refers to the percentage of the total program content that the user has gone through.

Number of Reviews

This metric refers to the number of messages that the supporter sent to the user to encourage use of the platform while providing feedback about the progress from the last review.

Number of Review Notes

This metric relates to the number of replies that the user left for their supporter after a review.

Data Analysis

Data analyses were completed using SPSS 24 (IBM corporation). In the first place, *t* tests and chi-square tests were computed to explore potential differences in sociodemographic and clinical variables at baseline between those participants who met the inclusion criteria for the trial and those who did not. Boxplots of the sample were computed to look for any extreme outliers (3 box lengths away from the edge of their box). One-way analysis of variance (ANOVA) analyses and *t* tests were conducted to explore differences in usage across different sociodemographic and clinical variables at baseline. Pairwise comparisons applying Bonferroni correction were conducted between the subgroups associated with each category. An assessment of reliable change (RC) was made using criteria of a change of ≥9 points or greater on pre-to-post treatment BDI-II scores. Similar criteria have been used in other studies of internet-delivered interventions for depression [19,32,33]. *t* tests were performed to explore differences in the usage metrics between those who obtained an RC and those who did not. Variables identified as significantly associated with RC were further examined; 2×8 repeated-measures ANOVA were conducted to explore the change in the scores across the 8-week intervention period, comparing those who obtained an RC and those who did not. When sphericity was violated, Greenhouse-Geisser correction was applied for ANOVA analyses. Pairwise comparisons applying Bonferroni correction for multiple comparisons were conducted within groups, comparing week 1 with further weeks, and between groups, comparing the outcomes between groups for each specific week. To obtain the optimal cutoff for achieving RC, a receiver operating characteristic (ROC) curve analysis [34] was performed using each of the 4 individual usage metrics (total time spent on the platform, number of sessions, number of activities, and percentage of the program viewed) as test variables against the *RC* state variable. ROC curves are constructed by plotting true positive rates (sensitivity) against the false positive rates (specificity). The optimal cutoff was determined using the *point of curve closest to the (0,1)* criteria, which uses the formula $d2=[(1–Sn)2+(1–Sp)2]$, where Sn = sensitivity and Sp = specificity, to calculate the distance of each point to the (0,1) point representing maximal sensitivity and specificity [35].

XSL•FO
RenderX

Results

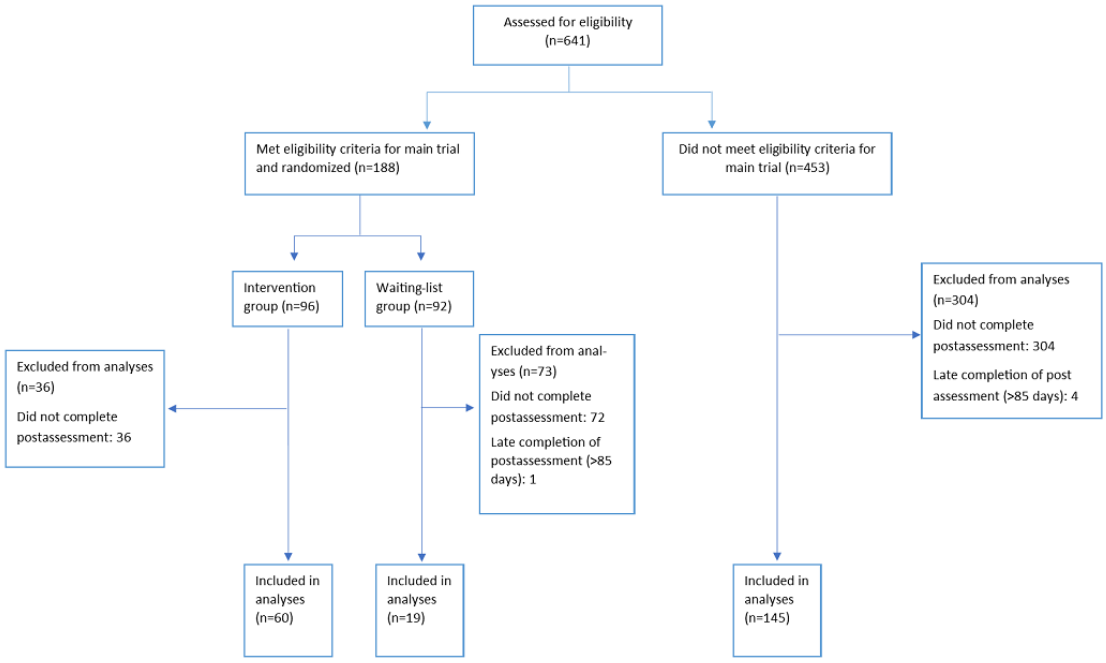
Overview

Of the 641 participants who showed interest in the study and provided consent, a total of 224 participants provided postintervention outcome data within a period of 85 days. Figure 1 illustrates the number of participants coming from each of the branches of the main trial. The number of participants included in this study that were excluded from the main RCT, and the reasons are as follows (participants can respond to more than one reason):

- BDI-II score<14 (n=52)
- BDI-II score>28 (n=56)
- Suicidal intent/ideation (n=3)
- Currently receiving psychological treatment for depression (n=30)
- Organic mental health condition (n=14)
- On medication for less than 1 month (n=26)
- Alcohol or drug misuse (n=14)
- Age<18 (n=1)
- Depressive symptoms that preceded or coincided with a diagnosed medical condition (n=46)

To ensure the homogeneity of the sample in terms of sociodemographic measures and usage metrics between those participants who met the inclusion criteria for the RCT (treatment group and waiting list) and those who were excluded, baseline differences in clinical and sociodemographic measures and usage metrics were examined. There were no differences between these groups in age, gender, marital status, BDI-II pre and postintervention, or any of the usage metrics included in the study. Before further statistical analyses, box plots were constructed for the different usage metrics to spot extreme outliers (3 box lengths away from the edge of their box). Owing to the high variability of the values on the usage metrics and to prevent the distribution of the sample to be very skewed, 8 participants were classified as extreme outliers, and they were excluded from further analyses, leaving a total sample of 216. To illustrate the type of outlier, one example was that of an individual user who spent a total time of 2526 min (42.11 hours) and completed 117 sessions and 321 activities, which is unrealistically far away from the average found in this study. In other words, these participants deviated substantially from the different usage patterns that an individual might take in this specific program.

Figure 1. Flowchart of the current study.



Platform Usage

Descriptive analyses of the total usage of the platform show that, on average, participants spent 339 min on the platform, they accessed the program 14 times, viewed 57% of the total program, completed 25 activities, got 7.5 reviews from the supporters, and left 1.62 messages to their supporters. ANOVA analyses were conducted to explore potential differences in the total usage of the platform across sociodemographic and clinical variables (Multimedia Appendix 1). Regarding age groups, univariate ANOVA models showed significant differences in the number of activities completed ($F_{3,212}$=4.03; P=.008) and activities per session by age group ($F_{3,212}$=2.95; P=.03). Pairwise comparisons showed that individuals in the age group of 31 to 40 years completed significantly more activities (mean difference=15.06, SE 5.65; P=.049) and more activities per session (mean difference=0.95, SE 0.34; P=.03) than those older than 50 years. Regarding depressive symptom severity at baseline, univariate ANOVA models showed significant differences between depression severity groups at baseline for total time spent ($F_{3,212}$=2.65; P=.049), number of sessions ($F_{3,212}$=5.75; P=.001), number of activities completed ($F_{3,212}$=2.87; P=.04), percentage viewed ($F_{3,212}$=2.82; P=.04), and number of reviews ($F_{3,212}$=3.7; P=.01). Pairwise

XSL•FO
RenderX

comparisons showed that users with minimal depressive symptoms had lower engagement overall. This group had significantly lower usage rates as compared with those with severe symptoms in terms of number of sessions (mean difference=6.31, SE 1.74; *P*=.002), number of activities (mean difference=14.19, SE 5.08; *P*=.03), and number of reviews (mean difference=0.78, SE=0.25; *P*=.01). It also had significantly lower number of sessions (mean difference=5.56, SE 1.66; *P*=.006) and percentage of program viewed (mean difference=0.15, SE 0.05; *P*=.04) as those with moderate depression, and it had significantly lower number of sessions than those with mild depression at baseline (mean difference= 6.12, SE 1.9; *P*=.009).

Platform Usage Associated With Reliable Change in Beck Depression Inventory 2nd Edition Scores

The sample of participants included in the study was divided between those who obtained an RC in the BDI-II (reduction of 9 points or greater on pre-to-post treatment BDI-II scores) and those who did not. Overall, 89 participants (41%, 89/216) reached an RC, and 127 (59%, 127/216) did not obtain an RC. *t* tests were run to determine whether there were differences in the usage of the platform between users of both groups. Results (see Table 1) showed significant differences in the total time spent on the platform (time spent), number of sessions, program viewed (percentage viewed), total activities completed (number of activities), and number of reviews in favor of those who obtained an RC, showing medium between-group effect sizes (Cohen *d*=0.45-0.61). No significant differences were obtained for min per session, activities per log-in, and number of review notes.

Table 1. Descriptive data and mean differences in usage metrics between those who reliably changed and those who did not.

Usage metrics	Reliable change (n=89), mean (SD)	No reliable change (n=127), mean (SD)	*t* test (df)	*P* value	Effect size (Cohen *d*)
Time spent	420.63 (280.77)	282.23 (253.24)	3.78 (214)	<.001	0.52
Sessions	17.63 (8.93)	12.22 (8.92)	4.38 (214)	<.001	0.61
Percentage viewed	67.56 (25.24)	50.46 (30.41)	4.50 (207.87)	<.001	0.61
Activities	33.28 (29.42)	19.62 (22.97)	3.67 (159.01)	<.001	0.52
Reviews	7.83 (0.77)	7.28 (1.57)	3.39 (194.60)	.001	0.45
Reviews notes	1.94 (2.29)	1.39 (2.10)	1.85 (214)	.06	0.25
Min per session	24.46 (13.43)	22.68 (16.86)	0.83 (214)	.41	0.12
Activities per session	1.84 (1.35)	1.53 (1.74)	1.41 (214)	.16	0.20

[a]Bonferroni correction applied (alpha=.05/8=.006).

Weekly Usage of the Platform

Weekly usage of the program was explored among the 4 usage variables that were significant in the previous analyses, namely time spent, number of sessions, number of activities, and percentage viewed (Figures 2-5). Number of reviews were not included, as all participants got 1 review per week, as this was established in the protocol, unless they dropped out from the treatment. To explore potential differences in these variables over time and between individuals who showed RC and individuals who did not among the 8-week intervention period, 2×8 repeated-measures ANOVA analyses were computed. With regard to the time spent, analyses showed significant time ($F_{4.58,980.08}$=21.91; *P*<.001) and interaction effects ($F_{4.58,980.08}$=2.77; *P*=.02). Pairwise comparisons between groups showed significant differences between groups in the time spent on the platform from week 1 to 4 and week 6. Pairwise comparisons within groups showed that, in the group of individuals with RC, the time spent in week 1 was not significantly different compared with weeks 2, 3, and 4, but it was significantly higher compared with week 5 and the following weeks, indicating that the time spent during the first 4 weeks was longer than the following 4 weeks.

Regarding the number of sessions, there was a significant difference in time ($F_{4.99,1067.45}$=21.03; *P*<.001); however, interaction effects were not significant ($F_{4.99,1067.45}$=1.82; *P*=.11), indicating that, altogether, there were no differences between conditions in the number of sessions performed. However, pairwise comparisons showed significant differences between groups in the number of sessions in each week across the 8-week period. Pairwise comparisons within the RC group showed nonsignificant differences in the number of sessions done in week 1 compared with weeks 2, 3, and 4. However, when comparing week 1 to 5 and following weeks, significant differences were observed, indicating again that the number of sessions was more similar from week 1 to 4, and there was a significant decrease in the number of sessions in the following weeks.

With regard to the percentage viewed, ANOVA analysis showed significant time ($F_{4.73,1012.57}$=43.96; *P*<.001) and interaction effects ($F_{4.73,1012.57}$=2.71; *P*=.02). Between-group pairwise comparisons showed significant differences in the percentage viewed in weeks 1, 2, and 4, and this percentage was nonsignificant for the other weeks. Pairwise comparisons within the RC group showed that the weekly percentage viewed in the first week was significantly higher compared with all the other weeks.

XSL•FO
RenderX

Figure 2. Time spent on the platform per week and divided between those who got a reliable change (reduction of 9 points or greater on pre-to-post treatment Beck Depression Inventory 2nd Edition) and those who did not. Significant differences were found in between-group comparisons at weeks 1-4 and week 6; significant within-group differences were found between week 1 and weeks 5-8 for the reliable change (RC) group; significant within-group differences were found between week 1 and weeks 6-8 for the no RC group.

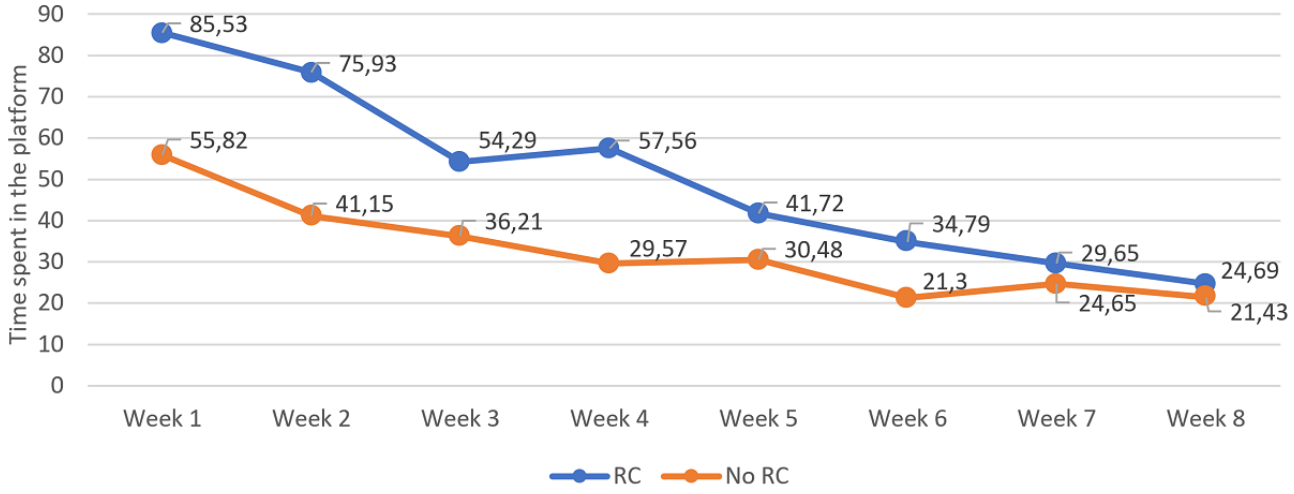
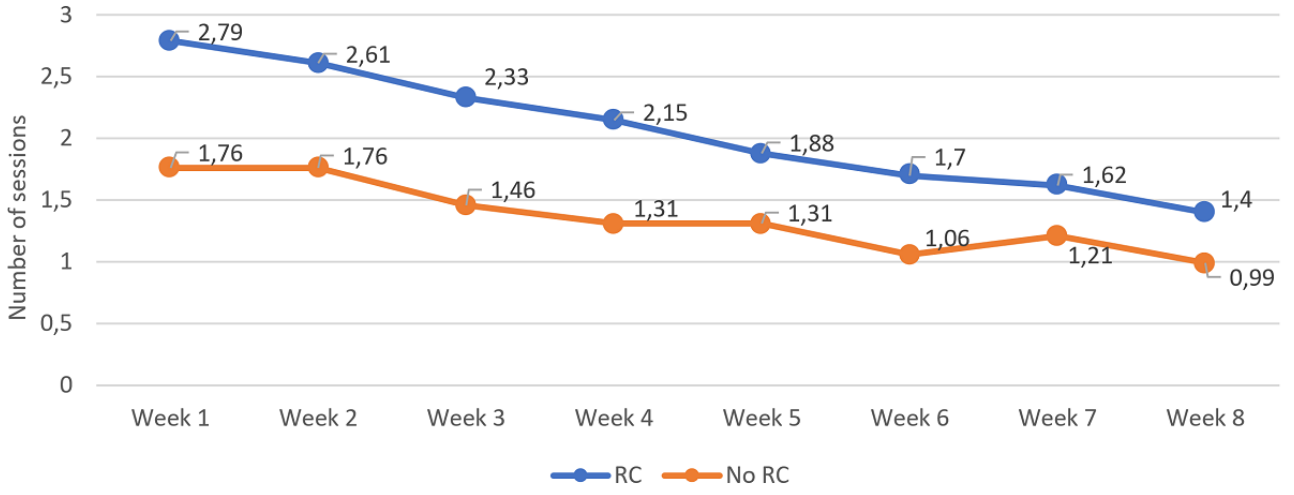


Figure 3. Number of sessions performed per week and divided between those who got a reliable change (reduction of 9 points or greater on pre-to-post treatment Beck Depression Inventory 2nd Edition) and those who did not. Significant differences were found in between-group comparisons at weeks 1-8; significant within-group differences were found between week 1 and weeks 5-8 for the reliable change (RC) group; significant within-group differences were found between week 1 and weeks 6 and 8 for the no RC group.



XSL•FO
RenderX

Figure 4. Percentage of the program viewed per week and divided between those who got a reliable change (reduction of 9 points or greater on pre-to-post treatment Beck Depression Inventory 2nd Edition) and those who did not. Significant differences were found in between-group comparisons at weeks 1, 2 and 4; significant within-group differences were found between week 1 and weeks 2-8 for the reliable change (RC) group; significant within-group differences were found between week 1 and weeks 2-8 for the no RC group.

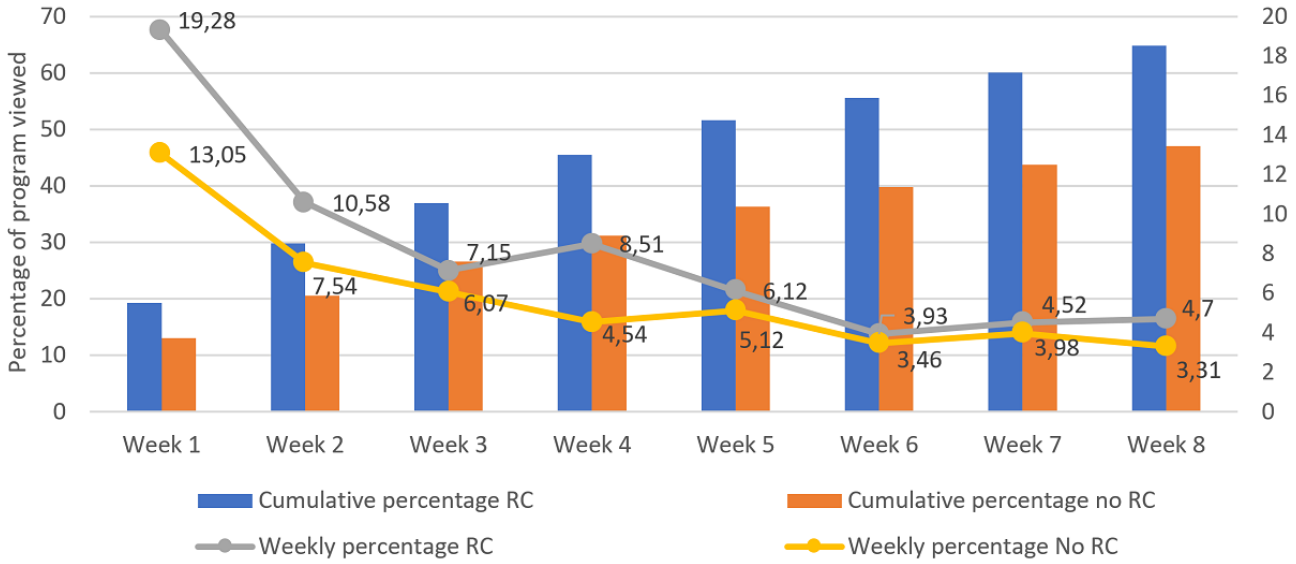
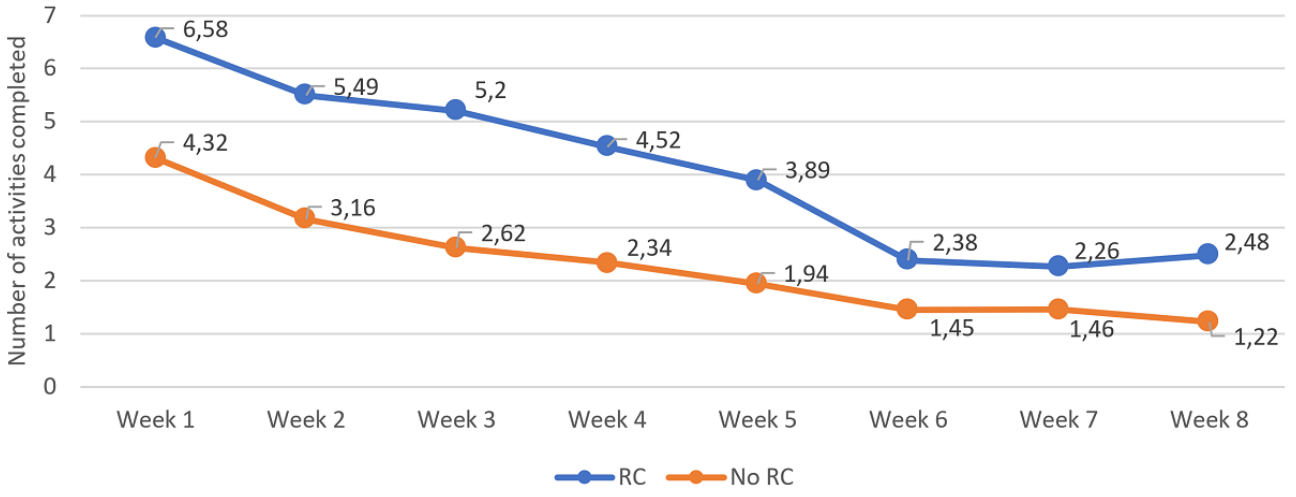


Figure 5. Number of activities completed per week and divided between those who got a reliable change (reduction of 9 points or greater on pre-to-post treatment Beck Depression Inventory 2nd Edition) and those who did not. Significant differences were found in between-group comparisons at weeks 1-6 and 8; significant within-group differences were found between week 1 and weeks 5-8 for the reliable change (RC) group; significant within-group differences were found between week 1 and weeks 6 and 8 for the no RC group.



Finally, regarding the number of activities, ANOVA analysis showed significant effects for time ($F_{5.12,1095.63}$=16.64; P<.001), but it did not show significant effects for the interaction between groups ($F_{5.12,1095.63}$=1.12; P=.35), indicating that there were no overall differences between conditions in the number of activities completed. Between-group pairwise comparisons showed significant differences in the number of activities completed in nearly all the weeks, with the exception of week 7. Within-group pairwise comparisons of the group with RC showed that there were significant differences only when comparing the number of tools used between week 1 and week 5 and the following weeks.

Overall, these analyses show that the weekly usage was consistently larger for the group with RC in the different metrics among the first half of the intervention period, and from here, the differences between groups started to reduce. Within the RC group, a pattern can be observed across nearly all of the metrics where the usage of the first week was not significantly different from the usage on weeks 2, 3, and 4, but it was significantly different from week 5 in advance.

Receiver Operating Characteristic Curve Analysis

A total of 4 separate ROC curve analyses were run with the following results: for total time spent on platform, the optimal cutoff was 420 min, with a specificity of 76% and sensitivity of 47%. For number of sessions, the optimal cutoff was 15, with a specificity of 71% and a sensitivity of 61%. For the percentage of the program viewed, the optimal cut-off was 50%, and this was related to a specificity of 54% and a sensitivity of 79%. Finally, for number of activities, the optimal cut-off was 30 tools used, which had a specificity of 77% and a sensitivity of 46%. The area under the curve values ranged from .65 to .68, with 95% CIs between .58 and .76, considered within an

XSL•FO
RenderX

acceptable range for using these variables to differentiate between those who reliably changed and those who did not show RC as defined in the study as a movement of ≥9 points on the BDI-II. When using these cutoffs as measures of intended use, those using the platform for at least 420 min or 7 hours (n=62) had a 58% RC rate, those logging in for at least 15 sessions (n=91) had a 59% RC rate, those using tools (activities) at least 30 times (n=61) had a 57% RC rate, and those using at least 50% of the program (n=115) had a 52% RC rate. All these values are significantly higher than the overall RC rate of 41% for the entire sample. Finally, of those who reached all 4 optimal cutoffs (n=34), the RC rate was 62%.

Discussion

Principal Findings

This study was intended to explore the relationship between usage of the platform and the outcomes in a sample of participants who used an iCBT intervention for depression. Overall, the results showed that those individuals who obtained an RC after the intervention had higher levels of exposure to the platform, in terms of time spent, number of sessions, percentage of the program viewed, and number of activities, compared with those who did not. Differences in program usage between those who improved and those who did not were observed from the first week, although these differences started to vanish during weeks 5 to 8 of the intervention. Furthermore, in the RC group, the usage of the platform during the first week was not significantly different in comparison to the first half of the intervention period (until week 4), and these differences turned significant when compared with week 5 in advance, indicating a significant reduction of the usage during the second half of the intervention period. Finally, independent ROC curve analyses showed that 7 hours of time spent on the platform, logging in 15 times, and completing 30 activities over the intervention course were associated with the achievement of RC. In other words, the likelihood of being in the RC group, compared with the likelihood of being in the no RC group, was highest at these thresholds.

With regard to the sociodemographic and clinical variables and their relation to program usage, our results showed that individuals older than 50 years completed fewer activities compared with younger cohorts. These results add up to the contradictory literature about age and adherence, where different studies are finding different directions in the relationship between age and adherence [36,37]. It may be the case that age on its own is not as important as the interaction of age with other factors, such as computer literacy, which makes a difference in this relationship. This study also showed that individuals with minimal depressive symptoms at baseline had significantly lower levels of usage compared with users with higher levels of symptomatology. This could be explained by the fact that the intervention is ideally intended for individuals with mild-to-severe depression, and this cohort of less depressed users did not require the same grade of exposure to get benefits, or the program indeed fell outside their needs, as it is a treatment intervention. In this sense, these participants might benefit more from preventive approaches, such as a resilience intervention. Our results here highlight the necessity and importance of delivering the most appropriate intervention at the right time. Still, it is worth mentioning that although the literature in this regard shows mixed results [36], a related study identified important benefits of the Space from Depression program for those with subclinical symptoms [38].

In this study, on average, the users utilized 57% of the program, which is similar to completion rates in other internet-based interventions for depression [36,39]. When comparing the usage between those who obtained an RC and those who did not, results showed that the former significantly spent more minutes, accessed more times, completed more activities, were exposed to more content, and received more reviews. It should be noted that the time spent on the platform by those who achieved an RC was around 7 hours, which is slightly higher but still similar to the 6 hours found in previous studies, and the time spent by those who did not achieve an RC is similar with previous results reporting between 4 and 5 hours [21,37]. However, the literature around the relationship between usage and outcomes in internet-based interventions for mental health shows contradictory results, as variables, such as time spent and number of sessions, are not consistently related to outcomes, and studies with greater statistical power need to be conducted to shed more light on this relationship [15]. Our results confirm the largely used statement *the more usage, the better* [5,16], although it is important to note that our variables account for *active* (ie, activities completed) and *passive* (ie, percentage viewed) engagement, and both elements are important to get the most out of these interventions [20]. On the other hand, composite measures, such as minutes per session and activities per session, were not significantly different between conditions, which contradicts results obtained by Donkin et al [21]. The nonlinear fashion of Space from Depression and the fact that this program is very focused on the usage of tools might explain the absence of differences between these groups in their behavior within sessions, where the amount of time and tools used was not significantly different between those who achieved an RC and those who did not.

With regard to the weekly usage, our results showed that participants who obtained an RC had significantly larger exposure levels to the intervention compared with those who did not improve, and these differences were more consistent among the first 4 weeks. Focusing on the usage over time of those who obtained an RC, the results showed similar usage levels among the first 4 weeks, but the results showed a significant subsequent decrease in the second half of the intervention. Overall, these results are in line with another study, where the program was mostly used during the first half of the intervention period [37]. In this sense, these results suggest that the usage levels during the first month might be key for improvement, and strategies for enhancing engagement at this stage could be beneficial. One of these strategies could be outcome and engagement monitoring, inbuilt in the feedback system, so that participants could be flagged if they were deviating from the expected results of the intervention, and causes for this could be explored and addressed [40]. In a similar vein, a recent study found that when therapists were given outcome feedback about patients who were deteriorating during

XSL•FO
RenderX

the intervention period, these patients had significantly less severe symptoms after treatment compared with similar patients assigned to therapists who were not receiving this feedback [41].

The ROC curve analyses are exploratory in nature, and the results do not allow to draw firmer conclusions about optimal usage levels; however, they can be understood as a first step toward determining specific thresholds that could be tested in controlled and experimental designs. As recommended by some authors, the optimal dose needs to consider the balance between user's burden and adherence to ensure that the effective (observed) dose is as close as possible to the efficacious dose, which is not dependent on adherence [23]. For this specific intervention, the maximal efficacious dose is to complete the 7 modules during an 8-week period at the pace of 1 module per week; however, our findings seem to indicate that the effective dose would not require the completion of all the modules. In 3 out of 4 measured variables, the specificity of the optimal dose was high, and the RC rates were close to 60% in those who reached this usage cutoff. The optimal cutoff for the percentage of program viewed variable showed low specificity, signifying that a minimal 50% use of the platform would have a high percentage of false positives, not ideal for determining intended use. Overall, in this population at least, 7 hours of platform usage spread out over 15 sessions and completion of 30 activities (these include repeat activities for learning key skills) over a maximum period of 12 weeks were associated with achieving a clinically significant change. Nevertheless, these results have to be considered within their context, and further studies with changes to settings, types of programs used, and other study features, such as population characteristics, may yield different results. However, the parameters found for this particular intervention are worth exploring further.

This paper contributes to the concept of adherence by providing an empirical justification of intended use, that is, "the extent to which individuals should experience the content to derive maximum benefit from the intervention, as defined or implied by its creators" [22]. As suggested by different authors, there is a need for demonstrating the dose-response relationship, and this paper has done so through the consideration of different metrics. In this sense, future studies should explore whether similar exposure levels are needed for RC in different platforms, and future studies should determine which are the actual tools that have been used as some studies have already done [37]. Although our results will help identify optimal levels of exposure to maximize the benefits from these interventions, future studies should be conducted to go deeper into the usage patterns, as they can take many different pathways, and different types of usage can lead to benefits [42]. Log data of those participants who reliably improved would be a way of understanding different successful patterns of usage and which tools or modules are most related to change.

Limitations

This study also has some limitations. First, although the sample of this study comes from an RCT, this substudy was observational in nature, and no manipulation of the variables related to usage was done. Given this, it is not possible to establish causal relationships between usage metrics and outcomes. Future analyses, such as cross-lagged models, could focus on whether higher usage rates lead to lower symptoms over time or whether decreasing symptoms is what drives higher usage. This may answer the question of causality in this association. RCTs could also look into comparing whether a group pushed and receiving recommendation for reaching certain usage levels will perform better than a group permitted to use an intervention freely. Moreover, the inclusion of only those participants who reported posttreatment outcomes might be limiting the generalizability of the results, as it has been found that participants who engage more are also more likely to respond to follow-up assessments [43]. Thus, it might be possible that levels of usage were lower for those who did not complete the postassessments. Another limitation relates to the number of reviews variable, as supporters were encouraged to offer reviews regardless of whether participants were actively using the platform. Future studies should provide different indications to the supporters, so that the supporters do not need to waste time writing reviews for users who are not logging in to the platform. For example, supporters could have 2 attempts to contact participants who are not using the platform and then discharge them from services if the answer is not given. Finally, the metric percentage viewed, which only accounts for new content viewed, not considering content reviewed, has been shown as a key element of internet-based interventions usage [39]. These results could also explain why the relationship of this variable with outcomes is not as clear as the others.

Conclusions

This study has used different ways to explore the relationship between usage of an iCBT intervention for depression and the outcomes achieved by participants. The results seem to reinforce the notion that *the more usage of a Web-based intervention, the better* but with some nuances. Thus, the usage during the first half of the intervention was significantly higher compared with the second half, which might have implications for engagement and how best to offer the support in the early stages. Furthermore, this study suggests that it may be possible to determine, at least preliminarily, an optimal dose that would need to be tested and replicated to draw firmer conclusions. If confirmed, these thresholds could be used to establish cutoffs of adherence to the intervention. Future studies should continue to explore the relationship between usage and outcomes to better understand how internet-delivered interventions work and how to make them more responsive to varying degrees of usage. The continuation of this line of research could lead us to a future where a responsive intervention takes into account usage levels, allowing for tailoring in real time to enhance the engagement of different participants and thus maximizing likelihood of a positive outcome.

XSL•FO
RenderX

Acknowledgments

The main trial was jointly funded from money and resources provided by SilverCloud Health Ltd and Aware Charity, Ireland. The current secondary analysis study was funded by SilverCloud Health Ltd.

Conflicts of Interest

AE, JP, HR, and DR are employees of SilverCloud Health, developers of computerized psychological interventions for depression, anxiety, stress, and comorbid long-term conditions.

Multimedia Appendix 1

Mean and SDs of usage metrics of those who completed posttreatment outcomes (n=216) divided by clinical and sociodemographic factors.

[PDF File (Adobe PDF File), 62KB - jmir_v21i8e12775_app1.pdf]

References

1. Andersson G, Cuijpers P. Internet-based and other computerized psychological treatments for adult depression: a meta-analysis. Cogn Behav Ther 2009;38(4):196-205. [doi: 10.1080/16506070903318960] [Medline: 20183695]
2. 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]
3. 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]
4. Wells MJ, Owen JJ, McCray LW, Bishop LB, Eells TD, Brown GK, et al. Computer-assisted cognitive-behavior therapy for depression in primary care: systematic review and meta-analysis. Prim Care Companion CNS Disord 2018 Mar 1;20(2):pii: 17r02196 [FREE Full text] [doi: 10.4088/PCC.17r02196] [Medline: 29570963]
5. Wright JH, Owen JJ, Richards D, Eells TD, Richardson T, Brown GK, et al. Computer-assisted cognitive-behavior therapy for depression: a systematic review and meta-analysis. J Clin Psychiatry 2019 Mar 19;80(2):pii: 18r12188 [FREE Full text] [doi: 10.4088/JCP.18r12188] [Medline: 30900849]
6. Bower P, Gilbody S. Stepped care in psychological therapies: access, effectiveness and efficiency. Narrative literature review. Br J Psychiatry 2005 Jan;186:11-17. [doi: 10.1192/bjp.186.1.11] [Medline: 15630118]
7. Nordgren LB, Hedman E, Etienne J, Bodin J, Kadowaki A, Eriksson S, et al. Effectiveness and cost-effectiveness of individually tailored internet-delivered cognitive behavior therapy for anxiety disorders in a primary care population: a randomized controlled trial. Behav Res Ther 2014 Aug;59:1-11 [FREE Full text] [doi: 10.1016/j.brat.2014.05.007] [Medline: 24933451]
8. Rollman BL, Belnap BH, Abebe KZ, Spring MB, Rotondi AJ, Rothenberger SD, et al. Effectiveness of online collaborative care for treating mood and anxiety disorders in primary care: a randomized clinical trial. JAMA Psychiatry 2018 Jan 1;75(1):56-64 [FREE Full text] [doi: 10.1001/jamapsychiatry.2017.3379] [Medline: 29117275]
9. Kolovos S, van Dongen JM, Riper H, Buntrock C, Cuijpers P, Ebert DD, et al. Cost effectiveness of guided internet-based interventions for depression in comparison with control conditions: an individual-participant data meta-analysis. Depress Anxiety 2018 Dec;35(3):209-219 [FREE Full text] [doi: 10.1002/da.22714] [Medline: 29329486]
10. Sieverink F, Kelders S, Poel M, van Gemert-Pijnen L. Opening the black box of electronic health: collecting, analyzing, and interpreting log data. JMIR Res Protoc 2017 Aug 7;6(8):e156 [FREE Full text] [doi: 10.2196/resprot.6452] [Medline: 28784592]
11. 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]
12. Christensen H, Mackinnon A. The law of attrition revisited. J Med Internet Res 2006 Sep 29;8(3):e20; author reply e21 [FREE Full text] [doi: 10.2196/jmir.8.3.e20] [Medline: 17032636]
13. 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]
14. Fuhr K, Schröder J, Berger T, Moritz S, Meyer B, Lutz W, et al. The association between adherence and outcome in an internet intervention for depression. J Affect Disord 2018 Dec 15;229:443-449. [doi: 10.1016/j.jad.2017.12.028] [Medline: 29331706]
15. Donkin L, Christensen H, Naismith SL, Neal B, Hickie IB, Glozier N. A systematic review of the impact of adherence on the effectiveness of e-therapies. J Med Internet Res 2011 Aug 5;13(3):e52 [FREE Full text] [doi: 10.2196/jmir.1772] [Medline: 21821503]

XSL•FO
RenderX

16. Sieverink F, Kelders SM, van Gemert-Pijnen JE. Clarifying the concept of adherence to ehealth technology: systematic review on when usage becomes adherence. J Med Internet Res 2017 Dec 6;19(12):e402 [FREE Full text] [doi: 10.2196/jmir.8578] [Medline: 29212630]
17. Bendelin N, Hesser H, Dahl J, Carlbring P, Nelson KZ, Andersson G. Experiences of guided internet-based cognitive-behavioural treatment for depression: a qualitative study. BMC Psychiatry 2011 Jun 30;11:107 [FREE Full text] [doi: 10.1186/1471-244X-11-107] [Medline: 21718523]
18. Han JY, Kim JH, Yoon HJ, Shim M, McTavish FM, Gustafson DH. Social and psychological determinants of levels of engagement with an online breast cancer support group: posters, lurkers, and nonusers. J Health Commun 2012;17(3):356-371 [FREE Full text] [doi: 10.1080/10810730.2011.585696] [Medline: 22085215]
19. de Graaf LE, Huibers MJ, Riper H, Gerhards SA, Arntz A. Use and acceptability of unsupported online computerized cognitive behavioral therapy for depression and associations with clinical outcome. J Affect Disord 2009 Aug;116(3):227-231. [doi: 10.1016/j.jad.2008.12.009] [Medline: 19167094]
20. 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 Dec;7(2):254-267 [FREE Full text] [doi: 10.1007/s13142-016-0453-1] [Medline: 27966189]
21. Donkin L, Hickie IB, Christensen H, Naismith SL, Neal B, Cockayne NL, et al. Rethinking the dose-response relationship between usage and outcome in an online intervention for depression: randomized controlled trial. J Med Internet Res 2013 Oct 17;15(10):e231 [FREE Full text] [doi: 10.2196/jmir.2771] [Medline: 24135213]
22. Kelders SM, Kok RN, Ossebaard HC, van Gemert-Pijnen JE. Persuasive system design does matter: a systematic review of adherence to web-based interventions. J Med Internet Res 2012 Nov 14;14(6):e152 [FREE Full text] [doi: 10.2196/jmir.2104] [Medline: 23151820]
23. Voils CI, King HA, Maciejewski ML, Allen KD, Yancy WS, Shaffer JA. Approaches for informing optimal dose of behavioral interventions. Ann Behav Med 2014 Dec;48(3):392-401 [FREE Full text] [doi: 10.1007/s12160-014-9618-7] [Medline: 24722964]
24. Baumeister H, Reichler L, Munzinger M, Lin J. The impact of guidance on internet-based mental health interventions — a systematic review. Internet Interv 2014 Oct;1(4):205-215. [doi: 10.1016/j.invent.2014.08.003]
25. Andrews G, Basu A, Cuijpers P, Craske MG, McEvoy P, English CL, et al. Computer therapy for the anxiety and depression disorders is effective, acceptable and practical health care: an updated meta-analysis. J Anxiety Disord 2018 Dec;55:70-78 [FREE Full text] [doi: 10.1016/j.janxdis.2018.01.001] [Medline: 29422409]
26. Aardoom JJ, Dingemans AE, Spinhoven P, van Ginkel JR, de Rooij M, van Furth EF. Web-based fully automated self-help with different levels of therapist support for individuals with eating disorder symptoms: a randomized controlled trial. J Med Internet Res 2016 Dec 17;18(6):e159 [FREE Full text] [doi: 10.2196/jmir.5709] [Medline: 27317358]
27. Klein B, Austin D, Pier C, Kiropoulos L, Shandley K, Mitchell J, et al. Internet-based treatment for panic disorder: does frequency of therapist contact make a difference? Cogn Behav Ther 2009;38(2):100-113. [doi: 10.1080/16506070802561132] [Medline: 19306149]
28. Mohr DC, Cuijpers P, Lehman K. Supportive accountability: a model for providing human support to enhance adherence to ehealth interventions. J Med Internet Res 2011 Mar 10;13(1):e30 [FREE Full text] [doi: 10.2196/jmir.1602] [Medline: 21393123]
29. Richards D, Timulak L, O'Brien E, Hayes C, Vigano N, Sharry J, et al. A randomized controlled trial of an internet-delivered treatment: its potential as a low-intensity community intervention for adults with symptoms of depression. Behav Res Ther 2015 Dec;75:20-31. [doi: 10.1016/j.brat.2015.10.005] [Medline: 26523885]
30. Richards D, Timulak L, Doherty G, Sharry J, Colla A, Joyce C, et al. Internet-delivered treatment: its potential as a low-intensity community intervention for adults with symptoms of depression: protocol for a randomized controlled trial. BMC Psychiatry 2014 May 21;14:147 [FREE Full text] [doi: 10.1186/1471-244X-14-147] [Medline: 24886179]
31. Beck AT, Steer RA, Brown GK. BDI-II, Beck Depression Inventory: Manual. San Antonio, TX: Psychological Corporation; 1996.
32. Meyer B, Berger T, Caspar F, Beevers CG, Andersson G, Weiss M. Effectiveness of a novel integrative online treatment for depression (Deprexis): randomized controlled trial. J Med Internet Res 2009 May 11;11(2):e15 [FREE Full text] [doi: 10.2196/jmir.1151] [Medline: 19632969]
33. Titov N, Andrews G, Davies M, McIntyre K, Robinson E, Solley K. Internet treatment for depression: a randomized controlled trial comparing clinician vs technician assistance. PLoS One 2010 Jun 8;5(6):e10939 [FREE Full text] [doi: 10.1371/journal.pone.0010939] [Medline: 20544030]
34. Greiner M, Pfeiffer D, Smith RD. Principles and practical application of the receiver-operating characteristic analysis for diagnostic tests. Prev Vet Med 2000 May 30;45(1-2):23-41. [doi: 10.1016/S0167-5877(00)00115-X] [Medline: 10802332]
35. Kumar R, Indrayan A. Receiver operating characteristic (ROC) curve for medical researchers. Indian Pediatr 2011 Apr;48(4):277-287. [doi: 10.1007/s13312-011-0055-4] [Medline: 21532099]
36. Christensen H, Griffiths KM, Farrer L. Adherence in internet interventions for anxiety and depression. J Med Internet Res 2009 Apr 24;11(2):e13 [FREE Full text] [doi: 10.2196/jmir.1194] [Medline: 19403466]

XSL•FO
RenderX

37. Mattila E, Lappalainen R, Välkkynen P, Sairanen E, Lappalainen P, Karhunen L, et al. Usage and dose response of a mobile acceptance and commitment therapy app: secondary analysis of the intervention arm of a randomized controlled trial. JMIR Mhealth Uhealth 2016 Jul 28;4(3):e90 [FREE Full text] [doi: 10.2196/mhealth.5241] [Medline: 27468653]
38. Earley C, Joyce C, McElvaney J, Richards D, Timulak L. Preventing depression: qualitatively examining the benefits of depression-focused iCBT for participants who do not meet clinical thresholds. Internet Interv 2017 Sep;9:82-87 [FREE Full text] [doi: 10.1016/j.invent.2017.07.003] [Medline: 30135841]
39. Kelders SM, Bohlmeijer ET, van Gemert-Pijnen JE. Participants, usage, and use patterns of a web-based intervention for the prevention of depression within a randomized controlled trial. J Med Internet Res 2013 Aug 20;15(8):e172 [FREE Full text] [doi: 10.2196/jmir.2258] [Medline: 23963284]
40. Shimokawa K, Lambert MJ, Smart DW. Enhancing treatment outcome of patients at risk of treatment failure: meta-analytic and mega-analytic review of a psychotherapy quality assurance system. J Consult Clin Psychol 2010 Jun;78(3):298-311. [doi: 10.1037/a0019247] [Medline: 20515206]
41. Delgadillo J, de Jong K, Lucock M, Lutz W, Rubel J, Gilbody S, et al. Feedback-informed treatment versus usual psychological treatment for depression and anxiety: a multisite, open-label, cluster randomised controlled trial. Lancet Psychiatry 2018 Dec;5(7):564-572. [doi: 10.1016/S2215-0366(18)30162-7] [Medline: 29937396]
42. Morrison C, Doherty G. Analyzing engagement in a web-based intervention platform through visualizing log-data. J Med Internet Res 2014 Nov 13;16(11):e252 [FREE Full text] [doi: 10.2196/jmir.3575] [Medline: 25406097]
43. Murray E, White IR, Varagunam M, Godfrey C, Khadjesari Z, McCambridge J. Attrition revisited: adherence and retention in a web-based alcohol trial. J Med Internet Res 2013 Aug 30;15(8):e162 [FREE Full text] [doi: 10.2196/jmir.2336] [Medline: 23996958]

Abbreviations

ANOVA: analysis of variance
BDI-II: Beck Depression Inventory 2nd edition
iCBT: internet-based cognitive behavioral therapy
RC: reliable change
RCT: randomized controlled trial
ROC: receiver operating characteristic

Edited by G Eysenbach; submitted 09.11.18; peer-reviewed by O Perski, R Schuster; comments to author 31.03.19; revised version received 04.06.19; accepted 10.06.19; published 01.08.19.

Please cite as:
Enrique A, Palacios JE, Ryan H, Richards D
Exploring the Relationship Between Usage and Outcomes of an Internet-Based Intervention for Individuals With Depressive Symptoms: Secondary Analysis of Data From a Randomized Controlled Trial
J Med Internet Res 2019;21(8):e12775
URL: https://www.jmir.org/2019/8/e12775/
doi:10.2196/12775
PMID:31373272

©Angel Enrique, Jorge E Palacios, Holly Ryan, Derek Richards. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 01.08.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

An Internet-Based HIV Self-Testing Program to Increase HIV Testing Uptake Among Men Who Have Sex With Men in Brazil: Descriptive Cross-Sectional Analysis

Raquel Brandini De Boni[1], MD, MSc, PhD; Valdilea Gonçalves Veloso[1], MD, MSc, PhD; Nilo Martinez Fernandes[1], MSc, PhD; Flavia Lessa[1], MSc; Renato Girade Corrêa[2], MSc; Renato De Souza Lima[2], BA; Marly Cruz[3], PhD; Juliane Oliveira[4], MD; Simone Muniz Nogueira[4], BA; Beto de Jesus[3], BSc; Toni Reis[5], PhD; Nena Lentini[6], BA; Raquel Lima Miranda[6], PhD; Trista Bingham[7], MPH, PhD; Cheryl C Johnson[8], MPH; Aristides Barbosa Junior[6], MD, PhD; Beatriz Grinsztejn[1], MD, MSc, PhD

[1]Instituto Nacional de Infectologia Evandro Chagas, Oswaldo Cruz Foundation, Rio de Janeiro, Brazil

[2]National Department of STI, AIDS, and Viral Hepatitis, Ministry of Health, Brailia, Brazil

[3]Escola Nacional de Saúde Pública, Oswaldo Cruz Foundation, Rio de Janeiro, Brazil

[4]Secretaria Municipal de Saúde Curitiba, Curitiba, Brazil

[5]Grupo Dignidade, Curitiba, Brazil

[6]Centers for Disease Control and Prevention-Brazil, Brasilia, Brazil

[7]Division of Global HIV & TB, Center for Global Health, Centers for Disease Control and Prevention, Atlanta, GA, United States

[8]HIV Department, World Health Organization, Geneva, Switzerland

Corresponding Author:
Raquel Brandini De Boni, MD, MSc, PhD
Instituto Nacional de Infectologia Evandro Chagas
Oswaldo Cruz Foundation
Av Brasil 4365
Rio de Janeiro, 20530-360
Brazil
Phone: 55 21 3865 9122
Email: raqueldeboni@gmail.com

Abstract

Background: Approximately 30% of people living with HIV worldwide are estimated to be unaware of their infection. HIV self-testing (HIVST) is a strategy recommended by the World Health Organization to increase access to and uptake of testing among key populations who are at high risk for HIV infection.

Objective: This study aimed to describe the development and feasibility of a free, anonymous, internet-based HIVST strategy designed for men who have sex with men in Curitiba, Brazil (electronic testing [e-testing]).

Methods: The project was developed under the scope of the "A Hora é Agora" (The Time is Now) program. Individuals aiming to request an HIVST package (two tests each) answered an anonymous 5-minute questionnaire regarding inclusion criteria and sexual risk behavior. Eligible individuals could receive one package every 6 months for free. Website analytics, response to online questionnaires, package distribution, and return of test results were monitored via a platform-integrated system.

Results: Between February 2015 and January 2016, the website documented 17,786 unique visitors and 3218 completed online questionnaires. Most individuals self-reported being white (77.0%), young (median age: 25 years, interquartile range: 22-31 years), educated (87.3% completed secondary education or more), and previously tested for HIV (62.5%). Overall, 2526 HIVST packages were delivered; of those, 542 (21.4%) reported a result online or by mail (23 reactive and 11 invalid). During the study period, 37 individuals who reported using e-testing visited the prespecified health facility for confirmatory testing (30 positive, 7 negative).

Conclusions: E-testing proved highly feasible and acceptable in this study, thus supporting scale-up to additional centers for men who have sex with men in Brazil.

XSL•FO
RenderX

(J Med Internet Res 2019;21(8):e14145) doi:10.2196/14145

KEYWORDS

HIV/AIDS; HIV self-testing; key populations; mobile health; men

Introduction

In 2016, it was estimated that 30% of people living with HIV worldwide were unaware of their infection [1]. This gap represents a major challenge to achieving the targets of the Joint United Nations Programme on HIV/AIDS (UNAIDS) "90-90-90," which aims to ensure that 90% of all people living with HIV are aware of their status by 2020 [2].

Unknown HIV status is associated with late entry into care and increased mortality rates and is a key driver of the epidemic [3-5]. An HIV test result is the first step in engaging in the HIV treatment continuum for those who test positive or in the HIV prevention continuum for those who test negative. Nonetheless, many barriers to expand HIV testing coverage persist, such as stigma, difficulty in accessing testing facilities, and inconvenient clinic hours [6,7].

HIV self-testing (HIVST) allows individuals to perform an HIV test and interpret their own results [8]. Different strategies for HIVST distribution worldwide have been proposed, such as through peers [9], nongovernment organizations (NGOs) [10], and mobile/social media apps [11]. Randomized controlled trials have shown that HIVST increases uptake and frequency of HIV testing overall, without adverse events, social harm, or increased risk behaviors [12]. Nevertheless, a review of HIVST implementation studies by Estem et al [13] concluded that despite the many potential benefits of HIVST, dissemination strategies need to be improved and the high cost of HIVST kits needs to be reduced to scale-up this strategy.

Key populations, including gay and other men who have sex with men (MSM), are greatly affected by HIV in Latin America and other regions [14,15]. Such populations have continued high HIV incidence and encounter multiple structural barriers, including stigma and discrimination, which increase their risk of HIV exposure and acquisition and inhibit access to evidence-based HIV prevention and treatment services [16]. In Brazil, the prevalence of HIV among MSM, estimated at 14% in 2009 [17] and 17.5% in 2016 [18], requires sustained prevention and treatment efforts. Furthermore, despite an increased proportion of MSM reporting annual testing (from 21.2% in 2009 to 43.3% in 2016) [19], innovations are needed to increase MSM testing coverage in the country. The convenience and discreet approach of self-testing would be particularly useful to increase first-time HIV testing and frequency among this vulnerable population.

Internet-based and mobile phone technologies are promising tools to overcome HIVST delivery barriers among MSM [20]. In the United Kingdom and United States, both online and social media–based promotion and delivery of free HIVST kits effectively increased the frequency of testing and reached MSM who were not previously tested [21,22]. In accordance with international data [23,24], a previous Brazilian study suggested that HIVST would be well accepted among MSM [25]. Thus, providing HIVST using these technologies may represent a step forward in increasing testing access in Brazil, where free HIV testing is already provided by the Brazilian Public Health System ("Sistema Unico de Saúde").

Considering the limited access to and uptake of HIV testing among MSM in Brazil and the evidence supporting its acceptability, this paper reports on the development and feasibility of an internet-based HIVST (electronic testing [e-testing]) approach. E-testing was designed to promote HIV prevention by providing free anonymous HIVST and to enhance linkage to HIV care for those with a confirmed HIV positive status. We describe the results from the e-testing program's first year, when HIVST was not available in Brazil outside the research studies.

Methods

Design

This cross-sectional analysis describes the results obtained from the e-testing strategy between February 2015 and January 2016.

E-testing was developed as part of the broader initiative "A Hora é Agora" (AHA; or "The Time is Now"), which included three community-based strategies to deliver rapid HIV testing: a mobile testing unit, an NGO site, and e-testing. AHA was implemented in partnership with the Oswaldo Cruz Foundation (FIOCRUZ); the Department of STI, AIDS, and Viral Hepatitis of the Brazilian Ministry of Health; the Municipal Health Secretariat of Curitiba; the Federal University of Paraná; Grupo Dignidade (a local Lesbian Gay Bisexual Transgender and Queer/Questioning NGO); and Centers for Disease Control and Prevention (CDC) - Brazil. Implementation of the AHA program was focused in Curitiba, a city with 1.75 million inhabitants in southern Brazil and an estimated HIV prevalence of 19.9% (95% CI 14.2-27.2) among MSM [18].

Recruitment

The AHA program counted on an extensive communications plan to increase HIV testing and target young MSM [26]. The appropriateness of the communications strategy was piloted in focus groups including the target population, conducted in the preformative research (which also mapped the main MSM gathering sites and possible community partners to the program). The communication plan was continuously monitored and evaluated during the entire program via independent focus groups conducted in 2016 and Google Analytics.

Briefly, the communications included printed handouts distributed via in-person outreach events in places where MSM socialize in the city; partnerships with gay and MSM-friendly establishments such as saunas, movie theatres, cafes, and bars; and virtual messages disseminated trough a Facebook page and gay online sites such as ManHunt and Grindr. Examples of the materials used for AHA are presented in Figure 1.

XSL•FO
RenderX

Figure 1. Sample of the "A Hora é Agora" communications strategy for Curitiba, Brazil (2015-2016).

Study Population

Participants were eligible to receive an HIVST package if they were male at birth, were at least 18 years of age, resided in Curitiba, were HIV-negative or had an unknown HIV status, had access to the internet, and agreed to participate in the study after reading the online informed consent form.

E-Testing Web Platform

The internet platform was made available to all at www.ahoraeagora.org and could be accessed via personal computer, smartphones, or tablets (apps were available for IOS and Android). This platform contained four modules (ie, website sections presenting different access credentials): general information, HIVST order, management, and monitoring. The general information module contained HIV prevention information targeted to MSM, a step-by-step video, written instructions for using HIVST, geocoded options for HIV testing locations in Curitiba, and a personal risk assessment calculator (Figure 2).

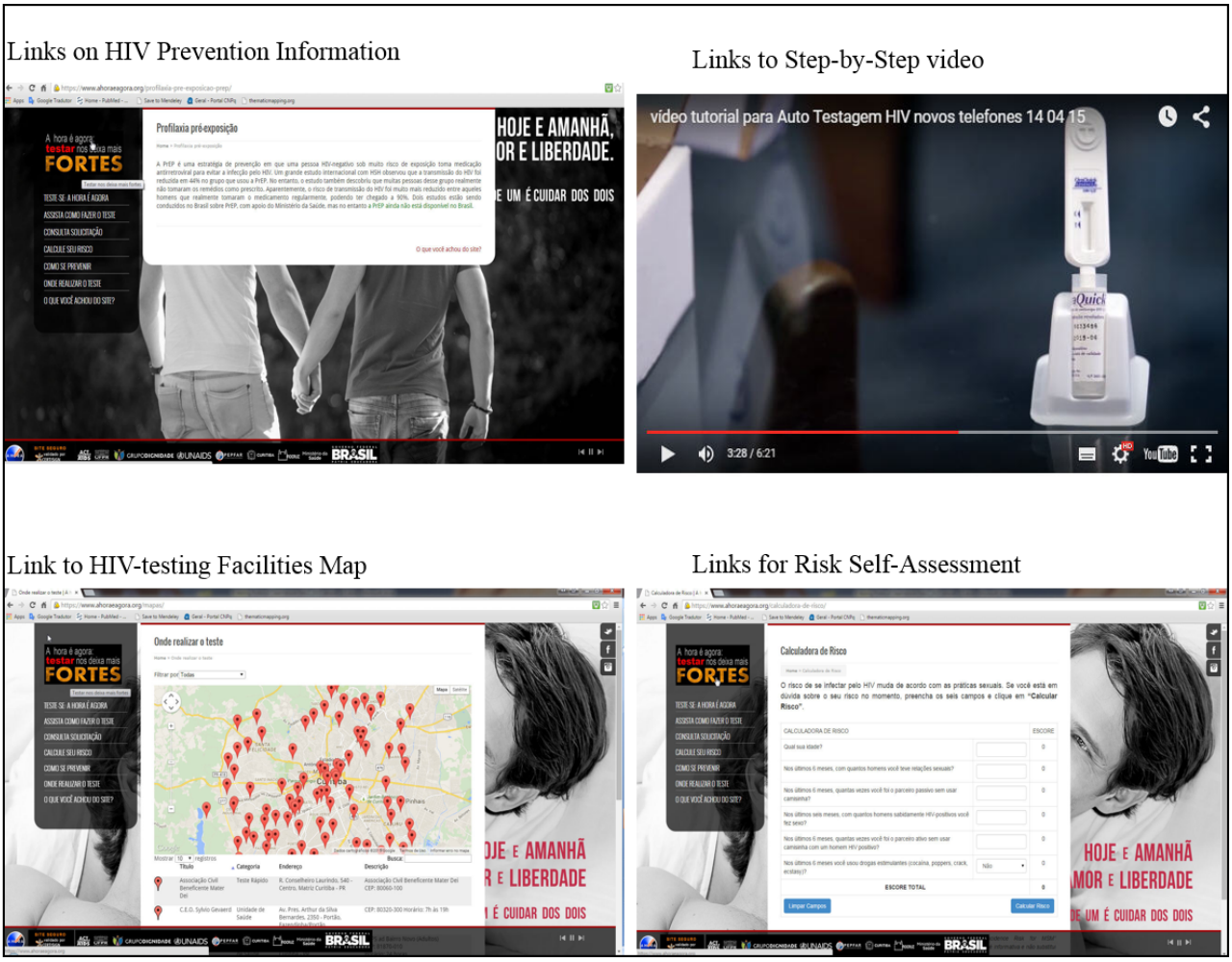
The risk assessment calculator was adapted from the HIV incidence risk for MSM scale, a 7-item questionnaire developed to predict HIV seroconversion among MSM [27] and recommended by the CDC to screen individuals for eligibility for pre-exposure prophylaxis [28]. Individuals could check their risk status and receive feedback, but data were not collected/saved in the platform. (Figure 3).

The HIVST order module included an online questionnaire and the option to request HIVST. This module evaluated eligibility criteria and enabled test kit delivery via standard Brazilian mail or self-pick up at a central pharmacy within 2 weeks. After completing the mandatory section of the questionnaire and selecting a delivery option, a random personal identification number (PIN) was assigned by the automated system. The PIN was used to track future HIVST requests, collect self-reported HIVST results, and monitor subsequent linkage to further testing and HIV care.

The management module encompassed the administration and logistics functions of HIVST package delivery (including central and pharmacy inventory control) and generated management reports for authorized staff. Finally, the fourth module monitored access and usability statistics of the AHA website (using Google Analytics), aggregated data from the order module, and generated routine reports.

XSL•FO
RenderX

Figure 2. E-testing Web platform: general information module.



XSL•FO
RenderX

Figure 3. E-testing Web platform: risk assessment calculator.

HIV Self-Testing Package

MSM who met the eligibility criteria and accessed the e-testing website could order an HIVST package. The package included two OraQuick Advanced HIV-1/2 HIVST kits, (OraSure Technologies, Inc, Bethlehem, PA, adapted and repackaged for self-testing in Brazil), condoms, lubricants, written self-test instruction, an anonymous prepaid card for returning test results (for those preferring to return the results by mail), and information about health facilities that provided free confirmatory HIV testing. Between March 7, 2015, and May 4, 2015, and between August 14, 2015, and February 29, 2016, only one HIVST was sent in each package. Using their PIN, individuals could order a new package every 6 months.

Measures

Website Analytics, HIV Self-Testing Distribution, and Platform Usability

Website analytics were measured using Google Analytics. Data on HIVST distribution were retrieved from the management module. Platform usability was measured using the following questions:

- Question: "Overall, did you find this platform easy to use?" Response: Yes/No
- Question: "What were the major difficulties you find in the platform?" Responses: "I was not able to find the links I was looking for" / "I would like to have access to more tests" / "I would like to retrieve the tests in other places" / "I did not find any difficulties" / "Other-specify"
- Question: "Did you find the testing instructions clear?" Response: Yes/No.

XSL•FO
RenderX

Characteristics of the Study Population

Individuals requesting an HIVST package were encouraged to answer the online questionnaire available in the HIVST order module. The questionnaire comprised two parts: the first was mandatory to initiate the HIVST request and included the eligibility criteria (eg, participant's age, city, and unknown or negative HIV status). The second part was optional and evaluated demographic characteristics (eg, schooling, color/race, gender and sexual orientation, and steady partner), risk perception (using the question "What is your likelihood of getting HIV in the next 12 months?") [29], and risk behavior for HIV infection (using the HIV incidence risk for the MSM scale) [27].

HIV Self-Testing Result Return

Individuals were asked to return their results either by directly entering result information into the website (using their PINs) or via the prepaid postal card. Additionally, when uploading results to the website, individuals were invited to answer a questionnaire regarding their experience and any possible social harm (using three dichotomous questions regarding coercion, feelings of shame, or violent reactions of partners).

Confirmatory Testing

Individuals with a reactive or invalid test result were instructed to seek confirmatory testing at Curitiba's reference center for HIV Counseling and Testing (COA) and to bring their PIN. The COA follows the HIV testing algorithm of the 2014 Brazilian Ministry of Health guidelines [30], and those confirmed to be HIV-positive were offered immediate HIV care. A question regarding the AHA testing strategies was included in the regular COA form, and staff reported the number of tests performed as a result of referrals from specific AHA project strategies.

Feasibility Outcomes

Project success was defined *a priori* as ≥60% individuals ordering HIVST after starting the online request process; ≥50% individuals picking up the HIVST at the pharmacy within 2 weeks; 1000 HIVST packages (2000 HIVST kits) distributed in 12 months; 20% HIVST results returned via mail or the online platform; and ≥50% individuals, who self-reported a reactive HIVST, accessing the COA for confirmatory testing and HIV care. There were no *a priori* outcomes related to the second HIVST kit included in the package.

Although there was scarce evidence to determine all the feasibility outcomes at the time of the study design (2014), a limited-advertised Brazilian Web survey was able to reach 629 MSM in around 10 days, of whom 82 were from the South region of the country. The vast majority (89.8%) of them reported high interest in using HIVST [25]. As the South of the country has three state capitals, and this project had a strong communications plan, we expected to distribute around 80 HIVST per month (rounded to 1000/year) in the selected South Capital (Curitiba). For the outcomes that we had no previous data on, we considered the most conservative estimate (50%) to be acceptable. Finally, regarding return of the testing results, we considered the general low response of mail and email surveys conducted in different areas of knowledge [31-34] and the additional absence of incentives; as such, we expected that 20% of the results would be returned to the project.

Statistical Analysis

We present the results of the online questionnaires and HIVST results returned using descriptive statistics, including both absolute counts and relative frequencies and 95% CIs. Data were analyzed using SPSS software (IBM Corp, Armonk, NY).

Ethical Issues

Details of informed consent were displayed in pop-up windows preceding the online questionnaire. No personally identifying information was collected in the informed consent forms or questionnaires to preserve users' anonymity.

Personal information necessary for HIVST delivery was saved in the Web server and was restricted to trained staff in charge of package mail delivery. All staff members signed a confidentiality agreement. Identification for HIVST delivery was not linked to individuals' questionnaire responses. To ensure information security and confidentiality, the website used verified access to the original site, data encryption between personal computers and the website, and encryption of database tables. Information was only available to users registered with a login and a safe password.

Trained counselors at a 24/7 hotline were available to assist in the event of any psychological crisis. Social harm, such as coercion or violent reactions of partners, was considered an adverse event and was assessed in the feedback questionnaire.

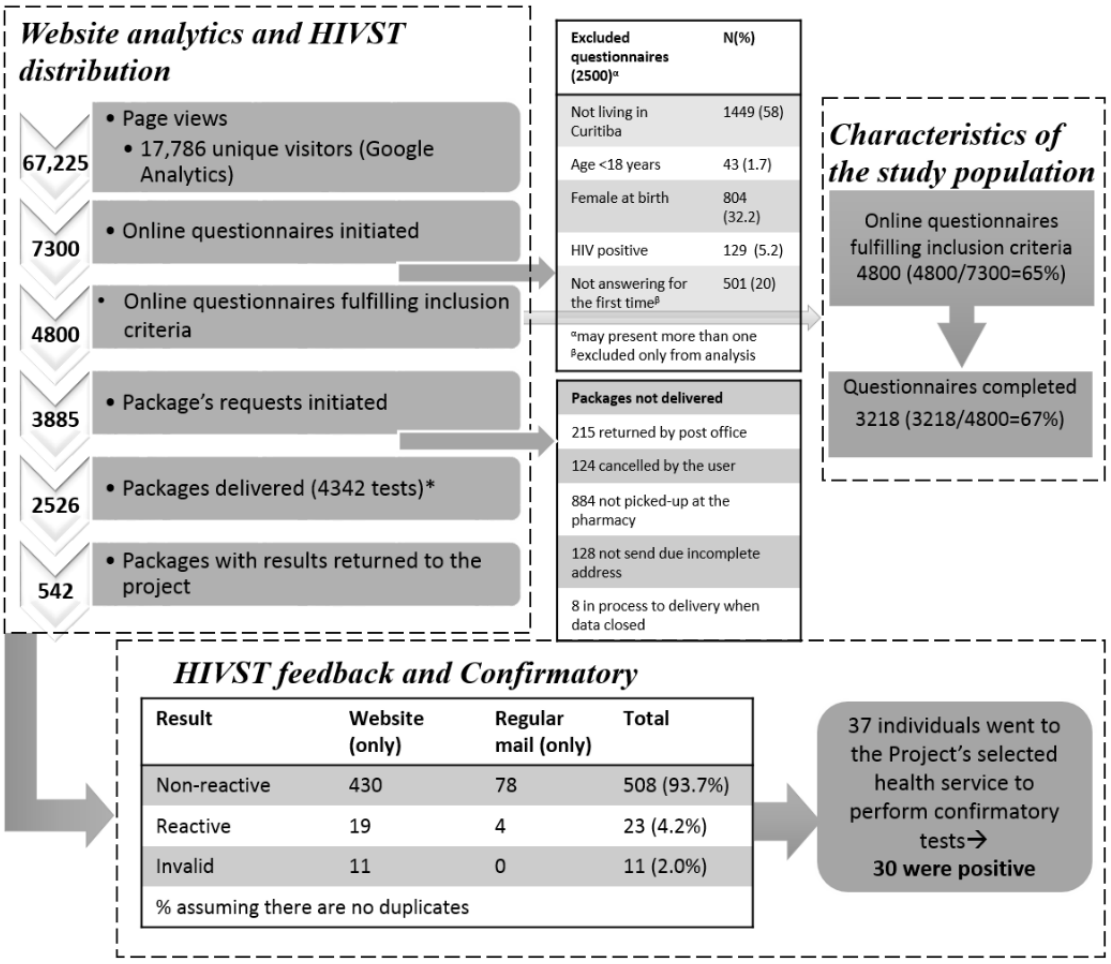
The e-testing study protocol was approved by the Evandro Chagas National Institute of Infectious Diseases/FIOCRUZ Institutional Review Board (CAAE number: 37848114.9.0000.5262). The study was reviewed according to the Centers for Disease Control and Prevention human research protection procedures and was determined to be research but the CDC was not engaged.

Results

The detailed frequencies and data sources of the e-testing study during its first year of implementation are provided in Figure 4.

XSL•FO
RenderX

Figure 4. E-testing project flowchart for Curitiba, Brazil (2015-2016). *Between March 7, 2015, and May 4, 2015, and between August 14, 2015, and February 29, 2016, only one HIVST was sent in each package; therefore, the total number of unique HIVST kits distributed is 4342. HIVST: HIV self-testing.



Website Analytics, HIV Self-Testing Distribution, and Platform Usability

During the study period, the AHA website had 67,225 page views from 17,786 unique visitors. The most frequently viewed pages were those in which users could order an HIVST package by mail (9.7%), all pages containing information on HIV prevention (6.6%), and the page on risk self-assessment (6.5%; Table 1). The e-testing app had 2571 downloads (1780 downloads at Google Play and 791 at Apple Store).

Overall, 3885 questionnaire respondents initiated an HIVST request, of which 2526 packages were delivered (215 requests were returned by mail, 124 were cancelled by the users, 884 were not picked up at the pharmacy within 2 weeks, 128 were not sent due to incomplete addresses, and 8 were in the process of delivery at the time the present dataset was closed; Figure 4). Most deliveries were made by the Brazilian mail (n=1943, 77%), and 583 (23%) packages were picked up at the pharmacy. A total of 66 PINs (of 3885, 1.7%) requested at least one additional HIVST package during the study period.

Of all individuals accessing the website, 362 answered questions about the usability of the website. Most respondents found the site "Very easy" (n=260, 71.8%) or "Easy" (n=71, 19.6%) to use. Furthermore, 72% (262/362) stated that they did not have difficulties navigating the website, and only 21 (5.8%) stated that they did not find the pages they were searching for. The vast majority (n=339, 93.6%) also found that the website's instructions for performing HIVST were clear.

XSL•FO
RenderX

Table 1. Pages viewed (n=67,225) on the A Hora é Agora e-testing website in Curitiba, Brazil (February 6, 2015, to January 31, 2016).

Pages viewed	Value, n (%)[a]
Any of the following links on HIV prevention information	4425 (6.6)
Viral load and treatment as prevention	337 (0.5)
Circumcision	750 (1.1)
Sexual transmitted diseases	280 (0.4)
Interrupted intercourse	746 (1.1)
Pre-exposure prophylaxis	418 (0.6)
Postexposure prophylaxis	583 (0.9)
Oral sex	946 (1.4)
Condom use	365 (0.5)
Link for risk self-assessment	4398 (6.5)
Link to HIV-testing facilities map	3420 (5.1)
Link to HIVST[b] request by mail	6546 (9.7)
Link to pick up HIVST in the pharmacy	1917 (2.9)
Link to step-by-step video	4118 (6.1)

[a]Proportion of hits in a specific link over total hits for the website.

[b]HIVST: HIV self-testing.

Characteristics of the Study Population

During the study period, 65% (4800/7300) of individuals who initiated an HIVST request met the eligibility criteria. Of the men who completed the online questionnaire (n=3218), most were young (median: 25 years; interquartile range [IQR]: 22-31 years), were of white race/ethnicity (n=2478, 77%), had completed at least secondary school (n=2810, 87%), and had never had an HIV test (n=1206, 37.5%; Table 2). The frequency of possible reasons for never testing were “I am afraid of having a positive result” (27.6%), “I feel ashamed” (21.7%), “It is inconvenient to go to a health care clinic” (15.7%), “I am not at risk of being infected” (10.4%), and “I feel lazy” (7.4%); 17.3% did not want to answer this question.

XSL•FO
RenderX

Table 2. Demographic and behavioral characteristics of men who completed the online questionnaire at the A Hora é Agora e-testing website in Curitiba, Brazil 2015-2016 (N=3218).

Characteristic	Value, n (%)
Demographics	
Age (years), median (IQR[a])	25 (22-31)
Race/ethnicity	
White	2478 (77.0)
Black	103 (3.2)
Mixed	495 (15.4)
Native	22 (0.7)
Asian	49 (1.5)
Don’t know/don’t want to answer	71 (2.2)
Schooling	
No schooling/incomplete primary school	85 (2.6)
Complete primary/incomplete secondary school	261 (8.1)
Complete secondary/incomplete college	1549 (48.1)
Complete college/graduate/professional school	1261 (39.2)
Don’t know/don’t want to answer	62 (1.9)
Sexual identity/orientation	
Gay/homosexual	1953 (60.7)
Bisexual	370 (11.5)
Heterosexual	649 (20.2)
Transvestite, transsexual, transgender	22 (0.7)
Other	25 (0.8)
Don’t know/don’t want to answer	199 (6.2)
Steady partner	
Yes	1334 (41.5)
No	1677 (52.1)
Don’t want to answer	207 (6.4)
Risk perception	
Perceived likelihood of acquiring HIV in the next 12 months	
None/little chance	1973 (61.3)
Some chance/high chance/certainly	266 (8.3)
Don’t want to answer	978 (30.4)
Had a previous HIV test in lifetime	
Yes	2012 (62.5)
No	1206 (37.5)
Risk behavior for HIV infection	
Number of male sexual partners (6 months)	
0-5	1486 (46.2)
6-10	345 (10.7)
>10	197 (6.1)
Don’t know/don’t want to answer	1190 (37.0)
Number of times participant was receptive to partner without condom (6 months)	

XSL•FO
RenderX

Characteristic	Value, n (%)
None	14 (0.4)
Once or more	1097 (34.1)
Don't know/don't want to answer	2107 (65.5)
Number HIV-positive partners (6 months)	
None	34 (1.0)
1	231 (7.2)
>1	37 (1.2)
Don't know/don't want to answer	2916 (90.6)
Number of times participant had condomless insertive anal sex with HIV-positive partner (6 months)	
0-4	419 (13.0)
≥5	149 (4.7)
Don't know/don't want to answer	2650 (82.3)
Stimulant use (6 months)[b]	
Yes	525 (16.3)
No	2466 (76.6)
Don't know/don't want to answer	227 (7.1)
STI[c] diagnosis in prior 6 months	
Yes	249 (7.8)
No	2728 (84.7)
Don't know/don't want to answer	241 (7.5)

[a]IQR, interquartile range.

[b]Stimulants include cocaine, crack, poppers, and ecstasy.

[c]STI: sexually transmitted infection.

Return of HIV Self-Testing Results

Of the 2526 packages delivered, 21.4% (n=542) returned an HIVST result on the website or by mail. Of these, 19 reactive and 11 invalid test results were entered into the website platform and 4 reactive results were returned by mail (as of February 29, 2016).

There were 91 responses to the feedback survey on self-testing, but we received no reports (questionnaire or hotline) of psychological crises or violent reactions from partners. However, one individual reported feeling ashamed because someone saw him self-testing.

Confirmatory Testing

As of February 2016, 37 individuals were tested at the COA reporting to have used the e-testing: 30 had HIV-positive results and 7 had negative results.

Feasibility Outcomes

Project success was achieved for most of the previously defined outcomes (Table 3).

Over the study period, the telephone hotline received 70 calls: two were questions regarding what to do following a reactive result, three were related to assistance with interpreting the self-test result, five were related to queries on test performance, six were related to problems with HIVST delivery, and the remaining were about HIV testing in general and not about self-testing/e-testing.

XSL•FO
RenderX

Table 3. Feasibility outcomes defined in the study protocol and achievements in the A Hora é Agora e-testing project in Curitiba, Brazil (2015-2016).

Measurement	Proportion	Indicator	*A priori* success indicator
Individuals who got an HIVST[a] package after starting the online request process	2526/3885	65.0%	≥60%
Individuals who collected the HIVST at the distribution pharmacy, within 2 weeks of ordering	544/1417	38.4%	≥50%
HIVST packages distributed during the initial 12-month timeframe (either sent by mail or retrieved at the pharmacy)[b]	Not applicable	2526[c]	≥1000[c]
Proportion of HIVST packages with a test result returned[d] via mail or the website up to February 29, 2016	(82[e]+460[f])/2526	21.4%	≥20%
Proportion of those with a positive result who accessed the Counseling and Testing Center for confirmatory testing up to February 29, 2016	30[g]/34[h]	88.2%	≥50%

[a]HIVST: HIV self-testing.

[b]Between March 7, 2015, and May 4, 2015, and between August 14, 2015, and February 29, 2016, only one HIVST was sent in each package; therefore, the total number of unique HIVST kits distributed is 4342.

[c]The N value is presented.

[d]Different date (February 29, 2016) was considered in this indicator to provide enough time for individuals to receive/retrieve tests and include the results on the platform.

[e]Overall, there were 90 results returned by mail, but 8 were also included on the site (n=82 test results returned by mail). Four unique reactive results were returned by mail.

[f]460 test results returned on the website.

[g]Overall, 37 individuals went to Counseling and Testing Center to access confirmatory testing: 30 were positive and 7 were negative. Given the absence of identification, we cannot confirm that these are the same individuals who reported their results into the site or by mail.

[h]19 reactive+11 invalid inserted in the platform+4 reactive returned by mail up to February 29, 2016.

Discussion

We describe an internet-based strategy to provide HIV prevention information and HIVST to MSM in Curitiba, Brazil. In the first year, we delivered 2526 HIVST packages, substantially more than predicted in the study design (1000 packages). Our findings are similar to reports from previous studies that have shown good uptake and use of websites containing information on HIV prevention, self-assessment of risk, and HIVST among MSM [35,36]. Such high uptake is likely due to the significant number of MSM who use social media and the internet to look for sexual partners [37]. A previous Web-based survey in 10 Brazilian cities [38] found that 1798 (35.6%) and 678 (13.4%) of responding MSM reported daily and weekend use of websites and smart phone apps to seek sexual partners.

The use of internet and social media approaches for health promotion and behavioral interventions has become more routine according to international literature and already includes smoking cessation [39] and heart disease prevention programs [40]. The internet and social media are particularly promising ways to offer testing to MSM at high ongoing risk of HIV [41], as those using social media and the internet to meet sexual partners also have high-risk behavior and increased HIV risk [42,43]. Thus, internet-based health initiatives can help overcome health disparities to access prevention and care, particularly among MSM, and can help facilitate uptake of HIV testing, prevention, and treatment services.

E-testing appeared to appeal to never-tested MSM who reported a fear of having a positive result, shame (which is probably related to stigma), and unfriendly health facilities. These results are consistent with reports showing that HIVST is highly acceptable to MSM [25,44] and can facilitate uptake and frequency of testing. This demand was likely driven by the privacy of e-testing, which systematic reviews have highlighted as a priority for MSM [23]. Similarly, the large number of HIVST packages delivered by mail, compared to packages retrieved in the public pharmacy, was likely related to introducing delivery models that addressed challenges such as travel distances, wait times at facilities, and confidentiality issues [45,46], all of which have been previously cited as barriers to testing.

Of the individuals who answered the entire online questionnaire, most identified as young, white, well-educated, and without a steady partner. Although these characteristics were similar to those reported in other studies among MSM in the country [38,47], there were some differences as compared to a recent household probability survey conducted among men from Curitiba [48]: In our study, participants were younger, more likely to be white (77% vs 64%), less likely to perceive high risk for HIV infection (8% vs 52%), and less likely to have ever had an HIV test (62.5% vs 75.7%) than found in the survey [48].

However, the latest Brazilian estimates have shown an increase in the HIV prevalence among young MSM [18], who also showed riskier sexual behavior than older MSM [19]. This new evidence reinforces the importance of strategies to successfully reach this key population. Our findings suggest that, at least for young MSM, important subpopulations are open to new HIV testing strategies and understand the need for frequent or risk-based HIV testing.

XSL•FO
RenderX

In our study, HIVST was provided for free but without the provision of any incentives (the Brazilian Ethic regulations do not allow the provision of incentives for research participants or patients). We received 21.4% of HIVST results (18.2% through the website and 3.2% by mail). This was a substantially lower proportion than presented in China by Zhong et al [49], who reported 97% of self-testers returning results when HIVST kit costs were refunded upon return of results. We believe that this difference is related to the provision of incentives, and not providing incentives was one of the reasons why we estimated a low return rate when designing this feasibility study. However, the World Health Organization recommends self-testing as an additional HIV testing tool [8], not as an HIV surveillance tool, meaning that a low return rate may not represent a major drawback. On the other hand, confirmatory testing and rapid linkage to care are essential after a reactive HIVST result. A study conducted in Brazil reported that of 131 MSM who tested positive, only 95 (72.5%) were linked to care [50]. Because of the anonymity of the e-testing strategy, the proportion of study participants who sought confirmatory testing is uncertain. Furthermore, although users were instructed to bring their PIN to the COA when seeking confirmatory testing, no one accessing services at COA provided a PIN. Therefore, we could not match the HIVST results and confirmatory results. Other programs seeking ways to track users' confirmatory testing and linkage to HIV services following HIVST should be aware of this limitation and the challenges with monitoring HIVST implementation. Other potential caveats of self-testing discussed in the literature are psychological crises and social harm. In our study, neither of these events were reported, which is consistent with the international literature [12,23].

E-testing was largely successful during its first year: E-testing achieved most of the previously defined study outcomes, and demand for HIVST was twice that initially anticipated. However, our study had limitations. The anonymous, self-reported nature of data collection and the low proportion of users who reported their HIVST results prevented us from validating individual HIVST results with confirmatory tests and from estimating the HIV prevalence (due to the inability to exclude multiple answers from the same individuals). Additionally, because we did not track the second test distributed in the package, it is possible that the test requestor gave the second test to a friend or partner, which would increase the number of individuals tested during the study period. We likely also missed MSM who may be fearful of HIV testing and those who did not have access to the internet, which limits the generalizability of our findings. Despite these limitations, we achieved our primary goal, which was to reach individuals who do not access traditional HIV testing services.

In conclusion, our findings show that an internet-based strategy to provide free, anonymous HIVST kits was acceptable and feasible in Brazil. Lessons learned from this study may also be applied to other countries in Latin America with similar cultural and epidemiological contexts. The discretion and convenience of HIVST may increase testing among MSM who are not reached through existing services. Future investigation is needed, however, to identify strategies to support linkage to prevention and treatment services and to determine the potential public health impact and cost-effectiveness of HIVST in Brazil.

Acknowledgments

This project was supported (in part) by the President's Emergency Plan for AIDS Relief (PEPFAR) through the CDC under the terms of the Key Population Implementation Science (KPIS) initiative (Grant number: NU2GGH01152). The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the funding agencies.

Conflicts of Interest

None declared.

References

1. UNAIDS. Geneva: Joint United Nations Programme on HIV/AIDS (UNAIDS); 2017. Ending AIDS: Progress towards the 90-90-90 targets URL: http://www.unaids.org/sites/default/files/media_asset/Global_AIDS_update_2017_en.pdf [accessed 2018-04-18]
2. UNAIDS. Geneva: Joint United Nations Programme on HIV/AIDS (UNAIDS); 2014. 90-90-90 An ambitious treatment target to help end the AIDS epidemic URL: https://www.unaids.org/sites/default/files/media_asset/90-90-90_en.pdf [accessed 2018-04-24]
3. Mocroft A, Lundgren JD, Sabin ML, Monforte AD, Brockmeyer N, Casabona J, Collaboration of Observational HIV Epidemiological Research Europe (COHERE) study in EuroCoord. Risk factors and outcomes for late presentation for HIV-positive persons in Europe: results from the Collaboration of Observational HIV Epidemiological Research Europe Study (COHERE). PLoS Med 2013;10(9):e1001510 [FREE Full text] [doi: 10.1371/journal.pmed.1001510] [Medline: 24137103]
4. Fisher M, Pao D, Brown A, Sudarshi D, Gill O, Cane P, et al. Determinants of HIV-1 transmission in men who have sex with men: a combined clinical, epidemiological and phylogenetic approach. AIDS 2010 Jul 17;24(11):1739-1747. [doi: 10.1097/QAD.0b013e32833ac9e6] [Medline: 20588173]

XSL•FO
RenderX

5. Rutstein S, Ananworanich J, Fidler S, Johnson C, Sanders E, Sued O, et al. Clinical and public health implications of acute and early HIV detection and treatment: a scoping review. J Int AIDS Soc 2017 Dec 28;20(1):21579 [FREE Full text] [doi: 10.7448/IAS.20.1.21579] [Medline: 28691435]
6. Golub S, Gamarel K. The impact of anticipated HIV stigma on delays in HIV testing behaviors: findings from a community-based sample of men who have sex with men and transgender women in New York City. AIDS Patient Care STDS 2013 Nov;27(11):621-627 [FREE Full text] [doi: 10.1089/apc.2013.0245] [Medline: 24138486]
7. Meyerson B, Barnes P, Emetu R, Bailey M, Ohmit A, Gillespie A. Institutional and structural barriers to HIV testing: elements for a theoretical framework. AIDS Patient Care STDS 2014 Jan;28(1):22-27. [doi: 10.1089/apc.2013.0238] [Medline: 24313812]
8. World Health Organization. 2016. Guidelines on HIV self-testing and partner notification URL: https://apps.who.int/iris/bitstream/handle/10665/251655/9789241549868-eng.pdf;jsessionid=375E7E7B8BC009D0836193B4E5961AAF?sequence=1 [accessed 2018-04-24]
9. Lippman S, Lane T, Rabede O, Gilmore H, Chen Y, Mlotshwa N, et al. High Acceptability and Increased HIV-Testing Frequency After Introduction of HIV Self-Testing and Network Distribution Among South African MSM. J Acquir Immune Defic Syndr 2018 Dec 01;77(3):279-287 [FREE Full text] [doi: 10.1097/QAI.0000000000001601] [Medline: 29210826]
10. Carey J. HIV self-testing: feasibilityacceptability of a large scale national service delivery by a community organization. 2017 Presented at: 9th International AIDS Society; 2017; Paris URL: https://www.iasociety.org/Web/WebContent/File/IAS2017_conference_report.pdf
11. Rosengren A, Huang E, Daniels J, Young S, Marlin R, Klausner J. Feasibility of using Grindr to distribute HIV self-test kits to men who have sex with men in Los Angeles, California. Sex Health 2016 May 23:na (forthcoming). [doi: 10.1071/SH15236] [Medline: 27209076]
12. Johnson C, Kennedy C, Fonner V, Siegfried N, Figueroa C, Dalal S, et al. Examining the effects of HIV self-testing compared to standard HIV testing services: a systematic review and meta-analysis. J Int AIDS Soc 2017 Dec 15;20(1):21594 [FREE Full text] [doi: 10.7448/IAS.20.1.21594] [Medline: 28530049]
13. Estem K, Catania J, Klausner J. HIV Self-Testing: a Review of Current Implementation and Fidelity. Curr HIV/AIDS Rep 2016 Apr;13(2):107-115. [doi: 10.1007/s11904-016-0307-y] [Medline: 26879653]
14. De Boni R, Veloso V, Grinsztejn B. Epidemiology of HIV in Latin America and the Caribbean. Curr Opin HIV AIDS 2014 Mar;9(2):192-198. [doi: 10.1097/COH.0000000000000031] [Medline: 24356327]
15. Ackers M, Greenberg A, Lin C, Bartholow B, Goodman A, Longhi M, et al. High and persistent HIV seroincidence in men who have sex with men across 47 U.S. cities. PLoS One 2012;7(4):e34972 [FREE Full text] [doi: 10.1371/journal.pone.0034972] [Medline: 22529964]
16. Beyrer C, Baral S, van Griensven F, Goodreau S, Chariyalertsak S, Wirtz A, et al. Global epidemiology of HIV infection in men who have sex with men. Lancet 2012 Jul 28;380(9839):367-377 [FREE Full text] [doi: 10.1016/S0140-6736(12)60821-6] [Medline: 22819660]
17. Kerr L, Mota R, Kendall C, Pinho AA, Mello M, Guimarães MDC, HIVMSM Surveillance Group. HIV among MSM in a large middle-income country. AIDS 2013 Jan 28;27(3):427-435. [doi: 10.1097/QAD.0b013e32835ad504] [Medline: 23291540]
18. Kerr L, Kendall C, Guimarães MDC, Salani Mota R, Veras M, Dourado M, et al. HIV prevalence among men who have sex with men in Brazil: results of the 2nd national survey using respondent-driven sampling. Medicine (Baltimore) 2018 May;97(1S Suppl 1):S9-S15 [FREE Full text] [doi: 10.1097/MD.0000000000010573] [Medline: 29794604]
19. Guimarães MDC, Kendall C, Magno L, Rocha G, Knauth D, Leal A, Brazilian HIV/MSM Surveillance Group. Comparing HIV risk-related behaviors between 2 RDS national samples of MSM in Brazil, 2009 and 2016. Medicine (Baltimore) 2018 May;97(1S Suppl 1):S62-S68 [FREE Full text] [doi: 10.1097/MD.0000000000009079] [Medline: 29912816]
20. LeGrand S, Muessig K, Horvath K, Rosengren A, Hightow-Weidman L. Using technology to support HIV self-testing among MSM. Curr Opin HIV AIDS 2017 Sep;12(5):425-431 [FREE Full text] [doi: 10.1097/COH.0000000000000400] [Medline: 28617712]
21. MacGowan R. The impact of HIV self-testing among internetrecruited MSM, eSTAMP. : IAS; 2017 Presented at: 9th International AIDS Society; 2017; Paris.
22. James C. HIV self-testing: feasibilityacceptability of a large scale national service delivered by a community organization. : IAS; 2017 Presented at: 9th International AIDS Society; 2017; Paris.
23. Figueroa C, Johnson C, Verster A, Baggaley R. Attitudes and Acceptability on HIV Self-testing Among Key Populations: A Literature Review. AIDS Behav 2015 Nov 9;19(11):1949-1965 [FREE Full text] [doi: 10.1007/s10461-015-1097-8] [Medline: 26054390]
24. Oldenburg C, Biello K, Perez-Brumer A, Rosenberger J, Novak D, Mayer K, et al. HIV testing practices and the potential role of HIV self-testing among men who have sex with men in Mexico. Int J STD AIDS 2017 Dec;28(3):242-249 [FREE Full text] [doi: 10.1177/0956462416641556] [Medline: 27020081]
25. Lippman S, Périssé ARS, Veloso V, Sullivan P, Buchbinder S, Sineath R, et al. Acceptability of self-conducted home-based HIV testing among men who have sex with men in Brazil: data from an on-line survey. Cad Saude Publica 2014 Apr;30(4):724-734 [FREE Full text] [Medline: 24896048]

XSL•FO
RenderX

26. De Boni RB, Lentini N, Santelli A, Barbosa JA, Cruz M, Bingham T, et al. Self-testing, communication and information technology to promote HIV diagnosis among young gay and other men who have sex with men (MSM) in Brazil. J Int AIDS Soc 2018 Jul;21 Suppl 5:e25116 [FREE Full text] [doi: 10.1002/jia2.25116] [Medline: 30033612]
27. Smith D, Pals S, Herbst J, Shinde S, Carey J. Development of a clinical screening index predictive of incident HIV infection among men who have sex with men in the United States. J Acquir Immune Defic Syndr 2012 Aug 01;60(4):421-427. [doi: 10.1097/QAI.0b013e318256b2f6] [Medline: 22487585]
28. Centers for Disease Control and Prevention. 2014. Preexposure Prophylaxis for the Prevention of HIV Infection in the United States? 2017 Update URL: https://www.cdc.gov/hiv/pdf/risk/prep/cdc-hiv-prep-provider-supplement-2017.pdf [accessed 2018-04-24]
29. Hoagland B, De Boni RB, Moreira R, Madruga J, Kallas E, Goulart S, PrEP Brasil Study Team. Awareness and Willingness to Use Pre-exposure Prophylaxis (PrEP) Among Men Who Have Sex with Men and Transgender Women in Brazil. AIDS Behav 2017 May;21(5):1278-1287. [doi: 10.1007/s10461-016-1516-5] [Medline: 27531461]
30. Manual técnico para o diagnóstico da infecção pelo HIV [in Spanish]. Brasilia: Ministério da Saúde; 2013:74.
31. Carpe Carpe B, Hernando Arizaleta L, Monteagudo Piqueras O, Abellán Perpiñán JM, Palomar Rodríguez JA, Alonso J. Response bias in a population postal survey on Health-Related Quality of Life (sf-12v2) [in Spanish]. Rev Esp Salud Publica 2014;88(3):327-337 [FREE Full text] [doi: 10.4321/S1135-57272014000300003] [Medline: 25028301]
32. Krska J, du Plessis R, Chellaswamy H. Views and experiences of the NHS Health Check provided by general medical practices: cross-sectional survey in high-risk patients. J Public Health (Oxf) 2015 Jun;37(2):210-217. [doi: 10.1093/pubmed/fdu054] [Medline: 25118218]
33. Nota S, Strooker J, Ring D. Differences in response rates between mail, e-mail, and telephone follow-up in hand surgery research. Hand (N Y) 2014 Dec;9(4):504-510 [FREE Full text] [doi: 10.1007/s11552-014-9618-x] [Medline: 25414613]
34. Yarris L, Fu R, Duby R, Frakes B, Brooks H, Norton R. Comparison of mailed vs. on-site emergency department patient satisfaction surveys. J Emerg Med 2014 Dec;47(6):702-709. [doi: 10.1016/j.jemermed.2014.06.057] [Medline: 25281169]
35. Anand T, Nitpolprasert C, Ananworanich J, Pakam C, Nonenoy S, Jantarapakde J, et al. Innovative strategies using communications technologies to engage gay men and other men who have sex with men into early HIV testing and treatment in Thailand. J Virus Erad 2015 Apr 01;1(2):111-115 [FREE Full text] [Medline: 27482400]
36. Zou H, Wu Z, Yu J, Li M, Ablimit M, Li F, et al. Internet-facilitated, voluntary counseling and testing (VCT) clinic-based HIV testing among men who have sex with men in China. PLoS One 2013;8(2):e51919 [FREE Full text] [doi: 10.1371/journal.pone.0051919] [Medline: 23418417]
37. Liau A, Millett G, Marks G. Meta-analytic examination of online sex-seeking and sexual risk behavior among men who have sex with men. Sex Transm Dis 2006 Sep;33(9):576-584. [doi: 10.1097/01.olq.0000204710.35332.c5] [Medline: 16540884]
38. Torres T, De Boni RB, de Vasconcellos MT, Luz P, Hoagland B, Moreira R, et al. Awareness of Prevention Strategies and Willingness to Use Preexposure Prophylaxis in Brazilian Men Who Have Sex With Men Using Apps for Sexual Encounters: Online Cross-Sectional Study. JMIR Public Health Surveill 2018 Jan 22;4(1):e11 [FREE Full text] [doi: 10.2196/publichealth.8997] [Medline: 29358160]
39. Bottorff J, Oliffe J, Sarbit G, Sharp P, Caperchione C, Currie L, et al. Evaluation of QuitNow Men: An Online, Men-Centered Smoking Cessation Intervention. J Med Internet Res 2016 Apr 20;18(4):e83 [FREE Full text] [doi: 10.2196/jmir.5076] [Medline: 27097991]
40. Devi R, Singh S, Powell J, Fulton E, Igbinedion E, Rees K. Internet-based interventions for the secondary prevention of coronary heart disease. Cochrane Database Syst Rev 2015 Dec 22(12):CD009386. [doi: 10.1002/14651858.CD009386.pub2] [Medline: 26691216]
41. Mustanski B, Lyons T, Garcia S. Internet use and sexual health of young men who have sex with men: a mixed-methods study. Arch Sex Behav 2011 Apr;40(2):289-300 [FREE Full text] [doi: 10.1007/s10508-009-9596-1] [Medline: 20182787]
42. Bull S, McFarlane M, Rietmeijer C. HIV and sexually transmitted infection risk behaviors among men seeking sex with men on-line. Am J Public Health 2001 Jun;91(6):988-989. [doi: 10.2105/ajph.91.6.988] [Medline: 11392947]
43. Garofalo R, Herrick A, Mustanski B, Donenberg G. Tip of the Iceberg: young men who have sex with men, the Internet, and HIV risk. Am J Public Health 2007 Jun;97(6):1113-1117. [doi: 10.2105/AJPH.2005.075630] [Medline: 17463378]
44. Volk J, Lippman S, Grinsztejn B, Lama J, Fernandes N, Gonzales P, et al. Acceptability and feasibility of HIV self-testing among men who have sex with men in Peru and Brazil. Int J STD AIDS 2016 Dec;27(7):531-536 [FREE Full text] [doi: 10.1177/0956462415586676] [Medline: 25971262]
45. Gonçalves VF, Kerr L, Mota, Macena RHM, Almeida RLD, Freire DG, et al. Incentives and barriers to HIV testing in men who have sex with men in a metropolitan area in Brazil. Cad Saude Publica 2016 May 31;32(5)::e00049015 [FREE Full text] [doi: 10.1590/0102-311X00049015] [Medline: 27253459]
46. Sharma M, Ying R, Tarr G, Barnabas R. Systematic review and meta-analysis of community and facility-based HIV testing to address linkage to care gaps in sub-Saharan Africa. Nature 2015 Dec 03;528(7580):S77-S85 [FREE Full text] [doi: 10.1038/nature16044] [Medline: 26633769]
47. Grinsztejn B, Hoagland B, Moreira R, Kallas E, Madruga J, Goulart S, PrEP Brasil Study Team. Retention, engagement, and adherence to pre-exposure prophylaxis for men who have sex with men and transgender women in PrEP Brasil: 48

XSL•FO
RenderX

week results of a demonstration study. Lancet HIV 2018 Dec;5(3):e136-e145. [doi: 10.1016/S2352-3018(18)30008-0] [Medline: 29467098]
48. Szwarcwald C, Damacena G, Miranda R, Pascom A, Júnior AB. HIV Testing among men in Curitiba, Brazil. AIDS Care 2018 Jan;30(1):56-58 [FREE Full text] [doi: 10.1080/09540121.2017.1363370] [Medline: 28934867]
49. Zhong F, Tang W, Cheng W, Lin P, Wu Q, Cai Y, et al. Acceptability and feasibility of a social entrepreneurship testing model to promote HIV self-testing and linkage to care among men who have sex with men. HIV Med 2017 Dec;18(5):376-382 [FREE Full text] [doi: 10.1111/hiv.12437] [Medline: 27601301]
50. Castro R, Ribeiro-Alves M, Corrêa RG, Derrico M, Lemos K, Grangeiro J, et al. The Men Who Have Sex with Men HIV Care Cascade in Rio de Janeiro, Brazil. PLoS One 2016;11(6):e0157309 [FREE Full text] [doi: 10.1371/journal.pone.0157309] [Medline: 27299524]

Abbreviations

AHA: "A Hora é Agora" (The Time is Now) program
CDC: Centers for Disease Control and Prevention
COA: HIV Counseling and Testing Center
FIOCRUZ: Oswaldo Cruz Foundation
HIVST: HIV self-testing
IQR: interquartile range
MSM: men who have sex with men
NGO: nongovernmental organization
PEPFAR: President's Emergency Plan for AIDS Relief
PIN: personal identification number
STI: sexually transmitted infection
UNAIDS: Joint United Nations Programme on HIV/AIDS
WHO: World Health Organization

Edited by G Eysenbach; submitted 27.03.19; peer-reviewed by K Wilson, J Opoku; comments to author 14.06.19; revised version received 25.06.19; accepted 27.06.19; published 01.08.19.

Please cite as:
De Boni RB, Veloso VG, Fernandes NM, Lessa F, Corrêa RG, Lima RDS, Cruz M, Oliveira J, Nogueira SM, de Jesus B, Reis T, Lentini N, Miranda RL, Bingham T, Johnson CC, Barbosa Junior A, Grinsztejn B
An Internet-Based HIV Self-Testing Program to Increase HIV Testing Uptake Among Men Who Have Sex With Men in Brazil: Descriptive Cross-Sectional Analysis
J Med Internet Res 2019;21(8):e14145
URL: https://www.jmir.org/2019/8/e14145/
doi:10.2196/14145
PMID:31373276

©Raquel Brandini De Boni, Valdilea Gonçalves Veloso, Nilo Martinez Fernandes, Flavia Lessa, Renato Girade Corrêa, Renato De Souza Lima, Marly Cruz, Juliane Oliveira, Simone Muniz Nogueira, Beto de Jesus, Toni Reis, Nena Lentini, Raquel Lima Miranda, Trista Bingham, Cheryl C Johnson, Aristides Barbosa Junior, Beatriz Grinsztejn. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 01.08.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Efficacy of a Self-Regulation–Based Electronic and Mobile Health Intervention Targeting an Active Lifestyle in Adults Having Type 2 Diabetes and in Adults Aged 50 Years or Older: Two Randomized Controlled Trials

Louise Poppe[1,2], PhD; Ilse De Bourdeaudhuij[1], PhD; Maïté Verloigne[1], PhD; Samyah Shadid[3], MD, PhD; Jelle Van Cauwenberg[4], PhD; Sofie Compernolle[1], PhD; Geert Crombez[2], PhD

[1]Department of Movement and Sports Sciences, Ghent University, Ghent, Belgium

[2]Department of Experimental Clinical and Health Psychology, Ghent University, Ghent, Belgium

[3]Department of Endocrinology, Ghent University Hospital, Ghent, Belgium

[4]Department of Public Health and Primary Care, Ghent University, Ghent, Belgium

Corresponding Author:
Louise Poppe, PhD
Department of Movement and Sports Sciences
Ghent University
Watersportlaan 2
Ghent,
Belgium
Phone: 32 9 264 63 63
Email: louise.poppe@ugent.be

Abstract

Background: Adopting an active lifestyle plays a key role in the prevention and management of chronic diseases such as type 2 diabetes mellitus (T2DM). Web-based interventions are able to alter health behaviors and show stronger effects when they are informed by a behavior change theory. *MyPlan 2.0* is a fully automated electronic health (eHealth) and mobile health (mHealth) intervention targeting physical activity (PA) and sedentary behavior (SB) based on the Health Action Process Approach (HAPA).

Objective: This study aimed to test the short-term effect of *MyPlan 2.0* in altering levels of PA and SB and in changing personal determinants of behavior in adults with T2DM and in adults aged ≥50 years.

Methods: The study comprised two randomized controlled trials (RCTs) with an identical design. RCT 1 was conducted with adults with T2DM. RCT 2 was performed in adults aged ≥50 years. Data were collected via face-to-face assessments. The participants decided either to increase their level of PA or to decrease their level of SB. The participants were randomly allocated with a 2:1 ratio to the intervention group or the waiting-list control group. They were not blinded for their group allocation. The participants in the intervention group were instructed to go through *MyPlan 2.0*, comprising 5 sessions with an interval of 1 week between each session. The primary outcomes were objectively measured and self-reported PA (ie, light PA, moderate-to-vigorous PA, total PA, number of steps, and domain-specific [eg, transport-related] PA) and SB (ie, sitting time, number of breaks from sitting time, and length of sitting bouts). Secondary outcomes were self-reported behavioral determinants for PA and SB (eg, self-efficacy). Separate linear mixed models were performed to analyze the effects of *MyPlan 2.0* in the two samples.

Results: In RCT 1 (n=54), the PA intervention group showed, in contrast to the control group, a decrease in self-reported time spent sitting (P=.09) and an increase in accelerometer-measured moderate (P=.05) and moderate-to-vigorous PA (P=.049). The SB intervention group displayed an increase in accelerometer-assessed breaks from sedentary time in comparison with the control group (P=.005). A total of 14 participants of RCT 1 dropped out. In RCT 2 (n=63), the PA intervention group showed an increase for self-reported total PA in comparison with the control group (P=.003). Furthermore, in contrast to the control group, the SB intervention group decreased their self-reported time spent sitting (P=.08) and increased their accelerometer-assessed moderate (P=.06) and moderate-to-vigorous PA (P=.07). A total of 8 participants of RCT 2 dropped out.

Conclusions: For both the samples, the HAPA-based eHealth and mHealth intervention, *MyPlan 2.0*, was able to improve only some of the primary outcomes.

XSL•FO
RenderX

Trial Registration: ClinicalTrials.gov NCT03291171; http://clinicaltrials.gov/ct2/show/NCT03291171. ClinicalTrials.gov NCT03799146; http://clinicaltrials.gov/ct2/show/NCT03799146.

International Registered Report Identifier (IRRID): RR2-10.2196/12413

(J Med Internet Res 2019;21(8):e13363) doi:10.2196/13363

KEYWORDS

eHealth; mHealth; physical activity; type 2 diabetes; self-regulation

Introduction

The prevalence of chronic diseases, such as type 2 diabetes mellitus (T2DM), cardiovascular disease, and cancer, is high and rising [1,2]. Adopting an active lifestyle (ie, increasing physical activity [PA] and reducing sedentary behavior [SB]) plays an important role in the prevention and management of these diseases [3,4]. Indeed, adults are recommended to accumulate 150 min of moderate-to-vigorous PA (MVPA) [5] and minimize periods of prolonged sedentary time [6]. However, the majority of adults do not meet the guidelines considering PA and accumulate high levels of sitting time [4]. Even people for whom adopting an active lifestyle is considered a cornerstone in the management of their disease, such as people with T2DM, show high levels of physical inactivity and sedentary time [7,8]. Consequently, interventions targeting increases in PA and decreases in SB in adults with T2DM as well as in adults from the general population are needed.

As the number of internet and mobile phone users increases, interest in electronic health (eHealth) and mobile health (mHealth) interventions is growing [9]. eHealth and mHealth interventions offer several advantages as they can deliver fast and tailored information to large groups of individuals in a cost-effective way. Research examining the common effect of eHealth or mHealth interventions targeting PA or SB reports trivial-to-small and short-term effects [10-12]. Important points for improvement are better reporting of the theoretical basis as well as active ingredients (ie, the implemented behavior change techniques) of the intervention [10,12,13] and adopting objective measures to assess PA and SB [12].

Theory-based interventions delivered via the internet show stronger effects than internet-based interventions making less extensive or no use of theory (median d_+=0.19) [14]. Self-regulation frameworks highlight the importance of bridging the intention-behavior gap by considering pre- as well as postintentional determinants of behavior change [15]. A review of Rhodes et al (2015) provides an overview of models incorporating pre- as well as postintentional determinants of PA [16]. The identified models showed considerable overlap in the proposed processes to bridge the intention-behavior gap. The results further showed that the health action process approach (HAPA) [17] was the most often used and independently tested framework. Indeed, the HAPA has been applied to alter the levels of PA and SB in clinical (including adults with T2DM [18]) and in nonclinical populations [19,20]. In recent years, this theoretical framework has also been used for developing Web-based behavioral interventions [21-24]. For example, *SmartMobiel*, an eHealth intervention informed by the HAPA-model was found to be effective in increasing PA in adults [25]. According to the HAPA, *risk perception*, *outcome expectancies*, *self-efficacy*, *intention*, *action planning*, *coping planning*, and *monitoring* are personal determinants playing a key role in behavior change.

MyPlan 2.0 is a stand-alone HAPA-based eHealth and mHealth intervention comprising (1) a website offering weekly sessions to create and evaluate personal goals and (2) an optional mobile app providing daily support [26]. The program offers a module targeting increases in PA and a module targeting reductions in SB. The users autonomously select which behavior they will focus on. This was done because of two reasons. First, PA and SB are considered distinct rather than opposite behaviors, each having a unique contribution to people's mental and physical health [27]. Indeed, one might reach the health norms regarding PA and still show high levels of SB and vice versa. Second, the self-regulation framework emphasizes the importance of goal ownership and highlights the need to let people select goals that they can relate with [15]. The program aims to alter behavior by targeting the HAPA-based personal determinants of behavior. As obtaining large changes in behavior might take longer than the length of the program, it is important to also assess whether the intervention altered the targeted personal determinants for behavior. These personal determinants could be altered on a shorter term and, according to the HAPA [28], changes in the personal determinants will result in changes in behavior.

The aim of this study was to test the efficacy of *MyPlan 2.0* to alter behavior (primary outcome) and behavioral determinants (secondary outcome) in adults with T2DM. The research protocol for the randomized controlled trial (RCT) in patients with T2DM was published [26]. However, we encountered difficulties in recruiting participants with T2DM. For that reason, it was decided to recruit an additional group of participants from the general population from a similar age cohort as the population with T2DM. Consequently, the participants of RCT 1 were adults diagnosed with T2DM and the participants of RCT 2 were adults aged 50 years or older.

Methods

Hypotheses

Similar hypotheses were formulated for both RCTs. Regarding PA, we hypothesized that *MyPlan 2.0* would have a positive effect on self-reported and objectively measured levels of total PA, MVPA, and light PA (LPA) in the PA intervention group compared with the control group. Regarding sedentary behavior, we hypothesized that *MyPlan 2.0* would reduce self-reported and objectively measured total sitting time in the SB intervention group compared with the control group. Furthermore, as the

XSL•FO
RenderX

intervention focused on limiting sedentary time as well as interrupting periods of prolonged sitting, we expected to find an increase in breaks from sedentary time and a decrease in the length of the sedentary bouts in the intervention group targeting SB compared with the control group. Regarding the personal determinants, we expected that *MyPlan 2.0* would increase the participants' self-efficacy, outcome expectations, intention, action planning, coping planning, and self-monitoring. No hypotheses regarding the participants' risk perception were made as *MyPlan 2.0* did not specifically target this personal determinant [26].

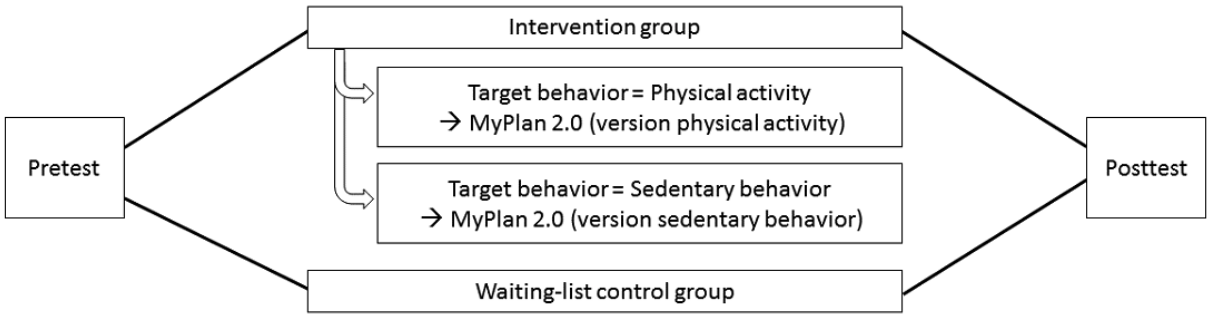
Study Design and Procedure

In total, 2 RCTs with a parallel group design were conducted to investigate the effect of *MyPlan 2.0* on PA, SB, and HAPA-based determinants. The protocol was preregistered [26]. The a priori power analysis suggested a sample size of 96 participants. Adults with T2DM were recruited via the Ghent University Hospital and the Damian General Hospital (Ostend). However, recruitment via the hospitals was slow. Therefore, in contrast with the recruitment process described in the protocol, we also advertised the study via the Flemish Diabetes Association and in adults with T2DM who participated in the previous research of the involved research groups. The sample of adults aged ≥50 years was recruited via advertisements in local newspapers and via snowball sampling. For both samples, the inclusion criteria were (1) being literate in the Dutch language to engage in the intervention, (2) being computer literate, (3) having internet access, and (4) not having participated in the qualitative study about *MyPlan 2.0*. Additional inclusion criteria to participate in RCT 1 were being diagnosed with T2DM since at least 1 month and being 18 years or older, whereas the participants of RCT 2 were required to be aged 50 years or older.

Figure 1 displays the design of the RCTs. After enrollment, the participants were visited by one of the researchers. During the home visit, the researcher explained the difference between PA and SB and asked the participants to select a target behavior (ie, increasing PA or decreasing SB). The participants completed questionnaires and their weight and waist circumference were assessed. The participants were instructed to wear an accelerometer for 10 consecutive days starting the day after the home visit. After these 10 days, the participants were allocated by LP to the intervention or the waiting-list control group using a 2:1 ratio. This was done via a random number generator. The participants allocated to the waiting-list control group were informed about their allocation and instructed to continue with their life as usual. The participants allocated to the intervention group received access to the MyPlan 2.0 website and the mobile app. The participants who selected to focus on their level of PA were guided to the version targeting PA (PA intervention group), whereas the participants who selected to alter their level of SB were guided to the version targeting SB (SB intervention group). They were instructed to go through each of the weekly sessions (5 in total) offered by the website. The involved researchers inspected the logfile of the website to check whether the participants logged in for each session. The participants who forgot to log in were contacted by a researcher via email and informed about the next session. If the participant did not respond, he or she was contacted via telephone. As having a smartphone was not an inclusion criterion, it was not obligatory to use the mobile app. To monitor any adverse effects (eg, hypoglycemia), all participants were weekly phoned by a member of the research team. No coaching took place during these phone calls.

After completing all sessions (PA and SB intervention groups) or the 5-week waiting period (control group), a second home visit was arranged. During this second home visit, the participants completed the same assessments as at baseline. The participants who decided to leave the study were contacted by one of the researchers and asked if they were willing to complete a questionnaire assessing potential reasons for attrition. Except during the pretest (the participants were allocated to a group after the pretest), neither the participants nor researchers assessing the outcome variables were blinded.

Figure 1. Design of the randomized controlled trials.



XSL•FO
RenderX

All data were collected between January and September 2018. No changes regarding bug fixes, downtimes, or content changes to the Web-based program occurred after trial commencement. The RCTs were approved by the Committee of Medical Ethics of the Ghent University Hospital (Belgian registration numbers: B670201732566 for RCT 1 and BE670201731996 for RCT 2).

MyPlan 2.0

MyPlan 2.0 is a free fully automated HAPA-based eHealth and mHealth intervention comprising a website and an optional mobile app. Its precursor, *MyPlan 1.0*, showed high levels of attrition. Several user-based studies were performed to better adapt *MyPlan 2.0* to the users' needs [29-32]. Multimedia Appendix 1 provides an overview of the lessons learned from each of these studies and describes how these findings guided the adaptations to *MyPlan 2.0*. The program offers a number of behavior-change techniques aiming to influence the users' HAPA-based personal determinants for change. The used techniques are mentioned below and labelled according to the taxonomy of behavior change techniques of Michie et al [33]. Multimedia Appendix 2 provides screenshots of the website and the mobile app.

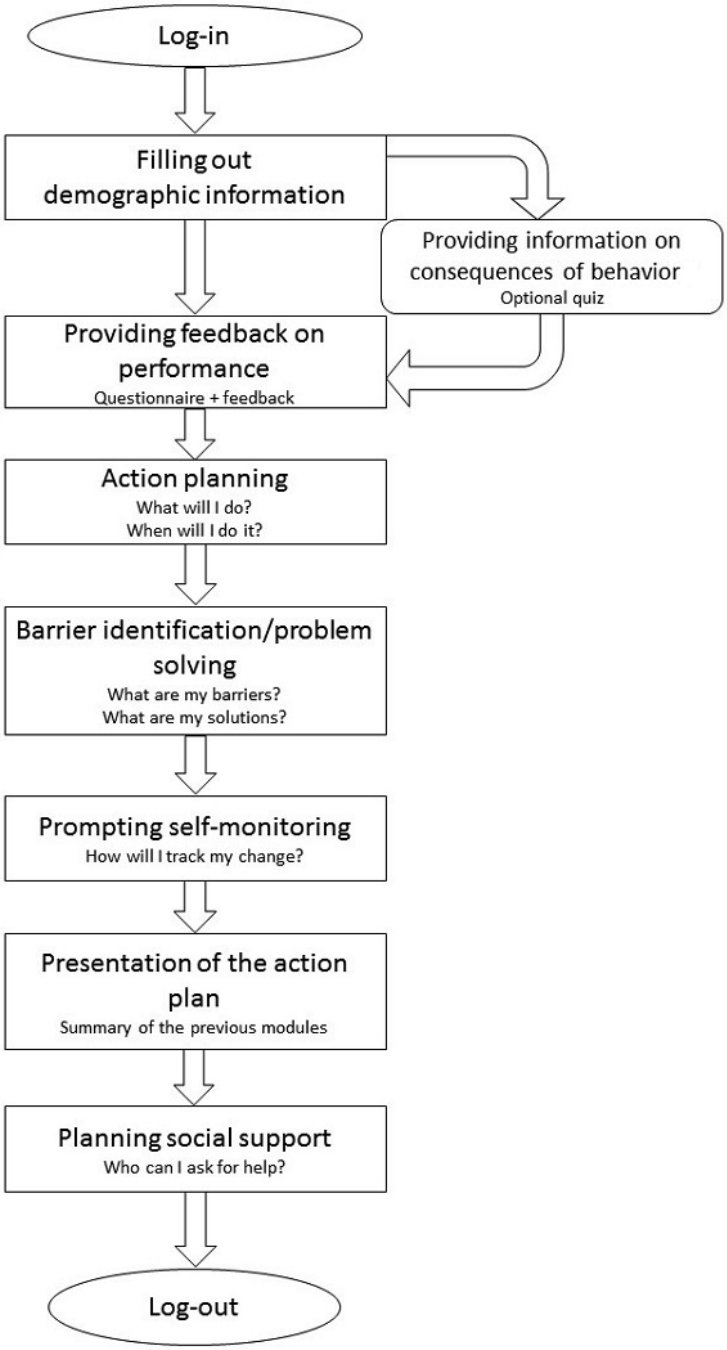
The Website

The website part of *MyPlan 2.0* was created using LifeGuide [34] and offers 5 sessions with a period of 1 week between each session. The two versions of the program (one targeting increases in PA and one targeting reductions in SB) have an identical structure and offer the same self-regulation techniques. During the first session, the users create a profile, complete an optional quiz regarding the benefits of the chosen health behavior (ie, increasing PA or reducing SB, *providing information on consequences of behavior)*, fill out a questionnaire assessing their current level of PA or SB and receive tailored feedback (*providing feedback on performance*), create a personal action plan to alter the chosen health behavior *(action planning)*, foresee potential barriers and search for solutions *(barrier identification/problem solving)*, and select how they will monitor their behavior *(prompting self-monitoring of behavior).* At the end of the first session, the users' answers are summarized in a printable action plan and they are offered optional information about how they can obtain support from their partner, friends, family, or colleagues *(exploring social support)*. Figure 2 shows the flow of the first session.

After 1 week, the users receive an email to start the second session. The follow-up sessions (ie, sessions 2-5) have a similar structure. After logging in, the users are asked to what extent they reached the goal set in the previous session *(prompting review of behavioral goals)* and whether they would like to keep or adapt this goal. When choosing the latter, the user is guided to the action planning section. All users again foresee potential barriers to reach the goal and search for solutions. Finally, their answers are summarized in a printable action plan and the users are optionally offered additional tips and tricks (eg, *try to take the stairs instead of using the elevator*) to become more physically active or less sedentary. Figure 3 depicts the flow of the follow-up sessions.

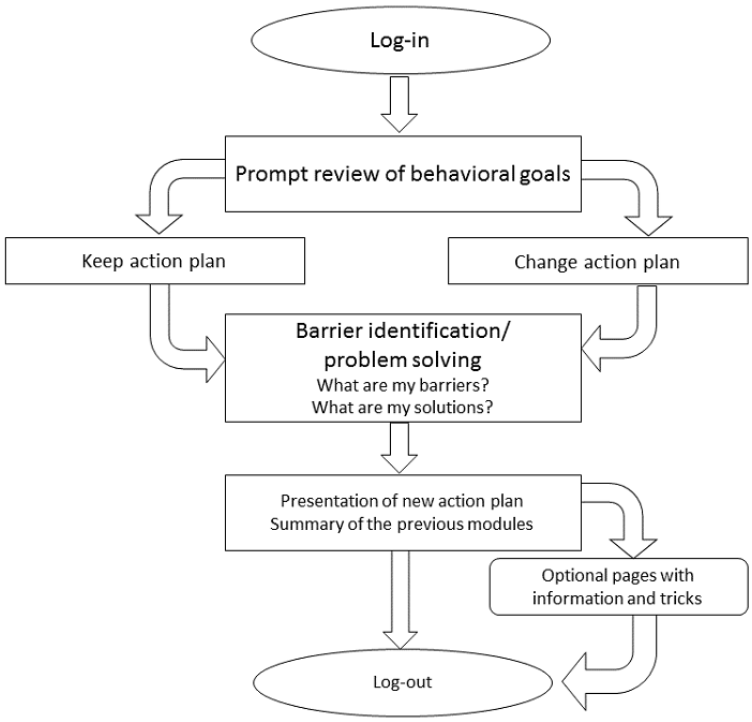
XSL•FO
RenderX

Figure 2. Flow of the first session.



XSL•FO
RenderX

Figure 3. Flow of the follow-up sessions.



The Mobile App

The mobile app comprises 5 modules through which the users can freely navigate. The first module supports users in monitoring their behavior *(prompting self-monitoring of behavior)*. Every evening, the users receive a notification to report the extent to which they were able to be more physically active or to sit less (eg, *not at all*, *not*, *a little*, *well*, and *very well*). These entries are then shown in a graph displaying all responses of the week. The second module allows the users to review their weekly goals (created on the website) and make adaptations to these goals *(action planning)*. The option to review potential problems and their solutions is offered in the third module *(barrier identification/problem solving)*. In the fourth module, the users can perform quizzes on the benefits of being more physically active or less sedentary *(providing information on the consequences of behavior)*. Finally, the users can collect points by visiting the website, completing quizzes, and monitoring their behavior. By collecting these points, they could earn the victory cups implemented in the mobile app. This gamification element was added to increase engagement with the mobile app.

Measurements

Participant Characteristics

An ad hoc questionnaire assessed age, sex, height, civil status, level of education, and time since diagnosis (only for the participants with T2DM). The participants who completed college or university were considered highly educated.

The participants' weight and waist circumference were assessed using a Seca weighting scale (type 813) and a Seca measuring tape. Waist circumference was measured at the lowest rib margin and the iliac crest at the midaxillary line. During each testing wave, the participants' weight and waist circumference were measured twice. In case the difference between the two measurements was >100 grams or >1 cm, the measurement was performed a third time. The mean of the measurements was calculated as the final score.

Primary Outcomes

The long version of the international physical activity questionnaire (IPAQ) [35] (translated into Dutch) assesses self-reported PA of the past week in 4 domains (work, transport, household, and leisure time) and provides indicators for work-related PA, transport-related PA, household-related PA, leisure-related PA, total PA, vigorous-intensity PA (VPA), and MVPA per week. The IPAQ has good reliability (intraclass range 0.46-0.96) and a fair-to-moderate criterion validity (Spearman rho between .30 and .37) [35]. As the IPAQ overestimates PA [36], the data were truncated according to the method described by Dubuy et al [37]. The LASA (Longitudinal Aging Study Amsterdam) sedentary behavior questionnaire [38], which has moderate reliability (intraclass=0.71) and moderate validity (Spearman rho=.35), was used to assess usual total sedentary time on weekdays. Data were truncated at a maximum of 16 hours of sitting time a day [39]. Both questionnaires were conducted via an interview by the visiting researcher.

ActiGraph accelerometers (type GT3X+), shown to be reliable and valid [40-43], were used to assess the participants' number of breaks from sedentary time, average length of the sedentary bouts, total sedentary time, number of steps, LPA, moderate PA (MPA), VPA, MVPA, and total PA. The participants were instructed to wear the accelerometer on the right hip during waking hours but to remove it for water-based activities (eg, showering). ActiLife 6.13.3 software (ActiGraph, Fort Walton Beach, FL, USA) was used to initialize the accelerometers and process the data. The epoch was set at 60 seconds and nonwear

XSL•FO
RenderX

time was calculated as ≥60 min of consecutive 0 counts. The participants' accelerometer data were included in the study when they had at least 4 valid days including 1 weekend day (with valid defined as ≥10 hours of wearing time) [44]. Using the cut points described by Freedson et al [45], each minute of wear time was categorized as sedentary (0-99 counts per min [CPM]), LPA (100-1951 CPM), MPA (1952-5724 CPM), VPA (5725-9498 CPM), or MVPA (≥1952 CPM). Total PA was calculated by combining LPA and MVPA. A bout of sedentary time was considered a period of at least 10 consecutive min <99 counts with zero tolerance allowed. A break from a sedentary bout was defined as a transition from <99 CPM to >99 CPM between 2 sedentary bouts.

Secondary Outcomes

The participants' HAPA-based personal determinants for behavior change (ie, self-efficacy, risk perceptions, outcome expectations, intention, action planning, coping planning, and self-monitoring) were measured using multiple items with a minimum of 3 items per determinant. To select these items, a large number of items measuring HAPA determinants were presented to 11 experts in the self-regulation framework. All experts indicated for each item whether or not it measured the presented HAPA determinant and how certain they were of their answer [46]. On the basis of their responses, a discriminant content validity method was used [46] and the best scoring items were selected. To assure comprehensibility of these items, cognitive interviews were conducted with 4 adults (mean age 58.3, SD 6.5, 3 women, 2 having T2DM, and 2 with a college degree or higher). On the basis of the results of these interviews, the final items were selected and adapted. Each item was assessed using 10 answer options ranging from *completely disagree* to *completely agree*. For each personal determinant, a mean score (potential range 1-10) was calculated.

Statistical Analysis

The data from both RCTs were analyzed separately using R version 3.2.5 [47]. Nevertheless, the analyses were similar for both the RCTs.

Group comparability at baseline between the two intervention groups (PA intervention group and SB intervention group) and the control group was investigated using a 1-way analysis of variance (for the quantitative variables) and chi-square tests (for the qualitative variables). *T* tests and chi square tests were used to perform the dropout analysis. Linear mixed models (2 levels: repeated measures clustered within the participants) were performed using the *lme4-package* [48] to investigate the intention-to-treat effect of *MyPlan 2.0* on levels of PA, SB, and the personal determinants [49]. In contrast to the multivariate analysis of variance, the linear mixed model can easily handle missing data in repeated measures [50]. Furthermore, mixed models without ad hoc imputation provide equal or more power than mixed models with ad hoc imputation [51]. In the protocol, we stated that we would consider the participants' choice of target behavior (ie, PA or SB) as moderator. However, because we were not able to recruit large enough samples, we decided to perform the analyses on the behavioral outcomes with a group variable (ie, the PA intervention group, the SB intervention group, and the control group).

All participants filled out one version of the HAPA-based determinants (ie, the version focusing on PA or the version focusing on SB). As described in the protocol, we planned to account for this issue by considering the choice of target behavior (ie, PA or SB) as moderator. However, considering the small sample sizes, we decided to combine the PA intervention group and the SB intervention group as one intervention group for analyzing the effect on the personal determinants. By doing so, we considered these outcome variables as personal determinants regarding the chosen health behavior rather than personal determinants regarding increasing PA or decreasing SB.

Owing to the low prevalence of accelerometer-based VPA (no VPA at baseline was detected in 93% (50/54) of the sample in RCT 1 and in 63% (40/63) of the sample in RCT 2), self-reported VPA (no self-reported VPA at baseline was detected in 80% (43/54) of the sample in RCT 1 and in 75% (47/63) of the sample in RCT 2), and self-reported work-related PA (no self-reported work-related PA at baseline in 69% (37/54) of the sample in RCT 1 and in 67% (42/63) of the sample in RCT 2) in both samples, these specific outcome variables were not analyzed.

Distribution of the dependent variables was first checked using Shapiro-Wilk tests. Normally distributed dependent variables were analyzed using the *lmer* function of the *lme4-package* [48]. For non-normally distributed variables, we compared models with different variance and link functions (ie, Gaussian with identity, gamma with log, gamma with identity, Poisson with log, and negative binomial with log) using the Bayesian information criterion (BIC). For each dependent variable, we selected the model providing the lowest BIC value. By exploring the interaction between time and group (ie, intervention vs control), the effect of the intervention on the dependent variable was assessed. The beta values for *time×group* reported in the results section describe the difference between the change in the intervention group and the change in the control group. Consequently, these values represent the intervention effect for each dependent variable. *P* values <.05 were considered statistically significant, whereas *P* values between .05 and .10 were considered borderline significant.

Effect sizes were calculated for each of the dependent variables in both samples [52]. As recommended by Morris [53], the pooled pretest standard deviation was used to estimate the effect sizes.

Results

The results of the two RCTs are reported separately. The first section will describe the results of the RCT with the sample with T2DM (RCT 1), whereas the second section will describe the results of the RCT with the sample aged ≥50 years (RCT 2).

Randomized Controlled Trial 1

Figure 4 shows the flow of the participants with T2DM. A total of 58 participants agreed to participate in the study. Of this sample, 18 participants were recruited via the Ghent University Hospital, 8 via the Damian General Hospital, 24 via the Flemish

XSL•FO
RenderX

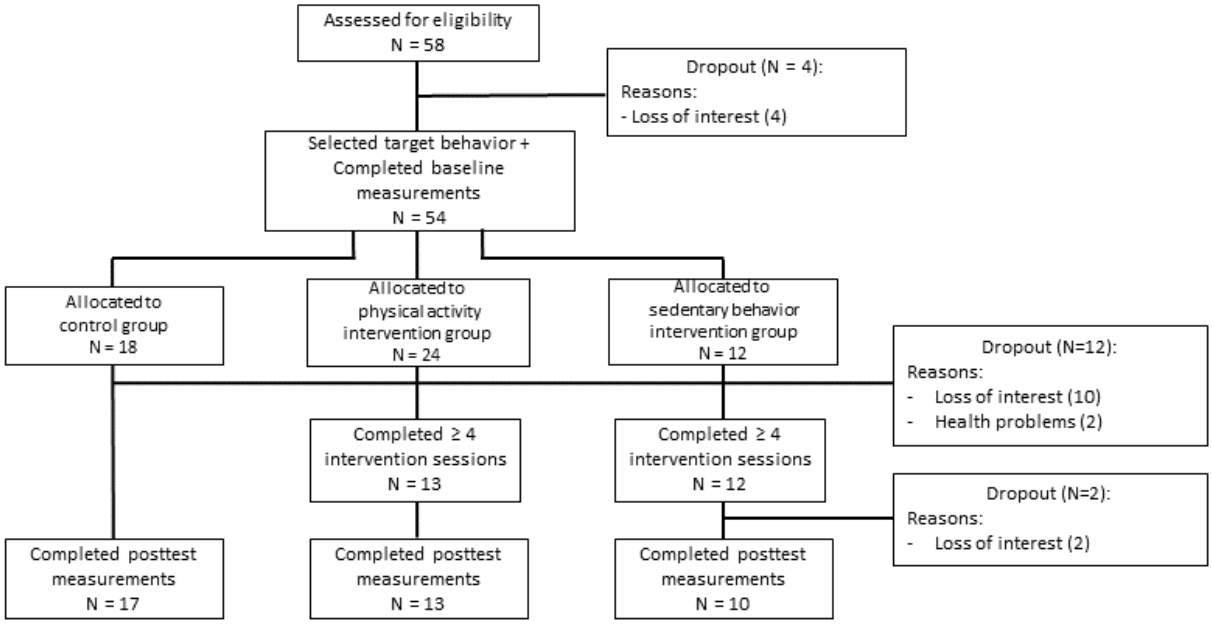
Diabetes Association, and 8 via previous studies. As we do not know how many patients saw the advertisements, the response rate could not be calculated. Out of them, 4 participants dropped out before completing the baseline measurements. Consequently, the data of 54 participants were analyzed. Of the 14 participants who dropped out before completing 4 sessions, only 3 participants (all belonging to the control group) were willing to complete the questionnaire assessing specific reasons for attrition. Among them, 1 participant indicated that he doubted to participate at the beginning of the study and 2 participants indicated that drastic changes in their life occurred while participating. Finally, 1 participant indicated that the high number of research-related questionnaires frustrated her.

The participants' baseline characteristics are provided in Table 1. At baseline, 32 participants decided to focus on PA (24 of these participants were later allocated to the intervention group) and 22 participants chose to focus on SB (12 of these participants were later allocated to the intervention group). Consequently, the PA intervention group comprised 24 participants and the SB intervention group comprised 12 participants. No significant baseline differences in sociodemographic characteristics were found among the PA intervention group, the SB intervention group, and the control group. Of the participants, 7 used the optional mobile app. The dropout analyses indicated that the participants allocated to the intervention group (χ^2_1=4.35, *P*=.04) were more likely to dropout. No significant differences between completers and dropouts were found for age, sex, level of education, body mass index (BMI), time since diagnosis, total PA at baseline (accelerometer-measured), or sedentary time at baseline (accelerometer-measured).

Table 2 displays the means and standard deviations for each of the behavioral outcomes in the three groups. Table 3 provides the time-by-group interactions and effect sizes for each of the behavioral outcomes. A borderline significant intervention effect favoring the PA intervention group was found for self-reported total daily sitting time (*P*=.09) and accelerometer-assessed MPA (*P*=.05) and MVPA (*P*=.049). A significant intervention effect favoring the SB intervention group was found for accelerometer-assessed daily breaks from sedentary time (*P*=.005). No intervention effects were found for self-reported total transport-related PA, self-reported total household-related PA, self-reported total leisure-related PA, self-reported total PA, self-reported MVPA, accelerometer-assessed length of the sedentary bouts, accelerometer-assessed sedentary time, accelerometer-assessed LPA, accelerometer-assessed total PA, or accelerometer-assessed daily steps.

Table 4 displays the time-by-group interactions and effect sizes for the personal determinants in RCT 1. Significant intervention effects favoring the control group were found for self-efficacy (*P*=.01) and risk perception (*P*=.03). A borderline significant intervention effect favoring the intervention group was found for action planning (*P*=.08). Finally, a significant time*group interaction effect favoring the intervention group was found for self-monitoring (*P*=.008). No intervention effects were found for outcome expectancies, coping planning, or intention.

Figure 4. Flow of the sample of randomized controlled trial 1.



XSL•FO
RenderX

Table 1. Baseline characteristics of the sample of randomized controlled trial 1.

Baseline characteristics	Total sample (N=54)	CG[a] (n=18)	IG[b]–PA[c] (n=24)	IG–SB[d] (n=12)	*F* or χ^2 (df)	*P* value
Age (years), mean (SD)	62.67 (8.40)	64.89 (8.62)	62.91 (7.16)	58.92 (9.52)	1.90[e] (2, 49)	.16
Males, n (%)	34 (63)	9 (50)	17 (71)	8 (67)	2.01[f] (2)	.37
University/college education, n (%)	29 (54)	12 (67)	10 (42)	7 (58)	1.85[f] (2)	.40
Body mass index (kg/m²), mean (SD)	30.84 (5.66)	30.51 (6.86)	30.86 (5.35)	31.25 (4.73)	0.06[e] (2, 48)	.94
Waist circumference (cm), mean (SD)	109.23 (14.08)	108.14 (18.38)	109.16 (11.09)	110.99 (13.06)	0.14[e] (2,51)	.87
Time since diagnosis (months), mean (SD)	129.78 (83.09)	157.69 (67.08)	100.18 (87.25)	146.83 (82.94)	2.73[e] (2,47)	.08

[a]CG: control group.

[b]IG: intervention group.

[c]PA: physical activity.

[d]SB: sedentary behavior.

[e]*F* value.

[f]χ^2 value.

Table 2. Means and standard deviations for each of the behavioral outcomes in the three groups in randomized controlled trial 1.

Behavioral outcomes	CG[a], mean (SD)		IG[b]–PA[c], mean (SD)		IG–SB[d], mean (SD)	
	Pre	Post	Pre	Post	Pre	Post
LASA[e] questionnaire						
Total sitting time (min/day)	553.06 (174.05)	567.94 (211.84)	592.83 (232.52)	470.38 (185.23)	599.17 (133.58)	579.44 (188.84)
IPAQ[f]						
Total transport-related PA (min/day)	18.02 (24.61)	31.30 (36.60)	13.90 (25.70)	29.34 (24.61)	35.54 (35.30)	26.50 (23.16)
Total household-related PA (min/day)	45.28 (71.89)	46.89 (62.34)	63.33 (71.60)	62.96 (56.18)	44.17 (78.15)	67.29 (92.89)
Total leisure-related PA (min/day)	16.35 (18.76)	38.45 (59.88)	19.08 (25.59)	30.61 (41.58)	38.63 (46.27)	49.86 (74.39)
Total PA (min/day)	86.07 (70.54)	124.50 (96.02)	113.81 (90.52)	143.83 (91.48)	137.98 (99.89)	150.07 (100.48)
Moderate-to-vigorous physical activity (min/day)	17.86 (24.37)	61.98 (75.83)	53.19 (72.03)	48.67 (41.35)	47.26 (67.18)	63.57 (76.93)
Accelerometer						
Number of breaks per day	16.63 (2.25)	15.65 (2.87)	15.51 (3.08)	14.64 (2.39)	16.88 (1.69)	17.50 (1.45)
Length of sedentary bouts (min/day)	22.90 (2.70)	23.46 (2.01)	22.34 (2.44)	22.82 (2.64)	22.80 (1.57)	23.22 (1.14)
Sedentary time (min/day)	544.52 (56.24)	537.77 (81.01)	528.09 (80.79)	498.09 (43.81)	551.39 (56.66)	555.21 (44.39)
Light physical activity (min/day)	239.34 (73.14)	231.51 (76.84)	238.13 (67.23)	244.46 (71.05)	218.87 (54.49)	215.69 (37.91)
Moderate physical activity (min/day)	23.14 (12.93)	19.36 (14.77)	17.02 (15.70)	25.50 (15.77)	20.15 (13.92)	19.13 (14.06)
Moderate-to-vigorous physical activity (min/day)	23.20 (12.92)	19.36 (14.77)	17.07 (15.68)	25.50 (15.77)	20.23 (13.87)	19.38 (13.81)
Total PA	262.54 (78.58)	267.17 (83.11)	255.19 (69.00)	269.95 (78.24)	239.10 (48.83)	235.07 (33.89)
Daily steps	6203.10 (2284.41)	6292.05 (2480.44)	5364.39 (2219.28)	6549.71 (2313.67)	6083.88 (1343.30)	6001.03 (1107.26)

[a]CG: control group.

[b]IG: intervention group.

[c]PA: physical activity.

[d]SB: sedentary behavior.

[e]LASA: Longitudinal Aging Study Amsterdam.

[f]IPAQ: international physical activity questionnaire.

XSL•FO
RenderX

Table 3. Time-by-group interactions and effect sizes for each of the behavioral outcomes in randomized controlled trial 1.

Behavioral outcomes	Time×group PA[a] (ref: pre×CG[b]), beta (SE)	ES[c] (IG[d]–PA vs CG)	Time×group SB[e] (ref: pre×CG), beta (SE)	ES (IG−SB vs CG)
LASA[f] questionnaire				
Total sitting time (min/day)[g]	−102.50 (59.32)[h]	−0.65	−4.61 (66.59)	−0.22
IPAQ[i]				
Total transport-related PA (min/day)[j]	0.19 (0.92)	0.09	−0.85 (1.05)	−0.76
Total household-related PA (min/day)[j]	−0.04 (0.96)	−0.03	0.39 (1.10)	0.29
Total leisure-related PA (min/day)[j]	−0.38 (0.99)	−0.46	−0.60 (1.13)	−0.33
Total PA (min/day)[j]	−0.14 (0.52)	0.23	−0.29 (0.60)	−0.31
Moderate-to-vigorous physical activity (min/day)[j]	−1.33 (0.85)	−0.85	−0.95 (0.97)	−0.60
Accelerometer				
Number of breaks per day[k]	0.28 (0.63)	0.04	1.93 (0.69)[l]	0.77
Length of sedentary bouts (min/day)[k]	−0.50 (0.59)	−0.03	−0.46 (0.62)	−0.06
Sedentary time (min/day)[k]	−19.71 (23.92)	−0.32	19.91 (25.00)	0.19
Light physical activity (min/day)[g]	7.02 (14.70)	0.20	−5.07 (14.94)	0.07
Moderate physical activity (min/day)[j]	0.37 (0.19)[h]	0.86	0.02 (0.92)	0.21
Moderate-to-vigorous physical activity (min/day)[j]	0.37 (0.19)[m]	0.84	0.06 (0.19)	0.39
Total PA[g]	11.09 (15.47)	0.002	−4.54 (15.73)	−0.13
Daily steps[g]	499.46 (543.12)	0.49	−302.11 (555.83)	−0.09

[a]PA: physical activity.
[b]CG: control group.
[c]ES: effect size.
[d]IG: intervention group.
[e]SB: sedentary group.
[f]LASA: Longitudinal Aging Study Amsterdam.
[g]Gaussian (identity).
[h]$P<.10$.
[i]IPAQ: international physical activity questionnaire.
[j]Gamma (log).
[k]Gamma (identity).
[l]$P<.01$.
[m]$P<.05$.

XSL•FO
RenderX

Table 4. Time-by-group interactions and effect sizes for the personal determinants in randomized controlled trial 1.

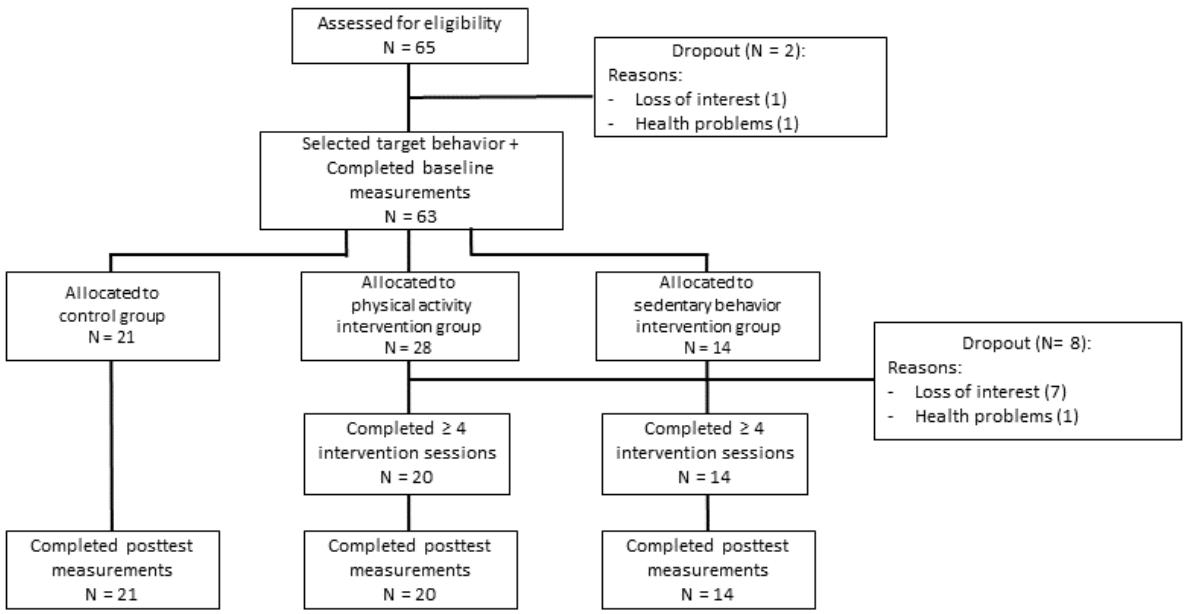
Personal determinants	CG[a], mean (SD)		IG[b], mean (SD)		Time×group (ref: pre×CG), beta (SE)	ES[c]
	Pre	Post	Pre	Post		
Self-efficacy[d]	5.68 (1.98)	6.88 (1.22)	6.35 (1.81)	7.23 (1.61)	−1.24 (0.48)[e]	−0.17
Outcome expectancies[d]	7.03 (1.51)	8.01 (1.10)	7.62 (1.63)	8.04 (0.96)	−0.09 (0.42)	−0.35
Risk perception[f]	4.46 (1.81)	5.10 (1.71)	5.16 (2.19)	4.74 (92.05)	−1.17 (0.53)[e]	−0.51
Action planning[f]	5.41 (1.99)	5.29 (2.21)	5.19 (2.32)	6.09 (2.05)	1.18 (0.65)[g]	0.46
Coping planning[d]	3.52 (2.43)	4.69 (2.22)	3.95 (2.66)	5.77 (2.32)	0.79 (0.54)	0.25
Intention[f]	6.87 (2.83)	7.88 (1.26)	7.82 (2.11)	8.06 (1.81)	−0.87 (0.76)	−0.32
Monitoring[h]	4.65 (3.21)	4.84 (2.98)	3.46 (2.61)	5.17 (2.41)	0.47 (0.18)[i]	0.54

[a]CG: control group.
[b]IG: intervention group.
[c]ES: effect size.
[d]Gamma (identity).
[e]*P*<.05.
[f]Gaussian (identity).
[g]*P*<.10.
[h]Gamma (log).
[i]*P*<.01.

Randomized Controlled Trial 2

Figure 5 shows the flow of the participants. A total of 65 participants agreed to participate in the study. As we do not know how many people saw the advertisements, the response rate could not be calculated. Of them, 2 participants dropped out before completing the baseline measurements. Consequently, the data of 63 participants were analyzed. Of the 8 participants who dropped out before completing 4 sessions, only 1 was willing to complete the questionnaire assessing specific reasons for attrition. The participant indicated that *MyPlan 2.0* did not meet her expectations and that her friends or family did not respond positively to her participation in the study. Furthermore, she indicated that the high number of research-related questionnaires frustrated her.

Figure 5. Flow of the sample of randomized controlled trial 2.



XSL•FO
RenderX

The baseline characteristics of the participants are provided in Table 5. At baseline, 46 participants decided to focus on PA (28 of these participants were later allocated to the PA intervention group) and 17 participants chose to focus on SB (14 of these participants were later allocated to the SB intervention group). Consequently, the PA intervention group comprised 28 participants and the SB intervention group comprised 14 participants. No significant baseline differences in sociodemographic characteristics were found among the PA intervention group, the SB intervention group, and the control group. Of the participants, 5 used the optional mobile app. The dropout analyses indicated that the participants with a lower level of education (ie, no college or university degree) (χ^2_1=3.2; *P*=.07) and those allocated to the intervention group (χ^2_1=3.0; *P*=.08) were more likely to drop out. No significant dropout effects were found for age, sex, BMI, total PA at baseline (accelerometer-measured), or sedentary time at baseline (accelerometer-measured).

Table 6 displays the means and standard deviations for each of the behavioral outcomes in the three groups. Table 7 provides the time-by-group interactions and effect sizes for each of the behavioral outcomes. A significant intervention effect favoring the PA intervention group was identified for self-reported total PA (*P*=.003). Borderline significant intervention effects favoring the SB intervention group were found for self-reported daily sitting (*P*=.08), MPA (*P*=.06), and MVPA (*P*=.07). No intervention effects were detected for the outcome variables self-reported total transport-related PA, self-reported total household-related PA, self-reported total leisure-related PA, accelerometer-assessed MVPA, accelerometer-assessed number of breaks per day, accelerometer-assessed length of the sedentary bouts, accelerometer-assessed sedentary time, accelerometer-assessed LPA, accelerometer-assessed total PA, or accelerometer-assessed daily steps.

Table 8 displays the time-by-group interactions and effect sizes for the personal determinants in RCT 2. As described above, the PA intervention group and the SB intervention group were considered as one group to analyze the effect on the personal determinants. For coping planning, a significant intervention effect favoring the intervention group was found (*P*<.001). Furthermore, borderline significant intervention effects favoring the intervention group were found for intention (*P*=.07), self-efficacy (*P*=.05), and monitoring (*P*=.09). No intervention effect was found for outcome expectancies, risk perception, or action planning.

Table 5. Baseline characteristics of the sample of randomized controlled trial 2.

Baseline characteristics	Total sample (N=63)	CG[a] (n=21)	IG[b]–PA[c] (n=28)	IG–SB[d] (n=14)	*F* or χ^2 (df)	*P* value
Age (years), mean (SD)	58.68 (7.76)	57.67 (7.18)	59.00 (7.98)	59.57 (8.55)	0.29[e] (2,60)	.75
Males, n (%)	16 (25)	6 (29)	4 (14)	6 (43)	4.19[f] (2)	.12
University/college, n (%)	38 (60.32)	13 (62)	14 (50)	11 (79)	3.22[f] (2)	.20
Body mass index (kg/m²), mean (SD)	25.91 (3.86)	25.29 (4.07)	26.14 (3.94)	26.34 (3.55)	0.40[e] (2,58)	.68
Waist circumference (cm), mean (SD)	89.08 (12.89)	89.03 (14.69)	89.26 (11.22)	91.51 (13.93)	0.48[e] (2,58)	.62

[a]CG: control group.

[b]IG: intervention group.

[c]PA: physical activity.

[d]SB: sedentary behavior.

[e]*F* value.

[f]χ^2 value.

XSL•FO
RenderX

Table 6. Means and standard deviations for each of the behavioral outcomes in the three groups in randomized controlled trial 2.

Behavioral outcomes	CG[a], mean (SD)		IG[b]–PA[c], mean (SD)		IG–SB[d], mean (SD)	
	Pre	Post	Pre	Post	Pre	Post
LASA[e] questionnaire						
Total sitting time (min/day)	414.21 (187.08)	421.52 (189.29)	378.39 (182.25)	335.25 (167.12)	615.00 (195.46)	549.69 (175.37)
IPAQ[f]						
Total transport-related PA (min/day)	13.30 (15.29)	11.29 (12.93)	16.17 (35.33)	17.8 (13.32)	28.98 (40.10)	12.40 (15.48)
Total household-related PA (min/day)	38.01 (41.71)	58.88 (70.27)	35.08 (36.66)	75.00 (78.41)	42.40 (51.65)	50.56 (37.63)
Total leisure-related PA (min/day)	45.99 (68.62)	39.08 (39.40)	19.24 (27.66)	28.04 (31.95)	26.58 (31.16)	36.33 (37.84)
Total PA (min/day)	134.57 (96.05)	117.55 (86.04)	109.39 (107.99)	168.93 (99.52)	98.42 (94.30)	104.18 (48.76)
Moderate-to-vigorous physical activity (min/day)	65.20 (73.66)	64.25 (75.19)	63.88 (81.80)	106.39 (78.42)	36.12 (39.59)	60.61 (38.70)
Accelerometer						
Number of breaks per day	13.61 (3.60)	12.81 (3.27)	13.01 (2.19)	12.02 (2.43)	15.02 (2.22)	14.10 (2.76)
Length of sedentary bouts (min/day)	20.78 (2.50)	20.60 (3.02)	20.77 (2.08)	20.42 (2.50)	22.19 (2.19)	21.95 (2.66)
Sedentary time (min/day)	482.41 (76.11)	460.11 (75.94)	472.52 (65.42)	450.09 (64.09)	512.94 (50.07)	486.94 (73.86)
Light physical activity (min/day)	306.08 (89.43)	316.74 (74.81)	337.39 (80.84)	348.74 (81.03)	262.25 (53.62)	270.39 (61.93)
Moderate physical activity (min/day)	29.13 (21.70)	22.70 (14.56)	24.35 (11.37)	17.14 (10.40)	26.26 (21.29)	28.27 (17.09)
Moderate-to-vigorous physical activity (min/day)	29.33 (22.07)	23.51 (14.75)	24.96 (12.26)	17.23 (10.53)	28.96 (23.40)	30.94 (17.62)
Total PA (min/day)	335.41 (91.72)	340.25 (73.81)	362.36 (82.76)	365.98 (87.64)	291.20 (66.53)	301.33 (73.21)
Daily steps	7929.68 (2976.07)	7779.78 (2147.86)	8271.67 (2464.25)	7663.54 (2797.72)	7809.16 (3231.18)	8479.68 (3343.09)

[a]CG: control group.

[b]IG: intervention group.

[c]PA: physical activity.

[d]SB: sedentary behavior.

[e]LASA: Longitudinal Aging Study Amsterdam.

[f]IPAQ: international physical activity questionnaire.

XSL•FO
RenderX

Table 7. Time-by-group interactions and effect sizes for each of the behavioral outcomes in randomized controlled trial 2.

Behavioral outcomes	Time×group PA[a] (ref: pre×CG[b]), beta (SE)	ES[c] (IG[d]–PA vs CG)	Time×group SB[e] (ref: pre×CG), beta (SE)	ES (IG−SB vs CG)
LASA[f] questionnaire				
Total sitting time (min/day)[g]	−0.06 (0.07)	−0.27	−0.14 (0.08)[h]	−0.37
International physical activity questionnaire				
Total transport-related PA (min/day)[g]	0.26 (0.77)	0.13	−0.69 (0.88)	−0.43
Total household-related PA (min/day)[g]	0.32 (0.65)	0.49	−0.26 (0.75)	−0.26
Total leisure-related PA (min/day)[g]	0.54 (0.76)	0.32	0.48 (0.87)	0.35
Total PA (min/day)[i]	73.85 (25.80)[j]	0.74	22.79 (28.92)	0.24
Moderate-to-vigorous physical activity (min/day)[g]	0.52 (0.66)	0.55	0.53 (0.76)	0.48
Accelerometer				
Number of breaks per day[i]	−0.30 (0.63)	−0.07	−0.36 (0.71)	−0.04
Length of sedentary bouts (min/day)[i]	−0.12 (0.59)	−0.08	−0.11 (0.66)	−0.03
Sedentary time (min/day)[i]	−4.76 (16.97)	−0.002	−8.90 (19.08)	−0.06
Light physical activity (min/day)[i]	2.12 (12.79)	0.008	0.70 (14.29)	−0.04
Moderate physical activity (min/day)[k]	2.36 (2.76)	−0.05	7.85 (4.17)[h]	0.40
Moderate-to-vigorous physical activity (min/day)[k]	1.64 (2.78)	−0.11	7.50 (4.21)[h]	0.34
Total PA (min/day)[i]	1.71 (14.43)	−0.02	7.23 (16.14)	0.07
Daily steps[i]	−91.16 (553.35)	−0.17	763.22 (619.53)	0.26

[a]PA: physical activity.
[b]CG: control group.
[c]ES: effect size.
[d]IG: intervention group.
[e]SB: sedentary group.
[f]LASA:Longitudinal Aging Study Amsterdam.
[g]Gamma (log).
[h]$P<.10$.
[i]Gaussian (identity).
[j]$P<.01$.
[k]Gamma (identity).

XSL•FO
RenderX

Table 8. Time-by-group interactions and effect sizes for the personal determinants in randomized controlled trial 2.

Personal determinants	CG[a], mean (SD)		IG[b], mean (SD)		Time×group (ref: pre×CG), beta (SE)	ES[c]
	Pre	Post	Pre	Post		
Self-efficacy[d]	5.84 (2.67)	5.66 (2.08)	6.13 (1.82)	6.71 (1.71)	0.76 (0.39)[e]	0.37
Outcome expectancies[d]	7.27 (1.32)	7.32 (1.59)	7.31 (1.41)	7.30 (1.21)	0.30 (0.27)	−0.04
Risk perception[f]	2.64 (1.82)	2.57 (1.84)	3.79 (2.09)	3.83 (2.30)	−0.07 (0.08)	0.05
Action planning[d]	5.49 (2.27)	5.40 (2.29)	5.63 (2.24)	5.25 (2.16)	0.05 (0.65)	−0.13
Coping planning[d]	3.84 (2.69)	3.32 (1.78)	3.32 (2.30)	5.65 (2.04)	2.59 (0.50)[g]	1.19
Intention[d]	7.19 (2.11)	6.81 (2.57)	7.83 (1.88)	7.78 (1.70)	0.93 (0.51)[e]	0.17
Monitoring[d]	2.46 (2.40)	2.16 (1.69)	2.65 (1.77)	3.60 (2.36)	0.57 (0.34)[e]	0.65

[a]CG: control group.

[b]IG: intervention group.

[c]ES: effect size.

[d]Gamma (identity).

[e]$P<.10$.

[f]Gamma (log).

[g]$P<.001$.

Discussion

Efficacy of MyPlan 2.0

This study investigated the effect of a self-regulation–based eHealth and mHealth intervention (*MyPlan 2.0*) targeting an active lifestyle in two samples: adults having T2DM and adults aged ≥50 years. The study comprised two RCTs with an identical design. Although the pattern of results was overall in line with our hypotheses, the analyses revealed that the intervention only altered some of the outcomes. Indeed, this effect might be because of a lack of statistical power caused by the small samples in both the trials. The RCTs described here should, therefore, be considered pilot RCTs providing preliminary information regarding the potential effect of a HAPA-based eHealth and mHealth intervention in adults with T2DM and in adults aged ≥50 years.

The HAPA describes a number of personal determinants influencing the behavior change process. *MyPlan 2.0* affected various of these determinants. In RCT 1, an intervention effect in favor of the intervention group was found for action planning (borderline) and self-monitoring, but significant intervention effects favoring the control group were detected for risk perceptions and self-efficacy. In the RCT 2, intervention effects favoring the intervention group were detected for self-efficacy (borderline), intention (borderline), coping planning, and self-monitoring (borderline).

Some of these findings require additional attention. First, although targeted in the intervention, no intervention effect was found for outcome expectancies. This finding might be explained by a ceiling effect caused by the high levels of positive outcome expectancies at baseline in both RCTs. Indeed, our qualitative studies indicated that the users often have an extensive knowledge of the benefits of adopting an active way of living [29,30]. Second, although *MyPlan 2.0* does not provide the users with a pedometer or wearable automatically tracking the users' behavior change, both RCTs identified intervention effects favoring the intervention group for monitoring. Avery et al found a negative effect of pedometer use on PA in people with T2DM and older adults, indicating that without additional support, these populations found it difficult to effectively reflect on the information provided by this self-monitoring tool [54]. Our results indicate that prompting the users to monitor their change and reviewing this change in the following session might be a feasible alternative to target self-monitoring in these samples. Third, the lack of effect for action planning in the RCT with adults aged ≥50 years was unexpected, as this determinant was targeted in each session. Sniehotta et al argued that action planning might play an important role for individuals who just started to put their intentions into actions, whereas coping planning would support individuals who moved further in the behavior change process to maintain their change under challenging conditions [55]. As the baseline levels of PA and SB of the RCT with the sample aged ≥50 years are quite close to the health norms [5], it is possible that this group already knew how to plan their actions and consequently, did not benefit from the action planning component. Similarly, considering the low levels of PA and high levels of SB at baseline in the RCT with adults with T2DM, the lack of evidence for coping planning could be explained by the fact that this group was not yet ready to optimally benefit from the coping planning component.

MyPlan 2.0 focused on altering the users' level of PA or SB. In RCT 1, borderline significant intervention effects favoring the PA intervention group were found for self-reported daily sitting and accelerometer-assessed MPA and MVPA. This is an important result as a previous study by Silfee et al, testing a self-regulation–based intervention targeting PA in adults with T2DM, did not show behavioral effects despite the positive

XSL•FO
RenderX

effect on personal determinants for change (including self-monitoring) [56]. In RCT 2, an intervention effect favoring the PA intervention group was found for self-reported total PA. This effect is in line with the previous research with *MyPlan 1.0* in recently retired older adults [57]. The lack of evidence for intervention effects on self-reported domain-specific PA in both RCTs is in line with our hypotheses and can be explained by the fact that *MyPlan 2.0* allows the users to select each session a different PA-domain that is at that moment most relevant to them rather than imposing a specific domain.

In RCT 1, an intervention effect favoring the SB intervention group was found for accelerometer-assessed daily breaks from sedentary time. To our knowledge, *MyPlan 2.0* is the first eHealth and mHealth intervention targeting sedentary behavior in adults with type 2 diabetes. Considering the health effects of breaking up periods of prolonged sitting in adults with T2DM [58], this result warrants further research regarding eHealth interventions targeting sedentary behavior in adults with type 2 diabetes. In RCT 2, an intervention effect favoring the SB intervention group was detected for self-reported daily sitting time (borderline). This finding is in line with the research by Stephenson et al, indicating that technology enhanced interventions are able to reduce sedentary behavior [59]. Although it is assumed that sedentary behavior will be replaced by LPA rather than MVPA [60], intervention effects favoring the SB intervention group were found for MPA (borderline) and MVPA. Similarly, Gardiner et al found that their intervention to reduce and break up sedentary time in older adults resulted in changes in sedentary time, breaks from sedentary time, LPA, and MVPA [61].

Overall, the lack of intervention effects reaching statistical significance could be interpreted as disappointing. However, one has to keep in mind the following issues that may have led to an underestimation of our effects. First, in keeping with the self-regulation literature, *MyPlan 2.0* motivated the users to set and pursue their own goals. Consequently, the set goals could differ strongly between as well as within the participants (ie, each session the participants could select a different goal) on 4 aspects: chosen behavior (eg, MVPA vs LPA), ambitiousness (eg, reaching 500 vs 5000 additional steps), setting (eg, leisure time vs transport), and time frame (eg, every day of the week vs in the weekend). This might have lowered the chance of finding an effect. However, this approach was believed to be better and more sustainable. It would lead to more success experiences and a greater willingness to continue with the process of behavioral change. From a methodological point of view, we may, therefore, recommend targeting one type of behavior (eg, decreasing sitting time) that can be performed in a wide variety of settings. This approach will allow (1) the users to create personal goals (ie, create a sense of goal ownership) and (2) the researchers to select the most appropriate measurement methods to detect alterations in the targeted behavior. Second, as accelerometers are not able to capture posture, these devices tend to have problems to distinguish between sedentary time and light-intensity PA [62]. This could imply that some of the accelerometer-assessed breaks do not automatically reflect posture change from sitting to standing. Furthermore, previous research already indicated that the agreement between self-reported and objective measurements of PA is limited [63]. Indeed, instead of creating a hierarchy of preferred measures, objective and self-report measures should be considered distinct rather than interchangeable [64]. Finally, our limited power caused by the small samples might have hindered a number of effects to reach statistical significance.

Attrition Levels in MyPlan 2.0

Web-based interventions are characterized by high levels of attrition [65]. More than 70% of *MyPlan 1.0* users did not complete the intervention [66,67]. In RCT 1, 36% (13/36) of the participants receiving *MyPlan 2.0* did not complete the intervention. In RCT 2, this was 19% (8/42). These massive reductions in attrition might be explained by the iterative adaptations that were made to the program to increase engagement and by the fact that the participants were phoned on a weekly basis. However, in both RCTs, we found that the participants receiving the intervention were still more likely to quit compared with those in the control group. Furthermore, in RCT 2, we found that dropout was higher in users with a lower level of education. These findings were disappointing as we, being aware of this issue, purposefully conducted a series of studies to adapt the intervention's content to this target population [29,30,32]. Yardley et al argue to make a distinction between the micro (engagement with intervention itself) and macro (engagement with the behavior change process to reach the set goals) level of engagement to create effective engagement (ie, *sufficient engagement with the intervention to reach the desired outcomes* [68]) rather than simply more engagement. This idea is in line with the hypothesis of Eysenbach stating that the users need to experience the added value of using the Web-based intervention to prevent attrition [65]. Consequently, not only investigating whether the users like the program itself but also identifying how they put the learned techniques into practice and which variables (eg, level of education) moderate this process might be a fruitful avenue to (1) decrease the level of attrition and (2) increase the effectiveness of Web-based interventions in the future.

Strengths and Limitations of the Study

This study has several strengths. First, several studies have assessed the effect of internet-based interventions on SB in the general population [12,59]. To our knowledge, this is the first study testing a Web-based intervention targeting SB in adults with T2DM. Second, by also assessing the HAPA-based determinants for change, we were able to check whether the implemented behavior change techniques effectively altered the users' personal determinants for change. Finally, by using self-report as well as objective measurements, a more nuanced view of the effects was presented. However, it should be acknowledged that the self-report and objective measures did not represent the same time frame.

There are also a number of limitations. First, no power analysis was conducted for RCT 2. Second, the small sample sizes made it difficult to detect statistically significant effects. Third, a waiting-list rather than a placebo control group was created. Consequently, we are not certain whether the detected intervention effects were actually caused by the active ingredients of the intervention. Furthermore, informing a

XSL•FO
RenderX

participant that he or she is allocated to a waiting-list control group might have influenced his/her behavior (eg, participants of the control group might have felt reluctant to alter their behavior as they knew they would receive support later). Indeed, previous research has shown that trials using a waiting-list control condition might overestimate treatment effects [69]. Fourth, to analyze the effect of *MyPlan 2.0* on the HAPA-based personal determinants, no distinction was made between the two intervention groups (ie, they were combined into one group). Consequently, it was not possible to investigate whether the intervention effects for the personal determinants altered according to the chosen behavior. Fifth, the users were contacted each week to assess potential negative effects (eg, hypoglycemia). Furthermore, the users who forgot to log in for the following session were contacted by the researcher to inform them about the awaiting session. These phone calls might have motivated the participants to stay in the study and to complete the intervention. Consequently, the detected attrition rates might be an underestimation of the actual attrition rates of the program. Finally, the effects reported here reflect short-term changes. However, a third wave of data collection at 10 months post baseline will be performed.

Conclusions

To conclude, this study suggests that a self-regulation–based Web-based intervention has the potential to alter levels of PA and SB in adults with T2DM and in adults aged ≥50 years. However, further research with larger samples is needed to confirm the consistency of these findings.

Acknowledgments

The authors would like to thank the Ghent University Hospital and the Damian General Hospital (Ostend) for their collaboration. They also thank Prof Dr Armand De Clercq for his support in developing *MyPlan 2.0*. LP and MV are funded by the Research Foundation-Flanders (FWO). The funder had no role in study design, data collection, and analysis, decision to publish, or preparation of the manuscript.

Conflicts of Interest

The authors of this manuscript were involved in the development of the evaluated intervention.

Multimedia Appendix 1

Overview of adaptations made to “MyPlan 2.0” based on the user-based studies.

[PDF File (Adobe PDF File), 230KB - jmir_v21i8e13363_app1.pdf]

Multimedia Appendix 2

Screenshots of “MyPlan 2.0”.

[PDF File (Adobe PDF File), 738KB - jmir_v21i8e13363_app2.pdf]

Multimedia Appendix 3

CONSORT-EHEALTH checklist (V 1.6.1).

[PDF File (Adobe PDF File), 2MB - jmir_v21i8e13363_app3.pdf]

References

1. Ogurtsova K, da Rocha Fernandes JD, Huang Y, Linnenkamp U, Guariguata L, Cho NH, et al. IDF Diabetes Atlas: global estimates for the prevalence of diabetes for 2015 and 2040. Diabetes Res Clin Pract 2017 Jun;128:40-50. [doi: 10.1016/j.diabres.2017.03.024] [Medline: 28437734]
2. Geneau R, Stuckler D, Stachenko S, McKee M, Ebrahim S, Basu S, et al. Raising the priority of preventing chronic diseases: a political process. Lancet 2010 Nov 13;376(9753):1689-1698. [doi: 10.1016/S0140-6736(10)61414-6] [Medline: 21074260]
3. Warburton DE, Nicol CW, Bredin SS. Health benefits of physical activity: the evidence. CMAJ 2006 Mar 14;174(6):801-809 [FREE Full text] [doi: 10.1503/cmaj.051351] [Medline: 16534088]
4. Dempsey PC, Owen N, Yates TE, Kingwell BA, Dunstan DW. Sitting less and moving more: improved glycaemic control for type 2 diabetes prevention and management. Curr Diab Rep 2016 Dec;16(11):114. [doi: 10.1007/s11892-016-0797-4] [Medline: 27699700]
5. World Health Organization. Global Recommendations On Physical Activity For Health. 2010. URL: https://www.who.int/dietphysicalactivity/factsheet_recommendations/en/
6. Australian Government: Department of Health. Australia's physical activity and sedentary behaviour Guidelines (adults) URL: http://www.health.gov.au/internet/main/publishing.nsf/content/health-pubhlth-strateg-phys-act-guidelines [accessed 2019-06-01]

XSL•FO
RenderX

7. Unick JL, Gaussoin SA, Hill JO, Jakicic JM, Bond DS, Hellgren M, et al. Four-year physical activity levels among intervention participants with type 2 diabetes. Med Sci Sports Exerc 2016 Dec;48(12):2437-2445 [FREE Full text] [doi: 10.1249/MSS.0000000000001054] [Medline: 27471785]
8. Guo VY, Brage S, Ekelund U, Griffin SJ, Simmons RK, ADDITION-Plus study team. Objectively measured sedentary time, physical activity and kidney function in people with recently diagnosed Type 2 diabetes: a prospective cohort analysis. Diabet Med 2016 Dec;33(9):1222-1229 [FREE Full text] [doi: 10.1111/dme.12886] [Medline: 26282583]
9. Müller AM, Maher CA, Vandelanotte C, Hingle M, Middelweerd A, Lopez ML, et al. Physical activity, sedentary behavior, and diet-related eHealth and mHealth research: bibliometric analysis. J Med Internet Res 2018 Apr 18;20(4):e122. [doi: 10.2196/jmir.8954] [Medline: 29669703]
10. 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]
11. Davies CA, Spence JC, Vandelanotte C, Caperchione CM, Mummery WK. Meta-analysis of internet-delivered interventions to increase physical activity levels. Int J Behav Nutr Phys Act 2012;9:52 [FREE Full text] [doi: 10.1186/1479-5868-9-52] [Medline: 22546283]
12. Direito A, Carraça E, Rawstorn J, Whittaker R, Maddison R. mHealth technologies to influence physical activity and sedentary behaviors: behavior change techniques, systematic review and meta-analysis of randomized controlled trials. Ann Behav Med 2017 Apr;51(2):226-239. [doi: 10.1007/s12160-016-9846-0] [Medline: 27757789]
13. Pal K, Eastwood SV, Michie S, Farmer AJ, Barnard ML, Peacock R, et al. Computer-based diabetes self-management interventions for adults with type 2 diabetes mellitus. Cochrane Database Syst Rev 2013;3:CD008776. [doi: 10.1002/14651858.CD008776.pub2] [Medline: 23543567]
14. 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]
15. Maes S, Karoly P. Self‐regulation assessment and intervention in physical health and illness: a review. Appl Psychol 2005 Apr;54(2):267-299 [FREE Full text] [doi: 10.1111/j.1464-0597.2005.00210.x]
16. Rhodes RE, Yao CA. Models accounting for intention-behavior discordance in the physical activity domain: a user's guide, content overview, and review of current evidence. Int J Behav Nutr Phys Act 2015 Feb 07;12:9 [FREE Full text] [doi: 10.1186/s12966-015-0168-6] [Medline: 25890238]
17. Schwarzer R, Lippke S, Luszczynska A. Mechanisms of health behavior change in persons with chronic illness or disability: the Health Action Process Approach (HAPA). Rehabil Psychol 2011 Aug;56(3):161-170. [doi: 10.1037/a0024509] [Medline: 21767036]
18. MacPhail M, Mullan B, Sharpe L, MacCann C, Todd J. Using the health action process approach to predict and improve health outcomes in individuals with type 2 diabetes mellitus. Diabetes Metab Syndr Obes 2014;7:469-479 [FREE Full text] [doi: 10.2147/DMSO.S68428] [Medline: 25342914]
19. Sui W, Prapavessis H. Standing Up for Student Health: an application of the health action process approach for reducing student sedentary behavior-randomised control pilot trial. Appl Psychol Health Well Being 2018 Dec;10(1):87-107. [doi: 10.1111/aphw.12105] [Medline: 29205909]
20. Maher JP, Sliwinski MJ, Conroy DE. Feasibility and preliminary efficacy of an intervention to reduce older adults' sedentary behavior. Transl Behav Med 2017 Dec;7(1):52-61 [FREE Full text] [doi: 10.1007/s13142-016-0394-8] [Medline: 27734281]
21. Reinwand D, Kuhlmann T, Wienert J, de Vries H, Lippke S. Designing a theory- and evidence-based tailored eHealth rehabilitation aftercare program in Germany and the Netherlands: study protocol. BMC Public Health 2013 Nov 19;13:1081 [FREE Full text] [doi: 10.1186/1471-2458-13-1081] [Medline: 24245493]
22. Plaete J, De Bourdeaudhui I, Verloigne M, Oenema A, Crombez G. A self-regulation eHealth intervention to increase healthy behavior through general practice: protocol and systematic development. JMIR Res Protoc 2015;4(4):e141 [FREE Full text] [doi: 10.2196/resprot.4835] [Medline: 26694101]
23. Haug S, Schaub MP, Venzin V, Meyer C, John U. Efficacy of a text message-based smoking cessation intervention for young people: a cluster randomized controlled trial. J Med Internet Res 2013;15(8):e171 [FREE Full text] [doi: 10.2196/jmir.2636] [Medline: 23956024]
24. Duan YP, Wienert J, Hu C, Si GY, Lippke S. Web-based intervention for physical activity and fruit and vegetable intake among Chinese university students: a randomized controlled trial. J Med Internet Res 2017 Apr 10;19(4):e106 [FREE Full text] [doi: 10.2196/jmir.7152] [Medline: 28396306]
25. Gomez QS, Walthouwer MJ, Schulz DN, de Vries H. mHealth or eHealth? Efficacy, use, and appreciation of a Web-based computer-tailored physical activity intervention for Dutch adults: a randomized controlled trial. J Med Internet Res 2016 Nov 9;18(11):e278 [FREE Full text] [doi: 10.2196/jmir.6171] [Medline: 27829576]
26. Poppe L, De Bourdeaudhuij I, Verloigne M, Degroote L, Shadid S, Crombez G. A self-regulation-based eHealth and mHealth intervention for an active lifestyle in adults with type 2 diabetes: protocol for a randomized controlled trial. JMIR Res Protoc 2019 Mar 22;8(3):e12413 [FREE Full text] [doi: 10.2196/12413] [Medline: 30901002]

XSL•FO
RenderX

27. Thivel D, Tremblay A, Genin PM, Panahi S, Rivière D, Duclos M. Physical activity, inactivity, and sedentary behaviors: definitions and implications in occupational health. Front Public Health 2018;6:288 [FREE Full text] [doi: 10.3389/fpubh.2018.00288] [Medline: 30345266]
28. Schwarzer R. Modeling health behavior change: how to predict and modify the adoption and maintenance of health behaviors. Appl Psychol 2008 Jan;57(1):1-29 [FREE Full text] [doi: 10.1111/j.1464-0597.2007.00325.x]
29. Poppe L, Crombez G, de Bourdeaudhuij I, van der Mispel C, Shadid S, Verloigne M. Experiences and opinions of adults with type 2 diabetes regarding a self-regulation-based ehealth intervention targeting physical activity and sedentary behaviour. Int J Environ Res Public Health 2018 Dec 10;15(5):E954 [FREE Full text] [doi: 10.3390/ijerph15050954] [Medline: 29748460]
30. Poppe L, van der Mispel C, Crombez G, de Bourdeaudhuij I, Schroé H, Verloigne M. How users experience and use an eHealth intervention based on self-regulation: mixed-methods study. J Med Internet Res 2018 Dec 1;20(10):e10412 [FREE Full text] [doi: 10.2196/10412] [Medline: 30274961]
31. van der Mispel C, Poppe L, Crombez G, Verloigne M, de Bourdeaudhuij I. A self-regulation-based eHealth intervention to promote a healthy lifestyle: investigating user and website characteristics related to attrition. J Med Internet Res 2017 Jul 11;19(7):e241 [FREE Full text] [doi: 10.2196/jmir.7277] [Medline: 28698168]
32. Poppe L, van der Mispel C, de Bourdeaudhuij BI, Verloigne M, Shadid S, Crombez G. Users' thoughts and opinions about a self-regulation-based eHealth intervention targeting physical activity and the intake of fruit and vegetables: a qualitative study. PLoS One 2017;12(12):e0190020 [FREE Full text] [doi: 10.1371/journal.pone.0190020] [Medline: 29267396]
33. 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]
34. Yang Y, Osmond A, Chen X, Weal M, Wills G, de Roure D, et al. Supporting the running and analysis of trials of web-based behaviour interventions: the LifeGuide. 2009 Presented at: 5th IEEE International Conference on e-Science; December 7-9, 2009; Oxford, United Kingdom.
35. Craig CL, Marshall AL, Sjöström M, Bauman AE, Booth ML, Ainsworth BE, et al. International physical activity questionnaire: 12-country reliability and validity. Med Sci Sports Exerc 2003 Aug;35(8):1381-1395. [doi: 10.1249/01.MSS.0000078924.61453.FB] [Medline: 12900694]
36. Bermúdez VJ, Rojas JJ, Córdova EB, Añez R, Toledo A, Aguirre MA, et al. International physical activity questionnaire overestimation is ameliorated by individual analysis of the scores. Am J Ther 2013;20(4):448-458. [doi: 10.1097/MJT.0b013e318235f1f2] [Medline: 23838635]
37. Dubuy V, de Cocker KA, de Bourdeaudhuij IM, Maes L, Metsu P, van Acker R, et al. '10 000 Steps Flanders': evaluation of the state-wide dissemination of a physical activity intervention in Flanders. Health Educ Res 2013 Jun;28(3):546-551. [doi: 10.1093/her/cyt039] [Medline: 23503570]
38. Visser M, Koster A. Development of a questionnaire to assess sedentary time in older persons--a comparative study using accelerometry. BMC Geriatr 2013 Jul 30;13:80 [FREE Full text] [doi: 10.1186/1471-2318-13-80] [Medline: 23899190]
39. Lakerveld J, Mackenbach JD, Horvath E, Rutters F, Compernolle S, Bárdos H, et al. The relation between sleep duration and sedentary behaviours in European adults. Obes Rev 2016 Jan;17 Suppl 1:62-67. [doi: 10.1111/obr.12381] [Medline: 26879114]
40. Trost SG, Loprinzi PD, Moore R, Pfeiffer KA. Comparison of accelerometer cut points for predicting activity intensity in youth. Med Sci Sports Exerc 2011 Jul;43(7):1360-1368. [doi: 10.1249/MSS.0b013e318206476e] [Medline: 21131873]
41. Aguilar-Farías N, Brown WJ, Peeters GM. ActiGraph GT3X+ cut-points for identifying sedentary behaviour in older adults in free-living environments. J Sci Med Sport 2014 May;17(3):293-299. [doi: 10.1016/j.jsams.2013.07.002] [Medline: 23932934]
42. Rothney MP, Apker GA, Song Y, Chen KY. Comparing the performance of three generations of ActiGraph accelerometers. J Appl Physiol (1985) 2008 Oct;105(4):1091-1097 [FREE Full text] [doi: 10.1152/japplphysiol.90641.2008] [Medline: 18635874]
43. Sasaki JE, John D, Freedson PS. Validation and comparison of ActiGraph activity monitors. J Sci Med Sport 2011 Sep;14(5):411-416. [doi: 10.1016/j.jsams.2011.04.003] [Medline: 21616714]
44. Troiano RP, Berrigan D, Dodd KW, Mâsse LC, Tilert T, McDowell M. Physical activity in the United States measured by accelerometer. Med Sci Sports Exerc 2008 Jan;40(1):181-188. [doi: 10.1249/mss.0b013e31815a51b3] [Medline: 18091006]
45. Freedson PS, Melanson E, Sirard J. Calibration of the Computer Science and Applications, Inc accelerometer. Med Sci Sports Exerc 1998 May;30(5):777-781. [Medline: 9588623]
46. Johnston M, Dixon D, Hart J, Glidewell L, Schröder C, Pollard B. Discriminant content validity: a quantitative methodology for assessing content of theory-based measures, with illustrative applications. Br J Health Psychol 2014 May;19(2):240-257. [doi: 10.1111/bjhp.12095] [Medline: 24628841]
47. R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2017. URL: https://www.r-project.org/ [accessed 2019-06-01]

XSL•FO
RenderX

48. Bates D, Maechler M, Bolker B, Walker S, Christensen R, Singmann H, et al. CRAN - The R Project for Statistical Computing. 2015 Aug 25. Linear Mixed-Effects Models using 'Eigen' and S4 URL: https://cran.r-project.org/web/packages/lme4/index.html [accessed 2019-06-01]
49. Verbeke G, Molenberghs G. Linear Mixed Models For Longitudinal Data. New York: Springer; 2000.
50. Maas C, Snijders T. The multilevel approach to repeated measures for complete and incomplete data. Qual Quant 2003;37(1):71-89. [doi: 10.1023/A:1022545930672]
51. Chakraborty H, Gu H. A Mixed Model Approach for Intent-to-Treat Analysis in Longitudinal Clinical Trials with Missing Values. RTI Press Publication No. MR-0009-0903. Research Triangle Park, NC: RTI Press; 2009.
52. Lenhard W, Lenhard A. Psychometrica. 2016. Calculation of Effect Sizes URL: https://www.psychometrica.de/effect_size.html [accessed 2019-06-01]
53. Morris SB. Estimating effect sizes from pretest-posttest-control group designs. Organ Res Methods 2007 Jul 23;11(2):364-386 [FREE Full text] [doi: 10.1177/1094428106291059]
54. Avery L, Flynn D, Dombrowski S, van Wersch A, Sniehotta F, Trenell MI. Successful behavioural strategies to increase physical activity and improve glucose control in adults with Type 2 diabetes. Diabet Med 2015 Aug;32(8):1058-1062. [doi: 10.1111/dme.12738] [Medline: 25764343]
55. Sniehotta F, Schwarzer R, Scholz U, Schüz B. Action planning and coping planning for long-term lifestyle change: theory and assessment. Eur J Soc Psychol 2005 Jul;35(4):565-576 [FREE Full text] [doi: 10.1002/ejsp.258]
56. Silfee V, Petosa R, Laurent D, Schaub T, Focht B. Effect of a behavioral intervention on dimensions of self-regulation and physical activity among overweight and obese adults with type 2 diabetes: a pilot study. Psychol Health Med 2016 Sep;21(6):715-723. [doi: 10.1080/13548506.2016.1139144] [Medline: 26785605]
57. van Dyck D, Plaete J, Cardon G, Crombez G, de Bourdeaudhuij I. Effectiveness of the self-regulation eHealth intervention 'MyPlan1.0.' on physical activity levels of recently retired Belgian adults: a randomized controlled trial. Health Educ Res 2016 Oct;31(5):653-664. [doi: 10.1093/her/cyw036] [Medline: 27422898]
58. Dempsey P, Sacre J, Larsen R, Straznicky N, Sethi P, Cohen N, et al. Interrupting prolonged sitting with brief bouts of light walking or simple resistance activities reduces resting blood pressure and plasma noradrenaline in type 2 diabetes. J Hypertens 2016 Dec;34(12):2376-2382. [doi: 10.1097/HJH.0000000000001101] [Medline: 27512975]
59. Stephenson A, McDonough S, Murphy M, Nugent C, Mair JL. Using computer, mobile and wearable technology enhanced interventions to reduce sedentary behaviour: a systematic review and meta-analysis. Int J Behav Nutr Phys Act 2017 Dec 11;14(1):105 [FREE Full text] [doi: 10.1186/s12966-017-0561-4] [Medline: 28800736]
60. Mansoubi M, Pearson N, Biddle S, Clemes S. The relationship between sedentary behaviour and physical activity in adults: a systematic review. Prev Med 2014 Dec;69:28-35. [doi: 10.1016/j.ypmed.2014.08.028] [Medline: 25193005]
61. Gardiner PA, Eakin EG, Healy GN, Owen N. Feasibility of reducing older adults' sedentary time. Am J Prev Med 2011 Aug;41(2):174-177. [doi: 10.1016/j.amepre.2011.03.020] [Medline: 21767725]
62. Ainsworth B, Cahalin L, Buman M, Ross R. The current state of physical activity assessment tools. Prog Cardiovasc Dis 2015;57(4):387-395. [doi: 10.1016/j.pcad.2014.10.005] [Medline: 25446555]
63. Prince SA, Adamo KB, Hamel ME, Hardt J, Connor GS, Tremblay M. A comparison of direct versus self-report measures for assessing physical activity in adults: a systematic review. Int J Behav Nutr Phys Act 2008;5:56 [FREE Full text] [doi: 10.1186/1479-5868-5-56] [Medline: 18990237]
64. Kelly P, Fitzsimons C, Baker G. Should we reframe how we think about physical activity and sedentary behaviour measurement? Validity and reliability reconsidered. Int J Behav Nutr Phys Act 2016 Mar 1;13:32 [FREE Full text] [doi: 10.1186/s12966-016-0351-4] [Medline: 26931142]
65. 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]
66. Plaete J, Crombez G, van der Mispel C, Verloigne M, van Stappen V, de Bourdeaudhuij I. Effect of the web-based intervention MyPlan 1.0 on self-reported fruit and vegetable intake in adults who visit general practice: a quasi-experimental trial. J Med Internet Res 2016 Feb 29;18(2):e47 [FREE Full text] [doi: 10.2196/jmir.5252] [Medline: 26929095]
67. Degroote L, Plaete J, de Bourdeaudhuij I, Verloigne M, van Stappen V, de Meester A, et al. The effect of the ehealth intervention 'MypPlan 1.0' on physical activity in adults who visit general practice: a quasi-experimental trial. Int J Environ Res Public Health 2018 Jan 30;15(2):E228 [FREE Full text] [doi: 10.3390/ijerph15020228] [Medline: 29385770]
68. Yardley L, Spring BJ, Riper H, Morrison LG, Crane DH, Curtis K, et al. Understanding and promoting effective engagement with digital behavior change interventions. Am J Prev Med 2016 Nov;51(5):833-842. [doi: 10.1016/j.amepre.2016.06.015] [Medline: 27745683]
69. Cunningham JA, Kypri K, McCambridge J. Exploratory randomized controlled trial evaluating the impact of a waiting list control design. BMC Med Res Methodol 2013 Dec 6;13:150 [FREE Full text] [doi: 10.1186/1471-2288-13-150] [Medline: 24314204]

Abbreviations

BIC: Bayesian information criterion

XSL•FO
RenderX

BMI: body mass index
CG: control group
CPM: counts per minute
ES: effect size
IG: intervention group
IPAQ: international physical activity questionnaire
LASA: Longitudinal Aging Study Amsterdam
LPA: light physical activity
MPA: moderate physical activity
MVPA: moderate-to-vigorous physical activity
PA: physical activity
RCT: randomized controlled trial
SB: sedentary behavior
T2DM: type 2 diabetes mellitus
VPA: vigorous-intensity physical activity

Edited by G Eysenbach; submitted 10.01.19; peer-reviewed by D Goodrich, MH Nguyen; comments to author 27.04.19; revised version received 18.06.19; accepted 19.06.19; published 02.08.19.

Please cite as:
Poppe L, De Bourdeaudhuij I, Verloigne M, Shadid S, Van Cauwenberg J, Compernolle S, Crombez G
Efficacy of a Self-Regulation–Based Electronic and Mobile Health Intervention Targeting an Active Lifestyle in Adults Having Type 2 Diabetes and in Adults Aged 50 Years or Older: Two Randomized Controlled Trials
J Med Internet Res 2019;21(8):e13363
URL: https://www.jmir.org/2019/8/e13363/
doi:10.2196/13363
PMID:31376274

©Louise Poppe, Ilse De Bourdeaudhuij, Maïté Verloigne, Samyah Shadid, Jelle Van Cauwenberg, Sofie Compernolle, Geert Crombez. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 02.08.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

A Multicomponent eHealth Intervention for Family Carers for People Affected by Psychosis: A Coproduced Design and Build Study

Jacqueline Sin[1,2], PhD; Claire Henderson[3], PhD; Luke A Woodham[4], MSc; Aurora Sesé Hernández[4], MSc; Steve Gillard[1], PhD

[1]Population Health Research Institute, St George's, University of London, London, United Kingdom

[2]School of Psychology and Clinical Language Sciences, University of Reading, Reading, United Kingdom

[3]Health Service and Population Research Department, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, United Kingdom

[4]Institute of Medical and Biomedical Education, St George's, University of London, London, United Kingdom

Corresponding Author:
Jacqueline Sin, PhD
Population Health Research Institute
St George's, University of London
Cranmer Terrace
London, SW17 0RE
United Kingdom
Phone: 44 07817027036
Email: jasin@sgul.ac.uk

Abstract

Background: Psychosis, including schizophrenia, is the most common severe mental illness affecting 1% of the population worldwide. A large number of people provide long-term support and care for a relative with psychosis. Although psychoeducational interventions, especially those delivered through a face-to-face group format, have an established evidence base for improving the caregiving experience, well-being, and health outcomes, large-scale implementation and access remain limited. There is a demand for such provision to be made through the internet for greater flexibility and wider access.

Objective: This study aimed to integrate participatory research methodologies by the public, patients, and carers into the eHealth (electronic health) intervention design and build process to improve the product's usability and acceptability.

Methods: We adapted a structured eHealth intervention build method to include participatory research activities involving key stakeholders and end users to co-design and coproduce our intervention. An expert advisory group (EAG) comprising public involvement members led the formative design and build work using an agile build process. Carers independent from the study were consulted on the evolving drafts of the intervention prototype through focus group meetings. These results were fed back into the intervention build work continuously to ensure end users' input inform every stage of the process.

Results: An EAG comprising individuals with lived experience of psychosis, carers, health care professionals, researchers, voluntary organization workers, and eLearning experts (n=14) was established. A total of 4 coproduction workshops were held over 1 year during which the alpha and beta prototypes were designed and built through the participatory research work. Alongside this, 2 rounds of focus group study with carers (n=24, in 4 groups) were conducted to seek consultation on end users' views and ideas to optimize the intervention design and usability. Finally, the EAG carried out a Web-based walk-through exercise on the intervention prototype and further refined it to make it ready for an online usability test. The final product contains multiple sections providing information on psychosis and related caregiving topics and interactive discussion forums with experts and peers for psychosocial support. It provides psychoeducation and psychosocial support for carers through the internet, promoting flexible access and individualized choices of information and support.

Conclusions: The participatory research work led to the coproduction of a eHealth intervention called COPe-support (Carers fOr People with Psychosis e-support). We believe the study methodology, results, and output have optimized the intervention design and usability, fitting the end users' needs and usage pattern. COPe-support is currently being tested for its effectiveness in promoting carers' health outcome through an online randomized controlled trial.

XSL•FO
RenderX

Trial Registration: ISRCTN Registry ISRCTN89563420; http://www.isrctn.com/ISRCTN89563420

(J Med Internet Res 2019;21(8):e14374) doi:10.2196/14374

KEYWORDS

eHealth; family caregivers; psychosis; mental health; participatory research; public and patient involvement; coproduction

Introduction

Family Caregiving

With ever-advancing health care technologies and growing longevity worldwide, a significant proportion of people provide substantial and sustained help and support to friends or family members suffering from a long-term illness [1]. In the United Kingdom and the United States, nearly one-fourth of the adult population identifies itself as a carer for a loved one who is ill, disabled, or elderly [2,3]. Many of these carers support a loved one affected by a severe and long-term mental illness such as psychosis [4,5]. Family caregiving often covers a huge amount of care and support ranging from emotional and psychosocial support (eg, engaging their loved one in social activities and sharing ups and downs) to financial and practical support (eg, provision of financial and practical help and monitoring of health and treatment compliance). Compared with paid or professional care workers, family carers also have unique advantages in knowing the individual's strengths and interests in addition to their needs. Most have a well-established emotional bond and are committed to use all these preexisting knowledge and relationships to support their loved one in their recovery [6-8]. Caregiving imparts paramount emotional and psychosocial benefits to the cared-for individuals such that people in receipt of support from their family or social network have better prognosis, fewer relapses, and higher quality of life compared with those without such support [9-12]. Collectively, caregiving by family members amounts to significant economic savings to the wider society [4,9].

Conversely, it is well established that caring demands can jeopardize carers' well-being [2,13,14]. For instance, population-level research data repeatedly show that carers experience higher levels of distress and poorer well-being when compared with age-matched counterparts in the general population [10,15,16]. Indeed, distress in carers frequently reaches clinical thresholds, and their psychiatric symptom scores (eg, depression and anxiety) are found to be consistently inversely associated with the amount of care they provide, that is, carers' mental health worsens with increasing demands on caregiving [2,10,15]. Research evidence also suggests that poor well-being may hamper carers' caregiving capacity. Carers who feel they are not supported and lack resources (eg, information about the illness condition and related management issues) to cope are less likely to engage in caring for their loved ones or more likely to exhibit critical or hostile behavior toward the cared-for individuals, albeit unintentionally [13,14,17]. This, in turn, can impact negatively on both the carers and the cared-for individuals, leading to a vicious cycle of poor health and quality of life for all concerned.

Psychosocial Interventions Targeting Carers

Consequently, a body of research has been undertaken to explore interventions, which can best support carers [14,18,19]. Among a range of psychosocial interventions targeting whole families and/or carers alongside the individual treatment regime, psychoeducation (ie, information giving on the illness condition and related caregiving and problem-solving strategies) has the strongest evidence base for its effectiveness in enhancing carers' knowledge and coping with their caring roles [5,20]. Commonly based on the stress-appraisal-coping theory as applied in family caregiving [14,21-23], it has been hypothesized that psychoeducation, with education as its core feature and prime aim, works directly to improve carers' knowledge about psychosis and related caregiving issues. Improved knowledge about coping strategies and available resources can lead to a more positive appraisal of their caregiving experiences and carers' perceived self-efficacy in coping with the demands [14,24]. These, in turn, can translate into a more supportive home environment for all and better prognosis and reduced relapses in the cared-for individuals. In addition to information and advice on psychosis and related caregiving strategies, carers also identify that sharing mutual support and learning with other carers (ie, peers) as particularly useful in reducing their sense of isolation [18,20]. Consequently, psychoeducational interventions, especially those delivered in a group format, as a discrete treatment on their own or augmenting other treatments (eg, family intervention or mutual support programs) are widely recommended and practiced around the world [14,18,20].

Carer-Specific eHealth Interventions

Carers have expressed their desire for interventions to support them to be delivered to them through a digital medium to fit with their caregiving and other commitments [25-27]. The internet offers the potential to deliver interventions, which are highly flexible, accessible, and yet adaptable to individualized needs and schedule. With the popularity of eHealth (electronic health; ie, health care practice delivered through the internet) and mobile Health (ie, through the mobile network) interventions targeting a wide range of common public mental health issues (eg, insomnia and stress management) growing fast, internet-based interventions targeting carers have gathered momentum and popularity in the recent decade [11,19,28]. eHealth interventions using an enriched online environment can integrate multiple components, especially educational and therapeutic information and network support with health care professionals and peers, and deliver such provisions to a critical mass of carers [11,29]. Carers, as the end users, particularly appreciate the autonomy that eHealth interventions offer, as they can decide which components or strategies resonate with them, how much time to spend accessing the intervention, and when to do so [28,30]. Systematic reviews on interventions targeting carers identify that eHealth interventions are

XSL•FO
RenderX

particularly widespread in the field of dementia, stroke, or cardiovascular diseases. More importantly, research evidence to date shows promising effectiveness results in improving carers' health outcomes [11,31]. Furthermore, these reviews suggest that effective Web-based interventions commonly have 3 essential ingredients. These include (1) multiple components such as information and well-being promotion strategies (including mindfulness and cognitive behavioral–oriented exercises), (2) psychoeducational content to enhance understanding of the illness condition and related caregiving issues, and (3) flexibility to self-pace and self-tailor the intervention to individual needs [11,19,31].

Carers for People With Psychosis

Notwithstanding the evidence base for eHealth interventions targeting carers and their expressed desire for such provision, development in the field of psychosis appears to be lagging with few empirical studies documented to date. Of these studies, 2 were conducted in the United States nearly a decade ago. Rotondi et al developed and evaluated an online psychoeducational intervention with a peer forum for individuals with schizophrenia and their carers [32], and Glynn et al developed an online multifamily group program for carers of people with schizophrenia providing synchronous and asynchronous group sessions [33]. More recently, 1 study was conducted in the United Kingdom, trialing a fully Web-based psychoeducation and peer-support intervention for siblings of individuals who developed first episode psychosis (The E Sibling Project) [34]. A further psychoeducational intervention using an enriched online environment was developed and tested for acceptability in Hong Kong, China, targeting carers of people with psychosis [35]. All these interventions were highly valued by the carers, albeit in either a pilot trial or usability study with a relatively small sample size (n=21, 26, 20, and 81, respectively) [32-35]. No further definitive studies on any of the aforementioned or other eHealth interventions targeting carers for those with psychosis are available to date (of note, a peer-facilitated online educational intervention study conducted in the United Kingdom is yet to be published [36]).

Coproduction of eHealth Interventions

As eHealth interventions are for autonomous use by the end users, development methodologies commonly integrate users into the process. In the agile methodology commonly used in constructing software, iterative work sequences of technological expert–led *sprints* of developing and delivering working software are based on a brief commissioned by the clients. Through the cycles of set sprints, the developing software is shown to the clients or consumers for feedback, which is used to tune and adjust the software until its completion [37,38]. Although the agile method has been commonplace in the software development field since the 1970s, more recently, the use of participative research methods, which involve key stakeholders and end users at the core of the intervention development process, rather than brokering out the build work with a commission brief, have gained popularity [39-44].

Within mental health research, there is an established tradition of service user–led and survivor-led research, with survivor researchers consistently arguing that, for example, the *closeness* of the researcher to the enquiry increases the validity of study findings [45]. Although such an approach might help ensure that the build process responds closely to the experiences of end users in this study, the literature on, specifically, carer-led research in mental health is extremely limited and seemingly absent in the field of eHealth [46]. In contrast, there is a growing literature documenting participatory design methodologies, including, but not limited to, involving patients, carers, and the public as end users in the core of the research team and research activities, to inform the design and development of interventions [29,41,47]. Contrary to conventional interventions delivered through a face-to-face medium, where the intervention development and delivery are often driven solely by the professionals, consumers of eHealth interventions take on a much more active role [48,49]. In most eHealth interventions, including those that provide guided support from a professional or health care provider, the end users take on responsibility in initiating contact and engaging with the intervention, working through the content, and undertaking self-reported outcome measures online [19,39]. These make it paramount that the end users are involved actively in designing the intervention, not only to make sure the content meets their needs but also to ensure that the way the intervention is delivered keeps them engaged.

Overall, 2 recent systematic reviews on the topic have repeatedly identified the use of participatory research methods in eHealth intervention development as the 1 key factor in determining the acceptability and usability of most eHealth (especially electronic mental health [e-mental health]) interventions [44,50]. Among publicly funded health studies conducted in the United Kingdom, there is evidence to suggest that involving patients and the public in participatory research is positively linked to study success, in terms of recruitment and retention rates [47,51]. It can be argued that the significance of participatory research could only be amplified when it is applied in the eHealth arena where the end users assume much more direct control in using and adhering to the interventions.

The E-Support for Families and Friends of Individuals Affected by Psychosis Project

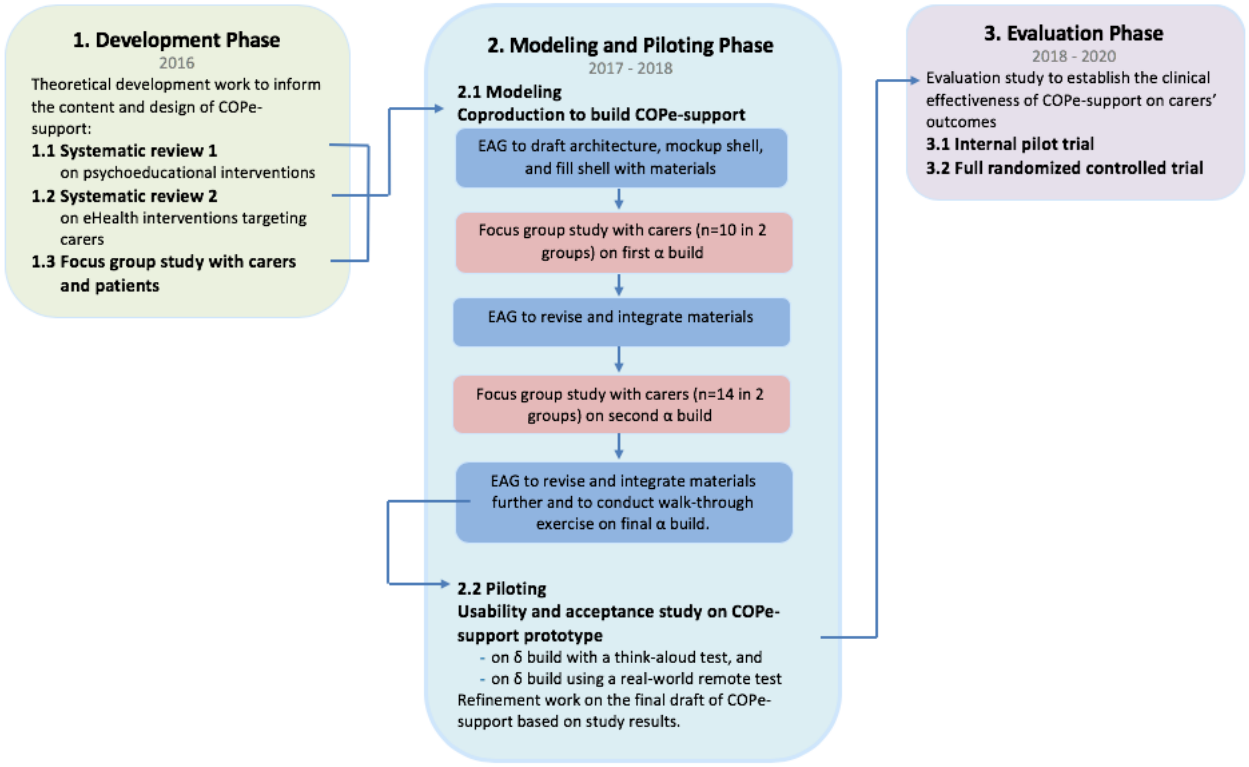
The E-support for Families and Friends of Individuals affected by Psychosis (EFFIP) Project was set up to develop and evaluate an eHealth intervention for carers supporting a relative with psychosis [52]. The overall EFFIP project lasts for 5 years spanning across the theoretical development work to the effectiveness evaluation of the end product on improving carers' health outcomes based on the Medical Research Council Complex Interventions Framework [53,54]. The design of the overall project is illustrated in Figure 1, focusing on this study.

This study reports the intervention building/modeling phase of the overall EFFIP project (see Figure 1) [53]. Before this phase, we conducted 3 studies in the theoretical development phase. These included 2 systematic reviews and a focus group study exploring research evidence and individuals with psychosis and carers' ideas and views for the optimal intervention design including essential ingredients, contact hours, and facilitation considerations ([14,19]; also JS et al, unpublished data, 2019). In the build and modeling phase of the study, reported here, we

XSL•FO
RenderX

aimed to integrate participatory research with end users alongside core research team activities to coproduce an eHealth intervention prototype to improve flexible access to high-quality psychoeducation and interactive support resources for carers [55,56]. We as a team decided to take this approach, rather than having carers lead the study, in part because of resource constraint and also because we considered that it was important that a range of consumer voices were included across all design decisions. Specific objectives of this study were to apply participatory methods with end users involvement to design and build the intervention, integrate iterative consultations with end users along the rapid prototyping work, report and consider the appropriateness of a participatory approach to eHealth intervention design, and test and refine the final draft of the intervention to get it ready for the usability evaluation, which is reported separately (JS et al, unpublished data, 2019).

Figure 1. Research methodologies used across the whole EFFIP project focusing on the intervention build and modeling phase. COPe-support: Carers fOr People with Psychosis e-support; EAG: expert advisory group; EFFIP: E-support for Families and Friends of Individuals affected by Psychosis; eHealth: electronic health.



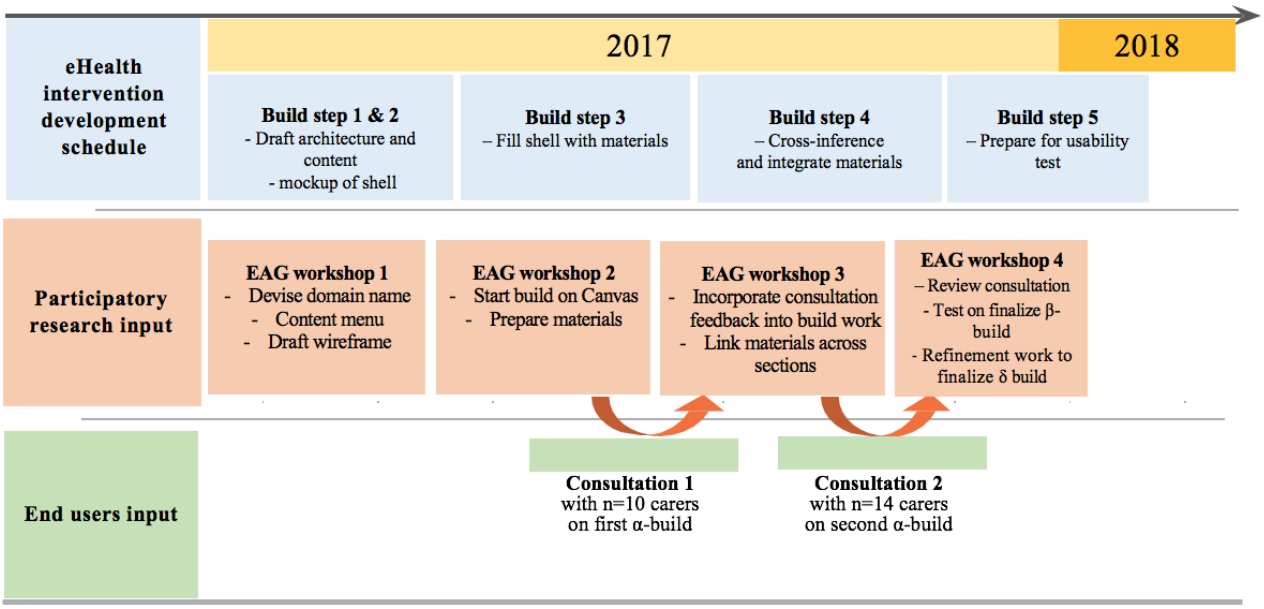
Methods

Design

The co-design and build process of the eHealth intervention followed the UK National Institute for Health Research online resource development cycle [57]. This build method was chosen as we were to develop and deliver an eHealth intervention through an existing platform, rather than developing software or the platform ourselves [37,38]. There are 5 build steps illustrated in Figure 2. These are as follows: (1) draft architecture and content, (2) mockup of shell, (3) fill shell with material, (4) cross-inference and integrate materials, and (5) prepare for piloting.

XSL•FO
RenderX

Figure 2. eHealth intervention build process integrating various inputs. EAG: expert advisory group; eHealth: electronic health.



Coproduction Workshops With an Expert Advisory Group

The co-design and build work were directed by an expert advisory group (EAG) comprising individuals and carers with lived experience of psychosis and professionals working in health, social, or voluntary sectors with the target population. The EAG membership was devised to address the principles of participatory research that value contributions from expertise through experience, and that should never be underrepresented within the bigger whole-team context [47,58]. We recruited EAG members with diverse demographic characteristics and varying degrees of ease and familiarity with digital communications from various clinical and voluntary service provider organizations across South East England. EAG members were paid a goodwill payment for meeting attendance and contributions and were also reimbursed for their travel expenses according to the INVOLVE payment scheme, a UK powerhouse that promotes patient, carers, and public involvement (PPI) in research [59].

The EAG comprised 3 individuals with lived experience of psychosis, 3 family members having different relationships with a loved one affected by psychosis in the family (including a cousin, a sister, and a daughter), 1 clinician who also has personal experience of family caregiving, and 1 voluntary service lead. The EAG worked with the core research team (comprising 2 clinical academics, 1 health services researcher, 2 eLearning experts, and 1 administrator) to lead the build work.

Over the build phase lasting 12 months, 4 EAG workshops (1-4) were held, equivalent to the sprint cycles used in the agile approach. We used the coproduction in mental health improvement work method to organize the workshops [55,56,60]. This approach emphasizes and values the hidden capacity and capability of end users, families, and significant others. At each workshop, all members assumed equal decision-making role, whereas we contributed with our respective strengths and expertise [60,61]. These participatory design workshops (and follow-up work) were mapped to fit the 5-step development process of eHealth products [57], and hence, each had specific aims and expected output, progressing from generating ideas and design to reviewing and refining product from rapid prototyping cycles.

After each participatory design workshop, the knowledge and ideas generated were translated to produce draft hand-sketched plans and wireframes, mockups of Web pages, and source materials for the intervention (eg, videos or textual information). The mockups and output produced between workshops and end users' feedback obtained from consultations (see below) were then presented at the next workshop, enabling content and broad design ideas to be critically discussed and reviewed and then further developed. Toward the end of the co-design and build work (build steps 4 and 5), the EAG undertook final development work before conducting a walk-through exercise of the online beta-build of the intervention [62-64]. Final revision and refinement work were undertaken in build step 5 to produce the intervention prototype (ie, the delta-build in Figure 2) ready for the real-world usability test.

Iterative Consultations With Carers

At the interfaces of build steps 3 and 4, we incorporated 2 rounds of iterative consultations with target end users [52]. For the consultations, we recruited carers who had no prior involvement with the EFFIP project to a focus group meeting lasting up to 2 hours. During the first consultation, we showed the carers the offline first alpha-build of the intervention and asked for their feedback on its design and likeability, the flow and readability of the content, and invited them to preempt potential usability issues and identify ways to promote end users' engagement. The second consultation followed a similar format with another set of carers, with an online second alpha-build of the intervention, which had been further developed by the EAG. The carers' views were fed back to the rapid prototyping cycles as overseen by the EAG. Figure 2 illustrates the inputs made by the focus groups.

XSL•FO
RenderX

The consultation was approved by the UK National Health Service Research Ethics Committee (REC) process (REC approval reference number: 16/LO/1300) and Health Research Authority (HRA IRAS project ID: 210571). Carers aged 18 years or older and who provided unpaid support and care for a loved one affected by psychosis were recruited from 3 mental health trusts in South London and Berkshire, South East England. Recruitment strategies for carers included the posting of study flyers at clinical areas and informing the carers support workers and mental health professionals at each trust to disseminate the study information to carers.

Analysis Strategies

All the workshops and consultation meetings were digitally recorded and transcribed verbatim. We took photos of the sketches produced at the workshops, on which the wireframes were developed later. The qualitative data were analyzed using the thematic analysis method [65,66], suiting the social constructionist and realist paradigms. We used an inductive approach to identify themes from the data without trying to fit it into a preexisting coding frame but grounded our understanding on the EAG members' and carers' perception and experiences of the evolving drafts of the intervention. As the study focused on designing and reviewing the build of the intervention, the data were analyzed descriptively to generate themes concerned broadly with the design and content of the intervention, general look and feel, usability factors, privacy and security, and ways to enhance engagement and usefulness.

Throughout the study, to ensure the rapid prototyping and build of the intervention was grounded in the data, the participatory design workshops, focus group consultation, and data analysis were performed in parallel with one another. This ensured timely and robust feedback to inform the evolving drafts of the intervention (see Figure 2).

Results

Expert Advisory Group Members and Consultation Group Participants

The EAG comprised 8 members, whereas the core team had 6 members. Over 1 year during which 4 participatory workshops were held, all EAG members and core team members attended all the workshops and selected follow-up work to create the materials as indicated.

During each round of consultations, 2 meetings involving 5 to 8 carers were held (see Figures 1 and 2). In total, 24 carers participated in the iterative consultations along the evolving build process: 10 on the first alpha-build, and 14 on the second alpha version. The participants comprised 10 men and 14 women. The range of the carers' ages was 22 to 83 years (mean 59 [standard deviation, SD 12.7]) with the median age being 61 years. Half of the carers had retired (n=9) or stayed at home being a full-time carer (n=3). The other half were in gainful employment: 5 working full time, 6 working part time, and 1 actively seeking employment. Most of the carers were a parent (19/24, 79%), and there were 2 spouses (8%, 2/24), 2 close friends (8%, 2/24), and 1 adult child (4%, 1/24). Just more than one-third of the carers (38%, 9/24) lived with their cared-for person at the time of the study. The gender mix of the cared-for persons was similar to that of the carers; 11 were male (46%, 11/24), and 13 were female (54%, 13/24). The ages of the cared-for persons ranged from 17 to 61 years (mean 35.3 [SD 15.8], median=32). In terms of diagnosis, half of the participants reported that their cared-for persons (50%, 12/24) had a diagnosis of psychosis, 9 were diagnosed with a schizophreniform disorder (45%, 9/24), and 3 had type 1 bipolar disorder (13%, 3/24). As reported by the participants, the cared-for persons had been unwell for less than 1 year to the longest for 36 years (mean 10.6 [SD=10]; median=6). The demographic characteristics and caring situation of the participants and their cared-for persons are summarized in Table 1.

XSL•FO
RenderX

Table 1. Summary of focus group consultation participant characteristics and caring situation.

Characteristics	Carer participants (n=24)	Their cared-for person (n=24)
Age		
Mean (SD[a])	59 (12.7)	35.3 (15.8)
Median (range)	61 (22-83)	32 (17-61)
Sex		
Male, n (%)	10 (41)	11 (46)
Ethnicity, n (%)		
White	18 (75)	—[b]
Black	1 (4)	—
Asian	2 (8)	—
Other	3 (13)	—
Work, n (%)		
Full-time work	5 (21)	—
Part-time work	6 (25)	—
Retired	9 (38)	—
Not working	1 (4)	—
Looking after home/family	3 (13)	—
Marital status, n (%)		
Single	1 (4)	—
Married/cohabiting	17 (71)	—
Other	6 (25)	—
Relationship with the cared-for person, n (%)		
Parent	19 (79)	—
Spouse/partner	2 (8)	—
Close friends	2 (8)	—
Child	1 (4)	—
Accommodation arrangement of carer, n (%)		
Living with cared-for person	9 (37)	—
Not living with cared-for person	15 (63)	—
Diagnosis of cared-for person, n (%)		
Psychosis	—	12 (50)
Schizophreniform disorder	—	9 (37)
Type 1 bipolar disorder	—	3 (13)

[a]SD: standard deviation.

[b]Not applicable.

Expert Advisory Group Coproduction Workshops and Iterative Consultations

Expert Advisory Group Workshop 1

The inaugural workshop was set to design the study website and the online intervention, including their (domain) name, content, and look. The EAG was provided with results obtained from prior studies conducted in the theoretical development phase of the EFFIP project ([14,19]; also JS et al, unpublished data, 2019) to aid their design decision.

Following an extended brainstorming exercise, we decided to call the eHealth intervention *COPe-support*, an acronym of Carers fOr People with Psychosis e-support. We subsequently secured the domain name of *cope-support.org* and hosted our project website on the World Wide Web. With carers as end users in mind, we set out to design a clean and user-friendly website giving information about the project and its research

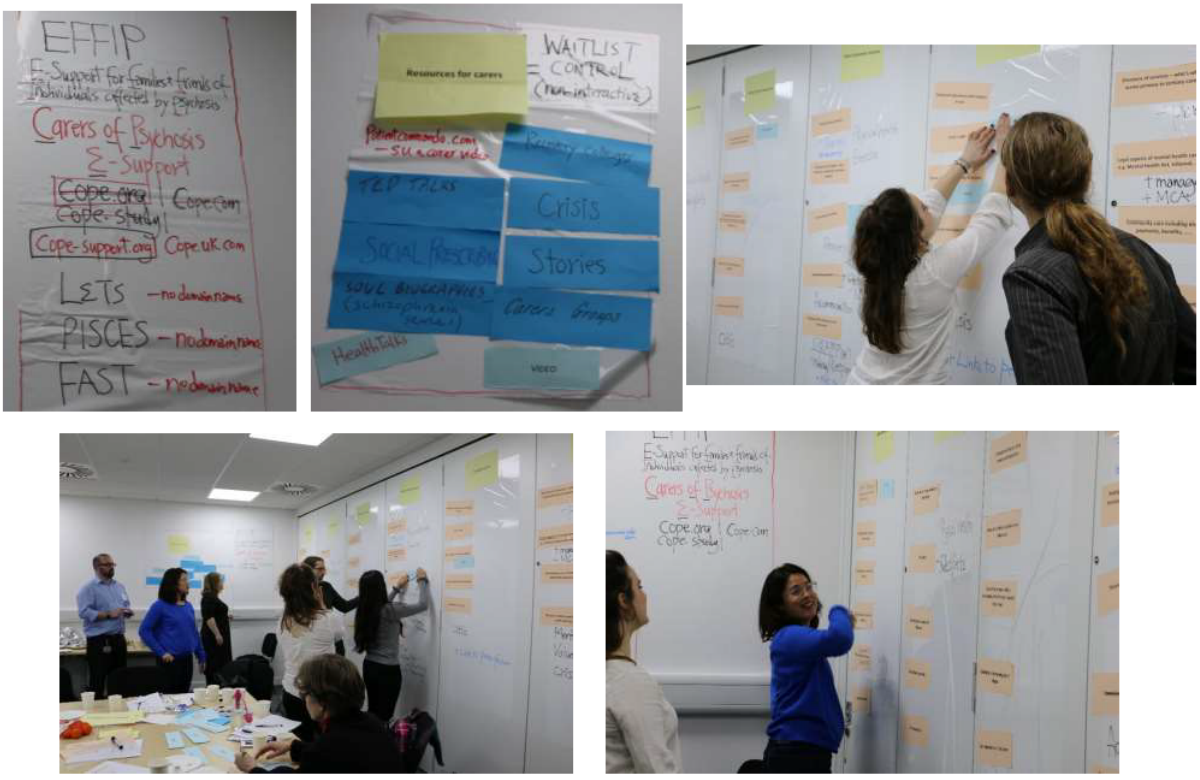
XSL•FO
RenderX

program. The EAG believed that the testimonials from other carers who helped develop the intervention would add credibility to the intervention; hence, these testimonials and photos of the EAG working together were added to the website (see Figure 3).

The EAG reviewed the essential ingredients as identified by our earlier development work (JS et al, unpublished data, 2019) to produce the *draft architecture and mockup of shell* (build step 1). We used nondigital design tools such as magic sheets, post-it notes, index-cards with content items written on them, papers, and pens to encourage a creative design atmosphere involving all members, regardless of their competency level of information and communication technology. EAG members did hand-drawn sketches of Web pages and organized the content and ingredients in a structure that they saw fit.

The EAG decided the master plan for the content and key functions of COPe-support. The intervention comprised 10 sections of psychoeducational materials, communication and problem-solving knowledge and skills, reflective exercises, and discussion points. There were 2 interactive discussion forums: one with a panel of expert members comprising professionals and experts through experience and another one with carers as peers. Figure 3 provides samples of early designs and sketches regarding the project website and the intervention.

Figure 3. Samples of early designs and sketches regarding the project website and the eHealth intervention devised by the EAG. COPe: Carers fOr People with Psychosis e-support; EAG: expert advisory group.



Expert Advisory Group Workshop 2

The second workshop was held 6 weeks later. For hosting and running COPe-support, our eLearning experts (LW and AS) identified a virtual learning environment (VLE) called Canvas [67]. The COPe-support intervention, accessible via the Canvas VLE, was designed to work on desktop (or laptop) Web browsers, as well as smartphones or tablets through a Canvas app.

Building on the framework and functions of Canvas, the EAG members worked together to *build shell with materials* at build step 2 [57]. We structured all the content items onto 1 platform hosting 12 sections. These include the following:

- Two modules on psychosis, common symptoms, and comorbid problems and evidence-based treatment for psychosis;
- Two modules on caring strategies for common symptoms and problems (eg, supporting your loved one with paranoid beliefs) and on ways to promote recovery;
- Two modules on wider social and service issues related to psychosis including ways to deal with stigma and discrimination and navigating the health and social care systems;
- Two modules focusing on well-being–promotion strategies for carers themselves;
- A virtual discussion forum and blog space for carers to share experiences and discuss commonly encountered issues;
- An *Ask the Experts* forum where participants can post questions to an expert panel comprising health and social care professionals and campaigners;
- A *Further Resources* section with supplementary Web links to relevant external resources; and

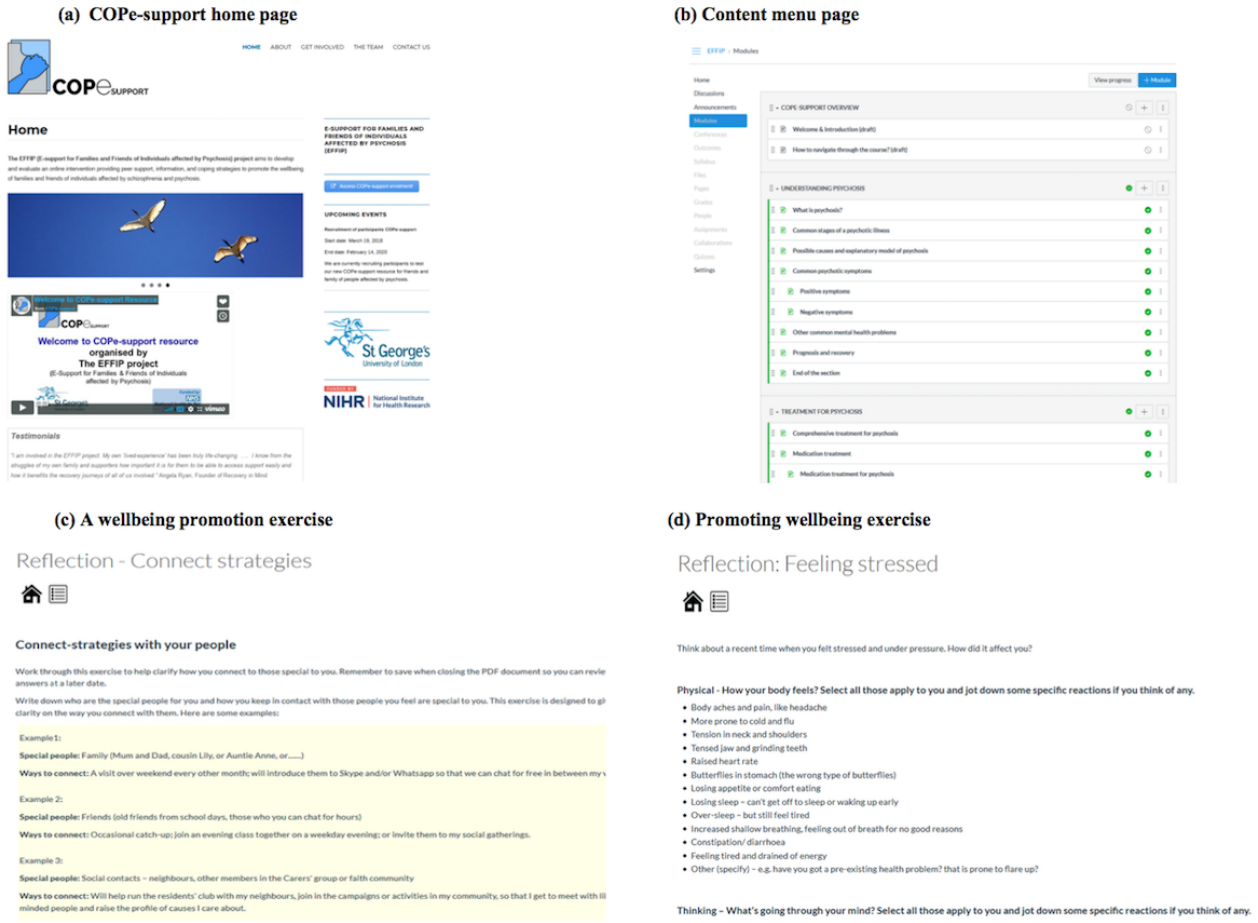
XSL•FO
RenderX

- A support page where carers can get in contact with the online facilitator for technical or emotional support directly.

During this workshop, the EAG members also discussed the best ways to present the different elements and topics. We devised a detailed work plan to source or produce the materials in various formats, ranging from textual documents to making videos with experts speaking on the specific topics and Weblinks to external sites (with explicit agreement sought).

The following 2 months after these workshops saw the EAG producing the materials and the sketch design of the intervention being turned into wireframes with materials developed filling the shell—the build step 3 [57]. Figure 4 provides screenshots of the first alpha build of COPe-support.

Figure 4. Screenshots of the first alpha-build of COPe-support. (a) COPe-support home page; (b) content menu page; (c) a well-being promotion exercise; (d) promoting well-being exercise. COPe: Carers fOr People with Psychosis e-support.



Consultation With End Users on the First Alpha-Build

Once the first alpha-build of COPe-support (including the home page where the intervention menu and content list sat and mockups of 3 modules and the 2 forums) was prepared, we ran the first consultation showing the carers its offline version to seek their feedback.

Feedback from carers indicated that the general presentation and the content of COPe-support were well received. However, carers found the linear program of sections and elements too prescriptive (see Figure 4), whereas they envisaged that carers as end users would prefer flexibility in choosing relevant content suiting their own caring situation. Although carers rated positively the differing elements including psychoeducational information, reflective exercises, and practice guides, which encourage the integration of skills learnt into their day-to-day life, they found the default terminology used by the Canvas VLE alienating. Some examples included the terms *quizzes* and *grades*. The findings obtained from the focus groups were fed back to the rapid prototyping cycle overseen by the EAG.

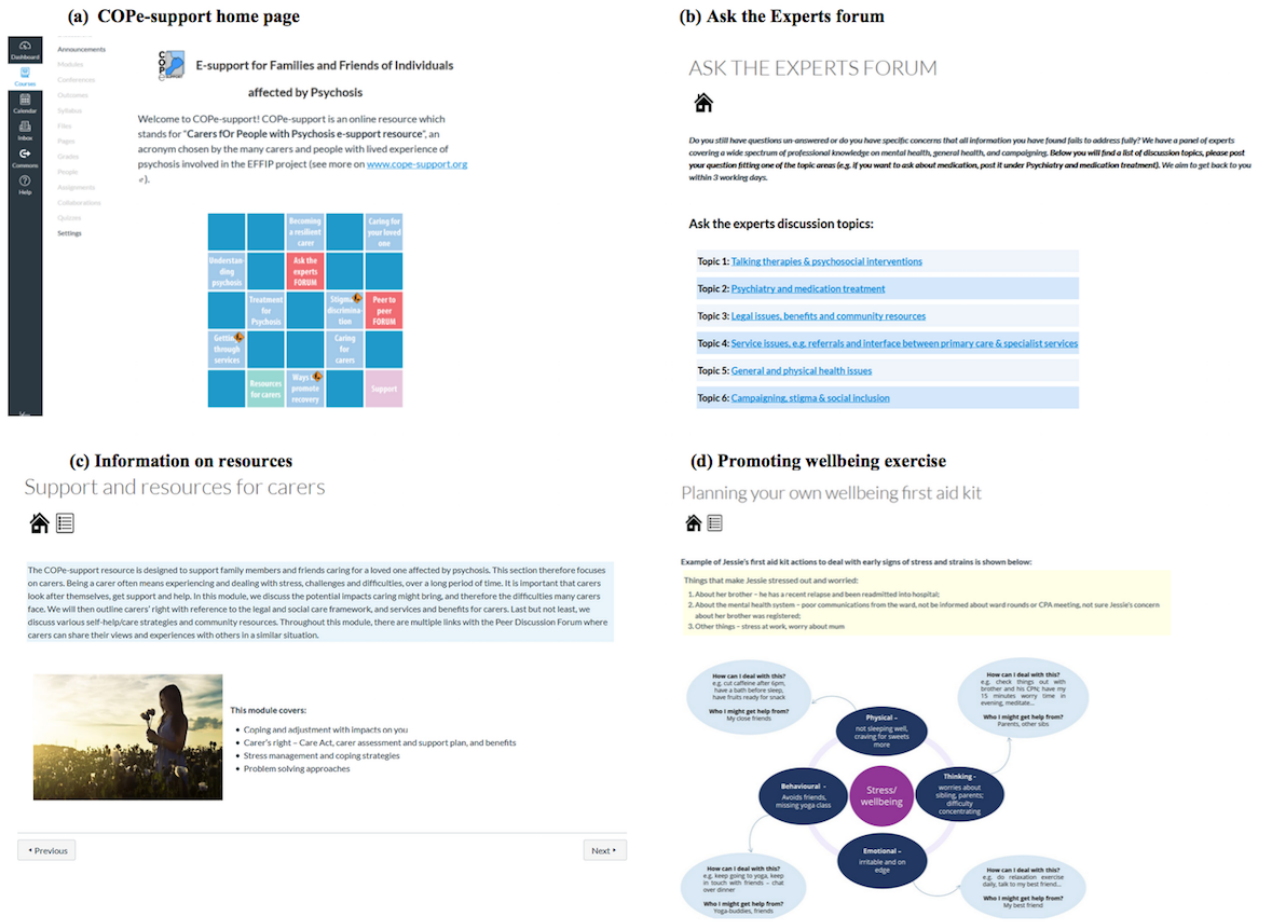
Expert Advisory Group Workshop 3

Although the rapid prototyping build work progressed onto build step 4—*cross-inferencing and integrating materials*, the consultation findings on the initial alpha-build of COPe-support were reviewed by the EAG at this workshop. The original linear program menu design was changed to a grid-based visualization, which implied no order for its content yet still provided the functionality to link to the relevant areas. The 3 key elements, that is, psychoeducational information, interactive forums, and further resources, were color-coded in the content menu. Explicit guidance notes were added to encourage participants to pick and choose the content relevant to their needs. We changed the terminology used by the Canvas VLE, which was originally designed for eLearning students, as much as possible to suit the carer population. These included labeling all the exercises and prompts for reflection and integration as *reflection* and leaving

XSL•FO
RenderX

out *submission or grading* and *sharing* of reflections through peer-to-peer forum as some participants might perceive that as a source of unintended pressure. In both forums, we added ground rules and notes to explain the privacy and security measures and provided examples in writing posts or questions without giving personally identifiable data away. We also developed further sections using these design principles in producing the second alpha-build of COPe-support. Figure 5 provides screenshots of the second alpha-build of the intervention.

Figure 5. Screenshots of the second alpha-build of COPe-support. (a) COPe-support home page; (b) ask the experts forum; (c) information on resources; (d) promoting well-being exercise. COPe: Carers fOr People with Psychosis e-support.



Consultation With End Users on the Second Alpha-Build

A second round of focus group consultation was subsequently organized when we showed the carers an online version of the second alpha-build of COPe-support (see Figure 5). Carers' feedback for this was positive in general. They found the home page welcoming and particularly appreciated the content menu design, which they found inviting and intuitive to use. The explicit display of the ground rules and guidance notes to ascertain privacy and security in the home page and again in forums was identified by carers as enhancing their sense of trust and ease of use. The layout of pages focusing on psychoeducational information given was well evaluated with the use of pictures and graphics to intercept the text. We asked carers to review the 2 forums and the discussion topics during the consultation. Most found that the topics gave some structures for the participants to focus their posts and felt confident and comfortable to use the forums with the guidance notes and the anonymous participation arrangement. Carers identified that there was a need for clearer instruction in how the participants can contact the COPe-support facilitator directly for support if necessary. They liked the cognitive behavioral–oriented reflections and exercises that promote self-care but would like more instruction on downloading the materials to allow practicing in their own time.

Expert Advisory Group Workshop 4

In the 2 months between the last consultation and the fourth coproduction workshop, further rapid prototyping work was carried out to address the focus group findings.

Workshop 4 concluded the build step 4— *cross-inference and integrating materials* —of the coproduction work and preceded the final build step 5— *ready for usability test*. During this workshop, the EAG members conducted a walk-through exercise as a group on the beta-build of COPe-support [62-64]. Through the walk-through exercise starting from *login* to *log-off*, the EAG members navigated through every section and tried out every function of COPe-support from an end user's perspective. A number of minor usability issues related to navigation and use of various functions were identified (eg, how to raise a new post?). We took note of these findings and carried out further refinement accordingly. We also captured the EAG members'

XSL•FO
RenderX

experience and feedback expressed through the exercise to inform the production of a navigation video (and a written navigation guide in parallel).

We discussed and confirmed the privacy and data security measures based on the exercise and the data output from it. In addition to the anonymous participation mechanism through the use of pseudonyms by all enrolled carers, the EAG considered it useful to state explicitly the monitoring, moderation, and facilitation provided by a qualified mental health professional (JS) daily during the week. We believed that it helped create and maintain a safe online environment as well as enhance the credibility of the intervention. The facilitator also posts weekly updates on COPe-support online forum to all carers with an aim to keep them engaged.

These considerations and final refinement work were undertaken in the final build step 5— *preparing for usability test* —with the delta-build of COPe-support developed by the end of the build and modeling phase.

Discussion

Principal Findings

Our study used an innovative approach to develop a Web-based intervention for psychosis carers. COPe-support provides psychoeducation and emotional support using health care professional contribution and peer support [52]. The participatory research design method ensured that carers as end users were fully involved in all phases of the design and build process [47,60,68,69]. We believe that carers, together with other key stakeholders such as individuals with lived experience of psychosis and health care professionals, provided some insightful foresight of target end users' expectations and usage pattern for the intervention. These were addressed and taken on board in the build process to optimize the matching of end users' expectations and needs and the final product design and delivery. Furthermore, these end users' inputs helped the core team, especially the eLearning experts, to understand how best to manipulate the technological remits as afforded by our chosen software to enhance the relevance and suitability of the intervention for the end users ([63,64]; also JS et al, unpublished data, 2019).

Our study combined 2 distinct participatory research coproduction methodologies with end users—the participatory research workshops to design and build the intervention and consultation meetings to provide feedback on drafts along the rapid prototyping cycles—in parallel and interacting with one another [29,52,58]. Furthermore, 2 different groups of end users and stakeholders were involved in the different methods and activities (carers with no involvement with the project joined the consultation meetings on initial and second alpha-builds, alongside an established set of EAG members). As we embedded the participatory research elements within a structured eHealth build model originally developed for e-mental health interventions [57], the repeated development cycle borrowed from the agile method enhanced the continuous generation of new ideas and feedback being inputted into the coproduction build process. Our approach integrating PPI into the agile build process, enabled end users to have hands-on involvement in producing and revising the developing intervention drafts; their experiential insights were not lost through interpretation by external technological experts which, we hope, the insights gained likely to increase end users' engagement with the intervention [47,60]. We would argue that our participatory approach was better suited than a carer-led method, as we were able to integrate a full range of stakeholders, including eLearning experts, clinicians, and researchers into the build process alongside carers.

Limitations

The participatory design and build work described in this paper have produced a final prototype of COPe-support. This study has produced a tangible output, that is, an eHealth intervention meeting the required and desired expectations of both EAG members and the carer participants who contributed to the iterative consultation alongside the prototype development. However, our results are descriptive in nature and limited in establishing the usability or acceptability of the intervention, and our sample of carers was limited by size and representativeness. The success of our approach will ultimately be tested in experimental evaluation of the intervention itself (see below).

Implications and Future Directions

This study illustrates the importance of coproduction of COPe-support, an eHealth intervention, which is designed to be used by carers autonomously. Another important contribution this study made to research in the field is the documentation of a rigorous and innovative build process, which combined intervention development method and participatory research methodologies throughout its life cycle [69]. Across the spectrum of eHealth intervention build methods, we recognize our participatory research method being in the middle ground of the 2 polar approaches: agile or technologist-driven and carer-led. Our approach establishes itself as a third way adopting the agile process and principles while integrating PPI with the technological and research core team conducting the build work as directed by the EAG.

Following this study with the prototype ready for feasibility and usability testing, a usability study of the final delta-build of COPe-support has been completed with both a remote usability trial and a think-aloud study (JS et al, unpublished data, 2019). The results were promising and provided further feedback from end users, in terms of facilitation and delivery strategies (JS et al, unpublished data, 2019). Following further refinement work as informed by the usability study (JS et al, unpublished data, 2019), COPe-support is currently being tested for its effectiveness in supporting carers through an online randomized controlled trial (RCT) [54]. The trial will also test levels of carers' engagement both with the intervention and the study and for any correlation of engagement or adherence with the effects on carers' health outcomes. These investigations will therefore enable us to consider, in the future, if our participatory research approach has been successful.

To the best of our knowledge, COPe-support is one of the few comprehensive Web-based interventions targeting carers for

XSL•FO
RenderX

psychosis patients. With the benefits of access, facilitation, and delivery completely through the internet, our product has the potential to provide an evidence-based psychoeducational intervention with its key ingredients [14,19]. These include providing health outcomes monitoring, psychoeducational information, and real-time interaction with health care professionals and other carers as peers. We expect the internet delivery will overcome some of the implementation and access barriers from both the service providers' and carers' perspective [19]. Dependent on the trial outcomes, we have considered several areas for further development and innovation in the future. These range from investigating further implementation strategies to optimize large-scale rolling out of COPe-support upon positive trial results to examining and incorporating additional health behavioral change techniques or methods used in eHealth interventions to enhance engagement and effects upon unclear effectiveness results [70,71]. Participatory research methods and principles will, no doubt, be an important approach integrated within such work.

Conclusions

We integrated participatory research methodologies with a structured eHealth intervention development process to develop COPe-support through this study. COPe-support is one of the few eHealth interventions dedicated for family carers of individuals affected by psychosis. It provides information and psychosocial support for carers through the internet, promoting flexible access and individualized choice. Following usability evaluation of the intervention prototype, we are currently undertaking an online RCT to evaluate its effectiveness in promoting carers' health outcomes.

Acknowledgments

The authors thank all the input and contributions from the expert advisory group members who helped design COPe-support. They are Ellen Harris, Jacqueline Marks, Angela Ryan, Storm Ryan, Lana Samuels, Dr Clive Travis, Leigh Wallbank, and Dr Elen Williams [72]. The authors are grateful to the carers who participated in the consultations and supported the development of COPe-support. For recruiting participants into the study, the project team acknowledges the support of the National Institute for Health Research (NIHR) through the Clinical Research Network (Division 4).

JS is funded by a National Institute for Health Research (NIHR) Post-Doctoral Research Fellowship for this research project. This paper represents independent research. The views expressed are those of the author(s) and not necessarily those of the National Health Service, the NIHR, or the Department of Health and Social Care.

Conflicts of Interest

None declared.

References

1. Shahly V, Chatterji S, Gruber MJ, Al-Hamzawi A, Alonso J, Andrade LH, et al. Cross-national differences in the prevalence and correlates of burden among older family caregivers in the World Health Organization World Mental Health (WMH) Surveys. Psychol Med 2013 Apr;43(4):865-879 [FREE Full text] [doi: 10.1017/S0033291712001468] [Medline: 22877824]
2. Smith L, Onwumere J, Craig T, McManus S, Bebbington P, Kuipers E. Mental and physical illness in caregivers: results from an English national survey sample. Br J Psychiatry 2014 Sep;205(3):197-203. [doi: 10.1192/bjp.bp.112.125369] [Medline: 25061119]
3. National Alliance for Caregiving. 2009. Caregiving in the US URL: http://www.caregiving.org/data/Caregiving_in_the_US_2009_full_report.pdf [accessed 2019-02-12] [WebCite Cache ID 77PTHQIpJ]
4. Rethink Mental Illness. 2012. The Abandoned Illness: A Report by the Schizophrenia Commission URL: https://www.rethink.org/media/2303/tsc_main_report_14_nov.pdf
5. NICE. London, UK: The National Institute for Health and Care Excellence; 2014. Psychosis and schizophrenia in adults: prevention and management URL: https://www.nice.org.uk/guidance/cg178
6. Smith J, Birchwood MJ. Relatives and patients as partners in the management of schizophrenia. The development of a service model. Br J Psychiatry 1990 May;156:654-660. [doi: 10.1192/bjp.156.5.654] [Medline: 2095942]
7. Sin J, Moone N, Wellman N. Developing services for the carers of young adults with early-onset psychosis--listening to their experiences and needs. J Psychiatr Ment Health Nurs 2005 Oct;12(5):589-597. [doi: 10.1111/j.1365-2850.2005.00883.x] [Medline: 16164510]
8. Sin J, Moone N, Harris P, Scully E, Wellman N. Understanding the experiences and service needs of siblings of individuals with first-episode psychosis: a phenomenological study. Early Interv Psychiatry 2012 Feb;6(1):53-59. [doi: 10.1111/j.1751-7893.2011.00300.x] [Medline: 21952020]
9. Carers Trust. 2014. Key Facts About Carers and the People They Care For URL: http://www.carers.org/key-facts-about-carers [accessed 2019-02-12]
10. Sin J, Murrells T, Spain D, Norman I, Henderson C. Wellbeing, mental health knowledge and caregiving experiences of siblings of people with psychosis, compared to their peers and parents: an exploratory study. Soc Psychiatry Psychiatr Epidemiol 2016 Dec;51(9):1247-1255 [FREE Full text] [doi: 10.1007/s00127-016-1222-7] [Medline: 27121259]

XSL•FO
RenderX

11. Sherifali D, Ali MU, Ploeg J, Markle-Reid M, Valaitis R, Bartholomew A, et al. Impact of internet-based interventions on caregiver mental health: systematic review and meta-analysis. J Med Internet Res 2018 Dec 3;20(7):e10668 [FREE Full text] [doi: 10.2196/10668] [Medline: 29970358]
12. Pharoah F, Mari J, Rathbone J, Wong W. Family intervention for schizophrenia. Cochrane Database Syst Rev 2010 Dec 8(12):CD000088 [FREE Full text] [doi: 10.1002/14651858.CD000088.pub2] [Medline: 21154340]
13. Cooper C, Blanchard M, Selwood A, Walker Z, Livingston G. Family carers' distress and abusive behaviour: longitudinal study. Br J Psychiatry 2010 Jun;196(6):480-485. [doi: 10.1192/bjp.bp.109.071811] [Medline: 20513860]
14. Sin J, Gillard S, Spain D, Cornelius V, Chen T, Henderson C. Effectiveness of psychoeducational interventions for family carers of people with psychosis: a systematic review and meta-analysis. Clin Psychol Rev 2017 Aug;56:13-24 [FREE Full text] [doi: 10.1016/j.cpr.2017.05.002] [Medline: 28578249]
15. Beaumont J, Thomas JH. Measuring National Well-Being: Health, 2012. London, United Kingdom: Office for National Statistics; 2012. URL: https://www.bl.uk/collection-items/measuring-national-wellbeing-health-2012
16. NHS Digital. 2011. Health Survey for England - 2010, Respiratory Health URL: https://digital.nhs.uk/data-and-information/publications/statistical/health-survey-for-england/health-survey-for-england-2010-respiratory-health
17. Kuipers E, Raune D. The early development of expressed emotion and burden in the families of first onset psychosis. In: Birchwood MJ, Fowler D, Jackson C, editors. Early Intervention in Psychosis: A Guide to Concepts, Evidence and Interventions. Chichester, England: John Wiley and Sons; 2000:128-140.
18. Sin J, Norman I. Psychoeducational interventions for family members of people with schizophrenia: a mixed-method systematic review. J Clin Psychiatry 2013 Dec;74(12):e1145-e1162 [FREE Full text] [doi: 10.4088/JCP.12r08308] [Medline: 24434103]
19. Sin J, Henderson C, Spain D, Cornelius V, Chen T, Gillard S. eHealth interventions for family carers of people with long term illness: a promising approach? Clin Psychol Rev 2018 Dec;60:109-125 [FREE Full text] [doi: 10.1016/j.cpr.2018.01.008] [Medline: 29429856]
20. Yesufu-Udechuku A, Harrison B, Mayo-Wilson E, Young N, Woodhams P, Shiers D, et al. Interventions to improve the experience of caring for people with severe mental illness: systematic review and meta-analysis. Br J Psychiatry 2015 Apr;206(4):268-274. [doi: 10.1192/bjp.bp.114.147561] [Medline: 25833867]
21. Lazarus RS, Folkman S. Stress, Appraisal and Coping. New York: Springer Publishing Company; 1984.
22. Szmukler GI, Burgess P, Herrman H, Benson A, Colusa S, Bloch S. Caring for relatives with serious mental illness: the development of the experience of caregiving inventory. Soc Psychiatry Psychiatr Epidemiol 1996 Jun;31(3-4):137-148. [doi: 10.1007/BF00785760] [Medline: 8766459]
23. Joyce J, Leese M, Kuipers E, Szmukler G, Harris T, Staples E. Evaluating a model of caregiving for people with psychosis. Soc Psychiatry Psychiatr Epidemiol 2003 Apr;38(4):189-195. [doi: 10.1007/s00127-003-0618-3] [Medline: 12664229]
24. Szmukler G. From family 'burden' to caregiving. Psychiatr Bull 1996;20(8):449-451. [doi: 10.1192/pb.20.8.449]
25. Rotondi AJ, Haas GL, Anderson CM, Newhill CE, Spring MB, Ganguli R, et al. A clinical trial to test the feasibility of a telehealth psychoeducational intervention for persons with schizophrenia and their families: intervention and 3-month findings. Rehabil Psychol 2005 Nov;50(4):325-336 [FREE Full text] [doi: 10.1037/0090-5550.50.4.325] [Medline: 26321774]
26. Rotondi AJ. Schizophrenia. In: Cucciare MA, Weingardt KR, editors. Using Technology to Support Evidence-Based Behavioural Health Practices: A Clinician's Guide. New York: Routledge; 2010:69-90.
27. Schook RM, Linssen C, Schramel FM, Festen J, Lammers E, Smit EF, et al. Why do patients and caregivers seek answers from the internet and online lung specialists? A qualitative study. J Med Internet Res 2014 Feb 4;16(2):e37 [FREE Full text] [doi: 10.2196/jmir.2842] [Medline: 24496139]
28. Powell J, Chiu T, Eysenbach G. A systematic review of networked technologies supporting carers of people with dementia. J Telemed Telecare 2008;14(3):154-156. [doi: 10.1258/jtt.2008.003018] [Medline: 18430288]
29. Vaughan C, Trail TE, Mahmud A, Dellva S, Tanielian T, Friedman E. Informal caregivers' experiences and perceptions of a web-based peer support network: mixed-methods study. J Med Internet Res 2018 Dec 28;20(8):e257 [FREE Full text] [doi: 10.2196/jmir.9895] [Medline: 30154074]
30. Chi NC, Demiris G. A systematic review of telehealth tools and interventions to support family caregivers. J Telemed Telecare 2015 Jan;21(1):37-44 [FREE Full text] [doi: 10.1177/1357633X14562734] [Medline: 25475220]
31. Boots LM, de Vugt ME, van Knippenberg RJ, Kempen GI, Verhey FR. A systematic review of internet-based supportive interventions for caregivers of patients with dementia. Int J Geriatr Psychiatry 2014 Apr;29(4):331-344. [doi: 10.1002/gps.4016] [Medline: 23963684]
32. Rotondi AJ, Anderson CM, Haas GL, Eack SM, Spring MB, Ganguli R, et al. Web-based psychoeducational intervention for persons with schizophrenia and their supporters: one-year outcomes. Psychiatr Serv 2010 Nov;61(11):1099-1105. [doi: 10.1176/ps.2010.61.11.1099] [Medline: 21041348]
33. Glynn SM, Randolph ET, Garrick T, Lui A. A proof of concept trial of an online psychoeducational program for relatives of both veterans and civilians living with schizophrenia. Psychiatr Rehabil J 2010;33(4):278-287. [doi: 10.2975/33.4.2010.278.287] [Medline: 20374986]
34. Sin J, Henderson C, Norman I. Usability of online psychoeducation for siblings of people with psychosis. Int J Technol Assess Health Care 2014 Oct;30(4):374-380. [doi: 10.1017/S0266462314000488] [Medline: 25394550]

XSL•FO
RenderX

35. Chan SK, Tse S, Sit HL, Hui CL, Lee EH, Chang WC, et al. Web-based psychoeducation program for caregivers of first-episode of psychosis: an experience of Chinese population in Hong Kong. Front Psychol 2016;7:2006 [FREE Full text] [doi: 10.3389/fpsyg.2016.02006] [Medline: 28082935]
36. Lobban F, Robinson H, Appelbe D, Barraclough J, Bedson E, Collinge L, et al. Protocol for an online randomised controlled trial to evaluate the clinical and cost-effectiveness of a peer-supported self-management intervention for relatives of people with psychosis or bipolar disorder: relatives education and coping toolkit (REACT). BMJ Open 2017 Dec 18;7(7):e016965 [FREE Full text] [doi: 10.1136/bmjopen-2017-016965] [Medline: 28720617]
37. Robertson EG, Wakefield CE, Cohn RJ, O'Brien T, Ziegler DS, Fardell JE. The development of Delta: using Agile to develop a decision aid for pediatric oncology clinical trial enrollment. JMIR Res Protoc 2018 May 4;7(5):e119 [FREE Full text] [doi: 10.2196/resprot.9258] [Medline: 29728347]
38. Sadasivam RS, Delaughter K, Crenshaw K, Sobko HJ, Williams JH, Coley HL, et al. Development of an interactive, web-delivered system to increase provider-patient engagement in smoking cessation. J Med Internet Res 2011 Oct 18;13(4):e87 [FREE Full text] [doi: 10.2196/jmir.1721] [Medline: 22011394]
39. Adenuga OA, Kekwaletswe RM, Coleman A. eHealth integration and interoperability issues: towards a solution through enterprise architecture. Health Inf Sci Syst 2015;3:1 [FREE Full text] [doi: 10.1186/s13755-015-0009-7] [Medline: 26167279]
40. Barry MJ, Edgman-Levitan S. Shared decision making—The pinnacle of patient-centered care. N Engl J Med 2012 Mar 1;366(9):780-781. [doi: 10.1056/NEJMp1109283] [Medline: 22375967]
41. Riper H, Andersson G, Christensen H, Cuijpers P, Lange A, Eysenbach G. Theme issue on e-mental health: a growing field in internet research. J Med Internet Res 2010 Dec 19;12(5):e74 [FREE Full text] [doi: 10.2196/jmir.1713] [Medline: 21169177]
42. Muller MJ, Kuhn S. Participatory design. Commun ACM 1993;36(6):24-28. [doi: 10.1145/153571.255960]
43. Sjöberg C, Timpka T. Participatory design of information systems in health care. J Am Med Inform Assoc 1998;5(2):177-183 [FREE Full text] [doi: 10.1136/jamia.1998.0050177] [Medline: 9524350]
44. Orlowski SK, Lawn S, Venning A, Winsall M, Jones GM, Wyld K, et al. Participatory research as one piece of the puzzle: a systematic review of consumer involvement in design of technology-based youth mental health and well-being interventions. JMIR Hum Factors 2015 Jul 9;2(2):e12 [FREE Full text] [doi: 10.2196/humanfactors.4361] [Medline: 27025279]
45. Rose D. Service user/survivor-led research in mental health: epistemological possibilities. Disabil Soc 2017 May 16;32(6):773-789. [doi: 10.1080/09687599.2017.1320270]
46. Richardson E, Laird SE. Involving carers from minority ethnic backgrounds in carer-led research. Disabil Soc 2012 Jul 12;28(1):67-80. [doi: 10.1080/09687599.2012.695525]
47. Staley K. National Institute for Health Research. 2013. A Series of Case Studies Illustrating the Impact of Service User and Carer Involvement on Research URL: https://www.nihr.ac.uk/nihr-in-your-area/mental-health/documents/Impact-of-service-user-and-carer-involvement%20PDF.pdf
48. Schmidt U, Wykes T. e-Mental health - a land of unlimited possibilities. J Ment Health 2012 Aug;21(4):327-331. [doi: 10.3109/09638237.2012.705930] [Medline: 22823092]
49. Trivedi P, Wykes T. From passive subjects to equal partners: qualitative review of user involvement in research. Br J Psychiatry 2002 Dec;181:468-472. [doi: 10.1192/bjp.181.6.468] [Medline: 12456515]
50. Simblett S, Greer B, Matcham F, Curtis H, Polhemus A, Ferrão J, et al. Barriers to and facilitators of engagement with remote measurement technology for managing health: systematic review and content analysis of findings. J Med Internet Res 2018 Dec 12;20(7):e10480 [FREE Full text] [doi: 10.2196/10480] [Medline: 30001997]
51. Ennis L, Wykes T. Impact of patient involvement in mental health research: longitudinal study. Br J Psychiatry 2013 Nov;203(5):381-386. [doi: 10.1192/bjp.bp.112.119818] [Medline: 24029538]
52. Sin J, Henderson C, Sese A, Woodham L, Krishnamoothy K, Gillard S. European Academy of Nursing Science Summer Conference 2018: Leadership in Nursing: challenges for the future—EFFIP (E-support for Families & Friends of Individuals affected by Psychosis): Developing & evaluating an online complex intervention for carers. In: BMC Nursing. 2018 Presented at: EANS'18; July 11-12, 2018; Ghent, Belgium p. 34. [doi: 10.1186/s12912-018-0301-3]
53. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M. ResearchGate. 2008. Developing and Evaluating Complex Interventions: New Guidance URL: https://tinyurl.com/y6c98kke
54. ISRCTN Registry. 2018. Randomised Controlled Trial of COPe-Support Online Resource for Carers URL: http://www.isrctn.com/ISRCTN89563420 [accessed 2019-02-12]
55. Cahn ES. No More Throw-Away People: The Co-Production Imperative. Washington DC: Essential Books; 2001.
56. Needham C, Carr S. Social Care Institute for Excellence. 2009. SCIE Research Briefing 31: Co-Production - An Emerging Evidence Base for Adult Social Care Transformation URL: https://www.scie.org.uk/publications/briefings/briefing31/
57. Pinfold V, Duggan A, Huxley P, Tomlin A, Rapaport J. NETSCC. 2010. The Development of an Online Training Resource for Mental Health Professionals to Involve Carers in Information Sharing: Report for the National Institute for Health Research Service Delivery and Organisation Programme URL: http://www.netscc.ac.uk/hsdr/files/project/SDO_FR_08-1711-160_V01.pdf
58. Green SA, Carnegie W, Barber S, Matthew D, Matthews R. Co-designing interventions within quality improvement initiatives: notes from the field. J Health Des 2018 Mar 22;3(1):49-56. [doi: 10.21853/JHD.2018.38]

XSL•FO
RenderX

59. INVOLVE. 2018. Guidance on Co-Producing a Research Project URL: https://www.invo.org.uk/wp-content/uploads/2019/04/Copro_Guidance_Feb19.pdf
60. Lewis A, King T, Herbert L, Repper J. ImROC - Implementing Recovery Through Organisational Change. 2018. Co-Production – Sharing Our Experiences, Reflecting On Our Learning URL: http://imroc.org/wp-content/uploads/2017/10/ImROC-co-pro-briefing-FINAL-4.pdf
61. Baines RL, de Bere SR. Optimizing patient and public involvement (PPI): identifying its 'essential' and 'desirable' principles using a systematic review and modified Delphi methodology. Health Expect 2018 Dec;21(1):327-335 [FREE Full text] [doi: 10.1111/hex.12618] [Medline: 28929554]
62. Boren T, Ramey J. Thinking aloud: reconciling theory and practice. IEEE Trans Profess Commun 2000;43(3):261-278. [doi: 10.1109/47.867942]
63. Nielsen J. Usability Engineering. Boston, MA: Academic Press; 1993.
64. Nielsen J. Nielsen Norman Group. 1994. 10 Usability Heuristics for User Interface Design URL: https://www.nngroup.com/articles/ten-usability-heuristics/ [accessed 2019-02-12]
65. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
66. Ritchie J, Lewis J, editors. Qualitative Research Practice: A Guide for Social Science Students and Researchers. Second Edition. Los Angeles: Sage Publications; 2014.
67. Canvas the Learning Management Platform. 2019. URL: http://www.canvasvle.co.uk [accessed 2019-02-12] [WebCite Cache ID 77PUdfSO7]
68. Blackburn H, Hanley B, Staley K. Turning the Pyramid Upside Down: Examples of Public Involvement in Social Care Research. Eastleigh, England: INVOLVE; 2010.
69. Hickey G, Richards T, Sheehy J. Co-production from proposal to paper. Nature 2018 Dec;562(7725):29-31. [doi: 10.1038/d41586-018-06861-9] [Medline: 30283117]
70. Webb TL, Joseph J, Yardley L, Michie S. Using the internet to promote health behavior change: a systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. J Med Internet Res 2010 Feb 17;12(1):e4 [FREE Full text] [doi: 10.2196/jmir.1376] [Medline: 20164043]
71. Eysenbach G, Powell J, Englesakis M, Rizo C, Stern A. Health related virtual communities and electronic support groups: systematic review of the effects of online peer to peer interactions. Br Med J 2004 May 15;328(7449):1166 [FREE Full text] [doi: 10.1136/bmj.328.7449.1166] [Medline: 15142921]
72. COPe-Support. URL: http://cope-support.org/team/

Abbreviations

COPe-support: Carers fOr People with Psychosis e-support
EAG: expert advisory group
EFFIP: E-support for Families and Friends of Individuals affected by Psychosis
eHealth: electronic health
e-mental health: electronic mental health
NIHR: National Institute for Health Research
PPI: patient, carers, and public involvement
RCT: randomized controlled trial
REC: Research Ethics Committee
VLE: virtual learning environment

Edited by G Eysenbach; submitted 12.04.19; peer-reviewed by J Amann, N Miyoshi, O(Danilina; comments to author 14.06.19; revised version received 28.06.19; accepted 28.06.19; published 06.08.19.

Please cite as:
Sin J, Henderson C, Woodham LA, Sesé Hernández A, Gillard S
A Multicomponent eHealth Intervention for Family Carers for People Affected by Psychosis: A Coproduced Design and Build Study
J Med Internet Res 2019;21(8):e14374
URL: https://www.jmir.org/2019/8/e14374/
doi:10.2196/14374
PMID:31389333

©Jacqueline Sin, Claire Henderson, Luke A Woodham, Aurora Sesé Hernández, Steve Gillard. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 06.08.2019. This is an open-access article distributed under the terms of the

XSL•FO
RenderX

Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Medium-Term Effects of a Tailored Web-Based Parenting Intervention to Reduce Adolescent Risk of Depression and Anxiety: 12-Month Findings From a Randomized Controlled Trial

Marie Bee Hui Yap[1,2], MPsych, PhD; Mairead C Cardamone-Breen[1], DPsych; Ronald M Rapee[3], PhD; Katherine A Lawrence[1], DPsych; Andrew J Mackinnon[4], PhD; Shireen Mahtani[1], DPsych; Anthony F Jorm[2], PhD, DSc

[1]School of Psychological Sciences and Turner Institute for Brain and Mental Health, Monash University, Melbourne, Australia

[2]Melbourne School of Population and Global Health, University of Melbourne, Melbourne, Australia

[3]Centre for Emotional Health, Macquarie University, Sydney, Australia

[4]Black Dog Institute, University of New South Wales, Sydney, Australia

Corresponding Author:
Marie Bee Hui Yap, MPsych, PhD
School of Psychological Sciences and Turner Institute for Brain and Mental Health
Monash University
18 Innovation Walk
Wellington Road, Clayton
Melbourne, 3168
Australia
Phone: 61 0399051250
Email: marie.yap@monash.edu

Related Article:

This is a corrected version. See correction statement: https://www.jmir.org/2019/8/e15915/

Abstract

Background: Prevention of depression and anxiety disorders early in life is a global health priority. Evidence on risk and protective factors for youth internalizing disorders indicates that the family represents a strategic setting to target preventive efforts. Despite this evidence base, there is a lack of accessible, cost-effective preventive programs for parents of adolescents. To address this gap, we recently developed the Partners in Parenting (PiP) program—an individually tailored Web-based parenting program targeting evidence-based parenting risk and protective factors for adolescent depression and anxiety disorders. We previously reported the postintervention outcomes of a single-blinded parallel-group superiority randomized controlled trial (RCT) in which PiP was found to significantly improve self-reported parenting compared with an active-control condition (educational factsheets).

Objective: This study aimed to evaluate the effects of the PiP program on parenting risk and protective factors and symptoms of adolescent depression and anxiety using data from the final assessment time point (12-month follow-up) of this RCT.

Methods: Parents (n=359) and adolescents (n=332) were recruited primarily from secondary schools and completed Web-based assessments of parenting and adolescent depression and anxiety symptoms at baseline, postintervention (3 months later), and 12-month follow-up (317 parents, 287 adolescents). Parents in the PiP intervention condition received personalized feedback about their parenting and were recommended a series of up to 9 interactive modules. Control group parents received access to 5 educational factsheets about adolescent development and mental health. Both groups received a weekly 5-min phone call to encourage progress through their program.

Results: Intervention group parents completed an average of 73.7% of their intended program. For the primary outcome of parent-reported parenting, the intervention group showed significantly greater improvement from baseline to 12-month follow-up compared with controls, with a medium effect size (Cohen *d*=0.51; 95% CI 0.30 to 0.72). When transformed data were used, greater reduction in parent-reported adolescent depressive symptoms was observed in the intervention group (Cohen *d*=−0.21; 95% CI −0.42 to −0.01). Mediation analyses revealed that these effects were mediated by improvements in parenting (indirect effect *b*=−0.08; 95% CI −0.16 to −0.01). No other significant intervention effects were found for adolescent-reported parenting

XSL•FO
RenderX

or adolescent depression or anxiety symptoms. Both groups showed significant reductions in anxiety (both reporters) and depressive (parent reported) symptoms.

Conclusions: PiP improved self-reported parenting for up to 9 months postintervention, but its effects on adolescent symptoms were less conclusive, and parent-reported changes were not perceived by adolescents. Nonetheless, given its scalability, PiP may be a useful low-cost, sustainable program to empower parents of adolescents.

Trial Registration: Australian Clinical Trials Registration Number (ACTRN): 12615000328572; http://www.anzctr.org.au/ACTRN12615000328572.aspx (Archived by WebCite at http://www.webcitation.org/6qgsZ3Aqj).

(J Med Internet Res 2019;21(8):e13628) doi:10.2196/13628

KEYWORDS

family; parenting; mental health; depression; anxiety; adolescent; internet; randomized controlled trial; preventive health services

Introduction

Background

Depression and anxiety disorders are common in young people, with lifetime prevalence rates of 18% and 38%, respectively, in adolescents aged 13 to 17 years [1]. The incidence of these disorders peaks during adolescence, and early-onset disorders tend to have a chronic and relapsing nature. In particular, these disorders forecast a cascade of deleterious long-term sequelae across multiple domains of functioning and increase suicide risk [2,3]. Moreover, a large proportion of the burden of disease from these disorders remains unavertable even with optimal treatment [4]. With emerging evidence suggesting an increase in the rates of depression and anxiety problems in children and young people internationally [5,6], there is an urgent need for effective preventive approaches to stem this global public health problem.

The family setting is a strategic target for implementing preventive approaches for adolescent depression and anxiety (also known as internalizing) disorders. As posited by interpersonal theories of developmental psychopathology, internalizing problems both result from and contribute to disruptions in developmentally salient interpersonal processes (starting with early parent-infant attachment), which in turn interfere with young people's need for relatedness [7]. From an etiological perspective, parents have an important influence on young people's risk for internalizing problems, in terms of both nature and nurture. Genetic research suggests that the heritability of liability to internalizing behaviors is high in 3-year-olds (76%) but reduces to 48% by the age of 12 years, whereas the shared environmental influence increases from zero at age 3 to 18% at age 12 [8]. More recently, a children-of-twins study found significant environmental transmission of anxiety from parents to their adolescents, but no evidence of significant genetic transmission [9], highlighting the influence of parental anxiety and associated parenting behaviors (eg, overprotection or overcontrol) in the etiology of adolescent internalizing problems. Meta-analyses of individual parenting behaviors associated with adolescent internalizing problems have found that parenting behaviors prospectively account for a small but significant amount of variance (1%-16%) [10,11]. Together, the evidence indicates that parenting behaviors are a promising target for the prevention of adolescent internalizing problems.

For the purpose of prevention, interventions need to target modifiable risk and protective factors [12]. A burgeoning body of literature has identified various risk and protective factors for adolescent depression and anxiety problems [13,14], including some that are potentially modifiable by parents [10,11,15]. These factors are posited to operate bidirectionally in the transactions between parent and child, especially as the child develops increasing autonomy during adolescence [16]. Factors that increase adolescents' risk for depression and anxiety include interparental conflict, overinvolvement (including psychological control), and aversiveness (including harsh parental criticism and parent-adolescent conflict). Protective factors include parental warmth and acceptance, monitoring, and autonomy granting [10,11]. This evidence base delineates the parenting factors that should be translated into preventive interventions for parents of adolescents to reduce the societal burden of youth internalizing disorders.

For decades, preventive parenting interventions have been developed to capitalize on the influence parents have on their child's development and adjustment based on the assumption that improving parenting will in turn yield benefits for the child's mental health [17]. Unfortunately, the translation of research evidence into preventive parenting interventions continues to lag far behind the abovementioned evidence base, with a recent review identifying only 3 preventive parenting interventions (defined as programs where more than 50% of the intervention is delivered to the parent) targeting parents of adolescents [18]. One of these 3 interventions used a universal prevention approach (invited all parents regardless of their child's level of risk; [19]), whereas the other two were selective prevention programs that targeted parents with an affective disorder [20] or HIV/AIDS [21]. In contrast, relatively more preventive parenting interventions have been developed and evaluated for parents of younger children [18]. Most existing interventions that are designed for parents of adolescents target behavioral problems not directly related to internalizing disorders, such as externalizing problems, substance use, and risky behaviors [17,22]. Importantly, preventive parenting interventions have demonstrated long-term benefits for child internalizing [18] and externalizing [17,22] outcomes that last up to 20 years after the intervention. Sandler et al [17] proposed a theoretical pathway by which parenting programs can have long-term benefits for child outcomes via program effects on parents. The most parsimonious program effect involves parents learning new skills from the program, and the use of these skills

XSL•FO
RenderX

is maintained by positive responses from their children (eg, improved parent-child relationship).

However, the public health impact of preventive parenting interventions (regardless of child age) is limited by poor uptake and engagement [23]. In part, this is because most existing interventions are face-to-face group programs, which encounter the common barriers of stigma and practical logistics such as timing or scheduling, cost, travel, and child care [24,25]. Web-based platforms have the potential to overcome some of these barriers because of the anonymity and accessibility they afford [26]. However, a recent systematic review of technology-assisted parenting programs to prevent mental health problems in children aged 0 to 18 years [27] identified only 1 Web-based preventive parenting intervention that targets adolescent internalizing problems, known as PiP.

The PiP Web-based parenting intervention was developed to address the abovementioned gaps in preventive parenting resources [26]. It is an evidence-informed intervention that incorporates (1) developmental theory (including the developmental psychopathology framework) [28] and research into the role of parents in adolescent development and adjustment [10,15], (2) preventive medicine and public health approaches that advocate the targeting of risk and protective factors for prevention [12,29], and (3) the use of persuasive technologies to influence behavior change [30]. PiP draws its content from the parenting guidelines *How to prevent depression and clinical anxiety in your teenager: Strategies for parents* (henceforth referred to as the Guidelines) [31]. These Guidelines were the product of a rigorous research translation methodology comprising a systematic review of modifiable parental factors associated with adolescent depression and anxiety [10] and a Delphi study of international expert consensus about parenting strategies that can reduce adolescents' risk of depression and anxiety disorders [32]. Using a consumer-engagement approach [33], the intervention was designed following the principles of Persuasive Systems Design [30] to be an interactive individually tailored program. The intervention has been evaluated in a randomized controlled trial (RCT) and found to produce significantly greater improvements in parent-reported parenting risk and protective factors for adolescent depression and anxiety (primary outcome) from baseline to postintervention (3 months later) compared with an active control condition (Cohen d=0.57). No significant group differences in changes over time were found for secondary outcomes of interest, including adolescent-reported parenting factors, and adolescent depression and anxiety symptoms, as reported by both parents and adolescents [34]. It is likely that changes on these secondary outcomes may emerge in the longer term, once the proximal intervention effects (eg, behavior changes in parents) have had time to influence the broader family system.

Indeed, in the broader parenting and family intervention literature, proponents of the developmental cascade model have argued that changing parent-related factors in the short term can lead to significant long-term benefits for the child through a progression of events over the course of development. For instance, an RCT of the New Beginnings Program for divorced families with children aged 9 to 12 years [35] found significant effects on parenting, parent-child relationship quality, and internalizing problems at postintervention, but these effects were not maintained at 6-month follow-up. However, significant long-term effects emerged at the 6-year [35] and 15-year [36] follow-up assessments on various functioning outcomes, including internalizing, externalizing, and substance use problems, and mediational analyses supported the cascade effects model. Similarly, an RCT of the Family Check-Up program for low-income parents of toddlers [37] found significant direct effects on maternal depression but not child internalizing problems in children aged 2 to 3 years, but an indirect effect (through reductions in maternal depression) on child internalizing symptoms emerged in middle childhood (age 7.5-8.5 years). These findings underscore the importance of long-term follow-up to examine the preventive effects of parenting interventions over the course of the child's development.

This Study

This study reports the findings from a medium-term (12-month) follow-up of families in this RCT. One primary outcome of interest was again parenting risk and protective factors. We hypothesized that the intervention effects observed at postintervention would be observed at the 12-month follow-up. Parent- and adolescent-reported symptoms were also examined as primary outcomes in this paper with the aim of investigating whether the intervention effects on parent-reported parenting factors would yield benefits in terms of adolescent depressive and anxiety symptoms by the 12-month follow-up. Specifically, we hypothesized that compared with the control group, the intervention group would show greater reductions in parent- and adolescent-reported symptoms from baseline to 12-month follow-up. We also hypothesized that parenting at postintervention would mediate adolescent symptoms at the 12-month follow-up, after accounting for parenting and symptom scores at baseline. Adolescent report of parenting was again examined as a secondary outcome measure. We predicted greater improvement in adolescent-reported parenting from baseline to 12-months follow-up in the intervention compared with the control group.

Methods

Design

This study was a parallel-group superiority RCT, with assessments conducted at baseline (preintervention), postintervention (3-months postbaseline), and 12-month follow-up (final assessment timepoint). The trial was prospectively registered with the Australian New Zealand Clinical Trials Registry (registration number ANZCTR12615000328572) and approved by the Monash University Human Research Ethics Committee (CF14/3887-2014002024). A detailed description of the methodology has been published by Yap et al [34]; however, the pertinent details will be described below.

Sample Size

The sample size was determined based on an a priori power analysis. This indicated that a sample size of 294 parent-adolescent dyads (147 per group) was required to detect

XSL•FO
RenderX

a small effect size (Cohen *d*=0.20), with an alpha level of .05, power of 0.80, and a repeated-measures design. To allow for approximately 15% attrition, we aimed to recruit 338 dyads (169 per group).

Settings, Participants, and Eligibility Criteria

Eligible parents had an adolescent in the target age range (12-15 years at baseline), resided in Australia, had regular internet access, and an email account. Only 1 parent-adolescent dyad per family was eligible to participate. Computer and internet literacy were implicit eligibility criteria. Given the universal approach taken in this trial, no exclusion criteria were specified. Recruitment was primarily via secondary schools across Australia as well as online networks, social media, and mental health organizations (eg, *beyondblue* and Mental Health First Aid Australia). Baseline assessments were completed between August 2015 and November 2016 (when the desired sample size was reached), and 12-month follow-up data collection concluded in December 2017.

Interventions

The Partners in Parenting Intervention

PiP is a fully automated Web-based parenting program consisting of 3 components [26]. First, parents complete a self-assessment scale (the Parenting to Reduce Adolescent Depression and Anxiety Scale; PRADAS [38]), which assesses their current parenting practices against the Guidelines. Second, based on their responses to the PRADAS, parents receive an individually tailored feedback report outlining their parenting strengths and areas for improvement. The feedback messages are designed to be brief, motivate behavior change, provide practical parenting strategies, and contain links to further information that parents can access if desired. Third, parents are recommended a series of interactive online modules, also based on their responses to the PRADAS. A total of 9 modules are available, and parents can further personalize their program by selecting additional modules (not initially recommended to them) or declining recommended modules. Modules include interactive activities, goal setting exercises, audio clips, vignettes, illustrations, and an end-of-module quiz with immediate feedback, designed to consolidate learning. Each module takes approximately 15 to 25 min to complete. One module is made available to parents every 7 days, to allow sufficient time to complete each module and work on weekly goals before progressing to the next module. Parents were sent automated emails each week to notify them that their next module was available to access via their personalized dashboard. Once parents had completed the initially selected modules, all 9 modules (including those not initially selected) were made available for the remainder of the RCT. Table 1 presents the content covered in the PiP modules and the corresponding sections in the PRADAS, feedback report, and Guidelines. Multimedia Appendix 1 presents screenshots of the intervention.

XSL•FO
RenderX

Table 1. Partners in Parenting modules, corresponding sections of the Parenting to Reduce Adolescent Depression and Anxiety Scale (PRADAS) and feedback report and Guidelines subheadings.

Module title and content	Corresponding section of the PRADAS and feedback report	Guidelines subheading[a]
Module: "Connect"—Acknowledges the challenge of connecting with adolescent children and provides specific tips on how to do this.	Your relationship with your teenager	Establish and maintain a good relationship with your teenager
Module: "Nurture roots and inspire wings"—Helps parents establish the important balance between staying involved and interested in their adolescent's life, while encouraging increasing age-appropriate autonomy.	Your involvement in your teenager's life	Be involved and support increasing autonomy
Module: "Good friends, supportive relationships"—Provides strategies for parents to support their adolescent's social skills development.	Your teenager's relationships with others	Encourage supportive relationships
Module: "Raising good kids into great adults: establishing family rules"—Highlights the importance of consistent and clear boundaries for adolescent behaviors and provides specific strategies to establish these.	Your family rules	Establish family rules and consequences
Module: "Calm versus Conflict"—Addresses the need for adaptive conflict management between parents and between parent and adolescent and provides specific strategies to do these.	Your home environment	Minimize conflict in the home
Module: "Good health habits for good mental health"—Provides strategies to help parents encourage good health habits in their adolescent, including a healthy diet, physical activity, good sleep habits, and abstinence from alcohol and drugs.	Health habits	Encourage good health habits
Module: "Partners in problem solving"—Provides strategies for parents to help their adolescent develop good problem-solving and stress management skills.	Dealing with problems in your teenager's life	Help your teenager to deal with problems
Module: "From surviving to thriving: helping your teenager deal with anxiety"—Provides strategies for parents to help their adolescent manage their everyday anxiety.	Coping with anxiety	Help your teenager to deal with anxiety
Module: "When things aren't okay: getting professional help"—Helps parents understand what depression and anxiety problems can look like in adolescents, and what they can do if their adolescent is or becomes unwell.	Getting help when needed	Encourage professional help seeking when needed

[a]Adapted from [34]. Note that 2 of the 11 sections of the Guidelines (*You can reduce your child's risk of developing depression and clinical anxiety* and *Don't blame yourself*) do not have specific corresponding sections in the PRADAS or PiP modules, but the key messages they present are included in the feedback report and across all modules.

Educational Factsheets (Active Control Condition)

Parents in the control group received access to a series of 5 educational factsheets about adolescent development and mental health. The factsheets provided general information, without individual tailoring or actionable parenting strategies (cf. the PiP intervention). The factsheets were intended to provide information already available to parents as part of a current health promotion approach, with materials adapted from the *Raising Children Network* website [39]. The factsheet topics were as follows: (1) Teen development: an overview, (2) The teenager's developing brain, (3) The teenager's changing body, (4) Resilience, and (5) Happy teenagers and teenage wellbeing. The delivery of the factsheets was designed to mirror the delivery of the PiP intervention; parents were emailed once per week with a link to access their next factsheet via their personal dashboard on the trial website, and they had access to the factsheets for the duration of the RCT. The weekly emails occurred at the beginning of the intervention phase (ie, first 5 weeks postbaseline).

Weekly Check-In Phone Calls

Parents in both groups received a weekly phone call from a member of the research team, commencing 7 days after completion of their baseline assessment. Intervention group parents received 1 phone call per module selected in their program, unless they selected less than 5 modules, in which case, parents received a minimum of 5 calls (to match the control group who received 5 calls). The purpose of the calls was to encourage progress, enhance engagement, provide technical assistance, and answer study-related questions. Research assistants were trained to make the phone calls following a standard script and did not provide individual advice or therapeutic support.

Measures

Primary Outcome Measure 1: The Parenting to Reduce Adolescent Depression and Anxiety Scale

The PRADAS is a self-reported measure of parenting practices across the parenting domains covered in the Guidelines [38] (see Table 1). As a criterion-referenced measure, the PRADAS assesses current parenting practices against specific recommendations in the Guidelines (the *criterion*) and scores parents as either concordant (1) or nonconcordant (0) with the recommendations. The 73 items of the PRADAS cover 8 of the 9 domains of the Guidelines, with one of the original 9 subscales (*relationships with others*) dropped from the scale during the validation process [38]. Most items are scored on a Likert-type frequency (never, rarely, sometimes, and often) or likelihood scale (very unlikely, unlikely, likely, and very likely; for hypothetical scenarios). The item scores are summed to form a total score, ranging from 0 to 73, with higher scores indicating

XSL•FO
RenderX

greater concordance with the Guidelines. In a validation study of 711 parents of adolescents aged 12 to 15 years, which included baseline data from the current RCT sample, the total score demonstrated high reliability (agreement coefficient=0.97), acceptable 1-month test-retest reliability (0.78), and convergent validity with 2 existing parenting measures [38]. In the current sample, the agreement coefficient was high at all 3 time points (baseline=0.97, postintervention=0.96, and 12-month follow-up=0.95).

Primary Outcome Measure 2: Short Mood and Feelings Questionnaire

The Short Mood and Feelings Questionnaire (SMFQ) is a widely used measure of depressive symptoms in children and adolescents [40]. Both the child (SMFQ-C) and parent (SMFQ-P) versions have 13 items assessing the frequency of depressive symptoms in the past 2 weeks, on a 3-point scale of *not true* (0), *sometimes true* (1), or *true* (2). Item scores are summed to form a total score, ranging from 0 to 26, with higher scores indicative of higher symptom levels. The scale has been documented to have high internal consistency, criterion validity, and convergent validity with other measures of depressive symptoms in children [40-42]. In our sample, both the parent- and child-reported versions had high reliability at all time points, as assessed by coefficient omega (SMFQ-C: baseline omega=0.93, postintervention omega=0.93, 12-month omega=0.95, SMFQ-P baseline omega=0.93, postintervention omega=0.92, 12-month omega=0.94). The correlations between parent and child reports on the SMFQ were r=0.48 at baseline, r=0.40 at postintervention, and r=0.53 at 12-month follow-up (all P<.001).

Primary Outcome Measure 3: Spence Children's Anxiety Scale

The Spence Children's Anxiety Scale (SCAS) is a 39-item child- (SCAS-C) and parent- (SCAS-P) reported measure of child and adolescent anxiety across 6 subscales: separation anxiety, social anxiety, obsessive compulsive symptoms, panic or agoraphobia, generalized anxiety, and fear of physical injury [43,44]. Items are scored on a 4-point scale from 0 (never) to 3 (always) and can be summed to form the 6 subscale scores and a total anxiety score. We calculated the total score, which ranges from 0 to 114, with higher scores representing more anxiety symptoms. The SCAS has been normed on several samples, including Australian school children within the same age range as our sample [45]. Both the parent- and child-reported versions have demonstrated high internal consistency reliability and acceptable test-retest reliability [44-46]. Internal consistency reliability for our sample was high for both versions across the 3 time points (SCAS-C: baseline omega=0.95, postintervention omega=0.96, 12-month omega=0.96, SCAS-P: baseline omega=0.93, postintervention omega=0.94, 12-month ω=0.95). The correlations between SCAS-C and SCAS-P were r=0.46 at baseline, r=0.43 at postintervention, and r=0.47 at 12-month follow-up (all P<.001).

Secondary Outcome Measure: The Parenting to Reduce Adolescent Depression and Anxiety Scale—Adolescent Report

The Parenting to Reduce Adolescent Depression and Anxiety Scale—Adolescent Report (PRADAS-A) assesses the adolescent's perspective on the same 8 parenting domains covered in the PRADAS (Cardamone-Breen et al, forthcoming). The scale has fewer items than the PRADAS (total of 43 items), as only items that were developmentally appropriate and could be assessed from the adolescent's perspective were included. During validation analyses, the *relationships with others* subscale was also removed from the PRADAS-A because of poor psychometric properties (Cardamone-Breen et al, forthcoming). Response options and scoring are similar to the PRADAS, with items assessed on Likert-type scales and scored as either concordant (1) or nonconcordant (0) with the Guidelines' recommendations. The total score therefore ranges from 0 to 43, with higher scores indicating greater Guidelines-concordant parenting practices. The total score has demonstrated high reliability (agreement coefficient=.97, 3-month test-retest reliability=.81) and moderate correlations with adolescent-reported depression and anxiety symptoms in a validation sample of 670 adolescents aged 12 to 15 years (baseline data from the current RCT was included in the validation study [Cardamone-Breen et al, forthcoming]). The agreement coefficient for the total score in our sample was high (0.97) at all 3 time points. The correlation between the PRADAS and PRADAS-A at each time point was as follows: baseline, r=0.26; postintervention, r=0.26; and 12-month follow-up, r=0.33 (all P<.01).

Intervention Adherence, Completion, and Access During Follow-Up Period

Intervention adherence was operationalized as the percentage of parents who completed their program as intended, calculated as 100% × [(number of parents whose observed usage equals their intended usage)/(number of parents who received the intervention)] [34,47]. Intervention completion was defined as the percentage of the intended program that was completed, that is, intervention completion=[100% × (observed usage)/(intended usage)]. For the intervention group, observed usage was defined as the number of modules completed, and intended usage as the number of modules initially selected. For the control group, intended usage was defined as reading all 5 factsheets, and observed usage was defined as the number of factsheets that had been opened by the parent (determined by timestamps stored in the system when parents clicked the link to open their factsheet). We also examined whether parents completed their program during the active intervention phase, which was defined as the time between the parent baseline assessment and the adolescent postintervention. If the adolescent did not complete the assessment, the date of the parent postintervention assessment was used. Finally, we examined the number of parents who accessed their program after the active intervention phase (ie, between postintervention and 12-month assessment time points).

XSL•FO
RenderX

Procedures

Registration and Baseline Assessments

Parent participants registered themselves and their adolescent via the dedicated trial website and provided consent and contact details for their adolescent (see Multimedia Appendix 2 for informed consent documentation). Email verification was required at this point. A member of the research team then phoned the adolescent to inform them of the study requirements and obtain assent (if they agreed to participate). Parents were not informed of their adolescent's decision to participate and could continue in the study regardless of adolescent participation. If the adolescent declined to take part, the adolescent assessments were cancelled so that parent participation could proceed as per protocol. Adolescents who agreed to participate were guided through completion of their online baseline assessment over the phone, with the researcher providing assistance as required. On completion of the adolescent baseline assessment (or cancellation of the assessment by a researcher), the trial website automatically generated an email invitation to the parent, inviting them to complete their baseline assessment. Parents were then required to log on to the website in their own time to complete their baseline assessment. Parent and adolescent assessments included their respective versions of the PRADAS, SCAS, and SMFQ.

Randomization and Blinding

Immediately following completion of the parent baseline assessment, parents were automatically allocated to the intervention or control condition using a computer-generated unblocked, unstratified randomization procedure, with a 1:1 allocation ratio. At this point, the intervention group parents were presented with their individually tailored feedback onscreen and were emailed a PDF copy of the feedback report. Control group parents were presented with their first factsheet. Therefore, parents were not blinded to their allocation nor were the researchers who spoke to parents during weekly check-in phone calls. Adolescents were not informed of their parent's allocation and, therefore, were assumed to be blinded. As all assessments were completed online via the dedicated trial website, blinding of assessor was not relevant.

Follow-Up Assessments

The procedure for 3- and 12-month follow-up assessments was similar to the baseline procedure. Both parents and adolescents were reimbursed with an Aus $15 electronic voucher for completion of each of the 3- and 12-month follow-up assessments.

Adolescent Symptom Elevation Procedure

At all time points, the participants were followed up by a member of the research team if both the parent and adolescent reported elevated symptoms on the SCAS or SMFQ, based on predetermined cutoff scores. For the SCAS, this was defined as a total score greater than or equal to 1.5 SDs above the mean based on Australian community sample norms [48]. For the SMFQ, scores greater than or equal to 8 were considered elevated [40]. Follow-up actions included email notifications to parents alerting them to their adolescent's elevated symptoms and providing avenues for seeking professional support (n=38 at baseline, n=25 at postintervention, and n=28 at 12-month follow-up; no significant differences between groups). Adolescents who reported particularly high scores on the SMFQ (SMFQ-C total score >20) were also phoned by a postgraduate clinical psychology student, who conducted a risk assessment and provided referral information as required (n=8 at baseline, n=5 at postintervention, and n=10 at 12-month follow-up; no significant differences between groups).

Statistical Methods

Less than 4% of participants had missing data on any measures. Item level missing data were replaced with the participant's mean response on the corresponding subscale for cases with less than 23% missing data on a given measure. This is considered an appropriate method of imputation for this amount of missing data [49]. For cases with greater than 23% of missing items on a measure, the measure was considered missing entirely and excluded from analyses.

Analyses were conducted in SPSS version 25 (IBM Corp), with an a priori alpha level of .05. To assess for potential attrition biases, we compared participant characteristics and scores on baseline and postintervention outcome measures between those who completed 12-month follow-up assessments and those who did not. Independent samples *t* tests (for continuous variables) and chi-square analyses (for categorical variables) were used for these analyses. We also examined potential group differences in follow-up actions taken for adolescents who reported elevated symptoms at postintervention.

Primary and secondary outcome analyses were conducted on an intention-to-treat basis, using mixed-model repeated measures (MMRM), with an unstructured covariance matrix. MMRM uses all the available data from all the participants, including those who withdrew from follow-up assessments [50]. It is a preferred analytic approach for repeated-measures designs when data are considered missing at random or missing completely at random [50,51]. As our hypotheses related to change from baseline to 12-month follow-up, we specified planned contrast tests of the group × measurement-occasion interaction from baseline to 12-month follow-up, within the overall group × measurement-occasion mixed model. This was the primary result of interest. Pairwise comparisons between groups at trial endpoint (12-month follow-up) were also examined. Cohen *d* effect sizes with 95% CIs are reported for all analyses.

Owing to the positive skew of model residuals for the SCAS and SMFQ at all occasions, a square-root transformation was applied, which improved the distribution of residuals. All analyses were repeated using the transformed data to check robustness of the results. For most analyses, the results did not change with transformation. When the results did differ, the overall conclusions were made by considering the findings based on both raw and transformed data. For ease of interpretation, raw data have been plotted, with footnotes to indicate where results differed with transformation.

To assess for potential mediation of intervention effects on adolescent symptoms by change in parenting, we conducted simple mediation analyses using the PROCESS macro for SPSS [52,53]. Separate mediation models were run for each symptom

XSL•FO
RenderX

measure (ie, outcome variable), with 5000 bootstrap samples for bias-corrected bootstrap 95% CIs. In each model, group (coded as 0= control, 1=intervention) was entered as the predictor variable, 12-month symptom measure score as the outcome variable, and postintervention PRADAS score as the mediator variable. Baseline PRADAS and baseline symptom measure score (corresponding to the outcome variable) were entered as covariates.

Finally, we conducted post hoc moderation analyses to explore moderation of intervention effects of adolescent age, gender, and baseline symptoms on outcomes. For age and baseline symptoms, the continuous moderator variable (ie, child age at registration or baseline symptom score) was added to the mixed model as a covariate, including a 3-way interaction term (ie, group × measurement-occasion × moderator) whose significance constituted a test of a differential effect of the moderator on outcome of the intervention. Significant moderation effects were interpreted using estimated marginal means plotted for values of the covariate (ie, moderator variable) at its 25th and 75th percentile at baseline in the sample. To minimize shared method variance effects for moderation analyses using baseline symptom measures, we used the symptom measure reported by the opposite informant as the moderator variable (eg, for the outcome of SMFQ-P, baseline SMFQ-C score was used as the moderator). For the parenting measures, we conducted 2 moderation analyses, with each of the symptom measures (SCAS and SMFQ, opposite informant to outcome measure) entered in separate models. Moderation by child gender was assessed in a similar manner using gender as an additional factor rather than as a covariate.

Results

Sample Characteristics

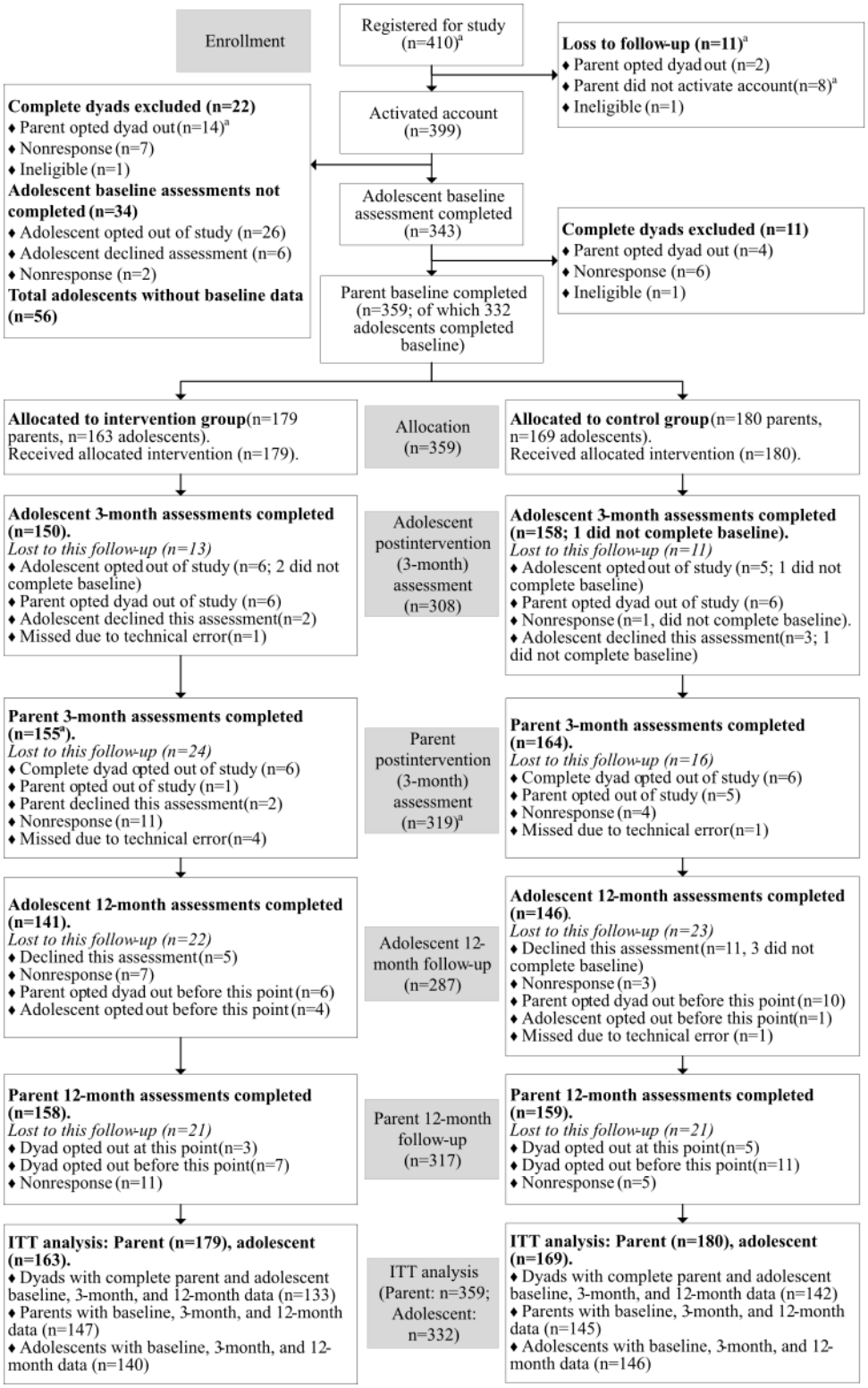
The sample's characteristics have been reported in detail elsewhere [34] and included in Multimedia Appendix 3. Parent participants were predominantly female (87.2%), were married or in a de facto relationship (76.6%), were employed full or part time (86.6%), spoke English as their primary language at home (84.1%), were from an intact family situation (70.5%), and had tertiary-level education (58.2%). Parents had a mean age of 45.15 years (SD 5.20) and their adolescents (55.4% male) had a mean age of 13.68 years (SD 1.06). In addition, 59.1% of the parents reported having a current or past history of mental illness, whereas less than a quarter of adolescents were reported by their parents to have a current (18.9%) or past (15.9%) mental health diagnosis.

Attrition

As shown in Figure 1, of the 359 parent participants who completed the baseline assessments and were randomized, 319 completed postintervention assessments and 317 went on to complete 12-month follow-up (intervention group n=158, control group n=159). The number of parents reported to complete postintervention differs from the original RCT outcome paper (previously reported as n=318) because of an error detected when preparing the 12-month data. A parent in the intervention group was excluded from the original paper because of missing individual item response data (missing 9 items [12.3%] of the postintervention PRADAS); however, in this paper, the missing items were imputed, allowing this participant to be included in the analyses. For adolescent participants, 332 completed baseline assessments, 308 completed postintervention assessments, and 287 completed 12-month follow-up. Therefore, the attrition rate at 12-month follow-up was 11.7% for parents and 13.6% for adolescents. This did not differ between conditions for parents (11.7% in each group) or adolescents (intervention group: 13.5%; control group: 13.6%). Figure 1 presents the participant flow diagram.

XSL•FO
RenderX

Figure 1. Consolidated Standards of Reporting Trials (CONSORT) participant flow diagram. (a) These numbers differ from those published in the postintervention paper because of errors detected when preparing the 12-month follow-up data.



We examined demographic characteristics and scores on baseline and postintervention measures between participants who completed 12-month follow-up and those who did not. There were no differences on any baseline measures between completers and noncompleters nor were there differences in parent or child demographic characteristics, with the exception of parent education level. Parents who completed the 12-month assessment were more likely to have tertiary-level education than those who did not (χ^2_1, [N=359]=9.9; *P*=.002). In addition, adolescents of parents who completed the 12-month follow-up had significantly higher PRADAS-A scores (mean 23.94, SD 6.20) at postintervention compared with adolescents of parents who did not complete 12-month follow-up (mean 21.05, SD 6.84; $t_{(306)}$=2.05; *P*=.04).

XSL•FO
RenderX

Time Interval Between Assessments

The mean time interval between parent baseline and 12-month follow-up was 385.71 days (SD 19.16, median 380.04, range 361-487 days), and the mean interval between parent postintervention and 12-month assessments was 266.66 days (SD 35.43, median 273.92, range 86-366 days). For adolescent assessments, the mean interval between baseline and 12-month follow-up was 376.59 days (SD 13.19, median 371.98, range 362-452) and between postintervention and 12-month follow-up was 266.41 days (SD 27.91, median 273.97, range 87-315). As reported in the postintervention paper [34], the wide range in time intervals was because of a technical error resulting in delayed postintervention assessments for 59 dyads. The time intervals did not differ significantly between groups (all $P>.05$). There was an average interval of 12.14 days between completion of the adolescent 12-month assessment and the parent 12-month assessment (SD 13.18, median 7.16, range 0-57). At 12-month follow-up, dyads in the control group had a significantly shorter interval between completion of adolescent and parent assessments (control group: mean 9.94 days, SD 12.22; intervention group: mean 14.47 days, SD 13.80; $t_{(265.54)}=-2.88$; $P=.004$).

Intervention Completion and Adherence

As reported previously, the mean intended program use by the intervention group (n=179) was 6.85 out of a possible 9 modules. The mean observed usage by the intervention group was 5.17 modules [34]. At the time of data extraction for this paper, parents in the intervention group had completed an average of 73.7% of their selected program. Intervention adherence in the intervention group was 44.1% (n=79 parents whose observed usage equaled their intended usage). During the follow-up period (from 3-12 months postbaseline), 12 of the 179 intervention-group parents (6.7%) accessed a mean of 2 modules (range 1-8, SD 2). In the control group, 33 of the 180 parents (18.3%) accessed a mean of 2 factsheets (range 1-4, SD 1.20).

Primary Outcome Measure 1: Parenting to Reduce Adolescent Depression and Anxiety Scale

Observed scores for all outcome measures at each measurement occasion are presented in Multimedia Appendix 4. Table 2 displays the planned contrast results of the group × measurement-occasion interaction from baseline to 12-month follow-up for all primary and secondary outcome analyses. As shown in Table 2, there was a significant group-by-time interaction for the PRADAS, $t_{328.17}=4.81$, $P<.001$, with the intervention group showing a significantly greater increase in PRADAS scores from baseline to 12-month follow-up compared with controls (see Figure 2). The effect size of the interaction was medium (Cohen $d=0.51$; 95% CI 0.30 to 0.72). Pairwise comparisons of the 2 groups at 12-month follow-up also indicated a significant group difference ($F_{1,347.77}=4.11$; $P=.04$), although the effect size was small (Cohen $d=0.21$; 95% CI −0.01 to 0.43).

Primary Outcome Measures 2 and 3: Adolescent Anxiety and Depression Symptoms

As shown in Table 2 and Figure 3, there was a significant group x measurement occasion interaction from baseline to 12-month follow-up for the SMFQ-P, when transformed data were used (Cohen $d=-0.21$; see Multimedia Appendix 5 for MMRM results on transformed data). However, when analyses were run on raw data, there was no significant interaction for any of the symptom measures (all $P>.05$). Using transformed data, there was a significant main effect for time over the 3 measurement occasions for the SCAS-P, SCAS-C, and SMFQ-P (all $P<.001$) but not for the SMFQ-C ($P=.12$). When raw data were used, there was a significant main effect for time for all symptom measures (SCAS-P and SMFQ-P, $P<.001$; SCAS-C, $P=.001$; SMFQ-C, $P=.009$; see Figure 3). For the parent-reported measures, both groups showed a significant reduction from baseline to postintervention, after which the groups appeared to diverge from postintervention to 12-month follow-up (see Figure 3). On the SMFQ-P, the increase from postintervention to 12-month follow-up was significant for the control group (only on raw data), whereas the intervention group remained stable (no significant difference between 3 and 12 months). The reduction from baseline to 12-month follow-up was significant for both groups on the SCAS-P, as well as for the intervention group on the SMFQ-P and the control group on the SCAS-C. When transformed, the reduction from baseline to 12-month follow-up was also significant for the control group on the SMFQ-P and the intervention group on the SCAS-C. On the SCAS-C, the control group demonstrated a significant reduction from baseline to postintervention (with both raw and transformed data). Pairwise comparisons indicated no significant group differences on any of the symptom measures at any time point, with both raw and transformed data (all $P>.05$).

XSL•FO
RenderX

Table 2. Mixed-model repeated measures planned contrast test of group × measurement-occasion interaction from baseline to 12-month follow-up for all primary and secondary outcome measures.

Outcome measure	Estimated marginal means (SE)		t test *(df)*[a]	*P* value	$d_{interaction}$ (95% CI)[b]	$d_{12\ months}$ (95% CI)[c]
	Intervention	Control				
Parenting to Reduce Adolescent Depression and Anxiety Scale						
Baseline	46.58 (0.57)	47.88 (0.57)	—[d]	—	—	—
12 months	51.68 (0.59)	49.99 (0.61)	4.81 (328.17)	<.001	0.51 (0.29 to 0.71)	0.21 (−0.01 to 0.43)
Parenting to Reduce Adolescent Depression and Anxiety Scale—Adolescent Report						
Baseline	24.44 (0.44)	24.89 (0.43)	—	—	—	—
12 months	23.38 (0.51)	24.16 (0.50)	−0.59 (297.53)	.56	−0.06 (−0.28 to 0.15)	−0.13 (−0.36 to 0.11)
Spence Children's Anxiety Scale—Parent Report						
Baseline	17.99 (0.90)	18.51 (0.90)	—	—	—	—
12 months	13.72 (0.95)	15.64 (0.94)	−1.28 (321.37)	.20	−0.14 (−0.37 to 0.08)	−0.12 (−0.34 to 0.10)
Spence Children's Anxiety Scale—Child Report						
Baseline	28.73 (1.36)	30.20 (1.33)	—	—	—	—
12 months	27.20 (1.52)	26.56 (1.49)	1.36 (294.14)	.18	0.16 (−0.07 to 0.39)	0.04 (−0.19 to 0.27)
Short Mood and Feelings Questionnaire—Parent Report						
Baseline	5.07 (0.40)	4.75 (0.40)	—	—	—	—
12 months	3.48 (0.40)	4.21 (0.40)	−1.88 (329.02)	.06[e]	−0.21 (−0.42 to 0.01)	−0.14 (−0.36 to 0.08)
Short Mood and Feelings Questionnaire—Child Report						
Baseline	6.16 (0.47)	6.40 (0.46)	—	—	—	—
12 months	7.06 (0.58)	7.08 (0.57)	0.33 (296.86)	0.74	0.04 (−0.19 to 0.27)	−0.01 (−0.24 to 0.22)

[a]*t* statistic of the planned contrast test of group × measurement-occasion interaction from baseline to 12-month follow-up, estimated under the group × measurement-occasion mixed model.

[b]Cohen *d* effect size of the group × measurement-occasion interaction from baseline to 12-month follow-up, calculated based on the *t* statistic of the planned contrast. Negative effect size indicates greater reduction in scores from baseline to 12-month follow-up in the intervention group compared with the control group.

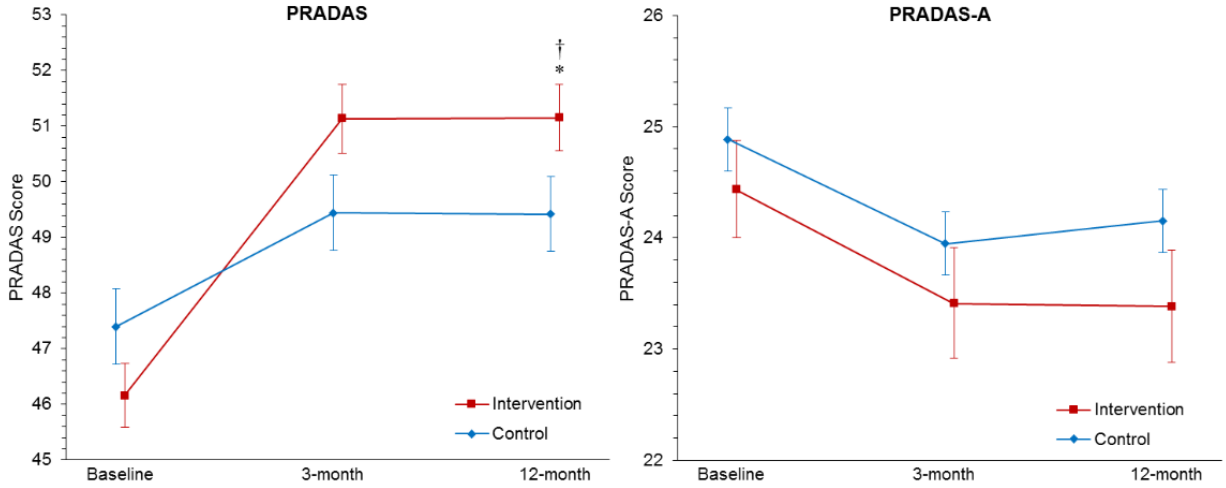
[c]Cohen *d* effect size of the difference between groups at 12-month follow-up. Negative effect size indicates lower scores in the intervention group compared with the control group.

[d]Not applicable.

[e]Becomes statistically significant with square-root-transformed data: $t_{(324.98)}$=−2.04; *P*=.04; *d*=−0.21 (95% CI −0.42 to −0.01).

XSL•FO
RenderX

Figure 2. Estimated marginal means for Parenting to Reduce Adolescent Depression and Anxiety Scale (PRADAS) and PRADAS—Adolescent (PRADAS-A) Report scores at baseline, postintervention (3-months postbaseline), and 12-month follow-up, estimated under the group × measurement-occasion mixed model. Error bars represent SEs. Higher scores on the PRADAS and PRADAS—Adolescent Report indicate greater concordance with the parenting guidelines (ie, more protective parenting factors and fewer parenting risk factors). Planned contrast of interaction (baseline to 12 months) was significant, P<.001. Pairwise comparison of group difference at 12-month follow-up was significant, P=.04.



Secondary Outcome Measure: Parenting to Reduce Adolescent Depression and Anxiety Scale—Adolescent Report

The planned contrast of the group × measurement-occasion interaction from baseline to 12-month follow-up was not significant for the PRADAS-A (see Table 2 and Figure 2). There was a significant main effect for time, with both groups reporting significantly reduced PRADAS-A scores over the 3 occasions: $F_{(2,297.99)}$=9.27; P<.001. However, the comparisons of group differences at each occasion were not significant (all P>.05).

Mediation Analyses

Mediation analyses revealed that the indirect effect of group on 12-month SMFQ-P via postintervention PRADAS was significant (indirect effect b=−0.08; 95% CI −0.16 to −0.01) when transformed symptom data were used. However, no significant mediation was found for any of the 4 symptom measures when using raw data (all P>.05; see Multimedia Appendix 6 and Table 1).

Post Hoc Moderation Analyses

Parental Concordance With the Guidelines (Parenting to Reduce Adolescent Depression and Anxiety Scale and Parenting to Reduce Adolescent Depression and Anxiety Scale—Adolescent Report), Moderated by Baseline Symptom Levels

Results of the PRADAS moderation analyses suggested that neither child anxiety (SCAS-C) nor depression (SMFQ-C) symptoms at baseline moderated intervention effects on parent reports of parenting (P>.05; see Multimedia Appendix 7). For adolescent reports of parenting (PRADAS-A), results suggested that baseline parent-reported anxiety (SCAS-P) moderated intervention effects on PRADAS-A: $F_{(2, 296.81)}$=10.23, P<.001. Figure 1 in Multimedia Appendix 7 presents the estimated marginal means of PRADAS-A, with baseline SCAS-P calculated at the 25th and 75th percentiles. As shown in this figure, among adolescents whose parents reported higher baseline SCAS-P scores, adolescents whose parents received PiP reported a greater reduction in PRADAS-A scores over time than controls. The analysis of PRADAS-A moderated by baseline SMFQ-P was not significant (P>.05). Multimedia Appendix 7 presents the results of these analyses.

Adolescent Depression and Anxiety Symptoms Moderated by Baseline Symptom Levels

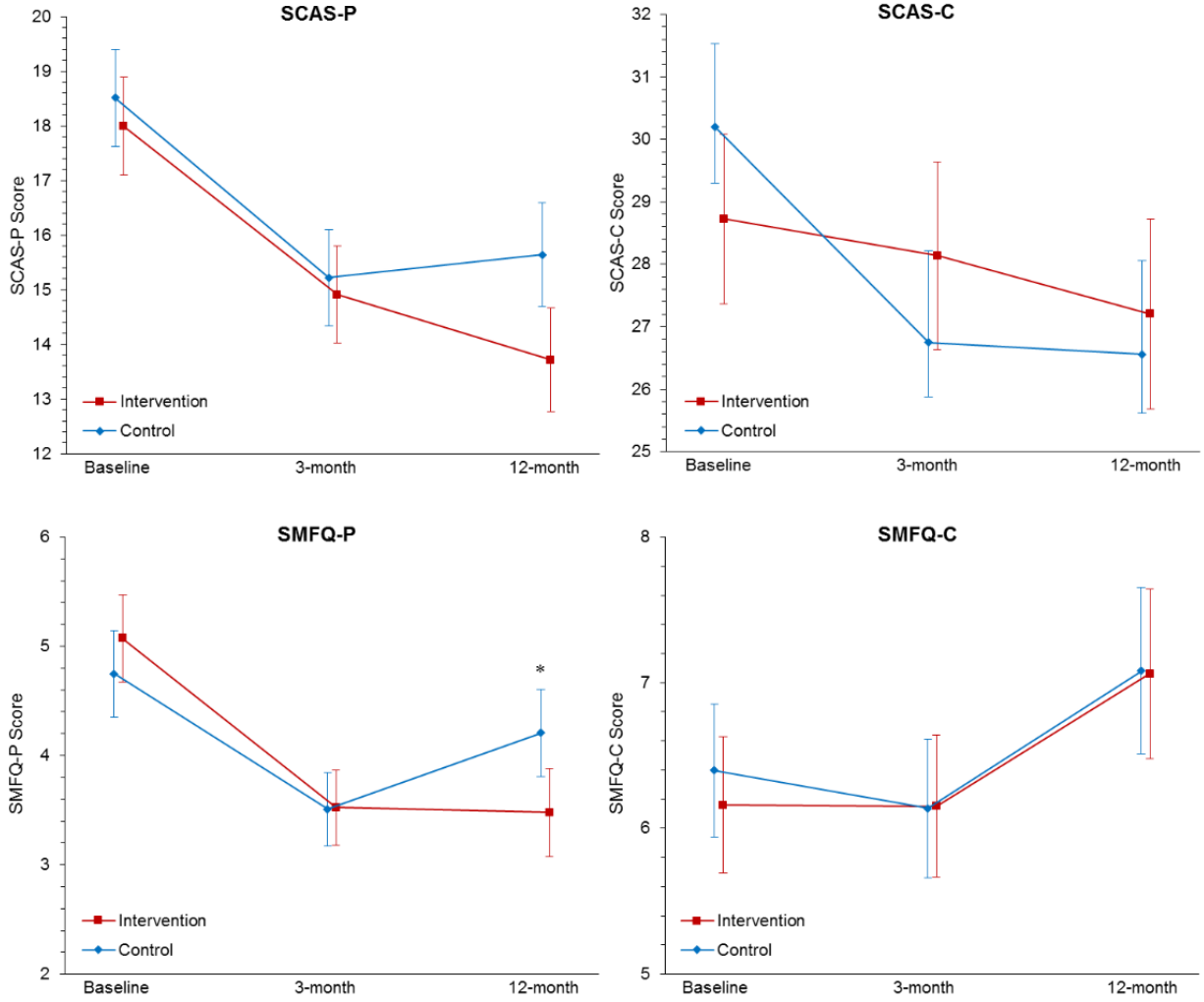
We conducted separate MMRMs for each of the symptom measures, moderated by baseline scores on the opposite informant of the same measure. Of the 4 analyses, only the model of SMFQ-P moderated by baseline SMFQ-C was significant: $F_{(2, 295.27)}$=4.40, P=.01. However, as shown in Multimedia Appendix 7 (Figure 2), this moderation effect may be attributable to baseline differences.

Adolescent Age and Gender as Moderators

We ran separate MMRMs for all 6 outcome measures, moderated by adolescent age at baseline and gender (in separate models). None of the 3-way interactions were significant (all P>.05; results available from the first author upon request).

XSL•FO
RenderX

Figure 3. Estimated marginal means for the SCAS—Parent Report (SCAS-P), SCAS—Child Report (SCAS-C), SMFQ—Parent Report (SMFQ-P), and SMFQ—Child Report (SMFQ-C) estimated under the group × measurement-occasion mixed model. Error bars represent SE. *Planned contrast of baseline to 12-month group × measurement-occasion interaction effect significant at $P<.05$ level, when square-root-transformed data were used. SCAS: Spence Children's Anxiety Scale; SMFQ: Short Mood and Feelings Questionnaire.



Discussion

Principal Findings

Parent Reports of Parenting

This RCT evaluated the medium-term effects of PiP. Consistent with our hypothesis, when compared with parents who received an educational-factsheet control intervention, parents who received PiP reported greater improvements in their parenting behaviors from baseline to 12 months later. These improvements represent reductions in parental risk factors and increases in parental protective factors for adolescent depression and anxiety [26,38]. Notably, the effect was medium in size, similar to that found at postintervention [34]. This suggests that the effect of PiP on parenting factors was maintained up to 9 months after the intervention and reduces the likelihood of social desirability and learning or priming effects (eg, receiving PiP alerts parents to *good parenting*, which influences their responses to the PRADAS at postintervention). The observed medium-term effect was despite minimal access to PiP modules by parents between the postintervention and 12-month assessments, so the maintenance of effects is unlikely because of *booster* effects.

Adolescent Symptoms

Contrary to expectations, at 12-month follow-up, there was no robust indication that PiP significantly reduced adolescent depression and anxiety symptoms as reported by either parents or adolescents, compared with the active-control intervention. This is largely consistent with findings at postintervention [34], suggesting that the effect of PiP on parent-reported parenting did not translate into significant reductions in adolescent symptoms compared with an active control. The effect of PiP on parent-reported depressive symptoms was significant (with a small effect size) when the analysis was conducted using transformed data, but it only approached significance when based on raw data. This suggests that PiP may lead to greater reduction in adolescent depressive symptoms than the control intervention, but given that this parent-reported finding was not verified by adolescent reports or other objective measures, the efficacy of PiP on adolescent depressive symptoms remains inconclusive. Nonetheless, there was a similar pattern of change

XSL•FO
RenderX

over time for both parent-reported depression symptoms and anxiety symptoms, whereby the intervention and control groups appeared to diverge after postintervention, with symptoms among the intervention group remaining stable, whereas symptoms in the control group started to increase over time. The pattern of change in the control group reflects an increase in depressive symptoms that is commonly seen in epidemiological research as adolescents approach mid-to-late adolescence, when the onset of depression and some anxiety disorders peaks [1]. A longer-term follow-up is required to ascertain whether these findings persist and demonstrate a long-term preventive effect of PiP.

It is possible that the use of an active control in this study, although a methodological strength, may have obscured the benefits of PiP on adolescent internalizing outcomes. Although the educational factsheets in the control intervention did not provide specific personalized parenting strategies in an engaging format (like PiP), they did provide credible information on adolescent development and general advice for parents about how they can best support their adolescent's development and well-being. For a highly educated and motivated sample of parents, this may be sufficient to produce the improvement in parenting observed in the control group, albeit to a weaker extent than the intervention group, which was maintained at 12-month follow-up. In turn, the improvement in parenting in both groups may explain the significant reduction in reported symptoms (except adolescent-reported depressive symptoms) over time in both groups and, hence, the nonsignificant group difference in change over time. This is notable because many preventive interventions for adolescent internalizing disorders that have demonstrated significant effects were compared with nonactive or no-intervention control groups [18,54]. Although this study's findings do not support the hypothesized superiority of PiP over an active control condition, it is notable that such brief, self-guided online parenting interventions were able to produce significant reductions in adolescent depressive and anxiety symptoms over time. These findings are promising given the scalability of PiP and the argument that even small reductions in mental health symptoms at the individual level could translate to significant population health benefits [55].

Adolescent Reports of Parenting

Contrary to expectation, a significant improvement in parenting was not seen in adolescent reports. This is consistent with findings at postintervention [34], suggesting that the effect of PiP on parent-reported parenting did not translate into adolescent perceptions of parenting. Though the discrepancy in findings involving parent-reported versus adolescent-reported parenting is notable, it concurs with most previous research involving parent and adolescent informants of parenting [56]. The PiP intervention components were tailored based on parents' responses to the PRADAS, without taking into account the adolescents' perspective (ie, their responses on the PRADAS-A). In particular, to personalize the feedback report for each parent and to identify the modules to recommend to the parent, PiP uses the parent's responses on the PRADAS to identify their strengths and areas from improvement. As such, insofar as the parent and adolescent perceptions of parenting differ, PiP may have targeted some parenting behaviors not perceived by adolescents to need improvement and failed to target other parenting behaviors that from the adolescent's perspective, need improvement. Post hoc analyses of 334 parent-adolescent dyads in this sample who completed the PRADAS and PRADAS-A support this possibility. On the basis of parent reports only, a mean of 6.71 modules (SD 1.53, range 1-9) were recommended, and a mean of 26.63 feedback messages (SD 7.81, range 7–60) were provided to parents. If the adolescents' perspectives were also taken into account, the mean number of modules recommended would have increased to 7.94 (SD 1.06, range 4-9), and a mean of 35.27 (SD 7.67, range 17-65) feedback messages would have been provided. Future research is required to examine whether tailoring PiP to both parent and adolescent perspectives of parenting can enhance the effects of PiP, especially on adolescent-reported parenting and symptom outcomes.

Divergence in parent and adolescent reports of parenting, especially during early adolescence, is well established and considered to be normative in the developmental literature [56]. From a developmental perspective, maturational processes during adolescence, such as autonomy seeking and individuation, may mean that parents and adolescents experience their interactions differently [57]. Nonetheless, it remains possible that parents' perceived improvements in their parenting were not observed by their adolescents up to 9 months after the intervention because parent-reported parenting changes did not translate into tangible behavioral changes observed by the adolescent. More recently, researchers have underscored the value of examining the degree of parent-adolescent discrepancy as a window into the dynamics of the parent-adolescent relationship and its associations with adolescent development and adjustment [58]. Further research is required to examine whether parent-adolescent discrepancies in perceptions of parenting may account for or moderate the effects of the PiP intervention on adolescent symptom outcomes.

Mediation and Moderation Analyses

Results from the mediation analyses were generally consistent with the above findings. Only the mediation hypothesis for parent-reported adolescent depressive symptoms was supported when using transformed data, whereby PiP compared with the control intervention led to greater improvements in parent-reported parenting from baseline to postintervention, which led to greater reductions in parent-reported adolescent depressive symptoms from baseline to 12-month follow-up. Consistent with the parenting pathway proposed by Sandler et al (2011), this mediation suggests that PiP's effects on parenting may account for a longer-term benefit on adolescent depressive symptoms. Future research is required to examine whether this preliminary finding would be corroborated by other measures of adolescent functioning, such as school engagement and academic performance.

In our post hoc moderation analyses, the moderation effect of adolescent-reported baseline depressive symptoms found at postintervention [34] emerged again, suggesting that among adolescents with elevated depressive symptoms at baseline, PiP was more effective than the control in reducing parent-reported depressive symptoms. However, upon probing, it appears that

XSL•FO
RenderX

the interaction effect may be largely accounted for by baseline differences between intervention and control groups. Hence, it remains to be ascertained whether the moderation effect supports the utility of PiP as an indicated prevention program. Adolescent age and gender did not significantly moderate intervention effects, suggesting that PiP's effects may be similar across early-to-mid-adolescence and for male and female adolescents.

Comparison With Previous Work

On the basis of the findings of a recent systematic review of technology-assisted preventive parenting interventions [27], PiP is the only online intervention aimed at preventing adolescent internalizing disorders. According to Yap et al's review [18], the only other universal preventive intervention for parents of adolescents is Tuning in to Teens (TINT), a group parenting program targeting emotion socialization [19]. In an RCT comparing TINT with a no-intervention control [19], TINT was found, at about 9 months postintervention, to produce a moderately large effect on parent-reported parenting, a small-to-medium effect on adolescent-reported parenting, small-to-medium effects on anxiety symptoms (small effect based on adolescent reports, small-to-medium based on parent reports), and a small-to-medium effect on parent-reported depressive symptoms (nonsignificant effect based on adolescent reports). A few observations are notable when comparing TINT with PiP, taking into consideration the differences in the comparison group (no-intervention versus active control, respectively) and modality of the intervention (face-to-face vs Web-based). First, the effects of PiP on parent-reported parenting compare favorably with those of TINT. Second, the Web-based self-guided modality of PiP may not provide an adequate intervention to produce changes in parenting behavior that are noticeable by adolescents, to in turn produce robust reductions in their internalizing outcomes. In contrast, the higher-intensity modality of TINT provides opportunities for parents to practice and discuss learned parenting behaviors with the facilitators and other parents, which may consolidate their learning into tangible changes in their behaviors when interacting with their adolescents. Further research is required to investigate whether a more-intensive, guided version of PiP (eg, providing additional coaching support via phone or videoconferencing; [26]) can yield stronger and more robust benefits on adolescent internalizing outcomes and adolescent-reported parenting.

The findings of PiP's effects to date share some similarities to those of other preventive parenting interventions that found developmental cascading effects at long-term follow-up (at least 5-6 years postintervention [35,37]). Specifically, intervention effects on parenting were found at postintervention and maintained into the 12-month follow-up; effects on adolescent depressive symptoms were observed at postintervention only among adolescents with elevated symptoms at baseline, but by the 12-month follow-up, the effect of PiP on depressive symptoms appeared to be emerging across the randomized sample, albeit not robustly. The earlier follow-up assessments for the New Beginnings [35] and Family Check-Up [37] RCTs had similarly promising but inconclusive findings, yet their subsequent follow-up assessments then found significant long-term benefits of the programs across various functioning domains. As such, a longer-term follow-up of this RCT is warranted to test whether PiP has developmental cascading effects on adolescent outcomes in late adolescence and early adulthood.

Strengths and Limitations

This study had various strengths. It evaluated a world-first tailored Web-based parenting intervention to prevent adolescent internalizing problems, using a rigorously designed RCT with an active control group, parent and adolescent informants on all outcomes of interest, low attrition rates that are balanced across groups, and high intervention completion. However, various limitations merit comment. First, despite successfully recruiting a large community sample, there was overrepresentation by mothers and highly educated parents. Although this is a limitation shared by most online preventive parenting interventions [27], it urgently needs to be redressed so that the dissemination of evidence-based online interventions does not perpetuate the exclusion of fathers and of parents from vulnerable or disadvantaged backgrounds, inadvertently contributing to the widening of social inequalities in health between families of higher versus lower socioeconomic positions [59]. Second, for reasons of parsimony, this study only included 1 parent and 1 adolescent per family. Nonetheless, we encouraged parent participants who had a co-parenting partner to share and discuss the resources they received to enhance consistency in co-parenting. It remains to be seen whether providing both parents with the intervention will yield synergistic benefits, as suggested by previous research [60]. Third, parents were allowed to participate in the trial if their adolescent declined participation, resulting in a small subset of our sample (37/359, 10.3%) without data on self-reported adolescent symptom outcomes. We chose to have inclusive eligibility criteria to capture a more representative and diverse sample, given the pragmatic design of the trial. Thus, the small loss of data is outweighed by the greater generalizability of our findings to a real-world implementation setting, where adolescents would not be required to participate with their parent. Fourth, this study did not include behavioral measures of parenting or other measures of adolescent functioning outcomes, including quality of life, school engagement, and academic achievements, as well as measures of cost-effectiveness. Future research using these measures will provide insights into whether PiP produces change in objectively measured parenting behaviors and yields broader benefits for adolescents beyond the reduction of internalizing symptoms. Evidence of its cost-effectiveness is also important for advancing the prevention agenda, given that prevention research and interventions are still largely underfunded even in developed countries, including Australia [61]. Fifth, we assumed that adolescents were unaware of their parent's group allocation, but we did not implement checks to verify the blinding nor did we ask parents to conceal their program from their adolescent. Finally, given that the onset of depression and some anxiety disorders peaks in mid-to-late adolescence, a longer-term follow-up examining cases of disorder is required to adequately test whether PiP can prevent the onset of depression and anxiety disorders across adolescence and into early adulthood. Such

XSL•FO
RenderX

evidence is notably lacking in preventive parenting intervention research [18,27].

Conclusions

Overall, this study found that the PiP intervention produced significantly greater improvements than an active control in parenting risk and protective factors associated with adolescent risk for depression and anxiety. The effects persisted for up to 9 months postintervention, reducing the likelihood that they were due to social desirability or short-term priming effects. Findings from the analyses using transformed data and from mediation analyses suggest that PiP may have some benefits for adolescent symptoms that need to be ascertained in future research. However, parent-reported changes in parenting were not reported by adolescents, and there were no robust findings with regard to reductions in adolescent symptoms. Nonetheless, significant reductions in adolescent symptoms were observed over time in both groups. Given these promising findings, the paucity of evidence-based resources for parents of adolescents, and the scalability of the Web-based platform, PiP may be useful as a low-cost, sustainable public health universal prevention program to empower parents for their adolescents' mental health.

Acknowledgments

The authors acknowledge funding from the National Health and Medical Research Council (NHMRC) for the Web development of the PiP intervention, and the partnership of *beyondblue, the national depression and anxiety initiative* in the development of the parenting guidelines. The authors received salary support from the NHMRC for a Career Development Fellowship (MBHY, APP1061744) and a Senior Principal Research Fellowship (AFJ, APP1059785), an Australian Research Council Laureate Fellowship (RMR, FL150100096), and a Monash University Postgraduate Publication Award (MCB). The RCT was supported by an Australian Rotary Health Research Grant and Monash University Advancing Women's Research Success Grant. None of the funding sources had any role in the conduct of publication of this study. The authors thank the reference group of parents and the focus group of students who contributed to the development of the PiP intervention. They also acknowledge the contributions of Jennifer Hanson-Peterson, Claire Nicolas, and Jacqueline Green in the development and testing of the intervention and in project management, as well as the research assistants and the schools and organizations that assisted with recruitment for the RCT.

Conflicts of Interest

MBHY, MCB, AFJ, and KAL are codevelopers of the PiP intervention, and MBHY and AFJ are cofounders of the broader Parenting Strategies online platform of parenting interventions, including PiP. None of the authors derive any personal financial benefit from these interventions.

Multimedia Appendix 1
Screenshots of the Partners in Parenting intervention.
[PDF File (Adobe PDF File), 486KB - jmir_v21i8e13628_app1.pdf]

Multimedia Appendix 2
Participant informed consent documentation.
[PDF File (Adobe PDF File), 714KB - jmir_v21i8e13628_app2.pdf]

Multimedia Appendix 3
Sample characteristics at baseline.
[PDF File (Adobe PDF File), 97KB - jmir_v21i8e13628_app3.pdf]

Multimedia Appendix 4
Observed scores for all measures at each measurement occasion.
[PDF File (Adobe PDF File), 78KB - jmir_v21i8e13628_app4.pdf]

Multimedia Appendix 5
Results of primary and secondary outcome mixed-model repeated measures, run on square-root-transformed data.
[PDF File (Adobe PDF File), 85KB - jmir_v21i8e13628_app5.pdf]

Multimedia Appendix 6
Results of mediation analyses.
[PDF File (Adobe PDF File), 80KB - jmir_v21i8e13628_app6.pdf]

Multimedia Appendix 7

XSL•FO
RenderX

Results of posthoc moderation analyses.
[PDF File (Adobe PDF File), 136KB - jmir_v21i8e13628_app7.pdf]

Multimedia Appendix 8
Consolidated Standards of Reporting Trials (CONSORT) E-Health (V 1.6.1) checklist.
[PDF File (Adobe PDF File), 2MB - jmir_v21i8e13628_app8.pdf]

References

1. Kessler RC, Petukhova M, Sampson NA, Zaslavsky AM, Wittchen HU. 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]
2. Woodward LJ, Fergusson DM. Life course outcomes of young people with anxiety disorders in adolescence. J Am Acad Child Adolesc Psychiatry 2001 Sep;40(9):1086-1093. [doi: 10.1097/00004583-200109000-00018] [Medline: 11556633]
3. Rao U, Ryan ND, Birmaher B, Dahl RE, Williamson DE, Kaufman J, et al. Unipolar depression in adolescents: clinical outcome in adulthood. J Am Acad Child Adolesc Psychiatry 1995 May;34(5):566-578. [doi: 10.1097/00004583-199505000-00009] [Medline: 7775352]
4. Andrews G, Issakidis C, Sanderson K, Corry J, Lapsley H. Utilising survey data to inform public policy: comparison of the cost-effectiveness of treatment of ten mental disorders. Br J Psychiatry 2004 Jun;184:526-533. [doi: 10.1192/bjp.184.6.526] [Medline: 15172947]
5. Bor W, Dean AJ, Najman J, Hayatbakhsh R. Are child and adolescent mental health problems increasing in the 21st century? A systematic review. Aust N Z J Psychiatry 2014 Jul;48(7):606-616. [doi: 10.1177/0004867414533834] [Medline: 24829198]
6. Collishaw S, Maughan B, Natarajan L, Pickles A. Trends in adolescent emotional problems in England: a comparison of two national cohorts twenty years apart. J Child Psychol Psychiatry 2010 Aug;51(8):885-894. [doi: 10.1111/j.1469-7610.2010.02252.x] [Medline: 20497281]
7. Rudolph KD, Lansford JE, Rodkin PC. Interpersonal theories of developmental psychopathology. In: Cicchetti D, editor. Developmental Psychopathology: Maladaptation and Psychopathology. Third Edition. Hoboken, New Jersey: John Wiley & Sons; 2016:1-69.
8. Boomsma DI, van Beijsterveldt CE, Hudziak JJ. Genetic and environmental influences on anxious/depression during childhood: a study from the Netherlands twin register. Genes Brain Behav 2005 Nov;4(8):466-481 [FREE Full text] [doi: 10.1111/j.1601-183X.2005.00141.x] [Medline: 16268991]
9. Eley TC, McAdams TA, Rijsdijk FV, Lichtenstein P, Narusyte J, Reiss D, et al. The intergenerational transmission of anxiety: a children-of-twins study. Am J Psychiatry 2015 Jul;172(7):630-637. [doi: 10.1176/appi.ajp.2015.14070818] [Medline: 25906669]
10. Yap MB, Pilkington PD, Ryan SM, Jorm AF. Parental factors associated with depression and anxiety in young people: a systematic review and meta-analysis. J Affect Disord 2014 Mar;156:8-23. [doi: 10.1016/j.jad.2013.11.007] [Medline: 24308895]
11. Pinquart M. Associations of parenting dimensions and styles with internalizing symptoms in children and adolescents: a meta-analysis. Marriage Fam Rev 2016 Oct 14;53(7):613-640. [doi: 10.1080/01494929.2016.1247761]
12. O'Connell ME, Boat T, Warner K. Preventing Mental, Emotional, and Behavioral Disorders Among Young People: Progress and Possibilities. Washington, DC: National Academic Press; 2009.
13. Cairns KE, Yap MB, Pilkington PD, Jorm AF. Risk and protective factors for depression that adolescents can modify: a systematic review and meta-analysis of longitudinal studies. J Affect Disord 2014 Dec;169:61-75. [doi: 10.1016/j.jad.2014.08.006] [Medline: 25154536]
14. Beesdo K, Knappe S, Pine DS. Anxiety and anxiety disorders in children and adolescents: developmental issues and implications for DSM-V. Psychiatr Clin North Am 2009 Sep;32(3):483-524 [FREE Full text] [doi: 10.1016/j.psc.2009.06.002] [Medline: 19716988]
15. Schleider JL, Weisz JR. Family process and youth internalizing problems: a triadic model of etiology and intervention. Dev Psychopathol 2017 Feb;29(1):273-301. [doi: 10.1017/S095457941600016X] [Medline: 27048767]
16. Sameroff AJ. The Transactional Model of Development: How Children and Contexts Shape Each Other. Washington, DC: American Psychological Association; 2009.
17. Sandler IN, Schoenfelder EN, Wolchik SA, MacKinnon DP. Long-term impact of prevention programs to promote effective parenting: lasting effects but uncertain processes. Annu Rev Psychol 2011;62:299-329 [FREE Full text] [doi: 10.1146/annurev.psych.121208.131619] [Medline: 20822438]
18. Yap MB, Morgan AJ, Cairns K, Jorm AF, Hetrick SE, Merry S. Parents in prevention: a meta-analysis of randomized controlled trials of parenting interventions to prevent internalizing problems in children from birth to age 18. Clin Psychol Rev 2016 Dec;50:138-158. [doi: 10.1016/j.cpr.2016.10.003] [Medline: 27969003]
19. Kehoe CE, Havighurst SS, Harley AE. Tuning in to teens: improving parent emotion socialization to reduce youth internalizing difficulties. Soc Dev 2013 Oct 15;23(2):413-431. [doi: 10.1111/sode.12060]

XSL•FO
RenderX

20. Bearslee WR, Wright EJ, Gladstone TR, Forbes P. Long-term effects from a randomized trial of two public health preventive interventions for parental depression. J Fam Psychol 2007 Dec;21(4):703-713. [doi: 10.1037/0893-3200.21.4.703] [Medline: 18179342]
21. Rotheram-Borus MJ, Stein JA, Lester P. Adolescent adjustment over six years in HIV-affected families. J Adolesc Health 2006 Aug;39(2):174-182. [doi: 10.1016/j.jadohealth.2006.02.014] [Medline: 16857528]
22. Sandler I, Ingram A, Wolchik S, Tein JY, Winslow E. Long-term effects of parenting-focused preventive interventions to promote resilience of children and adolescents. Child Dev Perspect 2015 Sep;9(3):164-171 [FREE Full text] [doi: 10.1111/cdep.12126] [Medline: 30854024]
23. Winslow EB, Bonds D, Wolchik S, Sandler I, Braver S. Predictors of enrollment and retention in a preventive parenting intervention for divorced families. J Prim Prev 2009 Mar;30(2):151-172 [FREE Full text] [doi: 10.1007/s10935-009-0170-3] [Medline: 19283483]
24. Morawska A, Sanders M. A review of parental engagement in parenting interventions and strategies to promote it. J Child's Serv 2006 Apr;1(1):29-40. [doi: 10.1108/17466660200600004]
25. Finan SJ, Swierzbiolek B, Priest N, Warren N, Yap M. Parental engagement in preventive parenting programs for child mental health: a systematic review of predictors and strategies to increase engagement. PeerJ 2018;6:e4676 [FREE Full text] [doi: 10.7717/peerj.4676] [Medline: 29719737]
26. Yap MB, Lawrence KA, Rapee RM, Cardamone-Breen MC, Green JM, Jorm AF. Partners in parenting: a multi-level web-based approach to support parents in prevention and early intervention for adolescent depression and anxiety. JMIR Ment Health 2017 Dec 19;4(4):e59 [FREE Full text] [doi: 10.2196/mental.8492] [Medline: 29258974]
27. Hansen A, Broomfield G, Yap MB. A systematic review of technology-assisted parenting programs to prevent mental health problems in youth aged 0-18 years: applicability to underserved Australian communities. Aust J Psychol 2019 Apr 1:1-31. [doi: 10.1111/ajpy.12250]
28. Rutter M, Sroufe LA. Developmental psychopathology: concepts and challenges. Dev Psychopathol 2000;12(3):265-296. [doi: 10.1017/S0954579400003023] [Medline: 11014739]
29. Mrazek PJ, Haggerty RJ, editors. Reducing Risks for Mental Disorders: Frontiers for Preventive Intervention Research. Washington, DC: National Academies Press; 1994.
30. Oinas-Kukkonen H, Harjumaa M. Persuasive systems design: key issues, process model, and system features. Commun Assoc Inf Syst 2009;24:485-500. [doi: 10.17705/1CAIS.02428]
31. Parenting Strategies: Protecting Your Child's Mental Health. 2013. How to Reduce Your Child's Risk of Depression and Clinical Anxiety: Strategies for Parents of Primary-School Aged Children URL: https://www.parentingstrategies.net/depression/media/pdf/Guidelines.pdf [WebCite Cache ID 6vCqbkNHX]
32. Yap MB, Pilkington PD, Ryan SM, Kelly CM, Jorm AF. Parenting strategies for reducing the risk of adolescent depression and anxiety disorders: a Delphi consensus study. J Affect Disord 2014 Mar;156:67-75. [doi: 10.1016/j.jad.2013.11.017] [Medline: 24359862]
33. Sanders MR, Kirby JN. Consumer engagement and the development, evaluation, and dissemination of evidence-based parenting programs. Behav Ther 2012 Jun;43(2):236-250 [FREE Full text] [doi: 10.1016/j.beth.2011.01.005] [Medline: 22440062]
34. Yap MB, Mahtani S, Rapee RM, Nicolas C, Lawrence KA, Mackinnon A, et al. A tailored web-based intervention to improve parenting risk and protective factors for adolescent depression and anxiety problems: postintervention findings from a randomized controlled trial. J Med Internet Res 2018 Jan 19;20(1):e17 [FREE Full text] [doi: 10.2196/jmir.9139] [Medline: 29351895]
35. McClain DB, Wolchik SA, Winslow E, Tein JY, Sandler IN, Millsap RE. Developmental cascade effects of the New Beginnings program on adolescent adaptation outcomes. Dev Psychopathol 2010 Nov;22(4):771-784 [FREE Full text] [doi: 10.1017/S0954579410000453] [Medline: 20883581]
36. Wolchik SA, Tein JY, Sandler IN, Kim HJ. Developmental cascade models of a parenting-focused program for divorced families on mental health problems and substance use in emerging adulthood. Dev Psychopathol 2016 Aug;28(3):869-888 [FREE Full text] [doi: 10.1017/S0954579416000365] [Medline: 27427811]
37. Reuben JD, Shaw DS, Brennan LM, Dishion TJ, Wilson MN. A family-based intervention for improving children's emotional problems through effects on maternal depressive symptoms. J Consult Clin Psychol 2015 Dec;83(6):1142-1148 [FREE Full text] [doi: 10.1037/ccp0000049] [Medline: 26302250]
38. Cardamone-Breen MC, Jorm AF, Lawrence KA, Mackinnon AJ, Yap MB. The parenting to reduce adolescent depression and anxiety scale: assessing parental concordance with parenting guidelines for the prevention of adolescent depression and anxiety disorders. PeerJ 2017;5:e3825 [FREE Full text] [doi: 10.7717/peerj.3825] [Medline: 28951815]
39. Raising Children Network. 2015. Teens (12-18 Years) URL: http://raisingchildren.net.au/teens/teens.html [accessed 2019-06-30] [WebCite Cache ID 6qz1JBTsj]
40. Angold A, Costello E, Messer S, Pickles A, Winder F, Silver D. Development of a short questionnaire for use in epidemiological studies of depression in children and adolescents. Int J Meth Psych Res 1995;5:237-249 [FREE Full text]

XSL•FO
RenderX

41. Rhew IC, Simpson K, Tracy M, Lymp J, McCauley E, Tsuang D, et al. Criterion validity of the short mood and feelings questionnaire and one- and two-item depression screens in young adolescents. Child Adolesc Psychiatry Ment Health 2010 Feb 9;4(1):8 [FREE Full text] [doi: 10.1186/1753-2000-4-8] [Medline: 20181135]
42. Thapar A, McGuffin P. Validity of the shortened mood and feelings questionnaire in a community sample of children and adolescents: a preliminary research note. Psychiatry Res 1998 Nov 16;81(2):259-268. [doi: 10.1016/S0165-1781(98)00073-0] [Medline: 9858042]
43. Nauta MH, Scholing A, Rapee RM, Abbott M, Spence SH, Waters A. A parent-report measure of children's anxiety: psychometric properties and comparison with child-report in a clinic and normal sample. Behav Res Ther 2004 Jul;42(7):813-839. [doi: 10.1016/S0005-7967(03)00200-6] [Medline: 15149901]
44. Spence SH. Structure of anxiety symptoms among children: a confirmatory factor-analytic study. J Abnorm Psychol 1997 May;106(2):280-297. [doi: 10.1037/0021-843X.106.2.280] [Medline: 9131848]
45. Spence SH, Barrett PM, Turner CM. Psychometric properties of the Spence children's anxiety scale with young adolescents. J Anxiety Disord 2003;17(6):605-625. [doi: 10.1016/S0887-6185(02)00236-0] [Medline: 14624814]
46. Spence SH. A measure of anxiety symptoms among children. Behav Res Ther 1998 May;36(5):545-566. [doi: 10.1016/S0005-7967(98)00034-5] [Medline: 9648330]
47. Kelders SM, Kok RN, Ossebaard HC, van Gemert-Pijnen JE. Persuasive system design does matter: a systematic review of adherence to web-based interventions. J Med Internet Res 2012 Nov 14;14(6):e152 [FREE Full text] [doi: 10.2196/jmir.2104] [Medline: 23151820]
48. Spence SH. Spence Children's Anxiety Scale. 2016. URL: http://www.scaswebsite.com/ [accessed 2019-06-30] [WebCite Cache ID 75x3DBSm5]
49. Bell ML, Fairclough DL. Practical and statistical issues in missing data for longitudinal patient-reported outcomes. Stat Methods Med Res 2014 Oct;23(5):440-459. [doi: 10.1177/0962280213476378] [Medline: 23427225]
50. Mallinckrodt CH, Sanger TM, Dubé S, DeBrota DJ, Molenberghs G, Carroll RJ, et al. Assessing and interpreting treatment effects in longitudinal clinical trials with missing data. Biol Psychiatry 2003 Apr 15;53(8):754-760. [doi: 10.1016/S0006-3223(02)01867-X] [Medline: 12706959]
51. Gueorguieva R, Krystal JH. Move over ANOVA: progress in analyzing repeated-measures data and its reflection in papers published in the archives of general psychiatry. Arch Gen Psychiatry 2004 Mar;61(3):310-317. [doi: 10.1001/archpsyc.61.3.310] [Medline: 14993119]
52. Hayes AF. PROCESS Macro. 2018. URL: http://www.processmacro.org/index.html [accessed 2018-09-01] [WebCite Cache ID 75x3L5ZTk]
53. Hayes AF. Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach. Second Edition. New York: The Guilford Press; 2018.
54. Hetrick SE, Cox GR, Witt KG, Bir JJ, Merry SN. Cognitive behavioural therapy (CBT), third-wave CBT and interpersonal therapy (IPT) based interventions for preventing depression in children and adolescents. Cochrane Database Syst Rev 2016 Aug 9(8):CD003380. [doi: 10.1002/14651858.CD003380.pub4] [Medline: 27501438]
55. Rose G. Rose's Strategy of Preventive Medicine. Oxford: Oxford University Press; 1992.
56. Branje SJ, Laursen B, Collins WA. Parent-child communication during adolescence. In: Vangelisti AL, editor. The Routledge Handbook of Family Communication. Second Edition. New York: Routledge; 2012:271-286.
57. Collins WA, Laursen B. Adolescent relationships: the art of fugue. In: Hendrick SS, Hendrick C, editors. Close Relationships: A Sourcebook. Thousand Oaks, CA: Sage Publications; 2000:59-69.
58. de los Reyes A, Ohannessian CM, Racz SJ. Discrepancies between adolescent and parent reports about family relationships. Child Dev Perspect 2019 Feb 9;13(1):53-58. [doi: 10.1111/cdep.12306]
59. Yap MB. Treatment of paediatric anxiety disorders: what next? Lancet Child Adolesc Health 2018 Nov;2(11):767-769. [doi: 10.1016/S2352-4642(18)30292-X] [Medline: 30241995]
60. Frank TJ, Keown LJ, Sanders MR. Enhancing father engagement and interparental teamwork in an evidence-based parenting intervention: a randomized-controlled trial of outcomes and processes. Behav Ther 2015 Nov;46(6):749-763. [doi: 10.1016/j.beth.2015.05.008] [Medline: 26520218]
61. Reavley NJ, Jorm AF. Mental health reform: increased resources but limited gains. Med J Aust 2014 Oct 6;201(7):375-376. [doi: 10.5694/mja13.00198] [Medline: 25296048]

Abbreviations

MMRM: mixed-model repeated measures
NHMRC: National Health and Medical Research Council
PiP: Partners in Parenting
PRADAS: Parenting to Reduce Adolescent Depression and Anxiety Scale
PRADAS-A: Parenting to Reduce Adolescent Depression and Anxiety Scale—Adolescent Report
RCT: randomized controlled trial
SCAS: Spence Children’s Anxiety Scale

XSL•FO
RenderX

SCAS-P: Spence Children's Anxiety Scale—Parent Report
SCAS-C: Spence Children's Anxiety Scale—Child Report
SMFQ: Short Mood and Feelings Questionnaire
SMFQ-P: Short Mood and Feelings Questionnaire—Parent Report
SMFQ-C: Short Mood and Feelings Questionnaire—Child Report
TINT: Tuning in to Teens

Edited by G Eysenbach; submitted 05.02.19; peer-reviewed by M Subotic-Kerry, A Murphy; comments to author 27.03.19; revised version received 04.06.19; accepted 12.06.19; published 15.08.19.

Please cite as:
Yap MBH, Cardamone-Breen MC, Rapee RM, Lawrence KA, Mackinnon AJ, Mahtani S, Jorm AF
Medium-Term Effects of a Tailored Web-Based Parenting Intervention to Reduce Adolescent Risk of Depression and Anxiety: 12-Month Findings From a Randomized Controlled Trial
J Med Internet Res 2019;21(8):e13628
URL: http://www.jmir.org/2019/8/e13628/
doi:10.2196/13628
PMID:31418422

©Marie Bee Hui Yap, Mairead C Cardamone-Breen, Ronald M Rapee, Katherine A Lawrence, Andrew J Mackinnon, Shireen Mahtani, Anthony F Jorm. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 15.08.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

Digital Health Apps in the Clinical Care of Inflammatory Bowel Disease: Scoping Review

Andrew Lukas Yin[1,2], BA; David Hachuel[2,3], BSc, MSc; John P Pollak[2], PhD; Ellen J Scherl[4], MD, FACP, FACG; Deborah Estrin[2], PhD

[1]Medical College, Weill Cornell Medicine, New York, NY, United States

[2]Cornell Tech, New York, NY, United States

[3]augGI Technologies, New York, NY, United States

[4]Jill Roberts Center for Inflammatory Bowel Disease, Weill Cornell Medicine, New York, NY, United States

Corresponding Author:
Andrew Lukas Yin, BA
Medical College
Weill Cornell Medicine
1300 York Avenue
New York, NY,
United States
Phone: 1 212 746 1067
Email: aly2011@med.cornell.edu

Abstract

Background: Digital health is poised to transform health care and redefine personalized health. As Internet and mobile phone usage increases, as technology develops new ways to collect data, and as clinical guidelines change, all areas of medicine face new challenges and opportunities. Inflammatory bowel disease (IBD) is one of many chronic diseases that may benefit from these advances in digital health. This review intends to lay a foundation for clinicians and technologists to understand future directions and opportunities together.

Objective: This review covers mobile health apps that have been used in IBD, how they have fit into a clinical care framework, and the challenges that clinicians and technologists face in approaching future opportunities.

Methods: We searched PubMed, Scopus, and ClinicalTrials.gov to identify mobile apps that have been studied and were published in the literature from January 1, 2010, to April 19, 2019. The search terms were (“mobile health” OR “eHealth” OR “digital health” OR “smart phone” OR “mobile app” OR “mobile applications” OR “mHealth” OR “smartphones”) AND (“IBD” OR “Inflammatory bowel disease” OR “Crohn's Disease” (CD) OR “Ulcerative Colitis” (UC) OR “UC” OR “CD”), followed by further analysis of citations from the results. We searched the Apple iTunes app store to identify a limited selection of commercial apps to include for discussion.

Results: A total of 68 articles met the inclusion criteria. A total of 11 digital health apps were identified in the literature and 4 commercial apps were selected to be described in this review. While most apps have some educational component, the majority of apps focus on eliciting patient-reported outcomes related to disease activity, and a few are for treatment management. Significant benefits have been seen in trials relating to education, quality of life, quality of care, treatment adherence, and medication management. No studies have reported a negative impact on any of the above. There are mixed results in terms of effects on office visits and follow-up.

Conclusions: While studies have shown that digital health can fit into, complement, and improve the standard clinical care of patients with IBD, there is a need for further validation and improvement, from both a clinical and patient perspective. Exploring new research methods, like microrandomized trials, may allow for more implementation of technology and rapid advancement of knowledge. New technologies that can objectively and seamlessly capture remote data, as well as complement the clinical shift from symptom-based to inflammation-based care, will help the clinical and health technology communities to understand the full potential of digital health in the care of IBD and other chronic illnesses.

(J Med Internet Res 2019;21(8):e14630) doi:10.2196/14630

XSL•FO
RenderX

KEYWORDS

digital health; mHealth; mobile health; mobile technology; smartphone; eHealth; review; inflammatory bowel disease; Crohn's disease; ulcerative colitis

Introduction

Digital health technologies—tools leveraging mobile phones, tablets, Web platforms, and wearables to improve health outcomes—are rapidly changing the practice of medicine and redefining approaches to health care. By the end of 2018, 67% of the global population (5.1 billion people) subscribed to mobile Internet services, a number expected to increase to 71% (5.8 billion people) by 2025 [1]. In 2013, the Pew Research Center showed that 72% of Internet users in the United States searched for health information the previous year, with 35% admitting to using the Internet to try to determine their own or someone else's medical condition [2].

More specifically, Makovsky's *Pulse of Online Health* in 2015 found that 66% of the US population reported interest in using mobile apps to manage their health [3]. Accordingly, in 2017 over 325,000 mobile health-related apps were commercially available for download, a 25% increase from 2016 [4]. Unfortunately, reviews of these commercially available health apps frequently lack an evidence base or adherence to guidelines, with very few going through clinical trials [5,6]. Moreover, clinical research has struggled to study and define clinical guidelines for the new data collected [7-9], even with exciting opportunities to engage hard-to-reach populations and provide innovative care [10]. Over 50% of individuals in developed countries have at least one chronic disease [11], and about one-fourth of them experience limitations in their activities of daily living [12].

One such chronic disease, inflammatory bowel disease (IBD), which is composed of Crohn's disease (CD) and ulcerative colitis (UC), is well-situated for technological intervention. By engaging, educating, and monitoring patients, technology can help us understand and improve care in a disease that presents uniquely in each individual. IBD has a 0.5% prevalence for both CD and UC in the western world [13], a number expected to increase globally, with specific spikes in certain populations [14]. IBD is relapsing and remitting in nature, with disease exacerbations (ie, flares) being a key driver of the acute need for medical care [15] and having a negative impact on quality of life (QoL) [16]. Treatment is unique to each patient's circumstance, including self-monitoring and behavioral interventions [17,18]. Some long-term intestinal complications of CD include strictures, fistulas, and abscesses [19,20]; some long-term intestinal complications of UC are perforation, colitis, colonic strictures, and colorectal cancer [20-23].

The Selecting Therapeutic Targets in Inflammatory Bowel Disease (STRIDE) recommendations and the CALM trial have sparked significant changes in clinical management, suggesting that resolution of symptoms alone is not a sufficient outcome; objective evidence, such as fecal calprotectin (FC) or C-reactive protein, is also necessary to guide clinical decisions [24,25]. Using the objective marker, FC, and patient symptoms to guide therapeutic decisions, 46% of patients in the CALM trial reached mucosal healing after 48 weeks compared to 30% of control subjects [25]. As a result, the goals of treatment are evolving from symptom-based to inflammation-based (ie, symptom and biomarker) control to slow or even reverse disease progression. Due to this shift, patients and providers are challenged to find new ways to engage in monitoring, set goals for treatment, and discuss longitudinal care. Digital health can help close these gaps.

Prior reviews have discussed both clinically tested and commercially available digital health interventions in IBD by listing out apps or discussing the various kinds of technological interventions [26-29]. This review includes some new digital health apps and will focus on how they have fit into clinical care. By highlighting areas where they have been tested and shown promise in the clinical flow, this review builds the foundation for clinicians and technologists contemplating new apps in IBD.

Methods

Scoping Review

Scoping reviews embrace systematic discovery and selection of literature to elucidate and summarize the depth and breadth of a field of interest. One of the benefits of creating this review is to allow commentary regarding potential gaps or areas for innovation. Given this background, we believed a scoping review was most appropriate, using the most current guidelines to (1) identify the research question, (2) identify relevant studies, (3) select relevant studies, (4) chart the data, and (5) collate and summarize the results [30,31].

Development of Research Questions

The research questions were as follows: What digital health apps have been described in the published literature in the setting of IBD? How have these apps complemented or been used in current clinical practice?

Identifying Apps Discussed in Published Literature

The literature was reviewed using PubMed, Scopus, and ClinicalTrials.gov for articles or trials published from January 1, 2010, through April 19, 2019; published articles and trials were accessed on April 19, 2019. The search terms were "mobile health" OR "eHealth" OR "digital health" OR "smart phone" OR "mobile app" OR "mobile applications" OR "mHealth" OR "smartphones" AND "IBD" OR "Inflammatory bowel disease" OR "Crohn's Disease" OR "Ulcerative Colitis" OR "UC" OR "CD". Only English-language articles were included. The Scopus search included titles, abstracts, and keywords. The PubMed search included the entire paper and excluded "UC" and "CD" in the search terms. The ClinicalTrials.gov search included studies containing these terms. All papers were reviewed by the first author of this paper (ALY), with consultation from other authors when necessary.

XSL•FO
RenderX

Identifying Commercially Available Apps

The iTunes iOS app store was reviewed to find a selection of commercial apps for the purpose of discussion and recognition of the large consumer market that patients face. As the review of commercial apps was not intended to be exhaustive, Google Play and other app stores were not explored. The search terms "inflammatory bowel disease," "crohn's disease," "ulcerative colitis," and "colitis" were used to identify potential apps; the iTunes iOS app store was last accessed on April 30, 2019. Considering the desired length and scope of this review, a restricted number of apps were included. These were selected based on top search *hits* in the app store, which reflect commercial apps that have historically been used most [29].

Article Selection

The titles and abstracts of all *hits* were reviewed. Articles were included if they explored the use of a digital health intervention in the care of IBD, or CD or UC specifically. Citations of included articles were then also screened for additional relevant articles or apps. Articles were excluded if they did not involve IBD, CD, or UC. They were excluded if they discussed teleconferencing or video chatting as the sole intervention, as this was decided to be out of the scope of the review.

Data Charting

The articles were organized according to the apps that they described. Each app was assessed for how it addressed any of five identified areas of clinical care: education, monitoring, treatment, follow-up, and patient satisfaction.

Collation and Summary

We summarize and present the relevant features of the apps by breaking them into the five areas of clinical care identified, providing a framework for each of these areas. The goal of the scoping review was to summarize the depth and breadth of digital health apps that have been used in IBD clinical care in order to provide a foundation for discussion for physicians and technologists to pursue future opportunities.

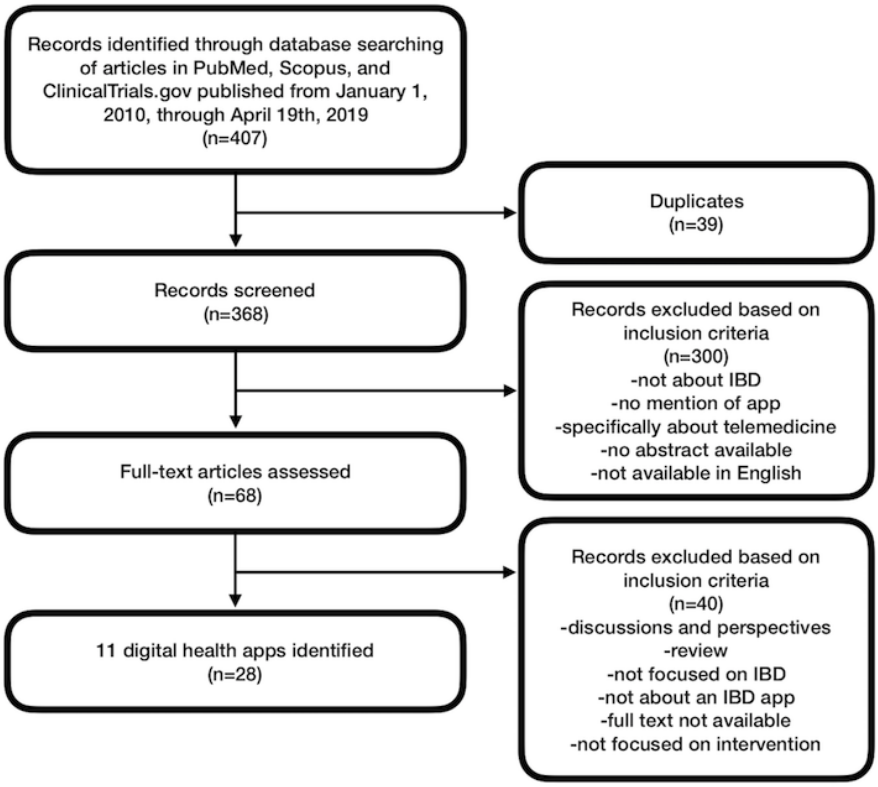
Results

Overview

The Scopus search yielded 227 *hits*. The PubMed search returned 168 *hits*. Search of ClinicalTrials.gov returned 12 *hits*. This totaled 407 *hits* from the three databases. A total of 39 duplicates were identified. The titles and abstracts of the remaining 368 articles were reviewed; 68 were identified to meet the inclusion criteria and were examined fully, after which 28 articles were selected (see Figure 1). Exploring citations revealed further information to be included. In total, 11 digital health apps and 4 commercial apps were included.

The five areas of clinical care discussed are education, monitoring, treatment, follow-up, and patient satisfaction. The sections below provide an overview of each area, a review of the clinically studied apps, and a review of a small selection of commercially available apps. An overview of each clinically studied app and the trials' significant findings can be found in Table 1 [26,32-55]. An overview of each commercial app can be found in Textbox 1 [56-59]. An overview of the features of each app as related to clinical care can be found in Tables 2 and 3.

Figure 1. Selection process for articles about inflammatory bowel disease (IBD) digital health apps.



XSL•FO
RenderX

Table 1. Overview of clinically studied apps.

App (sample size)	Feature highlights	Significant findings from trials
Constant Care (n=333, n=95, and n=17) [32-35]	• Weekly disease activity: SCCAI[a] and FC[b] results are mapped to "traffic light" color score • Educational modules • Portal for providers	• Fewer outpatient and acute care visits • Shorter flares and improved QoL[c] and medication adherence • Timely administration of medications • Individualized infliximab timing without compromising QoL
Young Constant Care (n=50) [36,55]	• Same as Constant Care, with one change: disease activity measured with Pediatric Crohn's Disease Activity Index (PCDAI) or Pediatric Ulcerative Colitis Activity Index (PUCAI)	• Individualized infliximab timing without compromise of QoL • Fewer doses given and longer average treatment interval
HealthPROMISE (n=320) [37-41]	• Biweekly disease activity (SIBDQ[d]) and QoL (EQ-5D[e]), graphically represented over time • Track health items such as vaccines, screenings, etc, to assess overall QoC[f] • Integrated into medical records • Patients can see the current plan and message their care team	• Improved QoL and QoC over 575 days • Fatigue and tension drivers of QoL • Patients report more equitable decision making • 75% still use at 6-month follow-up
IBD-live (n=84) [42,43]	• *Flarometer* portal collects PUCAI or PCDAI and FC at 1- or 3-month intervals • Communication module directly connected to local IBD[g] health team • Questionnaire module with surveys about QoL, absenteeism, and health care use • Patients and specialists alerted to out-of-range results	• No difference in flare occurrence • Improvement in QoL, although not reaching significance • Fewer face-to-face interactions with providers
myIBDcoach (n=909) [44,45]	• Weekly or monthly monitoring modules based on disease severity • *MonitorIBD at Home* questionnaire to assess disease activity • eLearning (electronic learning) module accompanies every session • Meeting disease activity thresholds will auto-alert health care providers • Patients can see their care plan and message their care team	• No change in QoL, flares, acute care visits, and QoC • Higher treatment adherence • Fewer outpatient visits and telephone calls to physicians • Fewer hospitalizations • 94% still using at 1-year follow-up
Telemonitoring of Crohn's Disease and Ulcerative Colitis (TECCU) (n=63) [46,47]	• Three-armed randomized controlled trial: control, nurse-assisted telephone care, and app group • App group: additional education, reminders, and prompted by text to report symptoms • App group: based on reported symptoms, developed alerts and treatment guidelines delivered to patients through the platform	• Disease activity and remission status improved most in app-monitored group • QoL and medical adherence improved in all groups • Differences in work productivity and activity impairment did not reach statistical significance • Fewer calls and outpatient visits in app group • Improved satisfaction in app and control groups
TELE-IBD (n=348 and n=219) [48-50]	• Text message about symptoms (HBI[h] and SCCAI) and medication side effects • Remote changes in management possible if alert thresholds of disease are met • Texts to share information about medications, dosing, and frequency	• Improved disease activity and QoL in controls and users • Decrease in hospitalizations, but increase in electronic encounters, phone calls, and noninvasive diagnostic tests • No significant change in knowledge as measured by the CCKNOW[i] after adjusting for confounding variables
TrueColours UC (n=66) [51]	• Daily monitoring (SCCAI) sent by email, biweekly QoL (EQ-5D), and monthly FC • "Traffic light" monitoring and presentation of disease state • Able to input blood, pathology, endoscopy, and histology results • Treatment guidance given regarding 5-ASA[j] and topical rectal medications	• Algorithm predicts escalation in therapy with 95% accuracy • Patients reported feeling empowerment, improved awareness, and communication

XSL•FO
RenderX

App (sample size)	Feature highlights	Significant findings from trials
UC HAT/HAT (n=25 and n=47) [52,53]	• Weekly reporting of symptoms, well-being, medications, side effects, and weight • Web-based clinician portal • Users are able to message their team and print an action plan	• No difference in QoL, disease activity, or medical adherence • 56% (14/25) completion in UC HAT[k] • 91% adherence to HAT[l] over 6 months and 86% report no interference with daily routines
UCLA eIBD (n=194 UC[m] and n=217 CD[n]) [26,54]	• Only available to UCLA[o] IBD patients • Disease activity monitored with UCLA-developed four-question PROs[p] (mHI[q]) with alarms built in at set thresholds • Integrates PROs into medical records • Value quotient measures patient value vs cost over time	• Correlation of mHI with HBI and partial Mayo score for CD and UC, respectively • Inverse correlation with QoL as measured by SIBDQ

[a]SCCAI: Simple Clinical Colitis Activity Index.

[b]FC: fecal calprotectin.

[c]QoL: quality of life.

[d]SIBDQ: Short Inflammatory Bowel Disease Questionnaire.

[e]EQ-5D: EuroQol-5 Dimension questionnaire.

[f]QoC: quality of care.

[g]IBD: inflammatory bowel disease.

[h]HBI: Harvey Bradshaw Index.

[i]CCKNOW: Crohn's and Colitis Knowledge Score.

[J]5-ASA: 5-aminosalicylate.

[k]UC HAT: home automated telemanagement in ulcerative colitis.

[l]HAT: home automated telemanagement.

[m]CD: Crohn's disease.

[n]UC: ulcerative colitis.

[o]UCLA: University of California, Los Angeles.

[p]PRO: patient-reported outcome.

[q]mHI: Mobile Health Index.

XSL•FO
RenderX

Textbox 1. An overview of commercial apps and their key features.

GI Monitor [56]

- Log symptoms, meals, bowel habits, medications, and mood
- Tries to help users make insights into behavior and symptoms
- Data can be exported to be shared with physicians

ibd.care [57]

- Website with basics of inflammatory bowel disease (IBD), tips for talking with a care team, advice about lifestyle, and insurance tips
- Parallel site for physicians to learn about insurance and IBD management
- Created by PRIME, a recognized provider of continuing medical education, reviewed by the Academy of Managed Care Pharmacy, and supported by the Case Management Society of America

myIBD [58]

- Mobile app built to assist pediatric patient transition to adult care
- Learn module with videos, articles, and short quizzes
- Journey module for input of features related to disease history: diagnosis, physicians, medications, allergies, major events, goals, symptoms, etc
- Access to disease activity questionnaires that can be saved
- Roadmap of the illness can be shared with new physicians

Oshi [59]

- Shares wellness and symptom scores and insights related to associations with flares
- Tracks symptoms and behavioral information (eg, exercise) and can sync to fitness devices
- *Learn* component has articles, IBD-friendly recipes, and personal stories
- *Ask* component provides a space to message Oshi health professionals

XSL•FO
RenderX

Table 2. An overview of the features of each clinically studied app as related to clinical care.

App (sample size)	Clinical care features				
	Educate	Monitor	Treatment	Follow-up	Patient sentiments
Constant Care (n=333, n=95, and n=17) [32-35]	• Disease-specific lecture at onboarding • eLearning (electronic learning) modules • eHealth nurse • Video clips	• SCCAI[a], HBI[b], and FC[c] • "Traffic light" system	• Can suggest treatment for flares and maintenance • Portal for providers to track patients	• N/A[d]	• First trial (n=333): 88.8% said system was feasible and preferred using it over standard clinical care • Second trial (n=95): 100% were satisfied • Third trial (n=17): high satisfaction
Young Constant Care (n=50) [36,55]	• eLearning modules • eHealth nurse • Video clips	• PUCAI[e] and PCDAI[f] • Height and weight • "Traffic light" system	• Timing of infliximab maintenance therapy • Portal for providers to track patients	• N/A	• 74% of surveys completed over the course of the trial
HealthPROMISE (n=320) [37-41]	• N/A	• SIBDQ[g] and EQ-5D[h] every 2 weeks • QoC[i] (vaccinations, and routine check-ups) • Integrated into EMR[j]	• N/A	• Can message care team through platform • App displays plan of care	• Maintained 75% use at 1-year follow-up
IBD-live (n=84) [42,43]	• N/A	• *Flarometer* (PUCAI or PCDAI + FC) every 1 or 3 months • Disease state shown to patient and team	• N/A	• Platform indicates when to see care team • Can message local IBD[k] team through platform	• 96% reported time-savings • 71% wished to continue • Highly compliant patients averaged €360 in annual savings
myIBDcoach (n=909) [44,45]	• Interactive eLearning modules	• Monthly modules about disease, medication, satisfaction, and side effects • Uses MIAH[l] survey	• At symptom threshold, red flag nudges health care worker to check in	• Can communicate with their health care team • Can view personal health plan	• In largest trial (n=909), 94% continued to use at 1-year follow-up • No significant difference in satisfaction vs controls
Telemonitoring of Crohn's Disease and Ulcerative Colitis (TECCU) (n=63) [46,47]	• Educational material created by the researchers received through the platform	• Symptoms assessed through texts and questionnaires • Disease activity (HBI for CD[m] and SCCAI for UC[n])	• Reminders sent through platform • Automatic alerts to care managers	• Could send messages through the platform to health team	• Significant increase in satisfaction (modified Client Satisfaction Questionnaire) • No perceived breaches in privacy
TELE-IBD (n=348 and n=219) [48-50]	• Educational tips delivered through text message, both about general health and IBD specifically	• N/A	• Remote changes in management if symptom burden is met	• N/A	• 13.9% did not complete every other week • 19% did not complete every week

XSL•FO
RenderX

App (sample size)	Clinical care features				
	Educate	Monitor	Treatment	Follow-up	Patient sentiments
TrueColours UC (n=66) [51]	• N/A	• SCCAI daily • EQ-5D every 2 weeks • FC monthly	• Guidance about 5-ASA[o] dose and topical rectal medications • Index predicting need for escalation of therapy	• N/A	• N/A
UC HAT/HAT (n=25 and n=47) [52,53]	• Weekly educational packages	• Weekly symptom diary, medication effects, and weight	• Customized action plan based on reported symptoms	• Option to message their health care team	• 56% (14/25) completed UC HAT[p] • 91% continued use of HAT[q] over 6 months and 86% report no interference with daily routines
UCLA eIBD (n=194 UC, 217 CD) [26,54]	• N/A	• Self-developed four-question PRO[r] questionnaires	• Automated message to coordinator at certain disease threshold	• Contact a coach through the app	• N/A

[a]SCCAI: Simple Clinical Colitis Activity Index.

[b]HBI: Harvey Bradshaw Index.

[c]FC: fecal calprotectin.

[d]N/A: not applicable.

[e]PUCAI: Pediatric Ulcerative Colitis Activity Index.

[f]PCDAI: Pediatric Crohn's Disease Activity Index.

[g]SIBDQ: Short Inflammatory Bowel Disease Questionnaire.

[h]EQ-5D: EuroQol-5 Dimension questionnaire.

[i]QoC: quality of care.

[j]EMR: electronic medical record.

[k]IBD: inflammatory bowel disease.

[l]MIAH: monitor IBD at home.

[m]CD: Crohn's disease.

[n]UC: ulcerative colitis.

[o]5-ASA: 5-aminosalicylate.

[p]UC HAT: home automated telemanagement in ulcerative colitis.

[q]HAT: home automated telemanagement.

[r]PRO: patient-reported outcome.

XSL•FO
RenderX

Table 3. An overview of the features of each commercial app as related to clinical care.

App	Clinical care features		
	Educate	Monitor	Follow-up
GI Monitor [56]	• N/A[a]	• Symptom diary: bowel movements, stress, meals, weight, pain, and medications	• Can export data to share with physician
ibd.care [57]	• Modules: overview of IBD[b], treatment options, choosing therapy, managing IBD, navigating insurance, and education for providers	• N/A	• N/A
myIBD [58]	• Teaching videos • Topic-focused articles • Self-assessments	• Can download and complete SIBDQ[c], SCCAI[d], HBI[e], and PHQ-9[f]	• Can share records in app with provider
Oshi [59]	• Educational articles written by MDs (ie, physicians), PhDs, and patients	• Symptom diary: bowel movements, abdominal pain, overall well-being, stress, diet, sleep, and exercise	• Offers to contact physician if symptoms change significantly • *Ask* component allows messaging of Oshi health professionals

[a]N/A: not applicable.

[b]IBD: inflammatory bowel disease.

[c]SIBDQ: Short Inflammatory Bowel Disease Questionnaire.

[d]SCCAI: Simple Clinical Colitis Activity Index.

[e]HBI: Harvey Bradshaw Index.

[f]PHQ-9: 9-item Patient Health Questionnaire.

Education

After initial diagnosis, patients may learn more about their disease through interactions with their health team, materials provided by clinical providers, discussions with family or friends, or support groups, but the Internet is also a common source of information. Some physicians encourage patients to use curated websites as sources of information. Patient education benefits QoL and continuity of care, reduces patient anxiety and complications from illness, and increases treatment adherence [60]. After diagnosis, up to 86% of individuals diagnosed with a chronic condition will turn to the Internet for information [61], but it is generally accepted that IBD information on the Internet may be too hard to read and have inaccuracies [62-65].

The app myIBDcoach utilizes interactive eLearning (electronic learning) modules about medications, adherence, smoking cessation, nutrition, symptom management, fatigue, work productivity, anxiety, and depression [44]. In addition to eLearning modules, Constant Care gives a disease-specific lecture during the onboarding process and provides access to an eHealth nurse and educational video clips. Constant Care researchers report that this has been valuable in empowering IBD patients to perform individualized, self-administered therapy [35]. At the end of 12 months, a study of 333 participants showed significant improvement in general knowledge about IBD (P<.001) and medications (P=.001) in the Danish cohort of the study compared to controls, as measured by the Crohn's and Colitis Knowledge Score (CCKNOW); however, the same effects were not observed in the Irish cohort [32,66]. Home automated telemanagement in ulcerative colitis (UC HAT) developed its own curriculum based on materials from the Crohn's and Colitis Foundation of America (CCFA), providing educational packages with each weekly check-in [52]. Telemonitoring of Crohn's Disease and Ulcerative Colitis (TECCU) produced its own educational and preventative materials that were incorporated and available on the platform [46].

TELE-IBD used a text message-based curriculum to send educational information at various frequencies, providing tips related to both general (ie, vaccinations and screenings) and specific (ie, IBD medication side effects) health information. In a study of 219 participants, patients were randomized to receive either TELE-IBD messages at different frequencies or standard of care. Results measured by CCKNOW showed significant improvement in patients receiving messages every other week as compared to controls (P=.03) and greater changes in scores in participants with lower baselines (P<.01); however, after adjusting for race, site, and baseline, researchers found no significant changes between control and TELE-IBD groups [48-50].

Of the commercial apps, myIBD, though mainly built for transition of care between providers, contains a robust educational component with videos, short articles on many related topics, and quizzes to assess knowledge [58]. The app ibd.care is created by PRIME, a recognized provider of continuing medical education, reviewed by the Academy of Managed Care Pharmacy, and supported by the Case Management Society of America. This app covers a broad base of information about IBD, including basics of disease, treatment, tips for talking with your care team, navigating insurance, and how to align lifestyle with IBD goals. The site similarly aims

XSL•FO
RenderX

to educate health care providers about managing patients and navigating insurance [57]. Many apps direct patients to the CCFA, where a large breadth of curated IBD information is available [20].

The apps approach education through learning modules, videos, and automated text messages covering a wide array of IBD topics. Besides the TELE-IBD study, these apps have acted as a repository for educational materials to be accessed by patients when needed or when the platforms are used, but the materials are not delivered proactively.

Although education is a common component across almost all apps, there is minimal assessment of the information quality or the value that the information provides to patients. It would be interesting to assess whether patients using these apps rely less on unverified sources of information and, as a result, experience improvements in treatment adherence, QoL, anxiety, and complications. With the varieties of technology available, technologists may be able to engage with physicians to create more interactive and engaging educational materials that increase health literacy and empower patients in managing their care.

Monitoring

Generally speaking, monitoring a patient's symptoms occurs at discrete time points at certain frequencies (ie, they can vary from every few months to every year) through outpatient office visits, when physicians and patients discuss symptoms experienced since a prior visit. The subjective reports from patients and laboratory data may both be collected at these time points to monitor the state of a patient's disease. Understanding the symptoms that patients experience is a foundational component of clinical care, but physicians are not always effective at collecting this information [67]. One study observed that health care providers frequently misinterpret reported symptoms, leading to differences in how the patient and the physician perceive a patient's current state [68]. As a further challenge, the concordance between a patient's memory of experience and actual experience of gastrointestinal (GI) symptoms has been observed to be generally poor [69]. The insufficient identification and management of IBD flares is one reason that many patients have poor outcomes [49].

Using a Web-based monitoring package, Constant Care maps symptoms reported through the Simple Clinical Colitis Activity Index (SCCAI) and Harvey Bradshaw Index (HBI) questionnaires, along with fecal calprotectin (FC) measurements, to a "traffic light" color system, which illustrates disease activity based on a total inflammatory burden score. Patients with recent flares take the surveys daily or weekly until they enter the *green zone* (ie, remission). Patients already in the *green zone* report symptoms monthly. Studies involving Constant Care have shown improved QoL, shorter flares, and decreased acute care and office visits [26,28,32].

HealthPROMISE tracks patient symptoms and QoL biweekly using the Short Inflammatory Bowel Disease Questionnaire (SIBDQ) and the EuroQol-5 Dimension questionnaire (EQ-5D), respectively. Providers can view the data on a Web-based dashboard integrated with this provider's electronic medical record (EMR). Due to the integration, researchers believe that in-person office visits can focus more on quality of care (QoC) as opposed to eliciting symptom history, allowing for more meaningful goal-focused discussions [37]. In their trial with 320 participants, the results showed *fatigue* and *tension* as the most important drivers of QoL [41], with QoL having significantly improved for study patients as compared to controls after 575 days of follow-up [39].

IBD-live is a Web-based app comprised of three modules that were tested on adolescents. The first module focuses on monitoring using a *flarometer*, where patients report disease activity through the Pediatric Ulcerative Colitis Activity Index (PUCAI) or the Pediatric Crohn's Disease Activity Index (PCDAI) and send an FC sample to track their status, as mapped to a "traffic light" system. Patients receive email reminders to report into the module. In the study, the 84 users of the platform had no difference in experiencing flares (33% vs 34% in controls) but showed some improvement in QoL (+1.32 vs -0.32 in controls, measured using the IMPACT-III questionnaire, P=.27) [42,43].

In TECCU, the 21 people in the intervention group were monitored via an IBD-modified version of NOMHADhome, a technological system designed for managing chronically ill patients and accessible on computers, tablets, or mobile phones. Patients reported symptoms by answering questions sent via text message and accessed questionnaires on the platform to monitor disease activity, adverse effects, and medication adherence. Monitoring frequency varied depending on the therapy plan for each patient, ranging from every week to every 4 weeks. Measurement of disease activity (HBI for CD or SCCAI for UC), QoL (Inflammatory Bowel Disease Questionnaire 9), and productivity and activity impairment (Work Productivity and Activity Impairment questionnaire) were done at the beginning and the end of the 24-week study. Researchers observed greater improvement in disease activity in the TECCU group compared to standard care, but the results did not reach statistical significance. QoL and social impairment improved significantly in all groups in the study, with no significant differences between them [46,47].

In a study with 47 participants, UC HAT used a symptom diary to elicit information on weekly symptoms, medications, side effects, and weight. The study did not find significant improvement in disease activity or QoL over standard care [52]. The app myIBDcoach has a monthly monitoring module to collect information about disease activity as measured with the Monitor IBD at Home questionnaire, medication use, treatment adherence, treatment satisfaction, and side effects. In a study with 909 participants using myIBDcoach, there was no significant difference in disease activity [44]. UCLA eIBD developed two specific, four-question, patient-reported outcome (PRO) questionnaires, named the Mobile Health Index (mHI), that are integrated into the EMR and used to create a patient value quotient. A study of 194 UC and 217 CD patients compared their mHI to standard UC and CD clinical disease activity measures, showing significant correlation for each (P<.001). Both had a strong inverse correlation with QoL as measured by the SIBDQ (P<.001) [29,54]. TrueColours UC emails users to report symptoms every day through the SCCAI

XSL•FO
RenderX

and to report on QoL every 2 weeks through the EQ-5D. Combined with a monthly FC test, users tracked disease activity using the system's "traffic light" severity tracker [51].

In the commercial market, symptom diaries represent about 57% of the IBD-related apps available for English speakers on the Google Play and Apple app stores [29]. Apps like Oshi, GI Monitor, and myIBD have various methods to track symptoms, bowel movements, medications, meals, emotional state (ie, stress and anxiety), sleep or fatigue, and physical activity. In return, these apps graphically represent the data to help users track and potentially discover associations between symptoms and reported information [56,58,59].

All but two of the apps included in this review have some approach to collecting PROs. Most apps going through trials tend to collect data at discrete, weekly, biweekly, or monthly time points, using standard clinical questionnaires for IBD. Conversely, commercial apps tend to aim for daily data entry and ask patients to report more granular data (eg, specific symptoms, activities, and emotions). No studies have reported decreases in QoL as a result of app use, and many—Constant Care, HealthPROMISE, and TECCU—have reported improved QoL.

Treatment

IBD management, for both UC and CD, includes the following: induction therapy and maintenance therapy. The goal is first to control inflammation quickly and then to sustain that control over time. Symptoms, side effects, laboratory data, and imaging inform the choice of therapy. Clinical treatment guidelines provide algorithms for physicians to follow based on patient symptoms, laboratory data, and prior treatments. The recommended treatments can vary greatly depending on these factors, and the treatments may not consider unique patient treatment preferences or individual characteristics [17,18].

Constant Care can provide recommendations for medication management based on reported symptoms. In one trial with Constant Care, patients with UC who reported acute symptoms indicative of a flare received recommendations to initiate 5-aminosalicylate (5-ASA) for a certain duration. Based on their own prior experience and current treatment guidelines, patients could also select additional treatments. In this trial with 333 participants, 100% of patients in the treatment group received 5-ASA in response to a flare as compared to 10% of control patients. Researchers attributed this benefit to the patient's ability to better recognize and understand the correct treatment for a flare [32,35]. Constant Care has also been studied in the context of down-titration of mesalazine therapy in mild-to-moderate UC patients in a trial with 95 participants. Patients were encouraged to decrease or maintain their dose or reach out to their care team based on the "traffic light" system. Results showed a significant increase in adherence to mesalazine (Visual Analog Scale 88 vs 100, Medical Adherence Rating Scale 23 vs 24, $P<.001$) in the group from baseline to the end of the study; results also showed significantly improved QoL ($P<.001$) as measured by the SIBDQ and 12-Item Short Form Health Survey (SF-12) from baseline to the end [33]. A third trial with 17 participants used Constant Care to determine the timing of infliximab maintenance therapy for CD, using weekly entries converted to the "traffic light" system. At the end of the trial, only 10% of individuals received therapy at 8-week intervals (ie, standard of care), 39% received treatment at shorter intervals, and 50% received treatment at longer intervals [34].

A similar randomized controlled trial of 50 patients used Young Constant Care—an adapted version of Constant Care—in a pediatric population. Based on symptom reporting and FC, patients reaching the *red* level or 2 weeks of the *yellow* level received an infusion of infliximab (ie, standard maintenance care). Otherwise, patients were allowed a maximum of 12 weeks between infusions. Using these methods, researchers observed a significant increase in the mean treatment interval (9.5 weeks vs 6.9 weeks, $P<.001$) and no change in QoL or levels of antibodies between treatment and controls who received standard care [36].

In TECCU, the responses provided through the platform were used to create individually tiered alerts and action plans. Patients received recommendations to adjust medication or visit their physician based on the alert level matching their reported symptoms. At the end of the trial, the HBI for CD or the SCCAI for UC were combined with partial Mayo scores from face-to-face visits to assess remission status. After 24 weeks, the 21 patients in the intervention group had no significant change in remission in UC (odds ratio [OR] 0.12, $P=.19$) or CD (OR 0.11, $P=.13$) or in mean improvement in measured FC (-0.90 mg/g, $P=.11$) [46,47].

In HealthPROMISE, patients have reported feeling that their care decision making has been more equitable and have experienced improved health outcomes (QoC increased +19% more than controls, $P<.01$) [40]. TrueColours UC used the data collected during its study of 66 participants to create a flare algorithm to help predict the need for escalation in therapy at outpatient visits and observed 95% accuracy of this algorithm [26].

Current evidence about the role of apps in IBD treatment comes largely from four different studies involving the Web-based platform Constant Care. Suggesting the inception or alteration of a medication through the use of IBD digital health apps has been viable. Certain studies have shown significantly increased QoL and have explored individualizing treatment timing. None of these studies have reported worsening QoL, QoC, flares, or treatment outcomes in participants.

The uses of Constant Care to titrate treatment for flares or maintenance therapy are exciting and have been recognized as an early step in the pursuit of pharmacokinetic monitoring [70]. One single treatment of a biologic therapy like infliximab can cost thousands of dollars and require an entire day in an infusion center [71]. The prospect of individualizing treatments according to when they are needed could benefit patients (ie, reduced medication exposure and time lost from daily life) as well as the health care system (ie, cost avoidance, fewer visits, and fewer side effects). Further studies are needed to validate this potential in IBD digital health apps.

Follow-Up

After diagnosis, patients with chronic diseases will regularly visit their physician to share updates about symptoms and side

XSL•FO
RenderX

effects. Unexpected or worsening symptoms may warrant scheduling additional visits to alter or add therapies. Severe changes may require visits to the emergency room or hospitalizations for immediate care. Higher numbers of outpatient visits can serve as a significant protective factor against IBD-related hospitalization in the following year [72], but nonadherence to medication is common in IBD and reported to be in the range of 40%-60%, with related adverse economic and clinical implications [73]. Appropriate continuity of medical care is associated with higher patient satisfaction, fewer hospitalizations, fewer emergency room visits, and improvement in receiving preventative services [74].

During testing of the Constant Care app, researchers observed an increase in the amount of contacts over the phone and through email, but a reduction in visits to the emergency department and the same number of hospitalizations [28]. The Constant Care trial also noted improved adherence to medication during flares (73% vs 42%, *P*=.005) [32]. UC HAT had a messaging option to a nurse coordinator through their Web portal but showed no changes in adherence and did not assess visits or hospitalizations [52]. Over 12 months, the trial of the myIBDcoach system showed a significant reduction in both visits to the gastroenterologist (1.55 vs 2.34 average visits over a year, *P*<.001) and telephone calls to the gastroenterologist (0.58 vs 0.84 average calls over a year, *P*<.001) compared to controls; however, there was no change in visits (0.29 vs 0.36, *P*=.17) or telephone calls (0.70 vs 0.74, *P*=.45) to the nurse. Users of the system had fewer hospitalizations (*P*=.046) and higher medication adherence, as measured by the Morisky Medication Adherence Scale (*P*<.001) [44].

In TECCU, patients could use the platform to send messages to their teams and receive direction about when to follow up or adjust medications. Researchers observed a significant increase in medication adherence in all groups in the trial, with significantly higher increase in the patients in the app group as measured by the Morisky-Green index (OR 0.0001). Researchers noted lower numbers of outpatient visits and telephone calls in the intervention group compared to controls [46,47].

Users of the IBD-live platform were triaged to different follow-up plans depending on their reported symptoms. Patients considered low risk would report symptoms again in 3 months, those at intermediate risk would report again in 1 month, and those at high risk would be directed to contact their physician. Patients were also able to contact their local health team through the platform. Among the 84 users of the Web-based system, there were significantly fewer face-to-face follow-ups as compared to controls (3.6 visits vs 4.3 in controls, *P*<.001) [42,43].

Among the commercial apps, MyIBD, aims to support the transition from a pediatric to an adult gastroenterologist. Problems with transition from pediatric to adult care can lead to treatment nonadherence, increased disease severity, and undue emotional and financial stress for patients [75]. Data show that 79% of adult gastroenterologists report inadequate preparation of adolescents coming from pediatric care [76]. The MyIBD app provides a *My Journey* module where patients can record all aspects of their own health record and grant access to new physicians they meet in their care [58]. Oshi has a feature where users can give the app permission to contact their physician if reported symptoms may indicate worsening disease [59].

The above apps provide another way for patients to interact with their care, whether through direct messaging or the ability to view information about their disease. Medication adherence is generally improved in the studied app users. These users also benefit from decreases in acute care and outpatient visits, although results are mixed about changes in telephone contact with providers. No studies reported increased hospitalizations or emergency visits or decreased medication adherence.

Patient Satisfaction

There is no globally agreed-upon formulation for patient satisfaction [77], and determinants of patient satisfaction vary across different studies, providing little explanation for the influencers of satisfaction [78-80]. One review of mobile health in managing digestive diseases determined that patient satisfaction ranged from 74% to 100% in the reviewed studies, with compliance ranging from 25.7% to 100% [81]. IBD patients want to be involved in decision making, with many reporting the desire for equitable collaboration with their physician [82]. In one set of focus groups, IBD patients reported a lack of understanding of how well their disease was being controlled, including a feeling that QoL was not discussed in many visits with their physicians. The patients reported an overuse of jargon and felt a lack of tangible goals or goal-setting in their care [83].

At the end of the largest Constant Care trial, 88.8% of individuals said the system was feasible and wanted to continue using it [32]. In another trial, 100% of individuals who finished the study reported being satisfied [33]. In a third trial, users reported high satisfaction with Web programs, education, and impact of the program on their disease [34]. The app myIBDcoach noted that 94% of users continued using the platform at the 1-year follow-up but saw no significant difference in satisfaction, as measured by a Visual Analog Scale in Web platform patients versus those receiving standard of care [44]. For UC HAT, only 14 of the 25 (56%) participants completed the study [52]; in the following version, home automated telemanagement (HAT), 86% reported that using the system did not interfere with their daily routines, 91% would consider using it in the future, and 91% were adherent to using the platform for 6 months [53]. In TELE-IBD, a posttrial qualitative study of both adherent and nonadherent patients identified benefits of understanding disease, monitoring symptoms, and feeling connected to their health provider. This study also noted that many participants had trouble remembering details of their action plans and that there were mixed results regarding timing, repetition, and technical aspects of the platform [84]. HealthPROMISE continued to have 75% adherence after 6 months of follow-up [37,41]. Participants of the Young Constant Care trial completed 74% of total desired survey entries [55]. In TECCU, patients in the control and intervention groups reported significant increases in satisfaction from baseline, as measured by an adapted version of the Client Satisfaction Questionnaire. Patients and providers also reported no perceived privacy breaches and minimal technical issues [46,47]. In IBD-live, 96% of users reported the platform to be

XSL•FO
RenderX

time-saving, 71% wished to continue, and highly compliant patients averaged €360 in annual savings [43].

Patients must remain the center of iteration and development. App adherence in the above studies usually means that users continue to use the app over the time frame stated, but researchers have not described the quantity of use and its possible relationship with patient progress. Although adherence has generally been adequate and satisfaction generally positive, studies have described little beyond these simple measures: a shortcoming observed across much of digital health [85]. As in the TELE-IBD posttrial study, narrative feedback will be important in improving specifics related to apps (ie, design and function) and answering larger questions about the perceived shortcomings of interventions. Clinical guidelines are able to advise specific treatments and medications but struggle to account for specific patient characteristics, such as treatment preference, access to care, childbearing interest, age, treatment history, etc. This can lead to disagreements or misunderstandings between patients and physicians. Digital health may be able to empower patients to better understand some of this context and help providers be more aware of patient preferences.

Discussion

Principal Findings

As in many chronic diseases, multiple influences—genetics, medications, behavior, social network, environment, psychological factors, and social determinants—play a role in the course of IBD, greatly increasing the number of variables and potential interventions available for study [8,86-88]. In such a large dimensional context, a randomized controlled trial (RCT) could assess whether an intervention, on average, has some effect, but it is unlikely to determine which of the components led to the observed effect. As seen in this review, adherence, QoL, QoC, and knowledge, among others, were valuable outcomes measured in the studies, but determining which components of each app contributed to these measurements poses a difficult challenge.

Given the breadth of variables and opportunities possible with digital health, RCTs have some limits. Microrandomized trials may be an interesting method to explore moving forward, as they randomly assign intervention options at relevant decision points. This allows for assessment of the effects of each intervention, including when and for whom it causes effects, as well as the examination of factors influencing these effects [89]. This may be a path to better understanding effects of individual components of increasingly complicated digital health apps.

New technologies have the potential to change the care of IBD. One such example is augGI, a technology company whose aim is to improve the management of chronic gut disorders. The company augGI is developing technology that uses computer vision and deep learning to characterize stool specimens from just an image. In particular, they are focusing on measuring stool consistency to better characterize motility changes [90]. Another area is the iteration on symptom monitoring. Many of the reviewed apps use written surveys converted into mobile text versions. Making these questionnaires visual, adaptive, or more specific to individuals could make the data more valuable and individually meaningful, as have been used in other areas [91-93]. On the horizon, toilets may be capable of collecting various data on urine and stool, making some data available more frequently [94-96]. Finally, apps for patients to assess FC at home (ie, IBDoc or Calpro) have been developed and have been positively validated, in general [28,97-99].

As promising as these technologies are, they bring to the forefront the lack of clinical guidelines for what types of new data should be collected and the appropriate frequency of data collection. In the CALM trial, FC was measured at 12-week intervals, already an increase in frequency from the norm. But what if FC is regularly measured every week or every day? As these questions arise, it is vital to utilize datasets that may already exist [100] and find new sets that lead to meaningful, cost-effective guidance for patients. Technologists will be challenged to design devices that elicit and present data streams with clinical relevance. Researchers will be challenged to build clinical guidelines and frameworks for translating these data streams into patient recommendations in real time as the data become available. One prospective study is exploring this dilemma using Fitbits to passively collect daily steps, heart rate, and sleep data and to determine if this data can help predict elevation in biomarkers. One of their early findings has shown that decreased physical activity, as measured by steps the week before, has occurred prior to the finding of active disease ($P<.001$) [101].

The clinical care of IBD is shifting from symptom-based to inflammation-based management. As digital health evolves, it becomes hard to ignore its potential to contribute to this shift. Digital health can help engage patients to track both the visible (ie, symptoms they experience) and invisible (ie, FC or other biomarkers) markers of disease. What a patient feels symptomatically is not fully descriptive of their disease state. Clinicians and technologists alike must pursue other lab or digitally trackable biomarkers to better describe the disease state. In tracking these, patients can receive treatment early in the course of flares, thereby reducing the need for acute care and the risk of long-term complications. Longitudinal assessment of FC or other markers will be able to not only reshape treatment but also allow for a more interactive, goal-focused dialogue between patients and providers.

Summary

This review discussed the role of studied IBD digital health apps and a small sample of commercial apps in clinical care. Significant benefits have been observed in education, QoL, QoC, treatment adherence, and medication management with the use of some apps. While digital health technologies have shown an ability to fit into, complement, and improve the standard clinical care of patients with IBD, research to further validate these findings from both a clinical and patient perspective is needed. As technologies change, research must expand to define new norms for using the different kinds of data that can be collected and integrated into clinical care. As the clinical management paradigm changes from symptom-based to inflammation-based care, it is an important time for all groups involved—patients, clinicians, technologists, insurers, etc—to

XSL•FO
RenderX

discuss and explore new opportunities to use digital health to improve understanding of disease, patient experience, and patient care.

Acknowledgments

We would like to acknowledge DE and EJS as the senior coauthors of this paper. We give a special thanks to the Small Data Lab at Cornell Tech as well as to those who supported the conception, discussion, and revision of the paper.

Conflicts of Interest

DH is cofounder and CEO of augGI Technologies Inc. EJS receives research funding from the CCFA and the UCSF-CCFA Clinical Research Alliance and also consults with the CCFA. The remaining authors have no conflicts of interest to declare.

References

1. The Mobile Economy: 2019. London, UK: GSM Association; 2019. URL: https://www.gsmaintelligence.com/research/?file=b9a6e6202ee1d5f787cfebb95d3639c5&download [accessed 2019-07-29]
2. Fox S, Duggan M. Pew Research Center. Washington, DC: Pew Internet & American Life Project; 2013 Jan 15. Health online 2013 URL: https://www.pewinternet.org/2013/01/15/health-online-2013/ [accessed 2019-07-29]
3. Makovsky. 2015 Feb 24. Fifth annual "Pulse of Online Health" Survey finds 66% of Americans eager to leverage digital tools to manage personal health URL: http://www.makovsky.com/news/fifth-annual-pulse-of-online-health-survey-2/ [accessed 2019-04-30] [WebCite Cache ID 781TqZHky]
4. mHealth App Economics 2017: Current Status and Future Trends in Mobile Health. Berlin, Germany: Research2Guidance; 2017 Nov. 325,000 mobile health apps available in 2017: Android now the leading mHealth platform URL: https://research2guidance.com/325000-mobile-health-apps-available-in-2017/ [accessed 2019-04-30] [WebCite Cache ID 781UkkFy4]
5. Con D, Jackson B, Gray K, De Cruz P. eHealth for inflammatory bowel disease self-management: The patient perspective. Scand J Gastroenterol 2017 Sep;52(9):973-980. [doi: 10.1080/00365521.2017.1333625] [Medline: 28598210]
6. Foster KR, Callans DJ. Smartphone apps meet evidence-based medicine: The future of medicine may (or may not) be in your smartphone. IEEE Pulse 2017;8(6):34-39. [doi: 10.1109/MPUL.2017.2750783] [Medline: 29155376]
7. Murray E, Hekler EB, Andersson G, Collins LM, Doherty A, Hollis C, et al. Evaluating digital health interventions: Key questions and approaches. Am J Prev Med 2016 Nov;51(5):843-851 [FREE Full text] [doi: 10.1016/j.amepre.2016.06.008] [Medline: 27745684]
8. 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 Nov;51(5):816-824. [doi: 10.1016/j.amepre.2016.05.001] [Medline: 27745681]
9. Bradway M, Ribu L, Hartvigsen G, Årsand E. The evolution of clinicians' preparedness for mHealth use (2013-2017) and current barriers. In: Proceedings from the 16th Scandinavian Conference on Health Informatics 2018. Linköping, Sweden: Linköping University Electronic Press; 2018 Presented at: 16th Scandinavian Conference on Health Informatics 2018; August 28-29, 2018; Aalborg, Denmark p. 45-50 URL: http://www.ep.liu.se/ecp/151/008/ecp18151008.pdf
10. Kane JM. Technology-based interventions in health care. Epidemiol Psychiatr Sci 2014 Dec;23(4):323-326. [doi: 10.1017/S2045796014000444] [Medline: 25154596]
11. Ward BW, Schiller JS. Prevalence of multiple chronic conditions among US adults: Estimates from the National Health Interview Survey, 2010. Prev Chronic Dis 2013 Apr 25;10:e65 [FREE Full text] [doi: 10.5888/pcd10.120203] [Medline: 23618545]
12. Burttorff C, Ruder R, Bauman M. Multiple Chronic Conditions in the United States. Santa Monica, CA: The RAND Corporation; 2017. URL: https://sbgg.org.br//informativos/29-06-17/1497877975_1_Chronic_Conditions.pdf [accessed 2019-04-30] [WebCite Cache ID 781dlzXsP]
13. Molodecky NA, Soon IS, Rabi DM, Ghali WA, Ferris M, Chernoff G, et al. Increasing incidence and prevalence of the inflammatory bowel diseases with time, based on systematic review. Gastroenterology 2012 Jan;142(1):46-54.e42; quiz e30. [doi: 10.1053/j.gastro.2011.10.001] [Medline: 22001864]
14. Shivashankar R, Tremaine WJ, Harmsen WS, Loftus EV. Incidence and prevalence of Crohn's disease and ulcerative colitis in Olmsted County, Minnesota from 1970 through 2010. Clin Gastroenterol Hepatol 2017 Jun;15(6):857-863 [FREE Full text] [doi: 10.1016/j.cgh.2016.10.039] [Medline: 27856364]
15. Feagins LA, Iqbal R, Spechler SJ. Case-control study of factors that trigger inflammatory bowel disease flares. World J Gastroenterol 2014 Apr 21;20(15):4329-4334 [FREE Full text] [doi: 10.3748/wjg.v20.i15.4329] [Medline: 24764669]
16. Bernklev T, Jahnsen J, Lygren I, Henriksen M, Vatn M, Moum B. Health-related quality of life in patients with inflammatory bowel disease measured with the short form-36: Psychometric assessments and a comparison with general population norms. Inflamm Bowel Dis 2005 Oct;11(10):909-918. [doi: 10.1097/01.mib.0000179467.01748.99] [Medline: 16189421]

XSL•FO
RenderX

17. Rubin DT, Ananthakrishnan AN, Siegel CA, Sauer BG, Long MD. ACG Clinical Guideline: Ulcerative colitis in adults. Am J Gastroenterol 2019 Mar;114(3):384-413. [doi: 10.14309/ajg.0000000000000152] [Medline: 30840605]
18. Lichtenstein GR, Loftus EV, Isaacs KL, Regueiro MD, Gerson LB, Sands BE. ACG Clinical Guideline: Management of Crohn's disease in adults. Am J Gastroenterol 2018 Apr;113(4):481-517. [doi: 10.1038/ajg.2018.27] [Medline: 29610508]
19. Peyrin-Biroulet L, Loftus EV, Colombel J, Sandborn WJ. Long-term complications, extraintestinal manifestations, and mortality in adult Crohn's disease in population-based cohorts. Inflamm Bowel Dis 2011 Jan;17(1):471-478. [doi: 10.1002/ibd.21417] [Medline: 20725943]
20. Crohn's & Colitis Foundation. URL: https://www.crohnscolitisfoundation.org/resources/ [accessed 2019-04-30] [WebCite Cache ID 781XKU8qW]
21. De Dombal FT, Watts JM, Watkinson G, Goligher JC. Local complications of ulcerative colitis: Stricture, pseudopolyposis, and carcinoma of colon and rectum. Br Med J 1966 Jun 11;1(5501):1442-1447 [FREE Full text] [doi: 10.1136/bmj.1.5501.1442] [Medline: 5933046]
22. Lutgens MW, van Oijen MG, van der Heijden GJ, Vleggaar FP, Siersema PD, Oldenburg B. Declining risk of colorectal cancer in inflammatory bowel disease: An updated meta-analysis of population-based cohort studies. Inflamm Bowel Dis 2013;19(4):789-799. [doi: 10.1097/MIB.0b013e31828029c0] [Medline: 23448792]
23. Viladomiu M, Kivolowitz C, Abdulhamid A, Dogan B, Victorio D, Castellanos JG, et al. IgA-coated E. coli enriched in Crohn's disease spondyloarthritis promote T17-dependent inflammation. Sci Transl Med 2017 Feb 08;9(376):1-27 [FREE Full text] [doi: 10.1126/scitranslmed.aaf9655] [Medline: 28179509]
24. Peyrin-Biroulet L, Sandborn W, Sands BE, Reinisch W, Bemelman W, Bryant RV, et al. Selecting therapeutic targets in inflammatory bowel disease (STRIDE): Determining therapeutic goals for treat-to-target. Am J Gastroenterol 2015 Sep;110(9):1324-1338. [doi: 10.1038/ajg.2015.233] [Medline: 26303131]
25. Colombel J, Panaccione R, Bossuyt P, Lukas M, Baert F, Vaňásek T, et al. Effect of tight control management on Crohn's disease (CALM): A multicentre, randomised, controlled phase 3 trial. Lancet 2018 Dec 23;390(10114):2779-2789. [doi: 10.1016/S0140-6736(17)32641-7] [Medline: 29096949]
26. Walsh A, Travis S. What's app? Electronic health technology in inflammatory bowel disease. Intest Res 2018 Jul;16(3):366-373 [FREE Full text] [doi: 10.5217/ir.2018.16.3.366] [Medline: 30090035]
27. Atreja A, Otobo E, Ramireddy K, Deorocki A. Remote patient monitoring in IBD: Current state and future directions. Curr Gastroenterol Rep 2018 Mar 07;20(2):6. [doi: 10.1007/s11894-018-0611-3] [Medline: 29516186]
28. Patil SA, Cross RK. Current landscape of telemedicine practice in inflammatory bowel disease. Inflamm Bowel Dis 2018 Sep;24(9):1910-1917. [doi: 10.1093/ibd/izy113] [Medline: 29718218]
29. Kelso M, Feagins LA. Can smartphones help deliver smarter care for patients with inflammatory bowel disease? Inflamm Bowel Dis 2018 Jun 08;24(7):1453-1459. [doi: 10.1093/ibd/izy162] [Medline: 29868764]
30. Arksey H, O'Malley L. Scoping studies: Towards a methodological framework. Int J Soc Res Methodol 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]
31. Levac D, Colquhoun H, O'Brien KK. Scoping studies: Advancing the methodology. Implement Sci 2010 Sep 20;5:69 [FREE Full text] [doi: 10.1186/1748-5908-5-69] [Medline: 20854677]
32. Elkjaer M, Shuhaibar M, Burisch J, Bailey Y, Scherfig H, Laugesen B, et al. E-health empowers patients with ulcerative colitis: A randomised controlled trial of the Web-guided 'Constant-care' approach. Gut 2010 Dec;59(12):1652-1661. [doi: 10.1136/gut.2010.220160] [Medline: 21071584]
33. Pedersen N, Thielsen P, Martinsen L, Bennedsen M, Haaber A, Langholz E, et al. eHealth: Individualization of mesalazine treatment through a self-managed Web-based solution in mild-to-moderate ulcerative colitis. Inflamm Bowel Dis 2014 Dec;20(12):2276-2285. [doi: 10.1097/MIB.0000000000000199] [Medline: 25248002]
34. Pedersen N, Elkjaer M, Duricova D, Burisch J, Dobrzanski C, Andersen NN, et al. eHealth: Individualisation of infliximab treatment and disease course via a self-managed Web-based solution in Crohn's disease. Aliment Pharmacol Ther 2012 Nov;36(9):840-849 [FREE Full text] [doi: 10.1111/apt.12043] [Medline: 22971016]
35. Ankersen DV, Carlsen K, Marker D, Munkholm P, Burisch J. Using eHealth strategies in delivering dietary and other therapies in patients with irritable bowel syndrome and inflammatory bowel disease. J Gastroenterol Hepatol 2017 Mar;32 Suppl 1:27-31. [doi: 10.1111/jgh.13691] [Medline: 28244677]
36. Carlsen K, Jakobsen C, Houen G, Kallemose T, Paerregaard A, Riis LB, et al. Self-managed eHealth disease monitoring in children and adolescents with inflammatory bowel disease: A randomized controlled trial. Inflamm Bowel Dis 2017 Mar;23(3):357-365. [doi: 10.1097/MIB.0000000000001026] [Medline: 28221247]
37. Atreja A, Khan S, Rogers JD, Otobo E, Patel NP, Ullman T, HealthPROMISE Consortium Group. Impact of the mobile HealthPROMISE platform on the quality of care and quality of life in patients with inflammatory bowel disease: Study protocol of a pragmatic randomized controlled trial. JMIR Res Protoc 2015 Feb 18;4(1):e23 [FREE Full text] [doi: 10.2196/resprot.4042] [Medline: 25693610]
38. Atreja A, Khan S, Rogers J, Ullman A, Grinspan A, Itzkowitz S, et al. Impact of real-world home-based remote monitoring on quality of care and quality of life in inflammatory bowel disease patients: One-year results of pragmatic randomized trial. J Crohns Colitis 2017 Feb 01;11(Suppl 1):S362-S363 [FREE Full text] [doi: 10.1093/ecco-jcc/jjx002.678]

XSL•FO
RenderX

39. Atreja A, Otobo E, Szigethy E, Kohli A, Shroff H, Chang H, et al. Improved quality of care and quality of life for IBD patients using mobile-based remote monitoring platform: A randomized control trial. Inflamm Bowel Dis 2018 Feb;24(Suppl 1):S21-S22. [doi: 10.1093/ibd/izy019.063]
40. Atreja A, Szigethy E, Otobo E, Chang HL, Keefer L, Rogers J, et al. Improved quality of care and quality of life for IBD patients using the HealthPROMISE app: A randomized, control trial. Gastroenterology 2018 May;154(6):S-6. [doi: 10.1016/s0016-5085(18)30506-7]
41. Atreja A, Otobo E, Chang H, Keefer L, Rogers J, Ullman T, et al. HealthPROMISE: Utilization of patient-reported outcomes to measure quality of life in inflammatory bowel disease. iProc 2017 Sep 22;3(1):e28. [doi: 10.2196/iproc.8452]
42. Heida A, Dijkstra A, Muller Kobold A, Rossen JW, Kindermann A, Kokke F, et al. Efficacy of home telemonitoring versus conventional follow-up: A randomized controlled trial among teenagers with inflammatory bowel disease. J Crohns Colitis 2018 Mar 28;12(4):432-441. [doi: 10.1093/ecco-jcc/jjx169] [Medline: 29228230]
43. Dijkstra A, Heida A, van Rheenen PF. Exploring the challenges of implementing a Web-based telemonitoring strategy for teenagers with inflammatory bowel disease: Empirical case study. J Med Internet Res 2019 Mar 29;21(3):e11761 [FREE Full text] [doi: 10.2196/11761] [Medline: 30924785]
44. de Jong MJ, van der Meulen-de Jong AE, Romberg-Camps MJ, Becx MC, Maljaars JP, Cilissen M, et al. Telemedicine for management of inflammatory bowel disease (myIBDcoach): A pragmatic, multicentre, randomised controlled trial. Lancet 2017 Sep 02;390(10098):959-968. [doi: 10.1016/S0140-6736(17)31327-2] [Medline: 28716313]
45. de Jong M, van der Meulen-de Jong A, Romberg-Camps M, Degens J, Becx M, Markus T, et al. Development and feasibility study of a telemedicine tool for all patients with IBD: MyIBDcoach. Inflamm Bowel Dis 2017 Apr;23(4):485-493. [doi: 10.1097/MIB.0000000000001034] [Medline: 28267047]
46. Aguas M, Del Hoyo J, Faubel R, Muñoz D, Domínguez D, Bastida G, et al. A Web-based telemanagement system for patients with complex inflammatory bowel disease: Protocol for a randomized controlled clinical trial. JMIR Res Protoc 2018 Dec 21;7(12):e190 [FREE Full text] [doi: 10.2196/resprot.9639] [Medline: 30578197]
47. Del Hoyo J, Nos P, Faubel R, Muñoz D, Domínguez D, Bastida G, et al. A Web-based telemanagement system for improving disease activity and quality of life in patients with complex inflammatory bowel disease: Pilot randomized controlled trial. J Med Internet Res 2018 Nov 27;20(11):e11602 [FREE Full text] [doi: 10.2196/11602] [Medline: 30482739]
48. Cross RK, Jambaulikar G, Langenberg P, Tracy JK, Collins JF, Katz J, et al. TELEmedicine for patients with Inflammatory Bowel Disease (TELE-IBD): Design and implementation of randomized clinical trial. Contemp Clin Trials 2015 May;42:132-144. [doi: 10.1016/j.cct.2015.03.006] [Medline: 25812483]
49. Cross RK, Langenberg P, Regueiro M, Schwartz DA, Tracy JK, Collins JF, et al. A randomized controlled trial of TELEmedicine for patients with Inflammatory Bowel Disease (TELE-IBD). Am J Gastroenterol 2019 Mar;114(3):472-482. [doi: 10.1038/s41395-018-0272-8] [Medline: 30410041]
50. Abutaleb A, Buchwald A, Chudy-Onwugaje K, Langenberg P, Regueiro M, Schwartz DA, et al. Inflammatory bowel disease telemedicine clinical trial: Impact of educational text messages on disease-specific knowledge over 1 year. Inflamm Bowel Dis 2018 Sep 15;24(10):2191-2197. [doi: 10.1093/ibd/izy149] [Medline: 29788137]
51. Walsh A. TrueColours: Real Time Data Collection in Patients With Ulcerative Colitis [doctoral thesis]. Oxford, UK: University of Oxford; 2017. URL: https://ora.ox.ac.uk/objects/uuid:a4ab55af-5364-4fa4-86ca-e84b917d7f70/download_file?file_format=pdf&safe_filename=DPhil%2BWALSH%2BTRUECOLOURS%2BFEB%2B2018.pdf&type_of_work=Thesis [accessed 2019-07-29]
52. Cross RK, Cheevers N, Rustgi A, Langenberg P, Finkelstein J. Randomized, controlled trial of home telemanagement in patients with ulcerative colitis (UC HAT). Inflamm Bowel Dis 2012 Jun;18(6):1018-1025 [FREE Full text] [doi: 10.1002/ibd.21795] [Medline: 21688350]
53. Cross RK, Finkelstein J. Feasibility and acceptance of a home telemanagement system in patients with inflammatory bowel disease: A 6-month pilot study. Dig Dis Sci 2007 Feb;52(2):357-364. [doi: 10.1007/s10620-006-9523-4] [Medline: 17211702]
54. van Deen WK, van der Meulen-de Jong AE, Parekh NK, Muyshondt Y, Kane E, Eimers L, et al. Remote monitoring of IBD disease activity using the mobile Health Index (mHI): A validation study. Gastroenterology 2015 Apr;148(4 Suppl 1):S-446. [doi: 10.1016/s0016-5085(15)31501-8]
55. 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 Sep;23(9):1473-1482. [doi: 10.1097/MIB.0000000000001170] [Medline: 28617758]
56. Apple App Store. GI Monitor by Vertical Health, LLC URL: https://itunes.apple.com/us/app/gi-monitor/id315539155?mt=8 [accessed 2019-07-29]
57. ibd.care. Fort Lauderdale, FL: PRIME Education URL: https://ibd.care/ [accessed 2019-05-02] [WebCite Cache ID 784QNbHxE]
58. Apple App Store. MyIBD URL: https://itunes.apple.com/us/app/myibd/id1383119954?mt=8 [accessed 2019-04-30]
59. Apple App Store. Oshi: IBD Tracker & Magazine by Oshi Health, Inc URL: https://itunes.apple.com/us/app/oshi-ibd-tracker-magazine/id1371752119?mt=8 [accessed 2019-04-30]
60. Bastable SB. Essentials Of Patient Education. 2nd edition. Burlington, MA: Jones & Bartlett Learning; 2017.

XSL•FO
RenderX

61. Fox S. Pew Research Center. Washington, DC: Pew Internet & American Life Project; 2007 Oct 08. E-patients with a disability or chronic disease URL: https://www.pewinternet.org/2007/10/08/e-patients-with-a-disability-or-chronic-disease/ [accessed 2019-04-30] [WebCite Cache ID 781VN7tLj]
62. Promislow S, Walker JR, Taheri M, Bernstein CN. How well does the Internet answer patients' questions about inflammatory bowel disease? Can J Gastroenterol 2010 Nov;24(11):671-677 [FREE Full text] [doi: 10.1155/2010/957264] [Medline: 21157582]
63. Siegel CA. Embracing the Internet for progress in shared decision-making. Inflamm Bowel Dis 2007 Dec;13(12):1579-1580. [doi: 10.1002/ibd.20259] [Medline: 17879279]
64. Azer SA, AlOlayan TI, AlGhamdi MA, AlSanea MA. Inflammatory bowel disease: An evaluation of health information on the Internet. World J Gastroenterol 2017 Mar 07;23(9):1676-1696 [FREE Full text] [doi: 10.3748/wjg.v23.i9.1676] [Medline: 28321169]
65. van Mierlo T, Fournier R, Fedorak R. Don't forget the doctor: Gastroenterologists' preferences on the development of mHealth tools for inflammatory bowel disease. JMIR Mhealth Uhealth 2015 Jan 21;3(1):e5 [FREE Full text] [doi: 10.2196/mhealth.3987] [Medline: 25608628]
66. Eaden JA, Abrams K, Mayberry JF. The Crohn's and Colitis Knowledge Score: A test for measuring patient knowledge in inflammatory bowel disease. Am J Gastroenterol 1999 Dec;94(12):3560-3566. [doi: 10.1111/j.1572-0241.1999.01536.x] [Medline: 10606319]
67. Fung CH, Hays RD. Prospects and challenges in using patient-reported outcomes in clinical practice. Qual Life Res 2008 Dec;17(10):1297-1302 [FREE Full text] [doi: 10.1007/s11136-008-9379-5] [Medline: 18709564]
68. Pittet V, Maillard MH, Simonson T, Fournier N, Rogler G, Michetti P. Differences in outcomes reported by patients with inflammatory bowel diseases vs their health care professionals. Clin Gastroenterol Hepatol 2018 Nov 22:1. [doi: 10.1016/j.cgh.2018.11.029] [Medline: 30471455]
69. Jones MP, Walter S, Faresjö Å, Grodzinsky E, Kjellström L, Viktorsson L, et al. Gastrointestinal recall questionnaires compare poorly with prospective patient diaries for gastrointestinal symptoms: Data from population and primary health centre samples. Eur J Gastroenterol Hepatol 2019 Feb;31(2):163-169. [doi: 10.1097/MEG.0000000000001296] [Medline: 30394943]
70. Dreesen E, Gils A. Pharmacodynamic monitoring of biological therapies in chronic inflammatory diseases. Ther Drug Monit 2019 Apr;41(2):131-141. [doi: 10.1097/FTD.0000000000000571] [Medline: 30883507]
71. Yokomizo L, Limketkai B, Park KT. Cost-effectiveness of adalimumab, infliximab or vedolizumab as first-line biological therapy in moderate-to-severe ulcerative colitis. BMJ Open Gastroenterol 2016;3(1):e000093 [FREE Full text] [doi: 10.1136/bmjgast-2016-000093] [Medline: 27195130]
72. Limsrivilai J, Stidham RW, Govani SM, Waljee AK, Huang W, Higgins PD. Factors that predict high health care utilization and costs for patients with inflammatory bowel diseases. Clin Gastroenterol Hepatol 2017 Mar;15(3):385-392.e2 [FREE Full text] [doi: 10.1016/j.cgh.2016.09.012] [Medline: 27645518]
73. Lakatos PL. Prevalence, predictors, and clinical consequences of medical adherence in IBD: How to improve it? World J Gastroenterol 2009 Sep 14;15(34):4234-4239 [FREE Full text] [doi: 10.3748/wjg.15.4234] [Medline: 19750566]
74. Cabana MD, Jee SH. Does continuity of care improve patient outcomes? J Fam Pract 2004 Dec;53(12):974-980. [Medline: 15581440]
75. Gray WN, Resmini AR, Baker KD, Holbrook E, Morgan PJ, Ryan J, et al. Concerns, barriers, and recommendations to improve transition from pediatric to adult IBD care: Perspectives of patients, parents, and health professionals. Inflamm Bowel Dis 2015 Jul;21(7):1641-1651. [doi: 10.1097/MIB.0000000000000419] [Medline: 25966837]
76. Sebastian S, Jenkins H, McCartney S, Ahmad T, Arnott I, Croft N, et al. The requirements and barriers to successful transition of adolescents with inflammatory bowel disease: Differing perceptions from a survey of adult and paediatric gastroenterologists. J Crohns Colitis 2012 Sep;6(8):830-844. [doi: 10.1016/j.crohns.2012.01.010] [Medline: 22398082]
77. Batbaatar E, Dorjdagva J, Luvsannyam A, Savino MM, Amenta P. Determinants of patient satisfaction: A systematic review. Perspect Public Health 2017 Mar;137(2):89-101. [doi: 10.1177/1757913916634136] [Medline: 27004489]
78. Naidu A. Factors affecting patient satisfaction and healthcare quality. Int J Health Care Qual Assur 2009;22(4):366-381. [doi: 10.1108/09526860910964834] [Medline: 19725209]
79. Vranceanu A, Ring D. Factors associated with patient satisfaction. J Hand Surg Am 2011 Sep;36(9):1504-1508. [doi: 10.1016/j.jhsa.2011.06.001] [Medline: 21794990]
80. Sitzia J, Wood N. Patient satisfaction: A review of issues and concepts. Soc Sci Med 1997 Dec;45(12):1829-1843. [doi: 10.1016/s0277-9536(97)00128-7] [Medline: 9447632]
81. Helsel BC, Williams JE, Lawson K, Liang J, Markowitz J. Telemedicine and mobile health technology are effective in the management of digestive diseases: A systematic review. Dig Dis Sci 2018 Jun;63(6):1392-1408. [doi: 10.1007/s10620-018-5054-z] [Medline: 29663265]
82. Baars JE, Markus T, Kuipers EJ, van der Woude CJ. Patients' preferences regarding shared decision-making in the treatment of inflammatory bowel disease: Results from a patient-empowerment study. Digestion 2010;81(2):113-119. [doi: 10.1159/000253862] [Medline: 20093836]

XSL•FO
RenderX

83. Khan S, Dasrath F, Farghaly S, Otobo E, Riaz MS, Rogers J, Health PROMISE Consortium Group. Unmet communication and information needs for patients with IBD: Implications for mobile health technology. Br J Med Med Res 2016;12(3):12119 [FREE Full text] [doi: 10.9734/BJMMR/2016/21884] [Medline: 28217532]
84. Quinn CC, Chard S, Roth EG, Eckert JK, Russman KM, Cross RK. The Telemedicine for patients with Inflammatory Bowel Disease (TELE-IBD) clinical trial: Qualitative assessment of participants' perceptions. J Med Internet Res 2019 Jun 03;21(6):e14165 [FREE Full text] [doi: 10.2196/14165] [Medline: 31162128]
85. Maramba I, Chatterjee A, Newman C. Methods of usability testing in the development of eHealth applications: A scoping review. Int J Med Inform 2019 Jun;126:95-104. [doi: 10.1016/j.ijmedinf.2019.03.018] [Medline: 31029270]
86. Loddo I, Romano C. Inflammatory bowel disease: Genetics, epigenetics, and pathogenesis. Front Immunol 2015;6:551 [FREE Full text] [doi: 10.3389/fimmu.2015.00551] [Medline: 26579126]
87. Gerbarg PL, Jacob VE, Stevens L, Bosworth BP, Chabouni F, DeFilippis EM, et al. The effect of breathing, movement, and meditation on psychological and physical symptoms and inflammatory biomarkers in inflammatory bowel disease: A randomized controlled trial. Inflamm Bowel Dis 2015 Dec;21(12):2886-2896. [doi: 10.1097/MIB.0000000000000568] [Medline: 26426148]
88. Matsuoka K, Kanai T. The gut microbiota and inflammatory bowel disease. Semin Immunopathol 2015 Jan;37(1):47-55 [FREE Full text] [doi: 10.1007/s00281-014-0454-4] [Medline: 25420450]
89. Klasnja P, Hekler EB, Shiffman S, Boruvka A, Almirall D, Tewari A, et al. Microrandomized trials: An experimental design for developing just-in-time adaptive interventions. Health Psychol 2015 Dec;34S:1220-1228 [FREE Full text] [doi: 10.1037/hea0000305] [Medline: 26651463]
90. auggi.: augGI Technologies Inc URL: http://www.auggi.ai/ [accessed 2019-04-30] [WebCite Cache ID 781Y4gfFq]
91. Selter A, Tsangouri C, Ali SB, Freed D, Vatchinsky A, Kizer J, et al. An mHealth app for self-management of chronic lower back pain (Limbr): Pilot study. JMIR Mhealth Uhealth 2018 Sep 17;6(9):e179 [FREE Full text] [doi: 10.2196/mhealth.8256] [Medline: 30224333]
92. Yang L, Freed D, Wu A, Wu J, Pollak J, Estrin D. Your activities of daily living (YADL): An image-based survey technique for patients with arthritis. In: Proceedings of the 10th European Alliance for Innovation (EAI) International Conference on Pervasive Computing Technologies for Healthcare. 2016 May 16 Presented at: 10th European Alliance for Innovation (EAI) International Conference on Pervasive Computing Technologies for Healthcare; May 16-19, 2016; Cancun, Mexico p. 41-44.
93. Pollak J, Adams P, Gay G. PAM: A photographic affect meter for frequent, in situ measurement of affect. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2011 May 07 Presented at: SIGCHI Conference on Human Factors in Computing Systems; May 7-12, 2011; Vancouver, BC p. 725-734.
94. Ranjitkar P. Toilet lab: Diagnostic tests on smart toilets? Clin Chem 2018 Jun 28;64(7):1128-1129. [doi: 10.1373/clinchem.2018.286567]
95. Wald C. Diagnostics: A flow of information. Nature 2017 Nov 08;551(7679):S48-S50. [doi: 10.1038/551S48a] [Medline: 29117161]
96. Alm E, Turgeman Y, Ratti C. WIRED. 2014 Mar 21. 'Smart toilets and sewer sensors are coming' URL: https://www.wired.co.uk/article/yaniv-j-turgeman [accessed 2019-04-30] [WebCite Cache ID 781Ybuv18]
97. Haisma S, Galaurchi A, Almahwzi S, Adekanmi Balogun JA, Muller Kobold AC, van Rheenen PF. Head-to-head comparison of three stool calprotectin tests for home use. PLoS One 2019;14(4):e0214751 [FREE Full text] [doi: 10.1371/journal.pone.0214751] [Medline: 30998692]
98. Heida A, Knol M, Kobold AM, Bootsman J, Dijkstra G, van Rheenen PF. Agreement between home-based measurement of stool calprotectin and ELISA results for monitoring inflammatory bowel disease activity. Clin Gastroenterol Hepatol 2017 Nov;15(11):1742-1749.e2 [FREE Full text] [doi: 10.1016/j.cgh.2017.06.007] [Medline: 28606846]
99. Vinding KK, Elsberg H, Thorkilgaard T, Belard E, Pedersen N, Elkjaer M, et al. Fecal calprotectin measured by patients at home using smartphones: A new clinical tool in monitoring patients with inflammatory bowel disease. Inflamm Bowel Dis 2016 Feb;22(2):336-344. [doi: 10.1097/MIB.0000000000000619] [Medline: 26535869]
100. Chung AE, Sandler RS, Long MD, Ahrens S, Burris JL, Martin CF, et al. Harnessing person-generated health data to accelerate patient-centered outcomes research: The Crohn's and Colitis Foundation of America PCORnet Patient Powered Research Network (CCFA Partners). J Am Med Inform Assoc 2016 May;23(3):485-490 [FREE Full text] [doi: 10.1093/jamia/ocv191] [Medline: 26911821]
101. Sossenheimer P, Yvellez O, Andersen MJ, Pearl TA, Jurdi KE, Rubin DB, et al. Wearable devices can predict disease activity in inflammatory bowel disease patients. Gastroenterology 2019 May;156(6):S-111 [FREE Full text] [doi: 10.1016/S0016-5085(19)37069-6]

Abbreviations

5-ASA: 5-aminosalicylate
CCFA: Crohn’s and Colitis Foundation of America
CCKNOW: Crohn’s and Colitis Knowledge Score

XSL•FO
RenderX

CD: Crohn's disease
EMR: electronic medical record
EQ-5D: EuroQol-5 Dimension questionnaire
FC: fecal calprotectin
GI: gastrointestinal
HAT: home automated telemanagement
HBI: Harvey Bradshaw Index
IBD: inflammatory bowel disease
mHI: Mobile Health Index
MIAH: monitor IBD at home
N/A: not applicable
OR: odds ratio
PCDAI: Pediatric Crohn's Disease Activity Index
PHQ-9: 9-item Patient Health Questionnaire
PRO: patient-reported outcome
PUCAI: Pediatric Ulcerative Colitis Activity Index
QoC: quality of care
QoL: quality of life
RCT: randomized controlled trial
SCCAI: Simple Clinical Colitis Activity Index
SF-12: 12-Item Short Form Health Survey
SIBDQ: Short Inflammatory Bowel Disease Questionnaire
STRIDE: Selecting Therapeutic Targets in Inflammatory Bowel Disease
TECCU: Telemonitoring of Crohn's Disease and Ulcerative Colitis
UC: ulcerative colitis
UC HAT: home automated telemanagement in ulcerative colitis
UCLA: University of California, Los Angeles

Edited by G Eysenbach; submitted 09.05.19; peer-reviewed by K Agarwal, J del Hoyo Francisco; comments to author 01.06.19; revised version received 02.07.19; accepted 04.07.19; published 19.08.19.

Please cite as:
Yin AL, Hachuel D, Pollak JP, Scherl EJ, Estrin D
Digital Health Apps in the Clinical Care of Inflammatory Bowel Disease: Scoping Review
J Med Internet Res 2019;21(8):e14630
URL: http://www.jmir.org/2019/8/e14630/
doi:10.2196/14630
PMID:31429410

©Andrew Lukas Yin, David Hachuel, John P Pollak, Ellen J Scherl, Deborah Estrin. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 19.08.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

The Association Between Web-Based or Face-to-Face Lifestyle Interventions on the Perceived Benefits and Barriers to Exercise in Midlife Women: Three-Arm Equivalency Study

Amanda Mary McGuire[1], PhD; Charrlotte Seib[1], PhD; Janine Porter-Steele[1,2], PhD; Debra Jane Anderson[1], PhD

[1]Menzies Health Institute Queensland, Griffith University, Southport, Australia

[2]Wesley Hospital, Brisbane, Australia

Corresponding Author:
Amanda Mary McGuire, PhD
Menzies Health Institute Queensland
Griffith University
Parklands Drive
Southport, 4215
Australia
Phone: 61 7 5552 8860
Email: a.mcguire@griffith.edu.au

Abstract

Background: Noncommunicable diseases pose a significant threat to women's health globally, with most diseases being attributed to modifiable risk factors such as physical inactivity. Women perceive a range of benefits and barriers to exercise; however, there is little evidence about the effect of different lifestyle intervention delivery modes on perceptions of exercise.

Objective: This study aimed to compare the effect of a multiple health behavior change (MHBC) intervention called the Women's Wellness Program. This intervention was delivered in 3 different modes on perceived exercise benefits, perceived exercise barriers, and actual physical activity and exercise in midlife women.

Methods: Women aged 45 to 65 years were recruited via the study website. They were assigned in blocks to 3 different treatment groups (A: Web-based independent; B: face-to-face with nurse consultations; and C: Web-based with virtual nurse consultations). All participants received the 12-week intervention that utilizes principles from social-cognitive theory to provide a structured guide to promote healthy lifestyle behaviors with an emphasis on regular exercise and healthy eating. Data were collected using a self-report Web-based questionnaire at baseline (T1) and postintervention (T2) including perceived exercise benefits and barriers and exercise and physical activity. A data analysis examined both within- and between-group changes over time.

Results: Participants in this study (N=225) had a mean age of 50.9 years (SD 5.9) and most were married or living with a partner (83.3%, 185/225). Attrition was 30.2% with 157 participants completing the final questionnaire. Women in all intervention groups reported a significant increase in positive perceptions of exercise (P<.05); a significant increase in exercise and overall physical activity (P<.01) with moderate-to-large effect sizes noted for overall physical activity (d=0.5 to d=0.87). Participants receiving support from registered nurses in the face-to-face and Web-based groups had a greater magnitude of change in benefit perceptions and physical activity than those in the Web-based independent group. There was no significant change in exercise barrier perceptions within or between groups over time.

Conclusions: The results of this study suggest that the (MHBC) intervention is effective in increasing exercise benefit perceptions, overall physical activity, and exercise in midlife women. Although Web-based programs are cost-effective and flexible and can be delivered remotely, providing a range of options including face-to-face group delivery and personalized electronic health coaching from registered nurses has the potential to enhance participant engagement and motivation.

(J Med Internet Res 2019;21(8):e10963) doi:10.2196/10963

KEYWORDS

exercise; physical activity; women; health behavior; behavioral medicine; health promotion; digital health; benefits and barriers

XSL•FO
RenderX

Introduction

Background

Noncommunicable diseases (NCDs) pose a significant threat to women's health globally [1,2]. Recent estimates suggest that 4 NCDs including cardiovascular disease, cancer, respiratory disease, and type 2 diabetes account for the majority of premature deaths in women between the ages of 30 and 70 years [3]. Although it is clear that regular exercise has many physical and mental health benefits and is an important component of good health, many health promotion programs fail to adequately address the often-correlated nature of many modifiable risk factors [4]. For example, physical inactivity is often associated with other modifiable lifestyle risk factors such as an unhealthy diet, tobacco smoking, and overweight and obesity, with many adults having multiple risk factors for NCDs [4,5].

Among women, midlife is a time when the risk of developing an NCD increases [3], particularly among those who do not adhere to recommended physical activity and healthy eating guidelines [6,7] and often experience menopause-related weight gain [8,9]. Although there is no established definition of *midlife*, women between the ages of 40 and 65 years have normally finished childbearing and experience a physiological transition to perimenopause and menopause. Many women in this age bracket continue in paid employment with the average age that Australian women intend to retire increasing over the last decade to 64.4 years in 2017 [10]. According to Mishra et al, midlife (particularly perimenopause) is also a sensitive period when the cumulative effects of unhealthy lifestyle behaviors have a greater impact on disease risk [11]. Therefore, engaging in regular physical activity, eating a healthy diet, and maintaining a healthy body weight (body mass index [BMI]=18.5-25.0) [8] in midlife are essential to reduce risk and ensure optimal health and well-being as women age.

There is evidence that multiple health behavior change (MHBC) interventions tailored for women are effective in changing behavior [12,13]. Furthermore, given the multiple and complex role demands and stressors reported by women in this age group [14], flexible health promotion interventions have the potential to yield greater success. Over the past decade, *Web-based* or *internet* interventions also show promising results in promoting physical activity and healthy eating [15]. Moreover, though Web-based interventions targeting multiple health behaviors are fewer in number, there is evidence that they can provide an effective, flexible, cost-effective means of promoting healthy lifestyle behaviors [16-19].

Despite the potential efficacy of MHBC interventions, women perceive a range of benefits and barriers to changing exercise behavior [20-22]. Research suggests that perceived benefits of exercise include physical health and fitness, improved mental health and stress reduction, and reduced risk of illness [23-25]. Women's perception of barriers to exercise are often complex and relate to a range of personal, social, and environmental factors such as lack of time, motivation, family support and care-giving responsibilities, climate and physical safety [23-25]. Arguably, these perceptions are very important to consider when designing health promotion interventions, with evidence that benefit and barrier perceptions are correlated with actual exercise behavior change [23,26]. In relation to behavior change theory, the concepts of *perceived benefits* and *perceived barriers* equate to *positive outcome expectations* and *impediments to change* that influence *self-efficacy* beliefs as described by Bandura in the social cognitive theory [27].

Women's Wellness Program Intervention

To date, there is a paucity of evidence about how different intervention delivery modes effect benefit and barrier perceptions. In this context, the Women's Wellness Program (WWP) is a 12-week MHBC intervention designed for midlife women (Multimedia Appendix 1), targeting a range of modifiable risk factors including regular physical activity and exercise and healthy eating [12]. The original WWP that included a paper-based journal and 2 face-to-face nurse consultations was evaluated in a previous randomized controlled trial (N=90), finding that the intervention was effective in increasing physical activity, decreasing smoking, BMI, and weight [12]. On the basis of the social cognitive theory [27], the WWP includes detailed health education to promote health literacy and knowledge about the benefits of regular physical activity and exercise and incorporates strategies to overcome impediments/barriers to change, including realistic goal setting and health coaching to promote self-efficacy for exercise. The original Program was designed to be delivered face-to-face in a community practice setting by registered nurses trained to deliver the intervention. In this study, the WWP was revised and adapted to be delivered as a Web-based self-directed program hosted on a purpose-built website. The website also includes a separate Web portal for health consultations. Furthermore, a revised Program book was published both as an eBook and a hard copy. In this study, participants receiving health coaching received 4 nurse consultations at 0, 4, 8, and 12 weeks. Results in relation to the effect of the intervention on the primary outcome measure of climacteric symptoms are reported elsewhere [28].

Objectives

This study addresses secondary outcome measures, where the effect of the WWP intervention delivered in 3 different modes/arms (Web-based independent; face-to-face with nurse consultations; and Web-based with virtual nurse consultations) on perceived exercise benefits, perceived exercise barriers, and actual physical activity and exercise in midlife women is investigated.

Methods

Participants and Procedure

Participants were Australian women aged between 40 and 65 years. Details of study participants, recruitment procedures, inclusion/exclusion criteria, and attrition are described elsewhere [28]. In short, following media publicity about the study, participants were recruited across Australia via the study website. Women with an existing diagnosis of an NCD or without computer/internet access were excluded.

XSL•FO
RenderX

Measures

A self-report Web-based questionnaire was used to collect data from participants at baseline (T1) and postintervention (T2), including (1) sociodemographic information (T1 only), (2) perceived exercise benefits and barriers [25], (3) exercise and physical activity (Seattle Physical Activity questionnaire) [29], and (4) height (cm) and weight (kg). The BMI was calculated by dividing the weight in kilograms by the height in meters squared (kg/m^2) [30].

This paper presents the pre- and postintervention perceived benefits and barriers measured using the Exercise Benefits and Barriers Scale (EBBS) [25]. The EBBS is a 46-item instrument using a forced response Likert-type scale. The scale contains 29 benefit items summed to calculate a total benefit subscale score ($EBBS_{BEN}$), with higher scores indicating higher benefit perceptions. Benefit items are then grouped and summed to create benefit subcategories: life enhancement, physical performance, psychological outlook, social interaction, and preventive health. Example benefit items include the following: 32. Exercise improves my self-concept; 15. Exercise increases my level of physical fitness; 2. Exercise decreases feelings of stress and tension; 11. Exercise lets me have contact with friends and persons I enjoy; and 13. Exercise will keep me from having high blood pressure.

The 46-item scale contains 14 barrier items summed to get a total barrier subscale score ($EBBS_{BAR}$), with higher scores indicating higher barrier perceptions. Barrier items are grouped and summed to create barrier subcategories: exercise milieu, time expenditure, physical exertion, and family encouragement. Example barrier items include: 9. Places for me to exercise are too far away; 4. Exercising takes too much of my time; 37. Exercise takes too much time from my family responsibilities; 6. Exercise tires me; and 33. My family members do not encourage me to exercise.

The EBBS demonstrates good reliability and internal consistency in studies that investigate exercise benefits and barriers in women [20-22]. In this study, subscale reliability was calculated with a Cronbach alpha of .94 for the benefit subscale and .87 for the barrier subscale, indicating high internal consistency.

Additional anecdotal feedback about what participants liked and disliked about the Program was obtained and invitation to make *other comments* was given through 3 open-ended questions asked postintervention via the Web-based questionnaire.

Intervention

The participating women completed baseline questionnaires before being assigned in blocks to one of the 3 different treatment modality groups. The 12-week program utilizes principles from the social-cognitive theory [27] to provide participants with a structured guide to promote healthy lifestyle behaviors with an emphasis on regular physical activity and exercise, healthy eating, healthy weight, stress management, and health screening behaviors. In relation to exercise and physical activity, the intervention provides detailed evidence-based information in plain language about the current guidelines for physical activity [6]. Information about the multiple health benefits of regular physical activity (aerobic exercise, strength training, and stretching) on physical and mental health and the reducing risk of chronic disease is provided in the Program book and website and reinforced in nurse consultations. Photographic illustrations of strength exercises are provided with practical advice on starting and maintaining a regular physical activity schedule, with daily walking recommended as the starting activity for participants who are sedentary or unfit. Over the 12 weeks, participants are encouraged to gradually increase the frequency and intensity of physical activity and exercise, with Program 1 strength training exercise introduced in Week 2 and Program 2 strength exercises with dumbbells introduced in Week 5. The Program book and website include weekly activity planning with participants invited to identify and reflect on their barriers to exercise behavior change through journal activities, reflections, and discussions with a registered nurse. Participants receiving nurse consultations are supported to develop personalized goals for exercise and diet that are specific, measurable, achievable, relevant, and time bound at 0, 6, and 12 weeks.

The WWP intervention including the Program book, the Program website, and health consultations is delivered in 3 different formats: (1) *Web-based independent* (Arm A) had access to the WWP website that contains all of the information provided in the book and an electronic copy of the Program book only; (2) *Face-to-face group supported* (Arm B) included a hard copy of the Program book and 4 30- to 60-min face-to-face consultations provided by a registered nurse at 0, 4, 8, and 12 weeks; and (3) *Web-based supported* (Arm C) were able to access the WWP website, download an electronic copy of the Program book, and were also provided 4 virtual consultations through a portal built into the website at 0, 4, 8, and 12 weeks. Intervention fidelity was maintained through provision of structured facilitator training, through consistent record keeping and auditing, and by employing one registered nurse to deliver all consultations.

Statistical Analysis

Analyses were performed using IBM® SPPS Statistics version 22 [31]. Descriptive data are expressed as counts and percentages or mean (SD), whereas inferential statistics were performed using *t* tests and analysis of covariance (ANCOVA) or their nonparametric equivalent. The statistical significance was set at alpha=.05. The effect size was also calculated using the Cohen *d* standard formula [32] to examine the meaning and magnitude of change in the 3 groups over the study period. Using Cohen guidelines [32], an effect size of 0.20 was deemed to be small, an effect size of 0.50 was moderate, and an effect size of 0.80 or more was considered to be large.

Ethical Approval

Before recruitment and data collection, ethical approval was obtained from the relevant Human Research and Ethics Committee (QUT HREC Approval No: 1300000048). Participation was voluntary, and women were able to withdraw from the study at any time. Furthermore, participants did not receive any rewards or incentives for participation; however, they were able to retain the Program book and materials and have continued access to the website resources.

XSL•FO
RenderX

Results

Sociodemographic Characteristics and Baseline Body Mass Index and Physical Activity

Participants in this study had a mean age of 50.9 years (SD 5.9) and most were married or living in a de-facto relationship (83.3%, 185/225). Overall, 79.1% of women were Australian born, most worked either full- or part-time (53.0%, 119/225 or 29.5%, 66/225 respectively), and almost two-thirds (68.9%, 155/225) of participants were university educated. According to the World Health Organization categories [30], where overweight is classified as BMI 25.0 to 29.9 and obesity BMI greater than 30.0, many of the participating women reported being overweight (35.5%, 79/225) or obese (32.9%, 74/225). When asked about general daily activity levels (including housework, gardening, shopping, caring for children, or activity at work), the proportion of all participants at baseline who were very active was 1.8% (4/225); moderately active, 32.2% (72/225); mildly active, 46.2% (104/225); and sedentary, 19.1% (43/225). Reported levels of aerobic exercise (15 min at a time in the past month including brisk walking, jogging, swimming, and cycling) for all women at baseline found 21.8% (49/225) getting no aerobic exercise; 36% (81/225) one or 2 times/week; 27.1% (61/225) 3 or 4 times/week; 10.2% (23/225) 5 or 6 times/week; and 4.4% (10/225) daily exercise.

Although 225 women completed the baseline questionnaire, an attrition rate of 30.2% (68/225) meant that 157 women completed the final questionnaire following the 12-week intervention. A comparison of retained participants and those lost to follow-up showed significantly more participants being lost from the Web-based independent Arm A (35.5%, 49/138) compared with Arm B (9.7%, 4/41) and Arm C (30.2%, 15/46). The comparison of baseline sociodemographic and health characteristics of the 3 intervention groups showed no significant within- or between-group differences [30].

Table 1 presents the mean total exercise benefits, barriers and benefits, and barriers subscale scores pre- and postintervention within each group. Although there was no statistically significant change in barrier scores over time, in contrast there was a significant increase in average total benefits, psychological, and social subscale scores within all 3 intervention groups. A one-way between-group ANCOVA showed no significant difference between the 3 intervention groups over time for both total exercise benefit scores ($F_{2,153}$=0.30; P=.74; partial eta squared=0.004) and total exercise barrier scores ($F_{2,153}$=0.65; P=.52; partial eta squared=0.01). Figures 1 and 2 present the change over time in average total perceived benefit and barrier scores within and between groups.

Table 1. Mean perceived benefits and barriers to exercise scores within and between groups pre- and postintervention (N=157). Mean scores are based on an ordinal scale representing the extent to which women strongly disagree/disagree/agree/strongly agree that the item is a benefit/barrier to exercise (range 1 to 4 with high scores representing higher agreement).

Variable	Arm A: Web-based independent (n=89), mean (SD)		Arm B: Face-to-face supported (n=37), mean (SD)		Arm C: Web-based supported (n=31), mean (SD)	
	Pre	Post	Pre	Post	Pre	Post
Perceived benefits[a]						
Total benefits	88.5 (10.8)	91.4 (13.4)[b]	91.8 (9.5)	95.1 (14.4)[b]	91.9 (11.1)	93.5 (10.4)[b]
Life enhancement	3.1 (0.5)	3.2 (0.5)	3.3 (0.4)	3.3 (0.5)	3.3 (0.5)	3.2 (0.4)
Physical	3.4 (0.4)	3.4 (0.4)	3.4 (0.4)	3.4 (0.5)	3.5 (0.4)	3.4 (0.4)
Psychological	2.9 (0.4)	3.2 (0.5)[b]	3.0 (0.3)	3.3 (0.6)[b]	3.0 (0.3)	3.4 (0.4)[b]
Social	2.3 (0.6)	2.4 (0.7)[b]	2.5 (0.6)	2.7 (0.7)[b]	2.5 (0.6)	2.5 (0.5)[b]
Preventive health	3.3 (0.5)	3.3 (0.5)	3.4 (0.4)	3.4 (0.5)	3.3 (0.6)	3.3 (0.4)
Perceived barriers[c]						
Total barriers	29.6 (6.1)	29.0 (6.9)	26.9 (6.5)	26.1 (5.3)	27.3 (5.6)	27.1 (6.2)
Exercise milieu	1.9 (0.5)	1.8 (0.5)	1.7 (0.5)	1.7 (0.4)	1.7 (0.5)	1.7 (0.5)
Time expenditure	2.2 (0.6)	2.2 (0.6)	2.0 (0.6)	1.9 (0.5)	2.1 (0.5)	2.1 (0.7)
Physical exertion	2.6 (0.6)	2.6 (0.6)	2.4 (0.6)	2.3 (0.5)	2.3 (0.5)	2.3 (0.5)
Family encourage	1.8 (0.7)	1.9 (0.9)	1.7 (0.8)	1.7 (0.7)	1.7 (0.5)	1.7 (0.5)

[a]Exercise Benefits Subscale.

[b]Within-group paired *t* test; P<.05.

[c]Exercise Barriers Subscale.

XSL•FO
RenderX

Figure 1. Change over time in average perceived benefits of exercise scores within and between groups.

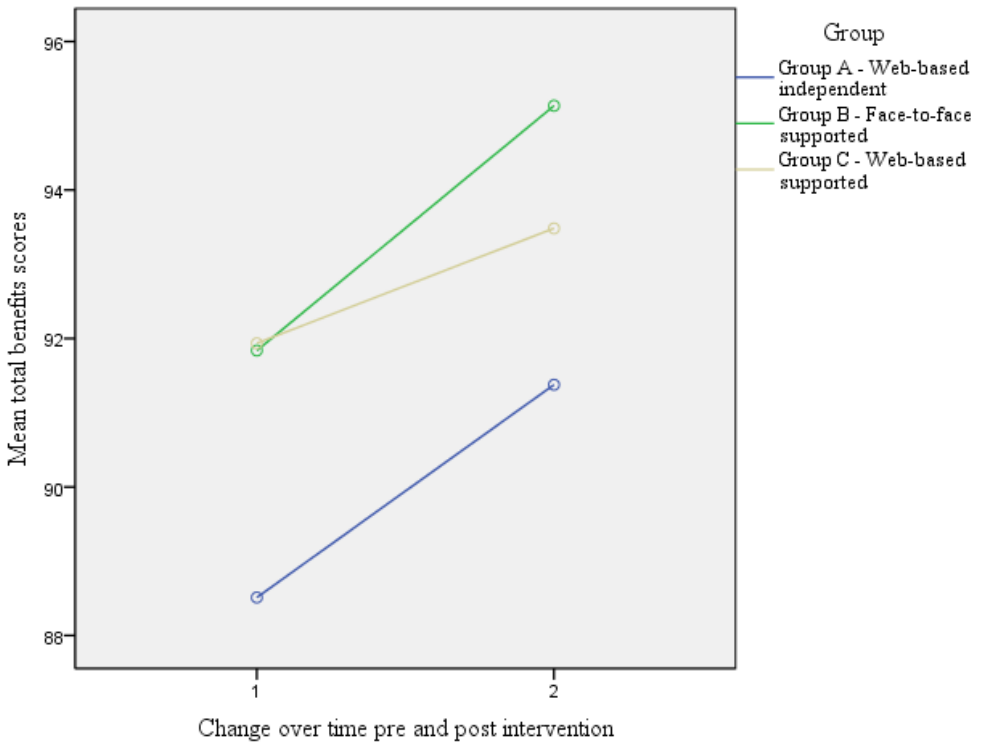
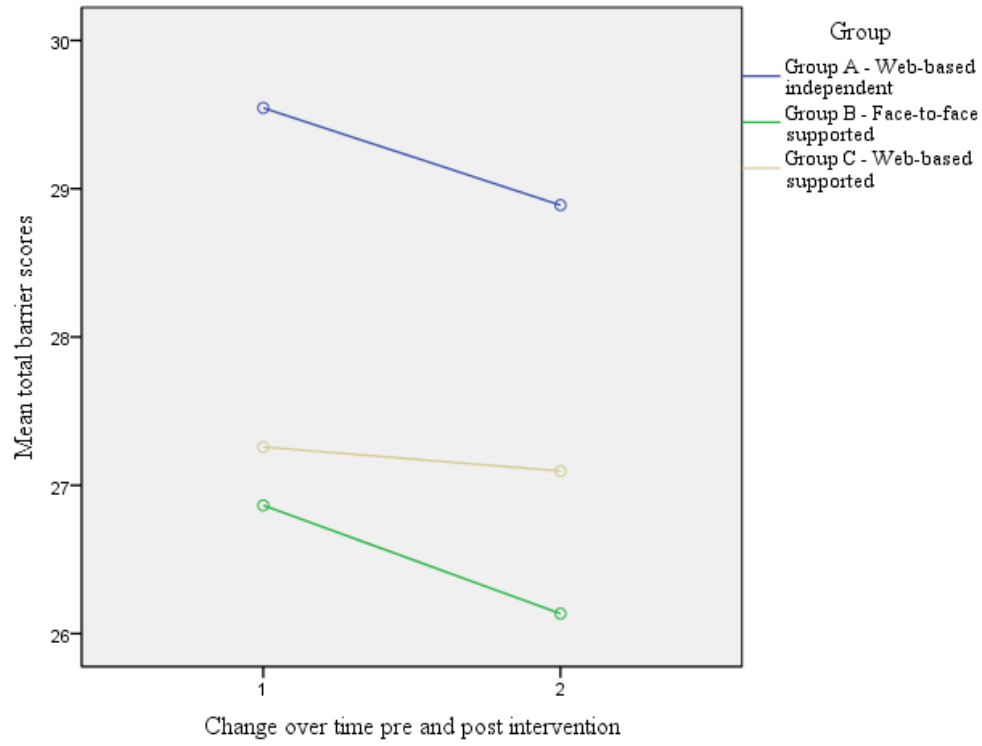


Figure 2. Change over time in average perceived barriers to exercise scores within and between groups.



Changes in overall physical activity, aerobic exercise, and general daily activity are presented in Table 2. There was a significant difference in all physical activity variables within all groups postintervention (P<.01). Between-group comparison of overall physical activity was close to statistical significance ($F_{2,148}$=45.1; P=.052; partial eta squared=0.04), with a greater increase in face-to-face supported Arm B and Web-based supported Arm C compared with Web-based independent Arm A.

XSL•FO
RenderX

Table 2. Comparison of overall physical activity, aerobic exercise, and general daily activity within groups pre- and postintervention (N=157).

Variable	Arm A: Web-based independent (n=89)		Arm B: Face-to-face supported (n=37)		Arm C: Web-based supported (n=31)	
	Pre	Post	Pre	Post	Pre	Post
Overall physical activity (PA)[a]						
Mean (SD)[b]	5.4 (1.8)	6.3 (2.1)[c]	5.8 (1.9)	6.9 (1.9)[c]	5.6 (1.9)	7.3 (1.5)[c]
Median[b]	6.0	7.0[c]	6.0	8.0[c]	6.0	8.0[c]
Weekly aerobic exercise , n (%)						
2 or less times weekly	55.7 (49)	33.7 (30)	59.5 (22)	24.3 (9)	64.5 (20)	29.0 (9)
3-4 times weekly	31.8 (28)	36.0 (32)	21.6 (8)	35.1 (13)	25.8 (8)	29.0 (9)
5+ times weekly	12.5 (12)	30.3 (27)	18.9 (7)	40.5 (15)	9.7 (3)	41.9 (13)
McNemar test	—[d]	21.3[c]	—	14.3[c]	—	14.5[c]
General daily activity , n (%)						
Sedentary	18.0 (16)	6.7 (6)	13.9 (5)	8.1 (3)	22.6 (7)	6.5 (2)
Mildly active	49.4 (44)	36.0 (32)	52.8 (19)	29.7 (11)	35.5 (11)	22.6 (7)
Moderately/very active	32.6 (29)	57.3 (51)	33.3 (12)	62.2 (23)	41.9 (13)	71.0 (22)
McNemar test	—	25.7[c]	—	9.3[c]	—	8.6[c]

[a]Overall physical activity including exercise and general daily activity measured on a scale of 0 to 10.

[b]Paired *t* test and Wilcoxon Signed Rank test.

[c]*P*<.01.

[d]Not applicable.

Effect Size

Table 3 presents results of the effect size analysis within and between groups over time in perceived benefits and barriers to exercise and overall physical activity. A small effect for perceived barriers to exercise was observed, with a small-to-moderate effect for perceived benefits to exercise within all 3 groups postintervention (Cohen d_{change}). A moderate-to-large effect was seen in overall physical activity, within all intervention groups. Using Cohen d_2 to compare the difference in effect size between Arm A (Web-based independent) and Arm B (face-to-face supported) and Arm C (Web-based supported), there was a small-to-moderate effect size observed for all variables.

To further illustrate and compare the magnitude of change over time within and between each of the 3 groups, Table 4 presents exercise benefits and barriers subscale variables and overall physical activity, grouped by effect size and study arm. Of note is the large effect size for psychological benefits in all groups and overall physical activity in the Web-based supported group C. There was a moderate effect size observed for overall physical activity in the Web-based independent (Arm A) and face-to-face supported (Arm B) groups.

XSL•FO
RenderX

Table 3. Effect size within and between groups over time in perceived benefits and barriers to exercise and overall physical activity (N=157).

Variables	Cohen d_{change}[a]			Cohen d_2[b]	
	Arm A: Web-based independent	Arm B: Face-to-face supported	Arm C: Web-based supported	A-B post	A-C post
Perceived benefits[c]					
Total benefits	0.27	0.30	0.14	0.28	0.16
Life enhancement	0.02	0.00	0.20	0.20	0.00
Physical	0.00	0.00	0.25	0.00	0.00
Psychological	0.70	1.0	1.3	0.20	0.40
Social	0.21	0.33	0.13	0.43	0.14
Preventive health	0.00	0.00	0.00	0.20	0.00
Perceived barriers[d]					
Total barriers	0.10	0.11	0.03	0.42	0.28
Exercise milieu	0.20	0.00	0.00	0.01	0.20
Time expenditure	0.00	0.20	0.00	0.50	0.20
Physical exertion	0.00	0.20	0.00	0.50	0.42
Family encourage	−0.14	0.00	0.00	0.22	0.22
Overall physical activity	0.50	0.60	0.87	0.31	0.48

[a]Cohen d_{change} compared the difference in the effect size within groups over time.

[b]Cohen d_2 compared the difference in effect size between groups postintervention.

[c]Exercise Benefits Subscale.

[d]Exercise Barriers Subscale.

Table 4. Comparison of magnitude of change postintervention within each group for exercise benefits and barriers subscale scores and overall physical activity.

Effect size[a]	Arm A: Web-based independent	Arm B: Face-to-face supported	Arm C: Web-based supported
Large >.7	Psychological benefits	Psychological benefits	Psychological benefits; Overall PA
Moderate .4 to .7	Overall PA	Overall PA	—[b]
Small .2 to .4	Total benefits score; Social benefits; Exercise milieu barriers	Social benefits; Total benefits score; Time barriers; Physical exertion barriers	Physical benefits; Life enhancement benefits
Very small <.2	Family barriers; Total barriers score; Life enhancement benefits	Total barriers score	Total benefits score; Social benefits; Total barriers score

[a]Cohen d_{change} compared the difference in effect size with groups over time.

[b]Not applicable.

Discussion

Principal Findings

This study has reported the results of a 3-arm MHBC intervention on perceived exercise benefits and barriers and self-reported physical activity and exercise in midlife women. Postintervention, women in all 3 arms reported a significant increase in overall exercise benefit perceptions and increased physical activity.

With regard to exercise benefits, there was a significant change in perceptions about the psychological and social benefits in particular, with large effect sizes being noted. This change was associated with a significant increase in overall physical activity and exercise, with moderate-to-large effect sizes across all 3 groups. What is striking about these results is the proportion of participants who moved from lower levels of exercise to reporting regular exercise on 5 or more days per week postintervention. These changes in benefit perceptions and actual physical activity are likely to be a result of specific program content and strategies, including detailed health promotion information contained in the Program book, individualized goal setting, and weekly exercise planning activities undertaken by all participants over the 12 weeks of the trial.

XSL•FO
RenderX

When comparing the magnitude of change between groups over time, the Web-based independent group (Arm A) was used as the comparison group. In comparison with the Web-based independent group, both the face-to-face supported group (Arm B) and Web-based supported group (Arm C) had greater change in benefit and barrier perceptions and overall physical activity postintervention. Women who received face-to-face support reported moderately higher social benefits, with the Web-based supported group reporting moderately higher psychological benefits. This is likely to be attributed to the additional support that Arm B and C received through nurse consultation and health coaching and the peer support available in Arm B. These results suggest that personalized and tailored health consultation provided by registered nurses with knowledge and skills in health behavior change theory and communication can facilitate positive behavior change.

In support of this, anecdotal feedback from participants indicated that having the opportunity to discuss personal health issues and work and family commitments and discuss strategies for change with a supportive health professional was highly valued by women in the face-to-face and Web-based supported groups. In contrast, feedback from participants in the Web-based independent group highlighted the lack of support being a barrier to engagement with the program, and it is likely that this contributed to the higher attrition rate in this group. However, participants in the Web-based independent group who remained in the study reported significant increases in physical activity, indicating that for some women undertaking a Web-based intervention independently is an effective option for undertaking a behavior change intervention.

Interestingly, despite a reported increase in physical activity and positive benefit perceptions, there was no statistically significant change in the average exercise barrier perceptions within or between groups postintervention. A possible explanation is that the average exercise barrier perceptions in all groups at baseline were relatively low, perhaps reflecting the fact that participants were motivated volunteers who self-selected to enroll in a health promotion program.

Comparison With Previous Work

In relation to the effect of a behavior change intervention on exercise benefit and barrier perceptions, there are limited studies to allow direct comparison. Our results are somewhat similar to a study on Latino American women that found an increase in total benefit perceptions following a 9-month biweekly education and exercise intervention [33]. In contrast, Kennedy et al [33] report a decrease in barrier perceptions. Other studies in women report no change in total benefit or barrier perceptions post exercise intervention [34,35]. One of these studies was a 7-week structured walking program designed for postmenopausal African American women [35] and the other, a 12-week group exercise program for mother and daughter pairs [34]. A more recent study investigated perceived benefits and barriers to exercise participation in n=43 overweight women with polycystic ovarian syndrome participating in a 3-arm 20-week lifestyle program, finding a significant improvement in increased benefit and decreased barrier perceptions [36]. Our study has appeared to be the first to report the effect of a 12-week MHBC intervention on exercise benefit and barrier perceptions in healthy midlife women.

There is a large body of literature in relation to the effect of Web-based physical activity interventions with our results aligning with systematic review [15] and meta-analysis [37] findings that indicate that the majority of internet PA interventions in adults lead to significant increases in physical activity. These studies both report that average effect sizes are usually small (d=0.14); in contrast, our study found moderate-to-large effect sizes across all intervention groups (d=0.5 to d=0.87, respectively). A possible explanation for these results is that the WWP intervention is specifically designed for midlife women with variable fitness levels and allows personal choice in type, intensity, and frequency of physical activity and exercise, facilitating incremental change over time. The intervention materials also explicitly address the multiple benefits of regular exercise for midlife women's health and healthy ageing and provide practical strategies for overcoming barriers to change, perhaps enhancing participant motivation to exercise. This is consistent with evidence that client-centered and personalized lifestyle interventions with ongoing support are likely to be more effective in changing behavior in both the short and medium term [38]. Results in this study are also consistent with our previous findings [12], where the original WWP was tailored for midlife women and included personalized support.

Similar to studies comparing different intervention delivery modes [39,40], our study found increases in physical activity in all groups. A similar study by Steele et al [39] investigated the effectiveness of delivery modes for a 12-week pedometer-based behavior change program (Health-*e* Steps), comparing face-to-face, internet-mediated, and internet-only delivery with an equivalent magnitude of increase in physical activity over time being reported. In contrast, our study found a greater magnitude of change in PA in the face-to-face and Web-based supported groups compared with the Web-based independent group. In contrast to the Web-based independent group, the face-to-face and Web-based supported groups both received health coaching including goal setting from a registered nurse. Recent systematic reviews indicate that behavioral counselling to promote PA and a healthy diet [41] and setting goals [42] are effective and important in facilitating behavior change. Building on our previous work [12], the results in this study including qualitative feedback from participants suggest that *one size does not fit all*. In addition to the tailored and personalized nurse-led intervention, the ability to offer women the choice of intervention delivery modes to fit their motivation levels, family and work commitments, and preferences is a positive outcome of this project.

Limitations

The results of this study need to be considered in light of the limitations. First, although women were recruited across Australia, volunteers were generally well-educated and from high-income groups. Although this is often the case for studies involving Web-based interventions with volunteers predominantly white, middle aged, and female [18], this is likely to limit the generalizability of results. In terms of attrition,

XSL•FO
RenderX

average barrier scores were higher in the Web-based independent group (Arm A) that had the highest level of attrition (35.5%), so data collected postintervention did not include those participants. Furthermore, this study had no control group, largely because it was designed to investigate the equivalency of the different intervention arms. Without a control group, it is difficult to definitively attribute posttest changes to the effects of the intervention. The study relied on self-report data that can be prone to response bias, with participants more likely to report positive outcomes for behavior change measures such as physical activity. Furthermore, data collection took place immediately after the intervention, so long-term effects might be lower.

Conclusions

Despite limitations, the results of this study suggest that the MHBC intervention (WWP) including personalized and tailored health coaching from registered nurses is likely to be effective in increasing exercise benefit perceptions, overall physical activity, and exercise in midlife women. The moderate-to-large effect sizes found in this study are particularly encouraging, indicating that for motivated participants, undertaking the program in different modes was beneficial in changing exercise behavior and positive perceptions of increased physical activity. Although Web-based interventions are cost-effective and flexible and can be delivered remotely, providing a range of options including face-to-face group delivery electronic health coaching from registered nurses has the potential to enhance participant engagement and motivation. With an increased focus on prevention of NCDs being evident, this study makes a timely contribution to the knowledge about exercise behavior change for prevention of NCDs in midlife women.

Acknowledgments

This study was funded by a Queensland University of Technology Bluebox Proof of Concept Grant awarded to DJA (Grant ID: WWP) and an Australian Postgraduate Award PhD Scholarship awarded to AMM. DJA and CB were PhD advisors with JP-S delivering the WWP intervention face-to-face and Web-based. All authors have been involved in preparation, review, and approval of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Women's Wellness Program Intervention.
[PDF File (Adobe PDF File), 394KB - jmir_v21i8e10963_app1.pdf]

References

1. GBD 2016 Risk Factor Collaborators. Global, regional, and national age-sex specific mortality for 264 causes of death, 1980–2016: a systematic analysis for the Global Burden of Disease Study 2016. The Lancet 2017 Sep 23;390(10100):1151-1210 [FREE Full text] [doi: 10.1016/S0140-6736(17)32152-9] [Medline: 28734670]
2. World Health Organization. Geneva; 2013. Global action plan for the prevention of and control of noncommunicable diseases 2013-2020 URL: https://apps.who.int/iris/bitstream/handle/10665/94384/9789241506236_eng.pdf;jsessionid=9C9891F90EF8AF16B130E04BE42AA433?sequence=1 [accessed 2018-04-01] [WebCite Cache ID 782C7wyB9]
3. Bennett JE, Stevens GA, Mathers CD, Bonita R, Rehm J, Kruk ME, et al. NCD Countdown 2030: worldwide trends in non-communicable disease mortality and progress towards Sustainable Development Goal target 3.4. Lancet 2018 Sep;392(10152):1072-1088. [doi: 10.1016/S0140-6736(18)31992-5]
4. GBD 2015 Risk Factors Collaborators. Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. Lancet 2016 Oct 8;388(10053):1659-1724 [FREE Full text] [doi: 10.1016/S0140-6736(16)31679-8] [Medline: 27733284]
5. Australian Institute of Health and Welfare. Canberra; 2018. Australia's Health 2018, in Australia's health series no 16 URL: https://www.aihw.gov.au/getmedia/7c42913d-295f-4bc9-9c24-4e44eff4a04a/aihw-aus-221.pdf.aspx?inline=true [accessed 2019-05-01] [WebCite Cache ID 782DWWafr]
6. Australian Government Department of Health. Australian Government Department of Health. Canberra: Commonwealth of Australia; 2014. Australian Physical Activity and Sedentary Behaviour Guidelines for Adults 18-64 years URL: http://www.health.gov.au/internet/main/publishing.nsf/Content/health-pubhlth-strateg-phys-act-guidelines [accessed 2019-04-30] [WebCite Cache ID 782E7HNRZ]
7. National Health and Medical Research Council. Canberra: NHMRC; 2013. Australian Dietary Guidlelines URL: https://www.nhmrc.gov.au/about-us/publications/australian-dietary-guidelines [accessed 2019-04-30] [WebCite Cache ID 782ETRszf]

XSL•FO
RenderX

8. Nejat EJ, Polotsky AJ, Pal L. Predictors of chronic disease at midlife and beyond—the health risks of obesity. Maturitas 2010 Feb;65(2):106-111. [doi: 10.1016/j.maturitas.2009.09.006] [Medline: 19796885]
9. Sun Q, Townsend MK, Okereke OI, Franco OH, Hu FB, Grodstein F. Adiposity and weight change in mid-life in relation to healthy survival after age 70 in women: prospective cohort study. Br Med J 2009 Sep 29;339:b3796 [FREE Full text] [Medline: 19789407]
10. Australian Bureau of Statistics. Canberra: ABS; 2017. Retirement and Retirement Intentions, Australia, July 2016 to June 2017 URL: https://www.abs.gov.au/ausstats/abs@.nsf/mf/6238.0 [accessed 2019-04-30] [WebCite Cache ID 782EsqUvv]
11. Mishra GD, Cooper R, Kuh D. A life course approach to reproductive health: theory and methods. Maturitas 2010 Feb;65(2):92-97 [FREE Full text] [doi: 10.1016/j.maturitas.2009.12.009] [Medline: 20079587]
12. Anderson D, Mizzari K, Kain V, Webster J. The effects of a multimodal intervention trial to promote lifestyle factors associated with the prevention of cardiovascular disease in menopausal and postmenopausal Australian women. Health Care Women Int 2006 Mar;27(3):238-253. [doi: 10.1080/07399330500506543] [Medline: 16524854]
13. Keyserling TC, Samuel Hodge CD, Jilcott SB, Johnston LF, Garcia BA, Gizlice Z, et al. Randomized trial of a clinic-based, community-supported, lifestyle intervention to improve physical activity and diet: the North Carolina enhanced WISEWOMAN project. Prev Med 2008 Jun;46(6):499-510. [doi: 10.1016/j.ypmed.2008.02.011] [Medline: 18394692]
14. Darling CA, Coccia C, Senatore N. Women in midlife: stress, health and life satisfaction. Stress Health 2012 Feb;28(1):31-40. [doi: 10.1002/smi.1398] [Medline: 22259156]
15. Joseph RP, Durant NH, Benitez TJ, Pekmezi DW. Internet-based physical activity interventions. Am J Lifestyle Med 2014 Dec;8(1):42-68 [FREE Full text] [doi: 10.1177/1559827613498059] [Medline: 25045343]
16. Duan YP, Wienert J, Hu C, Si GY, Lippke S. Web-based intervention for physical activity and fruit and vegetable intake among Chinese university students: a randomized controlled trial. J Med Internet Res 2017 Apr 10;19(4):e106 [FREE Full text] [doi: 10.2196/jmir.7152] [Medline: 28396306]
17. Schulz DN, Kremers SP, Vandelanotte C, van Adrichem MJ, Schneider F, Candel MJJM, et al. Effects of a web-based tailored multiple-lifestyle intervention for adults: a two-year randomized controlled trial comparing sequential and simultaneous delivery modes. J Med Internet Res 2014 Jan;16(1):e26 [FREE Full text] [doi: 10.2196/jmir.3094] [Medline: 24472854]
18. Spring B, Pellegrini C, McFadden HG, Pfammatter AF, Stump TK, Siddique J, et al. Multicomponent mHealth intervention for large, sustained change in multiple diet and activity risk behaviors: the make better choices 2 randomized controlled Trial. J Med Internet Res 2018 Jun 19;20(6):e10528 [FREE Full text] [doi: 10.2196/10528] [Medline: 29921561]
19. Storm V, Dörenkämper J, Reinwand DA, Wienert J, de Vries H, Lippke S. Effectiveness of a web-based computer-tailored multiple-lifestyle intervention for people interested in reducing their cardiovascular risk: a randomized controlled trial. J Med Internet Res 2016 Apr 11;18(4):e78 [FREE Full text] [doi: 10.2196/jmir.5147] [Medline: 27068880]
20. Adams M, McCrone S. Predictors of exercise in midline and older women based on the health promotion model. Women Sport Phys Act J 2011 Oct;20(2):65-75. [doi: 10.1123/wspaj.20.2.65]
21. Gatewood JG, Litchfield RE, Ryan SJ, Geadelmann JD, Pendergast JF, Ullom KK. Perceived barriers to community-based health promotion program participation. Am J Health Behav 2008;32(3):260-271. [doi: 10.5555/ajhb.2008.32.3.260] [Medline: 18067466]
22. Lovell GP, El Ansari W, Parker JK. Perceived exercise benefits and barriers of non-exercising female university students in the United Kingdom. Int J Environ Res Public Health 2010 Dec;7(3):784-798 [FREE Full text] [doi: 10.3390/ijerph7030784] [Medline: 20617003]
23. McGuire A, Seib C, Anderson D. Factors predicting barriers to exercise in midlife Australian women. Maturitas 2016 May;87:61-66. [doi: 10.1016/j.maturitas.2016.02.010] [Medline: 27013289]
24. McGuire AM, Anderson DJ, Fulbrook P. Perceived barriers to healthy lifestyle activities in midlife and older Australian women with type 2 diabetes. Collegian 2014;21(4):301-310. [Medline: 25632727]
25. Sechrist KR, Walker SN, Pender NJ. Development and psychometric evaluation of the exercise benefits/barriers scale. Res Nurs Health 1987 Dec;10(6):357-365. [doi: 10.1002/nur.4770100603] [Medline: 3423307]
26. Gho SA, Munro BJ, Jones SC, Steele JR. Perceived exercise barriers explain exercise participation in Australian women treated for breast cancer better than perceived exercise benefits. Phys Ther 2014 Dec;94(12):1765-1774. [doi: 10.2522/ptj.20130473] [Medline: 25060956]
27. Bandura A. Health promotion by social cognitive means. Health Educ Behav 2004 Apr;31(2):143-164. [doi: 10.1177/1090198104263660] [Medline: 15090118]
28. Anderson D, Seib C, McGuire A, Porter-Steele J. Decreasing menopausal symptoms in women undertaking a web-based multi-modal lifestyle intervention: the Women's Wellness Program. Maturitas 2015 May;81(1):69-75. [doi: 10.1016/j.maturitas.2015.02.263] [Medline: 25818771]
29. Woods NF, Mitchell ES. Sleep symptoms during the menopausal transition and early postmenopause: observations from the Seattle Midlife Women's Health Study. Sleep 2010 Apr;33(4):539-549 [FREE Full text] [Medline: 20394324]
30. World Health Organization. Geneva: WHO; 2018. Obesity and overweight URL: http://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight [accessed 2018-03-01] [WebCite Cache ID 782FhJCnz]

XSL•FO
RenderX

31. IBM. Chicago: IBM; 2013. Statistical Package for Social Sciences URL: https://www.ibm.com/au-en/analytics/spss-statistics-software [accessed 2019-05-01] [WebCite Cache ID 782GGnFm5]
32. Cohen JW. Statistical Power Analysis for the Behavioural Sciences. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.
33. Kennedy C, Devoe D, Skov J, Short-Degraff M. Attitudinal changes toward exercise in mexican american women. Occup Ther Health Care 1998;11(3):17-28. [doi: 10.1080/J003v11n03_02] [Medline: 23948043]
34. Ransdell LB, Detling N, Hildebrand K, Lau P. Can physical activity interventions change perceved exercise benefits and barriers? Am J Health Stud 2004;19(4):195-204 https://scholarworks.boisestate.edu/kinesiology_facpubs/81/.
35. Williams BR, Bezner J, Chesbro SB, Leavitt R. The effect of a walking program on perceived benefits and barriers to exercise in postmenopausal African American women. J Geriatr Phys Ther 2006;29(2):43-49. [Medline: 16914065]
36. Thomson RL, Buckley JD, Brinkworth GD. Perceived exercise barriers are reduced and benefits are improved with lifestyle modification in overweight and obese women with polycystic ovary syndrome: a randomised controlled trial. BMC Womens Health 2016 Mar 9;16:14 [FREE Full text] [doi: 10.1186/s12905-016-0292-8] [Medline: 26960762]
37. Davies CA, Spence JC, Vandelanotte C, Caperchione CM, Mummery WK. Meta-analysis of internet-delivered interventions to increase physical activity levels. Int J Behav Nutr Phys Act 2012 Apr 30;9:52 [FREE Full text] [doi: 10.1186/1479-5868-9-52] [Medline: 22546283]
38. Zubala A, MacGillivray S, Frost H, Kroll T, Skelton DA, Gavine A, et al. Promotion of physical activity interventions for community dwelling older adults: a systematic review of reviews. PLoS One 2017;12(7):e0180902 [FREE Full text] [doi: 10.1371/journal.pone.0180902] [Medline: 28700754]
39. Steele R, Mummery WK, Dwyer T. Using the internet to promote physical activity: a randomized trial of intervention delivery modes. J Phys Act Health 2007 Jul;4(3):245-260. [Medline: 17846455]
40. Vandelanotte C, Duncan MJ, Plotnikoff RC, Mummery WK. Do participants' preferences for mode of delivery (text, video, or both) influence the effectiveness of a Web-based physical activity intervention? J Med Internet Res 2012 Feb 29;14(1):e37 [FREE Full text] [doi: 10.2196/jmir.1998] [Medline: 22377834]
41. Patnode CD, Evans CV, Senger CA, Redmond N, Lin JS. Behavioral counseling to promote a healthful diet and physical activity for cardiovascular disease prevention in adults without known cardiovascular disease risk factors: Updated evidence report and systematic review for the US Preventive Services Task Force. J Am Med Assoc 2017 Jul 11;318(2):175-193. [doi: 10.1001/jama.2017.3303] [Medline: 28697259]
42. Epton T, Currie S, Armitage CJ. Unique effects of setting goals on behavior change: systematic review and meta-analysis. J Consult Clin Psychol 2017 Dec;85(12):1182-1198. [doi: 10.1037/ccp0000260] [Medline: 29189034]

Abbreviations

ANCOVA: analysis of covariance
BMI: body mass index
EBBS: Exercise Benefits and Barriers Scale
MHBC: multiple health behavior change
NCD: noncommunicable disease
WWP: Women's Wellness Program

Edited by G Eysenbach; submitted 04.05.18; peer-reviewed by S Lippke, D Reinwand; comments to author 03.09.18; revised version received 19.11.18; accepted 30.03.19; published 21.08.19.

Please cite as:
McGuire AM, Seib C, Porter-Steele J, Anderson DJ
The Association Between Web-Based or Face-to-Face Lifestyle Interventions on the Perceived Benefits and Barriers to Exercise in Midlife Women: Three-Arm Equivalency Study
J Med Internet Res 2019;21(8):e10963
URL: https://www.jmir.org/2019/8/e10963/
doi:10.2196/10963
PMID:31436162

©Amanda Mary McGuire, Charrlotte Seib, Janine Porter-Steele, Debra Jane Anderson. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 21.08.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Association Between the Use of a Mobile Health Strategy App and Biological Changes in Breast Cancer Survivors: Prospective Pre-Post Study

Mario Lozano-Lozano[1,2,3,4], MSc; Lucia Melguizo-Rodríguez[3,5], PhD; Carolina Fernández-Lao[1,2,3,4], PhD; Noelia Galiano-Castillo[1,2,3,4], PhD; Irene Cantarero-Villanueva[1,2,3,4], PhD; Lydia Martín-Martín[1,2,3,4], PhD; Manuel Arroyo-Morales[1,2,3,4], MD, PhD

[1]Department of Physical Therapy, Faculty of Health Sciences, University of Granada, Granada, Spain

[2]Sport and Health Joint University Institute, Granada, Spain

[3]Biohealth Research Institute, Granada, Spain

[4]"Cuidate" Support Unit for Oncology Patients, Granada, Spain

[5]Biomedical Group (BIO277), Department of Nursing, Faculty of Health Sciences, University of Granada, Granada, Spain

Corresponding Author:
Lydia Martín-Martín, PhD
Department of Physical Therapy, Faculty of Health Sciences
University of Granada
Avda de la Ilustracion, 60
Granada, 18016
Spain
Phone: 34 958242070
Email: lydia@ugr.es

Abstract

Background: There is a bidirectional relationship between chronic low-grade inflammation and cancer. Inflammatory markers, such as interleukin-6 (IL-6), have been associated with both the malignant transformation of epithelial cells and tumor progression, thus linking low-grade inflammation with a higher risk of cancer and recurrence in the survival phase. Therefore, they are considered valuable prognostic biomarkers. Knowing and finding appropriate primary prevention strategies to modify these parameters is a major challenge in reducing the risk of cancer recurrence and increasing survival. Different therapeutic strategies have shown efficacy in the modification of these and other biological parameters, but with contradictory results. There are apparently no strategies in which telemedicine, and specifically mobile health (mHealth), are used as a means to potentially cause biological changes.

Objective: The objectives of this study were to: (1) check whether it is feasible to find changes in inflammation biomarkers through an mHealth strategy app as a delivery mechanism of an intervention to monitor energy balance; and (2) discover potential predictors of change of these markers in breast cancer survivors (BCSs).

Methods: A prospective quasi-experimental pre-post study was conducted through an mHealth energy balance monitoring app with 73 BCSs, defined as stage I-IIIA of breast cancer and at least six months from the completion of the adjuvant therapy. Measurements included were biological salivary markers (IL-6 and C-reactive protein [CRP]), self-completed questionnaires (the European Organization for Research and Treatment of Cancer Quality of Life Questionnaire Core 30, the user version of the Mobile Application Rating Scale [uMARS] and an ad hoc clinical and sociodemographic questionnaire) and physical objective measures (accelerometry, weight and height). In addition, using the logging data of the mHealth app, the rate of use (in days) was recorded during the entire experimental phase of the study. Using Stata software, a paired two-tailed t test, Pearson and Spearman correlations, and a stepwise multiple regression analysis were used to interpret the data.

Results: Analyzing changes in inflammatory biomarker concentrations after using the mHealth app, differences between preassessment CRP (4899.04 pg/ml; SD 1085.25) and IL-6 (87.15 pg/ml; SD 33.59) and postassessment CRP (4221.24 pg/ml; SD 911.55) and IL-6 (60.53 pg/ml; SD 36.31) showed a significant decrease in both markers, with a mean difference of –635.25 pg/ml (95% CI –935.65 to –334.85; P<.001) in CRP and –26.61 pg/ml (95% CI –42.51 to –10.71; P=.002) in IL-6. Stepwise regression analyses revealed that changes in global quality of life, as well as uMARS score and hormonal therapy, were possible

XSL•FO
RenderX

predictors of change in CRP concentration after using the mHealth app. In the same way, the type of tumor removal surgery conducted, as well as changes in weight and pain score, were possible predictors of change in IL-6 concentration after using the app.

Conclusions: In conclusion, through the results of this study, we hypothesize that there is a possible association between an mHealth energy balance monitoring strategy and biological changes in BCSs. These changes could be explained by different biopsychosocial parameters, such as the use of the application itself, quality of life, pain, type of tumor removal surgery, hormonal treatment or obesity.

(J Med Internet Res 2019;21(8):e15062) doi:10.2196/15062

KEYWORDS

mHealth; interleukin-6; C-reactive protein; breast cancer survivors; low-grade inflammatory

Introduction

There is a bidirectional relationship between chronic low-grade inflammation and breast cancer, as a tumor can produce an inflammatory environment and therefore a systemic immune response, but chronic inflammation can also both precede and promote the development of cancer [1]. There is even talk of considering inflammation to be an enabling feature of breast cancer, or the seventh hallmark of the disease along with the six hallmarks already identified by Hanahan and Weinberg [2,3]. Inflammation is a process, or bodily response, secondary to infection or sudden injury, and it is associated with the activation of various molecular mechanisms [4]. This response can be local or systemic, depending on the severity, and both indicate an imbalance of the metabolism of the affected tissues. This metabolic imbalance in the lesion is produced by an increase of immune cells as well as inflammatory parameters of great clinical importance, such as C-reactive protein (CRP) and its inducer interleukin-6 (IL-6) [4,5].

Once the inflammatory response ends, tissue metabolism is normalized. If this process of remission is interrupted by some circumstance, such as pathogens, toxins or other stimuli, healthy tissue could be damaged and produce what is known as persistent low-grade inflammation, or chronic inflammation [6]. It is the result of an immune system that overreacts so that the concentrations of inflammatory factors are higher than in a healthy population [5]. This systemic and chronic inflammation is widely associated with chronic diseases [6] and even symptomatology, as there is a positive association between increased levels of CRP and excess of adipose mass (excess weight and obesity), which is a factor that could be potentially modified with physical activity and diet [4,7]. Moreover, inflammatory markers such as IL-6 have also been associated with the malignant transformation of epithelial cells and tumor progression, associating low-grade inflammation with a higher risk of cancer and recurrence in the survival phase. Thus, these factors are considered valuable prognostic biomarkers in the population of those with cancer [1,8,9]. Therefore, knowing and finding appropriate primary prevention strategies to modify these parameters is a major challenge in the field, so as to reduce the risk of cancer recurrence and increase survival.

Different therapeutic strategies have shown efficacy in the modification of these and other biological parameters, but with contradictory results. The beneficial effects of physical exercise as a means of controlling low-grade inflammation have been amply demonstrated [10,11], even in breast cancer survivors (BCSs). A study conducted by Jones et al, in which they used a physical exercise program in BCSs, found a significant reduction of IL-6 [12]. Additional studies have evaluated other strategies, such as manual therapy [13], tai chi [14], mindfulness [15], or yoga [16], to reduce inflammation markers in different cancer populations. However, scientific evidence about strategies based on telemedicine are scarce, and they are practically nonexistent for cancer. A study conducted by Haggerty et al assessed two technology-based, 6-month, lifestyle interventions (telemedicine or text messaging) in obese women with endometrial hyperplasia, showing a reduction of some biomarkers such as IL-6 after the intervention [17]. Another study by Frank et al examined the effectiveness of telehealth coaching promoting nutrition and exercise in soldiers, evaluating biomarkers of bone health [18]. There are also some clinical trial protocols with no published results at present [19-21]. Therefore, at the moment there are no strategies in which telemedicine, and specifically mobile health (mHealth), are used as a delivery mechanism for interventions that could cause biological changes.

Low-grade inflammation is highly influenced by aspects such as obesity, fatigue or a sedentary lifestyle [22-24], and its relationship with chronic pathologies has been demonstrated. However, the issue of association, or the factors that influence its regulation through nonpharmacological and distance-based intervention strategies, remain unresolved [6]. In the biopsychosocial context that encompasses a subject with cancer, promoting changes through mHealth strategies in psychological, physical or social aspects is not entirely complicated (eg, quality of life) [25]. However, biological parameters have a high intersubject variability and are not usually addressed in these types of studies [24]. Therefore, understanding what factors can influence these parameters can help to develop mHealth-based strategies, thus increasing patient empowerment in regard to their health.

To the best of our knowledge, scientific evidence is scarce in regard to mHealth-based strategies related to tracking biomarkers of inflammation, and the importance of low-grade inflammation in cancer recurrence has already been demonstrated. Thus, the objectives of this study were to:

1. Check whether it is feasible to find changes in inflammation biomarkers through an mHealth strategy as a delivery mechanism of an intervention to monitor energy balance

XSL•FO
RenderX

2. Discover potential predictors of change of these markers in breast cancer survivors

Methods

Study Design, Participants, and Description of the Mobile Health App

A prospective quasi-experimental pre-post study was conducted through an mHealth app to monitor energy balance (BENECA mHealth app) with 73 BCSs, defined as stage I-IIIA of breast cancer and at least six months from the completion of adjuvant therapy (only hormonal therapy was allowed). Participants were recruited from the oncology units of San Cecili University Hospital and Virgen de las Nieves University Hospital, both in Granada, Spain, through their reference oncologists. All participants received oral and written information about the assessment protocols, mHealth app characteristics, and risks and benefits of the study, and then written consent was obtained from all of them. The Ethics Committee on Human Research (CEIH) from Granada province, Spain, approved this study (FIS, PI14-01627), which was performed in accordance with the Declaration of Helsinki [26]. The inclusion and exclusion criteria for this study are shown in Textbox 1.

After the initial assessment, all participants were invited to use the mHealth app for two months. In summary, the mHealth app was developed to help breast cancer survivors overcome energy balance challenges and aimed to both motivate and sensitize breast cancer survivors to adhere to fully personalized physical exercise programs and nutritional plans, in compliance with the international guidelines for cancer survivors. On first use, the users of the app recorded their personal and anthropometric data such as weight, height, age, and type of cancer. They were then asked to record what they ate (every item) and what they did (in terms of physical activity) the day before. Regarding food intake, BENECA uses a dietary record questionnaire structured with 6 consumption times. Regarding physical activity, patients could record the activities they completed during the day (intensity and duration) from 3 possible time periods (morning, afternoon, and evening). BENECA only records those activities that have a duration of at least 10 minutes.

Based on all this information, the mHealth app provided automatic feedback about a person's energy balance or imbalance as well as nutritional information about what was ingested. In the presence of energy imbalance, it provided useful and simple tips to improve this imbalance. All these straightforward, daily notifications were based on the guidelines of the World Cancer Research Fund International [27], the strategies for physical activity and diet in patients with cancer from the American College of Sports Medicine [28], and the recommendations of the American Cancer Society [29,30]. The mHealth app was developed based on the theory of Learning, on Goal-Setting Theory, and on Social Cognitive Theory so as to include techniques such as reinforcement, facilitation, self-monitoring, goal setting, feedback on performance and reviewing goals, which have proven to be promising in increasing physical activity in different populations [31,32]. The technical characteristics of the mobile application [33], as well as validation of the energy balance monitoring system [34] and its feasibility [35], have been previously published.

Participants were able to contact a researcher at any time via WhatsApp, in case of technical problems or to discuss any doubts they had. In addition, an online video tutorial was available at any time.

Textbox 1. Inclusion and exclusion criteria.

Inclusion criteria:

- Between 30 and 75 years of age.
- Body mass index >25 kg/m^2.
- Stage I-IIIA of breast cancer.

Exclusion criteria:

- No medical clearance to participate.
- Any physical health condition that prevents them from walking.
- Any physical or mental health condition that prevents them from participating in assessments.
- No access to any mobile device or tablet with an internet connection.

Outcomes Measures

To assess changes after use of the mHealth app, all measurements were taken at baseline and 8-weeks after having used it. Participants were called via phone for pre and postassessments and invited to Cuidate Support Unit for Oncology Patients, a clinical research center from the University of Granada, Spain. Measurements taken included biological markers, self-completed questionnaires, and both anthropometric and physical objective measures. In addition, using the logging data of the mHealth app, the rate of use (in days) was recorded during the entire experimental phase of the study.

Biological Markers: Main Outcomes

Two salivary inflammatory markers were obtained: IL-6 and CRP. Salivary biomarkers have previously demonstrated the potential to be used for screening and research purposes [36].

XSL•FO
RenderX

Sample Handling and Preparation: Salivary Interleukin-6 and C-reactive Protein Concentrations

On the day of sample collection, the participants were informed of the requirements: they were not allowed to eat, drink or brush their teeth during the two hours prior to the collection, and they were not allowed to visit the dentist 24 hours before sampling, with the aim of reducing the risks of contamination. They were also not allowed to consume alcohol during the 12 hours prior to the collection of the sample, or to take acidic or high sugar foods. The saliva sample collection was done between 10:00 and 11:30 in the morning, and an attempt was made to match the time in the postassessment. Ten minutes before the collection of the sample, participants were asked to rinse their mouths with water. Saliva was collected by unstimulated passive drool for 3 minutes using a polypropylene vial. Participants were instructed to lean their heads forward, allowing the saliva to accumulate on the floor of the mouth. Immediately after collection, the sample was centrifuged at 3000-3500 rpm for 15 minutes (to remove mucins and other particles that might interfere with the results), and then the supernatant was stored in 200 µL tubes (total of 5 per participant). Finally, it was frozen and stored at –80°C for no longer than 3 months.

Sample Analysis (Enzyme Linked Immunosorbent Assay Procedures)

Once the sampling was completed, it was thawed completely until reaching room temperature prior to the completion of the solutions. The necessary sample was pipetted into dissolution tubes, and the residual saliva not analyzed was frozen again. The following enzyme linked immunosorbent assay (ELISA) kits were chosen: the Salimetrics C-Reactive Protein ELISA Kit (Kit number 1-3302, which is an enzyme-linked immunoassay specifically designed and validated for the quantitative measurement of salivary CRP), and the Salimetrics IL-6 ELISA Kit (Kit number 1-3602, which is a sandwich immunoassay specifically designed and validated for the quantitative measurement of salivary IL-6). Both have been designed and optimized for salivary research in humans. All analyses and calculations were performed following the manufacturer's protocol, as described by Salimetrics. A total of 15 µL and 60 µL of saliva were required for the analyses of CRP and IL-6, respectively.

Once the reagents were prepared, we designed the plate where 100 µl of the samples were added, as well as the successive dilutions of the standard of each marker that would be used for the design of the standard curve. The sample was covered with an adhesive and incubated for two hours at room temperature before mixing at the mix plate at 500 rpm. Then, the plate was washed 4 times with wash buffer by filling and emptying the wells to remove the solution by either aspiration or plate inversion. After washing, antibody conjugate solution was added (100 µL/well) and then diluted in blocking buffer in a series of twofold dilutions. Then the plate was sealed and incubated for 2 hours at room temperature. After the incubation, we repeated the washing as described above. Once the wash was completed, the substrate solution was added, and the plate was incubated in the dark at room temperature for 30 minutes before then mixing for 5 minutes on a plate rotator at 500 rpm. Then, we stopped the reaction by adding the stop solution (50 µl). The solution was mixed at the plate rotator for 3 minutes at 500 rpm. The absorbance was then measured with a spectrophotometer (Biotek ELx800) at 450 nm, according to kit manufacturers. Results were compared with a standard curve that was previously designed. All standards, controls and samples were analyzed in duplicate.

Self-Completed Questionnaires

The European Organization for Research and Treatment of Cancer Quality of Life Questionnaire Core 30 (EORT QLQ-C30) version 3.0 was used to measure quality of life of the participants. It is a questionnaire specifically designed to evaluate general aspects of quality of life of patients with cancer. It is composed of a global scale of health status, five functional scales (in which the higher the score, the higher the quality of life reported) and eight symptom scales (in which the higher the score, the greater the symptoms reported). This instrument has shown adequate reliability [37,38].

The user version of the Mobile Application Rating Scale (uMARS) was used to measure the satisfaction and quality of usage of the mHealth app. This questionnaire is composed of 23 elements grouped into different sections, each of them evaluated independently through a Likert scale of 1 to 5 points (5 being excellent). Finally, the average score is calculated. This scale has been validated and has proven to be simple, objective and reliable [39].

An ad hoc questionnaire was used to collect clinical and sociodemographic characteristics of participants, including the stage of breast cancer, the type of tumor removal surgery, and the medical treatment and hormonal therapy. The stage of breast cancer could be I, II or III-A, the type of surgery was categorized in increasing order according to invasion of the surgery method (lumpectomy, quadrantectomy, unilateral mastectomy and bilateral mastectomy), the medical treatment was either a neoadjuvant or adjuvant treatment, and the hormonal therapy was registered as either taking or not taking hormonal treatment, as well as its typology.

Anthropometric and Physical Objective Measures

A preprogrammed triaxial accelerometer (ActiGraph GT3X+, Pensacola, Florida) was used to collect data on participants' physical activity over 8 consecutive days, together with a questionnaire diary based on a previously published protocol of use and analysis [40,41]. Only the records of more than 4 days, and of at least 10 hours per day, were included in the analysis. Minutes of vigorous-to-moderate physical activity (MVPA) were recorded.

Weight (kg) and height (cm) were measured with light clothing and without shoes. Weight was measured using an electronic scale (model SECA 869, Hamburg, Germany), and height was measured in the Frankfort plane using a stadiometer (model SECA 213).

Statistical Analysis

Measures of central tendency and dispersion were used for continuous variables with a normal distribution. Categorical variables were reported as proportions (%). The

XSL•FO
RenderX

Kolmogorov-Smirnov test was used to check the normal distribution of the data. To evaluate the differences in the biological variables (CRP and IL-6) between baseline and after 8 weeks of use of the app, a paired two-tailed *t* test was used. To analyze the correlation between the different variables, Pearson and Spearman correlation were applied as appropriate. In this correlation analysis, the change variable (difference between postassessment and preassessment) was used with quantitative variables: biological variables, quality of life (EORT QLQ-C30 global score, fatigue and pain), satisfaction with the app (uMARS global score), use of the app (in days), MVPA (accelerometry) and weight. The change variable was also used to measure changes in clinical variables such as type of tumor removal surgery, stage of breast cancer, medical treatment and hormonal therapy. Dispersion diagrams were used to study the assumptions of normality, linearity and homoscedasticity. To determine which variables could explain the variation in CRP and IL-6 concentrations, a stepwise multiple regression analysis was used. For the regression model with the dependent variable of CRP, the score changes in general quality of life, hormonal treatment, quality and satisfaction were considered independent variables.

For the regression model with the dependent variable IL-6, type of tumor removal surgery, and score changes in both perceived pain and weight were considered independent variables. To be included in the multiple regression analysis, the independent variables had to have a correlation coefficient of $r>0.20$ between the dependent variable and the independent variable, and they had to be significant [42]. The possible collinearity between the independent variables was studied, and then the final model was validated using bootstrapping (the start-up method was carried out with repeated samples of the same size to replace the original samples). Two thousand repetitions were produced to estimate the confidence intervals accelerated and corrected for the starting bias. For statistical analyses, the level of significance was set at $P<.10$. All analyses were performed using the software Stata version 14 (Statacorp, College Station, Texas). At least two experiments were performed in all assays.

Results

User Statistics and Clinical Characteristics

Participants were, on average, 51.35 (SD 8.58) years of age, with a body mass index (BMI) of 28.86 (SD 8.58). A total of 64% of the BCSs listed their civil status as married and 21% as single, with 41% having an educational status of higher education and 31% having unfinished studies or primary school. Table 1 summarizes clinical and sociodemographic participants' characteristics.

Participants showed moderate quality of satisfaction (score range=0-5) with the mHealth app (mean 3.71 points; SD 0.47 points), high app usage (mean 47.9 days; SD 10.40; max=56 days), and moderate to low scores (range 0-100) in general quality of life (mean 57.6; SD 14.07), fatigue (mean 23.14; SD 15.46), and pain (mean 45.66; SD 25.91). Finally, mean weight was 72.56 kg (SD 10.85 kg) and the mean MVPA was 47.27 (SD 23.41). Analyzing changes in inflammatory biomarker concentrations after using the mHealth app, differences between preassessment CRP (4899.04 pg/ml; SD 1085.25) and IL-6 (87.15 pg/ml; SD 33.59) and postassessment CRP (4221.24 pg/ml; SD 911.55) and IL-6 (60.53 pg/ml; SD 36.31) showed a significant decrease in both of them, with a mean difference of –635.25 pg/ml (95% CI –935.65 to –334.85; $P<.001$) in CRP and –26.61 pg/ml (95% CI; –42.51 to –10.71; $P=.002$) in IL-6.

XSL•FO
RenderX

Table 1. Participants' demographics (N=73).

Variables	Participants
Age (years), mean (SD)	51.35 (8.58)
Marital status, n (%)	
Single	15 (20.6)
Married	47 (64.4)
Divorced	7 (9.6)
Other	4 (5.5)
Education, n (%)	
No education	1 (1.4)
Primary studies	22 (30.1)
Secondary studies	20 (27.4)
Higher education	30 (41.1)
Employment, n (%)	
Housewife	18 (24.7)
Employee	28 (38.4)
Low	10 (13.7)
Unemployed by the disease	17 (23.3)
Cancer stage, n (%)	
I	8 (11.3)
II	37 (52.1)
IIIA	26 (36.6)
Surgery, n (%)	
Lumpectomy	24 (32.8)
Quadrantectomy	12 (16.4)
Unilateral mastectomy	26 (35.6)
Bilateral mastectomy	11 (15.1)
Medical treatment, n (%)	
None	5 (6.9)
Radiation therapy alone	6 (8.2)
Chemotherapy alone	5 (6.9)
Chemotherapy and radiation therapy	48 (65.8)
Adjuvant chemotherapy	6 (8.2)
Neoadjuvant chemotherapy	3 (4.1)

Correlation Analyses

Significant negative correlations were found between changes in CRP concentration and EORT QLQ C30 general quality of life (r=–0.281; P=.03), with hormonal therapy (r=–0.235; P=.07), with uMARS score (r=–0.284; P=.02) and with mHealth app usage (r=–0.263; P=.04). In addition, significant positive correlations were found between change in IL-6 concentration and EORT QLQ C30 pain (r=0.404; P=.01), with weight (r=0.301; P=.06) and with type of tumor removal surgery (r=0.311; P=.05).

In addition, significant correlations existed among the independent variables (Table 2) but was only high between uMARS score and mHealth usage (r=0.907; P<.001). Therefore, considering multicollinearity possible (defined as r>0.70), only uMARS score was included in the regression analyses.

XSL•FO
RenderX

Table 2. Pearson product-moment correlation matrix for study variables.

Variable	Δ[a]CRP[b]	ΔIL-6[c]	ΔC30[d] QoL[e]	ΔC30 Fatigue	ΔC30 Pain	ΔWeight	ΔMV-PA[f]	Age	Stage BC[g]	Surgery type	Hormonal therapy	uMARS[h]	mHealth[i] use
ΔCRP	1.00	__[j]	—	—	—	—	—	—	—	—	—	—	—
ΔIL-6	0.191	1.00	—	—	—	—	—	—	—	—	—	—	—
P value	.28	—	—	—	—	—	—	—	—	—	—	—	—
ΔC30 QoL	–0.281	–0.168	1.00	—	—	—	—	—	—	—	—	—	—
P value	.03	.30	—	—	—	—	—	—	—	—	—	—	—
ΔC30 fatigue	0.054	0.208	–0.527	1.00	—	—	—	—	—	—	—	—	—
P value	.68	.20	<.001	—	—	—	—	—	—	—	—	—	—
ΔC30 pain	0.153	0.404	–0.35	0.678	1.00	—	—	—	—	—	—	—	—
P value	.24	.01	.002	<.001	—	—	—	—	—	—	—	—	—
ΔWeight	0.088	0.301	–0.183	0.088	0.04	1.00	—	—	—	—	—	—	—
P value	.50	.06	.12	.46	.73	—	—	—	—	—	—	—	—
ΔMVPA	0.099	0.011	–0.023	0.135	0.187	–0.116	1.00	—	—	—	—	—	—
P value	.47	.95	.85	.29	.14	.36	—	—	—	—	—	—	—
Age	–0.139	–0.04	0.259	–0.314	–0.301	0.20	–0.304	1.00	—	—	—	—	—
P value	.28	.8	.03	.01	.01	.09	.01	—	—	—	—	—	—
Stage BC	0.075	0.143	–0.055	0.079	–0.083	–0.045	–0.101	0.143	1.00	—	—	—	—
P value	.57	.38	.65	.51	.49	.71	.43	.23	—	—	—	—	—
Surgery type	–0.158	0.311	0.092	0.023	0.009	–0.023	–0.191	0.110	0.282	1.00	—	—	—
P value	.22	.05	.44	.84	.93	.85	.13	.35	.02	—	—	—	—
Hormonal therapy	–0.235	–0.001	0.142	–0.161	–0.161	0.088	0.018	0.285	0.052	0.247	1.00	—	—
P value	.07	>.99	.24	.18	.18	.46	.88	.02	.66	.04	—	—	—
uMARS	–0.284	–0.086	–0.105	0.09	0.188	–0.028	0.101	–0.309	–0.167	0.147	0.049	1.00	—
P value	.02	.57	.38	.45	.11	.81	.42	.01	.16	.23	.68	—	—
mHealth use	–0.263	–0.127	–0.101	0.112	0.184	–0.026	0.081	–0.402	–0.122	0.144	0.014	0.907	1.00
P value	.04	.43	.40	.35	.12	.82	.52	<.001	.31	.22	.91	<.001	—

[a]Δ: change between postassessment and preassessment.

[b]CRP: C-reactive protein.

[c]IL-6: interleukin-6.

[d]C30: EORT QLQ C-30 questionnaire.

[e]QoL: quality of life.

[f]MVPA: minutes of vigorous-to-moderate physical activity.

[g]BC: breast cancer.

[h]uMARS: user version of the Mobile Application Rating Scale.

[i]mHealth: mobile health.

[j]Not applicable.

Regression Analyses

Stepwise regression analyses revealed that changes in global quality of life, as well as uMARS score and hormonal therapy, were possible predictors of change in CRP concentration after using the app (Table 3). In the same way, the type of tumor removal surgery, as well as changes in weight and pain score, were possible predictors of change in IL-6 concentration after using the app (Table 4). For both tables, r^2 denotes the variability of change in biomarker concentration explained by the predictors in percent.

XSL•FO
RenderX

Table 3. Summary of stepwise regression analyses to determine predictors of change in C-reactive protein concentration (r^2=19%).

Independent variables	Unstandardized coefficients, β[a]	95% CI for β	Bootstrap BCA[b], 95% CI	Bootstrap, β	Standardized coefficients, β	*t*	*P* value
Interceptil	2496.949	218.504-4775.395	4.641-4989.268	2496.949	__[c]	2.2	.03
Hormonal therapy	–110.304	–286.447 to 65.839	–272.774 to 52.166	–110.304	–0.155	–1.25	.22
uMARS[d] score	–728.786	–1338.675 to –118.898	–1396.04 to –61.533	–728.785	–0.289	–2.39	.02
Δ[e]Global QoL[f]	–18.601	–36.253 to –0.945	–37.022 to –0.177	–18.605	–0.261	–2.11	.04

[a]β: regression coefficient.

[b]BCA: bias-corrected and accelerated.

[c]Not applicable.

[d]uMARS: user version of the mobile application rating scale.

[e]Δ: change between postassessment and preassessment.

[f]QoL: quality of life.

Table 4. Summary of stepwise regression analyses to determine predictors of change in IL-6 concentration (r^2=26%).

Independent variables	Unstandardized coefficients, β[a]	95% CI for β	Bootstrap BCA[b], 95% CI	Bootstrap, β	Standardized coefficients, β	*t*	*P* value
Interceptil	–36.498	–68.898; –4.098	–71.773; –1.223	–36.498	__[c]	–2.28	.03
Type of surgery	8.219	–4.604; 21.04	–5.383; 21.821	8.22	0.194	1.3	.20
Δ[d]Weight	4.456	–1.263; 10.176	–0.582; 9.495	4.455	0.23	1.58	.08
ΔPain score	0.667	0.054; 1.279	0.003; 1.33	0.667	0.328	2.21	.03

[a]β: regression coefficient.

[b]BCA: bias-corrected and accelerated.

[c]Not applicable.

[d]Δ: change between postassessment and preassessment.

Discussion

The objective of this study was to determine the preliminary results of the possible association between the use of an mHealth strategy app as a delivery mechanism to monitor energy balance in cancer and the reduction of systemic inflammation markers, as well as to suggest possible predictors of this change. Current findings suggest that after two months of use of the app, a significant reduction of these markers can be observed. Thus, there could be a possible association between the two. In addition, the change in weight, pain and quality of life, as well as the type of tumor removal surgery, hormone therapy and the uMARS score, can have a contribution in the changes found in the concentrations of CRP and IL-6.

A system of monitoring energy balance through an mHealth app seems to reduce the biological parameters of systemic inflammation (CRP and IL-6). Our results suggest that after two months of use of the mHealth app, based on the monitoring of energy balance (in terms of diet and physical activity), the concentration of CRP and IL-6 are significantly reduced in BCSs. In fact, this change has a moderate effect size in CRP (Cohen *d*=–0.640; 95% CI –0.985 to –0.293) and a high effect size in IL-6 (Cohen *d*=–0.805; 95% CI –1.225 to –0.379). A study by Skogstad et al was the only one found with a design similar to ours, as it used a virtual internet physical activity motivation strategy in which some biological parameters were measured [43]. In this study, participants were included in a motivational physical activity program in which they measured their steps using a wrist-band accelerometer. However, unlike our study, no differences were found in CRP concentration after the intervention, perhaps because their study target population were healthy workers without pathology. The effect size reported for both our biomarkers supports the hypothesis that these changes are not due to time, but it is important to remark that the quasi-experimental pre-post design of our study does not allow us to affirm that the changes found are only attributable to the use of the app. Therefore, a controlled and randomized clinical trial should be carried out in the future.

Biological parameters of systemic inflammation can be mediated or modified by lifestyles changes such as physical activity and diet [4,6,7]. This study is the first to examine clinical and anthropometric factors that affect changes in these biological parameters, after using a mobile strategy to monitor energy balance in breast cancer survivors. Because rehabilitation strategies focus on face-to-face or distance physical activity and diet programs, understanding the potential determinants of reducing inflammation markers can help design more effective intervention strategies.

The results of our study show that possible moderators of a reduction in CRP concentration include not receiving hormonal

XSL•FO
RenderX

therapy, as well as having higher satisfaction and changes in quality of life (the higher quality of life change, the lower the CRP concentration). The role of estrogen in inflammation is poorly understood, the mechanism is not well studied, and its relationship is very complex [44], and different studies show contradictory results depending on different pathologies, with some showing they are associated with an inflammatory activity, while others show a proinflammatory role [44]. The differences found in pre and postmenopausal women suggest that the peripheral production of estrogens plays an important role in these differences [44-47]. Our results provide new evidence in this regard, since not having received hormone therapy may be a predictor of a greater reduction in CRP concentration in female survivors of breast cancer. However, estrogen's relationship with quality of life has been considered from another point of view. We understand that it is not that a higher perception of quality of life is a predictor of a reduction of proinflammatory markers but rather the other way around, that the diminished inflammatory state is associated with an increase in the quality of life [48,49]. Ultimately, our results suggest a higher score in uMARS as a predictor of the change in CRP concentration. In addition, there is a strong association between satisfaction and quality with the amount of time spent on the mHealth app. If women with the highest score in uMARS use the mHealth app more, then the reduction in inflammatory markers could be due to the direct relationship caused by a healthier lifestyle [6,50,51].

The results of our study also show possible moderators of the reduction in IL-6 include the type of tumor removal surgery (less invasive surgery), as well as changes in both weight and pain (the greater the reduction of these factors, the greater the reduction of IL-6). These results are consistent with the known bidirectional relationship between obesity and low-grade inflammation, which contributes to systemic metabolic dysfunction that is associated with obesity-linked disorders [4]. In the same way, an inflammatory reaction is also mediated by the classic cardinal signs of inflammation (eg, pain) [52]. Therefore, it is logical to think that a reduction in pain reported by breast cancer survivors can be a predictor of a reduction in IL-6 concentration such as that observed in our results. Finally, there is a lot of scientific evidence to support the use of minimally invasive surgical techniques since they don't raise inflammatory reactants as much, and these findings may support the relationship between IL-6 and the type of tumor removal surgery found in our results [53-56].

It is worth highlighting some strengths and limitations of the present study. The main strength lies in the nature of the study. To the best of our knowledge, this is the first study that proposes a mobile strategy to monitor energy balance as a mediator in the reduction of proinflammatory markers in BCSs. If future research supports our results, we will have found another support strategy for cancer survivors that is low cost and accessible to everyone and which could reduce markers highly related to the risk of recurrence. However, there are also many limitations to be noted. The main limitation lies in the design of the study, as well as the sample size, which prevents us from speaking in terms of causality and effectiveness. In addition, the r^2 obtained in the multiple regression models was low. However, we must not forget that we are trying to explain biological parameters with nonbiological variables. Our results may support the biopsychosocial model, since it shows how biology can be modified through these variables. Other biological parameters that can justify the rest of the variability that has remained to be explained should be taken into account in the future.

In conclusion, through the results of this study, we hypothesize that there is a possible association between an mHealth energy balance monitoring strategy app and biological changes in BCSs. These changes could be explained by different biopsychosocial parameters such as the use of the application itself, quality of life, pain, type of tumor removal surgery, hormonal treatment or obesity. Future studies should be carried out with a specific focus on all these biological parameters, and with an appropriate study design, to support these findings.

Acknowledgments

The study was funded by the Spanish Ministry of Economy and Competitiveness (Plan Estatal de I+D+I 2013-2016), Fondo de Investigación Sanitaria del Instituto de Salud Carlos III (PI14/01627), Fondos Estructurales de la Unión Europea (FEDER), and by the Spanish Ministry of Education (FPU14/01069 and FPU17/00939). This study took place thanks to additional funding from the University of Granada, Plan Propio de Investigación 2016, Excellence Actions: Units of Excellence; Unit of Excellence on Exercise and Health. This work was part of a PhD thesis conducted in the Clinical Medicine and Public Health Doctoral Studies of the University of Granada, Spain.

Conflicts of Interest

None declared.

References

1. Allin K, Bojesen S, Nordestgaard B. Inflammatory biomarkers and risk of cancer in 84,000 individuals from the general population. Int J Cancer 2016 Oct 01;139(7):1493-1500 [FREE Full text] [doi: 10.1002/ijc.30194] [Medline: 27194008]
2. Colotta F, Allavena P, Sica A, Garlanda C, Mantovani A. Cancer-related inflammation, the seventh hallmark of cancer: links to genetic instability. Carcinogenesis 2009 Jul;30(7):1073-1081. [doi: 10.1093/carcin/bgp127] [Medline: 19468060]
3. Hanahan D, Weinberg R. Hallmarks of cancer: the next generation. Cell 2011 Mar 04;144(5):646-674 [FREE Full text] [doi: 10.1016/j.cell.2011.02.013] [Medline: 21376230]

XSL•FO
RenderX

4. Ouchi N, Parker J, Lugus J, Walsh K. Adipokines in inflammation and metabolic disease. Nat Rev Immunol 2011 Feb;11(2):85-97 [FREE Full text] [doi: 10.1038/nri2921] [Medline: 21252989]
5. Pietzner M, Kaul A, Henning A, Kastenmüller G, Artati A, Lerch M, et al. Comprehensive metabolic profiling of chronic low-grade inflammation among generally healthy individuals. BMC Med 2017 Nov 30;15(1):210 [FREE Full text] [doi: 10.1186/s12916-017-0974-6] [Medline: 29187192]
6. Minihane A, Vinoy S, Russell W, Baka A, Roche H, Tuohy K, et al. Low-grade inflammation, diet composition and health: current research evidence and its translation. Br J Nutr 2015 Oct 14;114(7):999-1012 [FREE Full text] [doi: 10.1017/S0007114515002093] [Medline: 26228057]
7. Bruun J, Helge J, Richelsen B, Stallknecht B. Diet and exercise reduce low-grade inflammation and macrophage infiltration in adipose tissue but not in skeletal muscle in severely obese subjects. Am J Physiol Endocrinol Metab 2006 May;290(5):E961-E967 [FREE Full text] [doi: 10.1152/ajpendo.00506.2005] [Medline: 16352667]
8. Chan D, Bandera EV, Greenwood DC, Norat T. Circulating C-Reactive Protein and Breast Cancer Risk-Systematic Literature Review and Meta-analysis of Prospective Cohort Studies. Cancer Epidemiol Biomarkers Prev 2015 Oct;24(10):1439-1449 [FREE Full text] [doi: 10.1158/1055-9965.EPI-15-0324] [Medline: 26224798]
9. Dossus L, Jimenez-Corona A, Romieu I, Boutron-Ruault M, Boutten A, Dupré T, et al. C-reactive protein and postmenopausal breast cancer risk: results from the E3N cohort study. In: Cancer Causes Control. C-reactive protein and postmenopausal breast cancer risk: Results from the E3N cohort study. Cancer Causes Control Internet Springer International Publishing; Apr 2014:533-539.
10. Mathur N, Pedersen B. Exercise as a mean to control low-grade systemic inflammation. Mediators Inflamm 2008;2008:109502 [FREE Full text] [doi: 10.1155/2008/109502] [Medline: 19148295]
11. Petersen A, Pedersen B. The anti-inflammatory effect of exercise. J Appl Physiol (1985) 2005 Apr;98(4):1154-1162 [FREE Full text] [doi: 10.1152/japplphysiol.00164.2004] [Medline: 15772055]
12. Jones S, Thomas G, Hesselsweet S, Alvarez-Reeves M, Yu H, Irwin M. Effect of exercise on markers of inflammation in breast cancer survivors: the Yale exercise and survivorship study. Cancer Prev Res (Phila) 2013 Feb;6(2):109-118 [FREE Full text] [doi: 10.1158/1940-6207.CAPR-12-0278] [Medline: 23213072]
13. Díaz-Rodríguez L, Fernández-Pérez AM, Galiano-Castillo N, Cantarero-Villanueva I, Fernández-Lao C, Martín-Martín LM, et al. Do Patient Profiles Influence the Effects of Massage? A Controlled Clinical Trial. In: Biol Res Nurs. Los Angeles: SAGE Publications; Oct 2016:489-497.
14. Campo R, Light K, O'Connor K, Nakamura Y, Lipschitz D, LaStayo P, et al. Blood pressure, salivary cortisol, and inflammatory cytokine outcomes in senior female cancer survivors enrolled in a tai chi chih randomized controlled trial. J Cancer Surviv 2015 Mar;9(1):115-125 [FREE Full text] [doi: 10.1007/s11764-014-0395-x] [Medline: 25164513]
15. Lengacher CA, Reich RR, Paterson CL, Shelton M, Shivers S, Ramesar S, et al. A Large Randomized Trial: Effects of Mindfulness-Based Stress Reduction (MBSR) for Breast Cancer (BC) Survivors on Salivary Cortisol and IL-6. Biol Res Nurs 2019 Jan;21(1):39-49. [doi: 10.1177/1099800418789777] [Medline: 30079756]
16. Bower J, Greendale G, Crosswell A, Garet D, Sternlieb B, Ganz P, et al. Yoga reduces inflammatory signaling in fatigued breast cancer survivors: A randomized controlled trial. Psychoneuroendocrinology 2014 May;43:20-29. [doi: 10.1016/J.PSYNEUEN.2014.01.019]
17. Haggerty A, Huepenbecker S, Sarwer D, Spitzer J, Raggio G, Chu C, et al. The use of novel technology-based weight loss interventions for obese women with endometrial hyperplasia and cancer. Gynecol Oncol 2016 Feb;140(2):239-244. [doi: 10.1016/j.ygyno.2015.11.033] [Medline: 26644265]
18. Frank L, McCarthy M. Telehealth Coaching: Impact on Dietary and Physical Activity Contributions to Bone Health During a Military Deployment. Military Medicine 2016 May;181(5S):191-198. [doi: 10.7205/milmed-d-15-00159]
19. Suchy C, Massen L, Rognmo , Van Craenenbroeck EM, Beckers P, Kraigher-Krainer E, et al. Optimising exercise training in prevention and treatment of diastolic heart failure (OptimEx-CLIN): rationale and design of a prospective, randomised, controlled trial. Eur J Prev Cardiol 2014 Nov;21(2 Suppl):18-25. [doi: 10.1177/2047487314552764] [Medline: 25354950]
20. Schakel L, Veldhuijzen D, van Middendorp H, Prins C, Joosten SA, Ottenhoff THM, et al. The effects of a psychological intervention directed at optimizing immune function: study protocol for a randomized controlled trial. Trials 2017 May 26;18(1):243 [FREE Full text] [doi: 10.1186/s13063-017-1983-7] [Medline: 28549484]
21. Dobson R, Whittaker R, Jiang Y, Shepherd M, Maddison R, Carter K, et al. Text message-based diabetes self-management support (SMS4BG): study protocol for a randomised controlled trial. Trials 2016 Apr 02;17:179 [FREE Full text] [doi: 10.1186/s13063-016-1305-5] [Medline: 27039300]
22. Orre I, Reinertsen KV, Aukrust P, Dahl AA, Fosså SD, Ueland T, et al. Higher levels of fatigue are associated with higher CRP levels in disease-free breast cancer survivors. J Psychosom Res 2011 Sep;71(3):136-141. [doi: 10.1016/j.jpsychores.2011.04.003] [Medline: 21843747]
23. Ariza-García A, Galiano-Castillo N, Cantarero-Villanueva I, Fernández-Lao C, Díaz-Rodríguez L, Arroyo-Morales M. Influence of physical inactivity in psychophysiological state of breast cancer survivors. Eur J Cancer Care (Engl) 2013 Nov 25;22(6):738-745. [doi: 10.1111/ecc.12101] [Medline: 23889104]
24. Barbaresko J, Koch M, Schulze M, Nöthlings U. Dietary pattern analysis and biomarkers of low-grade inflammation: a systematic literature review. Nutr Rev 2013 Aug;71(8):511-527. [doi: 10.1111/nure.12035] [Medline: 23865797]

XSL•FO
RenderX

25. Galiano-Castillo N, Cantarero-Villanueva I, Fernández-Lao C, Ariza-García A, Díaz-Rodríguez L, Del-Moral-Ávila R, et al. Telehealth system: A randomized controlled trial evaluating the impact of an internet-based exercise intervention on quality of life, pain, muscle strength, and fatigue in breast cancer survivors. Cancer 2016 Oct 15;122(20):3166-3174 [FREE Full text] [doi: 10.1002/cncr.30172] [Medline: 27332968]
26. Government of Spain. [Law 14/2007, 3 July, on biomedical research (BOE, 4 July 2007)]. Rev Derecho Genoma Hum 2007(26):283-325. [Medline: 18201045]
27. Stull V, Snyder D, Demark-Wahnefried W. Lifestyle interventions in cancer survivors: designing programs that meet the needs of this vulnerable and growing population. J Nutr 2007 Jan;137(1 Suppl):243S-248S. [doi: 10.1093/jn/137.1.243S] [Medline: 17182834]
28. Schmitz K, Courneya K, Matthews C, Demark-Wahnefried W, Galvão DA, Pinto B, American College of Sports Medicine. American College of Sports Medicine roundtable on exercise guidelines for cancer survivors. Med Sci Sports Exerc 2010 Jul;42(7):1409-1426. [doi: 10.1249/MSS.0b013e3181e0c112] [Medline: 20559064]
29. Demark-Wahnefried W, Rogers L, Alfano C, Thomson C, Courneya K, Meyerhardt J, et al. Practical clinical interventions for diet, physical activity, and weight control in cancer survivors. CA Cancer J Clin 2015;65(3):167-189 [FREE Full text] [doi: 10.3322/caac.21265] [Medline: 25683894]
30. Kushi L, Doyle C, McCullough M, Rock C, Demark-Wahnefried W, Bandera EV, American Cancer Society 2010 NutritionPhysical Activity Guidelines Advisory Committee. American Cancer Society Guidelines on nutrition and physical activity for cancer prevention: reducing the risk of cancer with healthy food choices and physical activity. CA Cancer J Clin 2012;62(1):30-67 [FREE Full text] [doi: 10.3322/caac.20140] [Medline: 22237782]
31. Finne E, Glausch M, Exner A, Sauzet O, Stölzel F, Seidel N. Behavior change techniques for increasing physical activity in cancer survivors: a systematic review and meta-analysis of randomized controlled trials. Cancer Manag Res 2018;10:5125-5143 [FREE Full text] [doi: 10.2147/CMAR.S170064] [Medline: 30464612]
32. Vollmer Dahlke D, 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 Mar 27;3(1):e31 [FREE Full text] [doi: 10.2196/mhealth.3861] [Medline: 25830810]
33. Lozano-Lozano M, Moral-Munoz J, Galiano-Castillo N, Martín-Martín L, Fernández-Lao C, Arroyo-Morales M, et al. Designing BENECA m-Health APP, A Mobile Health Application to Monitor Diet and Physical Activity in Cancer Survivors. In: Rojas I, Joya G, Catala A, editors. Advances in Computational Intelligence. IWANN 2017. Lecture Notes in Computer Science. Cham: Springer; 2017:70.
34. Lozano-Lozano M, Galiano-Castillo N, Martín-Martín L, Pace-Bedetti N, Fernández-Lao C, Arroyo-Morales M, et al. Monitoring Energy Balance in Breast Cancer Survivors Using a Mobile App: Reliability Study. JMIR Mhealth Uhealth 2018 Mar 27;6(3):e67 [FREE Full text] [doi: 10.2196/mhealth.9669] [Medline: 29588273]
35. Lozano-Lozano M, Cantarero-Villanueva I, Martin-Martin L, Galiano-Castillo N, Sanchez M, Fernández-Lao C, et al. A Mobile System to Improve Quality of Life Via Energy Balance in Breast Cancer Survivors (BENECA mHealth): Prospective Test-Retest Quasiexperimental Feasibility Study. JMIR Mhealth Uhealth 2019 Jun 25;7(6):e14136 [FREE Full text] [doi: 10.2196/14136] [Medline: 31237570]
36. Rathnayake N, Akerman S, Klinge B, Lundegren N, Jansson H, Tryselius Y, et al. Salivary biomarkers for detection of systemic diseases. PLoS One 2013;8(4):e61356 [FREE Full text] [doi: 10.1371/journal.pone.0061356] [Medline: 23637817]
37. Aaronson N, Ahmedzai S, Bergman B, Bullinger M, Cull A, Duez N, et al. The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. J Natl Cancer Inst 1993 Mar 03;85(5):365-376. [doi: 10.1093/jnci/85.5.365] [Medline: 8433390]
38. Shih C, Chen C, Sheu C, Lang H, Hsieh C. Validating and improving the reliability of the EORTC qlq-c30 using a multidimensional Rasch model. Value Health 2013;16(5):848-854. [doi: 10.1016/j.jval.2013.05.004] [Medline: 23947980]
39. 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]
40. Cain K, Sallis J, Conway T, Van Dyck D, Calhoon L. Using accelerometers in youth physical activity studies: a review of methods. J Phys Act Health 2013;10(3):437-450 [FREE Full text] [Medline: 23620392]
41. Carbonell-Baeza A, Ruiz J, Aparicio V, Ortega F, Munguía-Izquierdo D, Alvarez-Gallardo IC, et al. Land- and water-based exercise intervention in women with fibromyalgia: the al-Andalus physical activity randomised controlled trial. BMC Musculoskelet Disord 2012 Feb 15;13:18 [FREE Full text] [doi: 10.1186/1471-2474-13-18] [Medline: 22336292]
42. Tabachnick B, Fidell L. Using Multivariate Statistics (7th Edition). California: Pearson; 2019.
43. Skogstad M, Lunde L, Skare , Mamen A, Alfonso J, Øvstebø R, et al. Physical activity initiated by employer and its health effects; an eight week follow-up study. BMC Public Health 2016 May 04;16:377 [FREE Full text] [doi: 10.1186/s12889-016-3035-8] [Medline: 27146485]
44. Straub R. The complex role of estrogens in inflammation. Endocr Rev 2007 Aug;28(5):521-574. [doi: 10.1210/er.2007-0001] [Medline: 17640948]
45. Schmidt M, Weidler C, Naumann H, Anders S, Schölmerich J, Straub R. Androgen conversion in osteoarthritis and rheumatoid arthritis synoviocytes--androstenedione and testosterone inhibit estrogen formation and favor production of

XSL•FO
RenderX

more potent 5alpha-reduced androgens. Arthritis Res Ther 2005;7(5):R938-R948 [FREE Full text] [doi: 10.1186/ar1769] [Medline: 16207335]
46. Simard J, Gingras S. Crucial role of cytokines in sex steroid formation in normal and tumoral tissues. Mol Cell Endocrinol 2001 Jan 22;171(1-2):25-40. [Medline: 11165008]
47. Purohit A, Reed M. Regulation of estrogen synthesis in postmenopausal women. Steroids 2002 Nov;67(12):979-983. [Medline: 12398994]
48. Kim J, Kim H, Song S. Associations among inflammation, mental health, and quality of life in adults with metabolic syndrome. Diabetol Metab Syndr 2018;10:66 [FREE Full text] [doi: 10.1186/s13098-018-0367-9] [Medline: 30186371]
49. de Almeida Roediger M, de Fátima Nunes Marucci M, Duim E, Santos J, de Oliveira Duarte YA, de Oliveira C. Inflammation and quality of life in later life: findings from the health, well-being and aging study (SABE). Health Qual Life Outcomes 2019 Feb 06;17(1):26 [FREE Full text] [doi: 10.1186/s12955-019-1092-2] [Medline: 30728031]
50. Ringseis R, Eder K, Mooren F, Krüger K. Metabolic signals and innate immune activation in obesity and exercise. Exerc Immunol Rev 2015;21:58-68 [FREE Full text] [Medline: 25825956]
51. Nimmo M, Leggate M, Viana J, King J. The effect of physical activity on mediators of inflammation. Diabetes Obes Metab 2013 Sep;15 Suppl 3:51-60. [doi: 10.1111/dom.12156] [Medline: 24003921]
52. Kapoor S, Fitzpatrick M, Clay E, Bayley R, Graham R, Wallace G, et al. Chapter 11. Metabolomics in the Analysis of Inflammatory Diseases. In: Roessner U, editor. Metabolomics. Rijeka (HR): InTech; 2012:269-288.
53. Scarpa M, Cavallin F, Saadeh L, Pinto E, Alfieri R, Cagol M, et al. Hybrid minimally invasive esophagectomy for cancer: impact on postoperative inflammatory and nutritional status. Dis Esophagus 2015 Sep 24;29(8):1064-1070. [doi: 10.1111/dote.12418]
54. Clinical Outcomes of Surgical Therapy Study Group, Nelson H, Sargent D, Wieand H, Fleshman J, Anvari M, et al. A comparison of laparoscopically assisted and open colectomy for colon cancer. N Engl J Med 2004 May 13;350(20):2050-2059. [doi: 10.1056/NEJMoa032651] [Medline: 15141043]
55. Veenhof A, Sietses C, von Blomberg BME, van Hoogstraten IMW, vd Pas MHGM, Meijerink W, et al. The surgical stress response and postoperative immune function after laparoscopic or conventional total mesorectal excision in rectal cancer: a randomized trial. Int J Colorectal Dis 2011 Jan;26(1):53-59 [FREE Full text] [doi: 10.1007/s00384-010-1056-9] [Medline: 20922542]
56. Targarona E, Pons M, Balagué C, Espert J, Moral A, Martínez J, et al. Acute phase is the only significantly reduced component of the injury response after laparoscopic cholecystectomy. World J Surg 1996 Jun;20(5):528-33; discussion 533. [Medline: 8661626]

Abbreviations

BCS: breast cancer survivor
BMI: body mass index
CEIH: Ethics Committee on Human Research
CRP: C-reactive protein
ELISA: enzyme linked immunosorbent assay
EORT QLQ-C30: the European Organization for Research and Treatment of Cancer Quality of Life Questionnaire Core 30
IL-6: interleukin-6
MVPA: minutes of vigorous-to-moderate physical activity
RCT: randomized controlled trial
uMARS: user version of the mobile application rating scale
WCRF: World Cancer Research Fund International

Edited by G Eysenbach; submitted 17.06.19; peer-reviewed by F Artacho-Cordon, A Lucas, N Azevedo; comments to author 16.07.19; revised version received 31.07.19; accepted 31.07.19; published 14.08.19.

Please cite as:
Lozano-Lozano M, Melguizo-Rodríguez L, Fernández-Lao C, Galiano-Castillo N, Cantarero-Villanueva I, Martín-Martín L, Arroyo-Morales M
Association Between the Use of a Mobile Health Strategy App and Biological Changes in Breast Cancer Survivors: Prospective Pre-Post Study
J Med Internet Res 2019;21(8):e15062
URL: http://www.jmir.org/2019/8/e15062/
doi:10.2196/15062
PMID:31414667

XSL•FO
RenderX

©Mario Lozano-Lozano, Lucia Melguizo-Rodríguez, Carolina Fernández-Lao, Noelia Galiano-Castillo, Irene Cantarero-Villanueva, Lydia Martín-Martín, Manuel Arroyo-Morales. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 14.08.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Comparison of the Effects of Coaching and Receipt of App Recommendations on Depression, Anxiety, and Engagement in the IntelliCare Platform: Factorial Randomized Controlled Trial

David C Mohr[1], PhD; Stephen M Schueller[2], PhD; Kathryn Noth Tomasino[3], PhD; Susan M Kaiser[1], MPH; Nameyeh Alam[1], MS; Chris Karr[4], MS; Jessica L Vergara[1], BA; Elizabeth L Gray[5], MS; Mary J Kwasny[5], PhD; Emily G Lattie[1], PhD

[1]Center for Behavioral Intervention Technologies, Northwestern University, Chicago, IL, United States

[2]Department of Psychological Science, University of California, Irvine, Irvine, CA, United States

[3]Department of Gastroenterology, Northwestern University, Chicago, IL, United States

[4]Audacious Software, Chicago, IL, United States

[5]Department of Preventive Medicine, Northwestern University, Chicago, IL, United States

Corresponding Author:
David C Mohr, PhD
Center for Behavioral Intervention Technologies
Northwestern University
750 N Lakeshore Drive
10th Floor
Chicago, IL, 60611
United States
Phone: 1 312 503 1403
Email: d-mohr@northwestern.edu

Abstract

Background: IntelliCare is a modular platform that includes 12 simple apps targeting specific psychological strategies for common mental health problems.

Objective: This study aimed to examine the effect of 2 methods of maintaining engagement with the IntelliCare platform, coaching, and receipt of weekly recommendations to try different apps on depression, anxiety, and app use.

Methods: A total of 301 participants with depression or anxiety were randomized to 1 of 4 treatments lasting 8 weeks and were followed for 6 months posttreatment. The trial used a 2X2 factorial design (coached vs self-guided treatment and weekly app recommendations vs no recommendations) to compare engagement metrics.

Results: The median time to last use of any app during treatment was 56 days (interquartile range 54-57), with 253 participants (84.0%, 253/301) continuing to use the apps over a median of 92 days posttreatment. Receipt of weekly recommendations resulted in a significantly higher number of app use sessions during treatment (overall median=216; P=.04) but only marginal effects for time to last use (P=.06) and number of app downloads (P=.08). Coaching resulted in significantly more app downloads (P<.001), but there were no significant effects for time to last download or number of app sessions (P=.36) or time to last download (P=.08). Participants showed significant reductions in the Patient Health Questionnaire-9 (PHQ-9) and Generalized Anxiety Disorder-7 (GAD-7) across all treatment arms (P s<.001). Coached treatment led to larger GAD-7 reductions than those observed for self-guided treatment (P=.03), but the effects for the PHQ-9 did not reach significance (P=.06). Significant interaction was observed between receiving recommendations and time for the PHQ-9 (P=.04), but there were no significant effects for GAD-7 (P=.58).

Conclusions: IntelliCare produced strong engagement with apps across all treatment arms. Coaching was associated with stronger anxiety outcomes, and receipt of recommendations enhanced depression outcomes.

Trial Registration: ClinicalTrials.gov NCT02801877; https://clinicaltrials.gov/ct2/show/NCT02801877

(J Med Internet Res 2019;21(8):e13609) doi:10.2196/13609

XSL•FO
RenderX

KEYWORDS

depression; anxiety; mHealth; clinical trial

Introduction

Background

Depression and anxiety are common mental health problems that impose a very high societal burden in terms of cost, morbidity, quality of life, and disability worldwide [1-4]. Most people experiencing these common mental health problems cannot access treatment because of a variety of barriers including the lack of availability of services, time constraints, transportation problems, and high cost [5,6]. A wide variety of digital mental health interventions have demonstrated efficacy. Adherence and outcomes appear to be stronger when coupled with human coaching through telephone or messaging than in self-guided digital interventions [7-9]. Mobile apps, in particular, have a number of advantages. Smartphones are becoming ubiquitous in developed countries and are increasingly common in developing nations [10]. As people keep their phones with them, app-based interventions can fit more seamlessly into the fabric of people's lives.

Most digital interventions for depression or anxiety are single Web-based and mobile apps. Very few of the publicly available apps have been rigorously tested, with reviews suggesting that the percentage of available apps for depression and anxiety with any evidence of effectiveness may be around 2.6% to 3.8% [11,12]. A common design approach is to adapt an effective psychotherapeutic model such as cognitive behavioral therapy (CBT) and digitize it into an app format. Such apps typically contain a multitude of features related to the treatment model on which they are based. For example, in CBT, these apps might contain psychoeducation, symptom tracking, activity monitoring, activity scheduling, and cognitive restructuring [13]. This approach of feature-rich apps does not recognize how most people currently tend to use apps and their smartphones. In general, digital technologies have core features that break into basic components, allowing users to piece together those components that are most useful. For example, people tend not to have 1 app to meet their transportation needs; rather, they usually have multiple apps that search for locations and map routes, and different apps for different modes of transportation. Different people likely use different apps according to their preferences and lifestyles. Typically, popular apps serve singular purposes, such as searching for restaurants or businesses, managing flights, or posting pictures. People tend to use apps in very short bursts of time, and sometimes frequently [14,15]. Thus, apps tend to support a single or limited set of related tasks through simple, quick interactions. Indeed, even when people do use existing mental health apps, they typically only use 1 or 2 features. For example, in an evaluation of PE Coach, which contains a number of features, it was found that clinicians and patients mostly used the app to audio-record sessions [16]. This gap between the design of mental health apps, which is typically based on complex psychotherapy models, and how people use their devices, which commonly occurs in short bursts for single purposes, likely is 1 reason for the low engagement with mental health apps seen in real-world settings [17].

The IntelliCare platform was designed to address user engagement problems by providing self-help strategies and skills training in a manner that is consistent with how people use mobile phones. Rather than a single app containing a comprehensive set of behavioral strategies, the IntelliCare platform currently comprises 12 apps, each of which is focused on a single psychological or behavioral strategy. The time required for each use is short, with most uses lasting less than 1 min [18]. Thus, apps can be integrated more seamlessly into a person's life. Users are able to select and use apps that they find helpful and ignore the ones they do not like. A Hub app, if downloaded, coordinates the user's experience and provides weekly recommendations to try new apps. The provision of recommendations appears to be an important component in maintaining engagement with the IntelliCare platform. Indeed, the receipt of these recommendations has been shown to increase the likelihood that an individual will download the recommended app [19,20].

IntelliCare apps have been available on the Google Play Store, beginning in 2014, and have been downloaded more than 100,000 times. An initial field trial, in which participants received 8 weeks of coaching primarily through text messages, showed substantial improvements in both depressive and anxiety symptoms' severity, and strong, consistent engagement of an average of 3 to 4 app launches each day over the full 8 weeks [18]. Interestingly, although the coached field trial showed participants used most of the apps, the average use within any individual app was substantially different from that observed among users who simply downloaded the apps through the Google Play Store, suggesting that coaches may have encouraged exploration of new apps but did not appear to influence continued use once users had downloaded and tried the apps. Thus, there appear to be 2 methods of encouraging exploration in app platforms such as IntelliCare: coaching and automated recommendations.

Although sustained behavioral engagement, or app usage, has been noted to be a major problem in digital mental health interventions [21,22], few investigations have systematically explored different methods of maintaining engagement. Moreover, 1 study explored a factorial design of 5 different engagement elements, including automated versus human support, text messages, tailoring of success stories, personalization of content, and multimedia and interactive materials [23]. Human support was the only element that improved outcomes during the intervention period, which is consistent with a large body of literature showing human support improves engagement and outcomes [7,24]. However, those who received automated support experienced more change during the postintervention period. Given the ubiquity of app store recommendations, this is a promising element to evaluate.

Objectives

This study aimed to examine the effect of 2 separate methods of maintaining engagement with the IntelliCare platform: coaching and receipt of weekly recommendations to try different

XSL•FO
RenderX

apps. We hypothesized that coaching would produce better engagement with the apps and greater reductions in symptoms of depression and anxiety than that associated with self-guided use, and those who received both coaching and app recommendations would have the greatest reductions in symptoms and highest use of apps.

Methods

Participants

Participants were recruited from July 5, 2016, to May 5, 2017, through a variety of digital (eg, Instagram, Facebook, and Reddit) and print (eg, advertisements on Chicago Transit Authority bus and train lines) sources as well as research registries (eg, ResearchMatch), commercial recruitment firms (eg, Focus Pointe Global), and media coverage using methods that have been previously described [25,26]. Participants were included if they met criteria for depression (Patient Health Questionnaire-9 [PHQ-9]≥10) [27] or anxiety (Generalized Anxiety Disorder-7 [GAD-7]≥8) [28], were aged 18 years or older (aged 19 years if in Nebraska, given age of consent), resided in the United States, could speak and read English, and had an Android phone with data and text plans. Participants were excluded if they (1) had visual, voice, motor, or hearing impairments that would prevent participation; (2) met diagnostic criteria for a severe psychiatric disorder such as psychotic or bipolar disorders for which study treatments would be inappropriate; (3) imminent suicidality that included both a plan and intent; (4) had initiated or modified antidepressant pharmacotherapy in the previous 14 days; or (5) had used any IntelliCare app more than 1 time in the 3 months before study screening.

All procedures were approved by the institutional review board of Northwestern University. Participants completed a Web-based consent form, and a research assistant reviewed the Web-based consent document to ensure comprehension questions were answered correctly and that the consent form was signed. Any questions or concerns were then reviewed with participants. The trial was monitored by an independent data safety monitoring board.

Treatments

This trial used a 2 × 2 factorial design (coached vs self-guided treatment and weekly app recommendations vs no recommendations), resulting in 4 treatment cells. A no treatment or waitlist control was not included because it would be impossible to prevent control participants from accessing the apps, which are freely available on the Google Play Store, and it could not be reliably determined which apps were accessed and when. The IntelliCare platform, coaching protocol, and the recommended system are described below.

IntelliCare Platform

All participants received access to the IntelliCare platform apps. At the time of this trial, the IntelliCare platform consisted of 13 apps [18,20]. This included 12 clinical apps, each of which was designed to target a specific behavioral or psychological treatment strategy (eg, goal setting, behavioral activation, social support, living one's values, cognitive restructuring, emotion regulation, positive self-affirmations, coping, exercise for mood, sleep hygiene, relaxation, and psychoeducation with reminders) and improve symptoms of depression and anxiety through efficacious treatment strategies. Most apps were designed to require less than 30 seconds to use. Apps included automated reminders to encourage engagement and for both app use and implementation of the strategies. The user's experience with the clinical apps was coordinated through a Hub app that consolidated automated notifications and provided app recommendations for those who were randomized to receive recommendations. No substantive changes were made to the apps during the course of the trial.

Coaching Versus Self-Guidance

Coaching was guided by the IntelliCare coaching manual [29], which is based on the supportive accountability model [30] and the efficiency model [31], and was aimed primarily at encouraging participants to try the apps, answering questions about how to use the tools represented in the apps and the rationale behind the skills taught by the apps, encouraging application of the skills in daily life, and providing some technical support as needed. Coaching began with an initial 30- to 45-min engagement phone call to explain the program, understand the participant's goals for mood and anxiety management, set expectations for the coach-participant relationship, build rapport, and ensure the Hub app was properly installed on the participant's phone. After the initial engagement call, participants received 2 to 3 text messages per week from their coach to provide support in using apps, offer encouragement, reinforce app use, and check-in on progress or challenges. Coaches also responded to all participant-initiated text messages within 1 working day. Coaches offered but did not require an additional 10-min call around midtreatment to support engagement. The coaches had a dashboard that provided information about the IntelliCare apps on each participant's phone, including which apps were installed, when they were downloaded, each time an app was used, and which apps were selected as *primary* in the Hub app. The dashboard also included a short message service text messaging tool, a section for brief notes, and an alert indicating when no IntelliCare app had been used for 3 days or when a participant sent a text message indicating they might be at risk for self-harm, which resulted in an automated safety response and prompted coaches to check-in. Coaches had at least a bachelor's degree in psychology or a related field and were trained and monitored by 1 of the coaching manual authors.

Participants assigned to self-guidance received the initial 10- to 15-min engagement call to ensure the Hub app was properly installed and that they understood how to use the IntelliCare platform but had no further contact with coaches.

Recommendations

Participants randomized to the recommendation arm received a weekly tray notification on their phone's home screen, which took them to the Hub app, where they were provided with weekly recommendations for new apps. Touching the recommendation button took them to the app store where they could download the recommended app. The recommendation engine used app usage data from approximately 100,000 users

XSL•FO
RenderX

who had downloaded the IntelliCare apps from the Google Play Store to identify, based on the individual user's app use profile, apps that the individual was more likely to use. Participants were asked to at least to try the newly recommended apps but were encouraged to use the apps they found most helpful.

Those participants who were in the no recommendation arm had the recommendation feature in the Hub app removed and were simply encouraged to explore the apps on the IntelliCare platform.

Outcome Assessment

Depression was measured using the PHQ-9 [27], and anxiety was measured using the GAD-7 [32]. These measures were administered as Web-based self-reports through Research Electronic Data Capture [33] at baseline, week 4, week 8 (end of treatment), and 3- and 6-month follow-up and completed by the participants themselves.

Engagement was defined using 3 commonly used behavioral engagement metrics [34]: *time to last use*, *number of app sessions*, and *number of apps downloaded*. Number of app sessions is a very common metric. *Time to last use* was defined as the time between the first launch and the last launch of any app during the 8-week trial or posttrial period. Posttrial app use data were truncated at 6 months to avoid biases related to time of entry into the trial. *Number of app sessions* is a commonly used metric. In this study, an app use session was defined as a sequence of user-initiated actions or events separated by less than 5 min between events. A new app launch (or session) was defined as a new activity after 5 min of no activity (we note that some apps have audio or video content that may last longer than 5 min, in which case, the running content is counted as activity). *Number of apps downloaded* is similar to the number of features or modules used in feature-rich applications [34]. Given the IntelliCare unbundled features into individual apps, this was defined here as the number of apps downloaded with at least one launch.

Randomization and Masking

A statistician provided a sequentially masked randomization scheme, created before the start of the trial, assigning participants to (1) coached with recommendations, (2) coached without recommendations, (3) self-guided with recommendations, and (4) self-guided without recommendations, stratified by current antidepressant medication status and psychotherapy status, with a block size of 4 within each stratum.

Statistical Analyses

Power was calculated on the assumption that approximately 50% of patients stop using the apps by week 7 based on previous electronic health work [35]. We powered for an effect size of 0.30 or a difference between 50% and 61.6% between groups. On the basis of a type I error rate of 5% and 80% power, power calculations using a log-rank test indicated a required sample size of 135 per arm (total sample of 270). Assuming 15 participants would be lost to follow-up in each arm, we aimed to recruit 150 participants in each arm. This provided power to detect effect sizes of 0.34 for clinical outcomes based on independent *t* tests. Week 7 was selected, rather than week 8, to avoid any *end-of-treatment* effects that might occur, such as participants ceasing or reinitiating engagement as a result of approaching end of treatment. Power calculations were performed using PASS 2008. No power calculations were performed to determine effect sizes detectable when examining the relationship between use metrics and patient-centered outcomes, as those were secondary aims.

Descriptive statistics are provided for baseline demographic variables across the 4 groups. Log-rank tests were performed to determine if the time to study dropout was different for each of the main effects, Kaplan-Meier plots are presented, and the engagement rate at 7 weeks. Cox proportional hazard models were used to compare the main effects while adjusting for randomization strata, medication use, age, and sex. Mean and standard deviation for PHQ-9 and GAD-7 over time and randomization groups are presented. Generalized linear mixed models were used to compare patient outcomes, adjusting for randomization strata, and baseline values of PHQ-9 or GAD-7, and assuming a heterogeneous unstructured covariance structure by randomization strata. First, 3-way interactions between time and the main effects of the 2X2 factorial were tested. If those were not significant, 2-way interactions with each main effect and time were modeled and, subsequently, models with main effects without the interaction. If interactions were not significant, within the main effect of time, least square means and differences adjusting for randomization strata, relative to baseline, were tested using Dunnett adjustment for multiple comparisons. Owing to the skewed nature of use data, app launches and time to last use were compared using nonparametric Kruskal-Wallis tests when comparing all 4 groups and Wilcoxon rank-sum test if comparing 2 groups. The relationship between use data and treatment outcome was examined using linear models adjusting for baseline PHQ-9 or GAD-7 and randomization strata. All analyses were run in R version 3.5.1 or SAS version 9.4 [36].

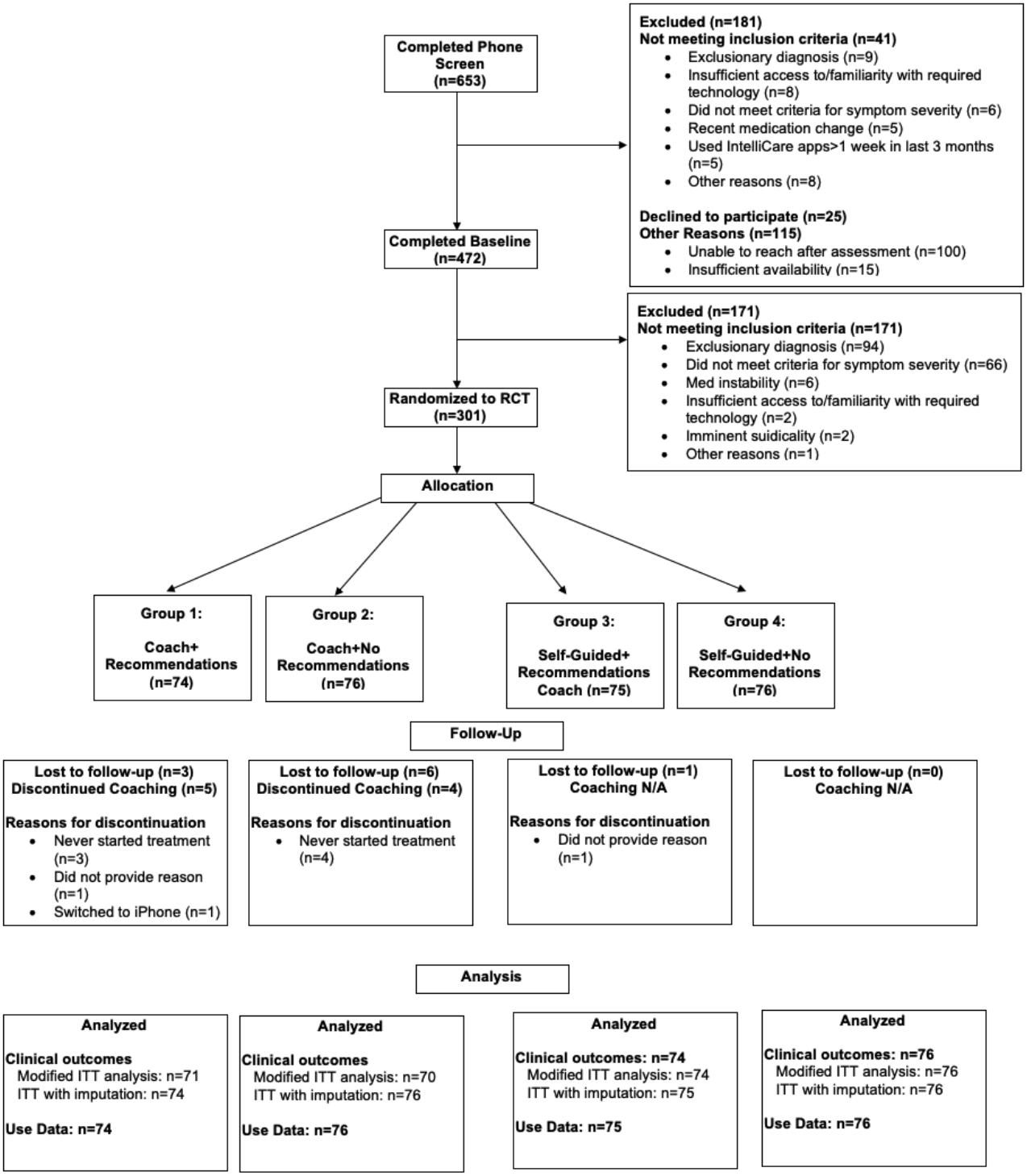
Results

Participants

The flow of participants through the study is depicted in Figure 1. Lost to follow-up rates differed significantly across treatment cells (P=.02). Table 1 summarizes the baseline demographics and psychiatric characteristics of the participants.

XSL•FO
RenderX

Figure 1. Flow of participants through the trial.



XSL•FO
RenderX

Table 1. Participants' characteristics.

Characteristics	Coached or recommendations (N=74)	Self-guided or recommendations (N=75)	Recommendations or coached (N=76)	No recommendations or self-guided (N=76)
Age (years), mean (SD)	37.57 (12.22)	36.17 (11.49)	37.09 (12.28)	35.34 (11.46)
Gender, n (%)				
Female	57 (77)	54 (72)	62 (81)	55 (72)
Male	15 (20)	21 (28)	14 (18)	21 (27)
Other	2 (2)	0 (0)	0 (0)	0 (0)
Race, n (%)				
White	57 (77)	63 (84)	57 (75)	60 (78)
Black	10 (13)	5 (6)	8 (10)	6 (7)
Asian	3 (4)	0 (0)	5 (6)	2 (2)
Other	4 (5)	7 (9)	6 (7)	8 (10)
Ethnicity, n (%)				
Non-Hispanic	62 (83)	67 (89)	68 (89)	71 (93)
Hispanic	10 (13)	8 (10)	7 (9)	5 (6)
Missing	2 (2)	0 (0)	1 (1)	0 (0)
Insurance=yes, n (%)	69 (93)	62 (82)	73 (96)	70 (92)
Marital status, n (%)				
Married/partnered	36 (48)	36 (48)	36 (47)	43 (56)
Single	30 (40)	29 (38)	31 (40)	26 (34)
Separated/divorced/widowed	8 (10)	10 (13)	9 (11)	7 (9)
Education, n (%)				
High school or less	4 (5)	4 (5.3)	4 (5)	3 (3)
Some college	16 (21)	18 (24)	11 (14)	15 (19)
College degree	53 (71)	53 (70)	61 (80)	58 (76)
Missing	1 (1)	0 (0)	0 (0.0)	0 (0.0)
Household income median, US$ (IQR[b])	58,000.00 (39,000.00-100,000.00)	50,000.00 (27,000.00-80,000.00)	60,000.00 (32,000.00-92,000.00)	57,500.00 (37,750.00-100,000.00)
Antidepressant status=yes, n (%)	35 (47)	31 (41)	37 (48)	37 (48)
Baseline GAD-7[c], mean (SD)	11.86 (4.05)	11.88 (3.85)	12.33 (4.51)	11.84 (3.66)
Baseline PHQ-9[d], mean (SD)	12.78 (4.45)	13.24 (4.55)	13.11 (4.75)	13.70 (4.79)

[a]IQR: interquartile range.

[b]GAD-7: Generalized Anxiety Disorder-7.

[c]PHQ-9: Patient Health Questionnaire-9.

Engagement Outcomes

Time to Last Use

The median time to last use of any app was 56 days (IQR 54-56). Figure 2 shows the Kaplan-Meier estimates for the time to last engagement by the groups. There was no significant difference in coached versus self-guided treatment (log-rank *P*=.94), and the difference between recommendation versus no recommendation did not reach significance (log-rank *P*=.06). The mean engagement percentages and 95% CIs for the number of last uses at 7 weeks or after for the coached and noncoached treatment were 90.7% (86.1%-95.4%) and 83.4% (77.7%-89.6%), and for recommendations versus no recommendations were 88.6% (83.6%-93.8%) and 85.5% (80.1%-91.3%), respectively.

XSL•FO
RenderX

Figure 2. Survival analysis: time to last app launch by treatment cell. Recs: recommendations.

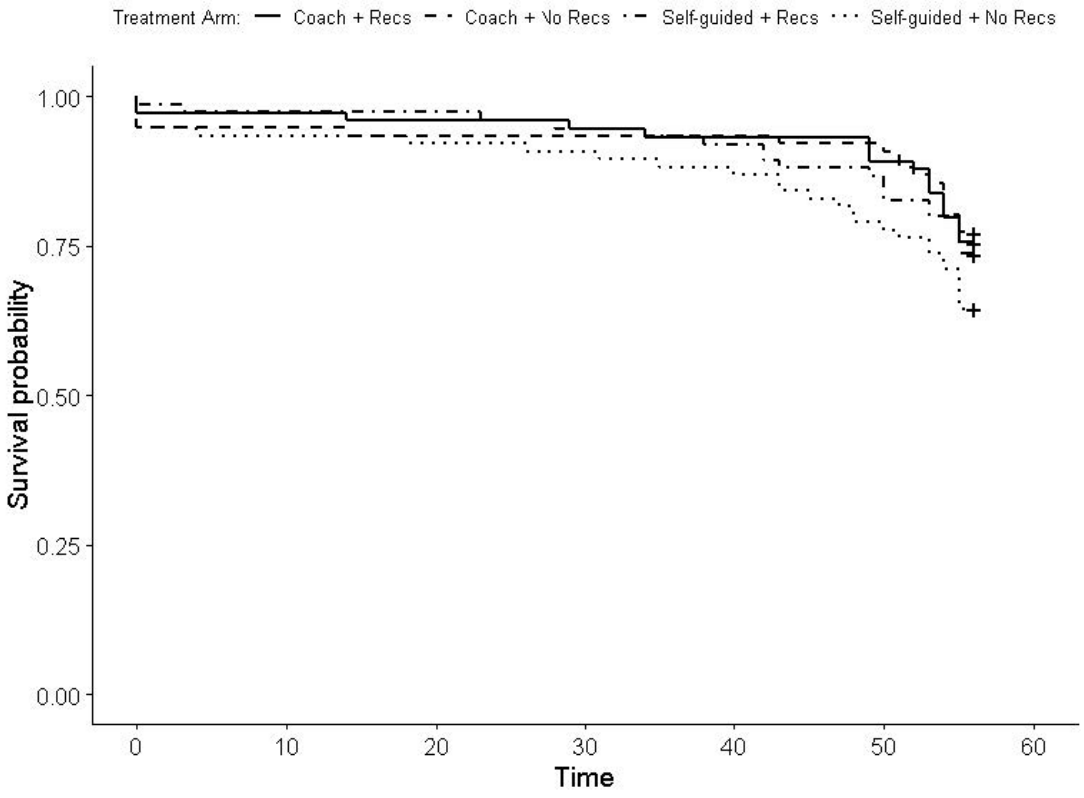


Table 2. Median (interquartile range) app sessions by week and treatment arm.

Week	Coached	Self-guided	Recommendations	No recommendations	All treatments
1	19 (12-27)	21 (7-46)	19 (9-29)	21.5 (11.75-40)	20 (10-35)
2	24.5 (14-38.5)	27 (11-43)	24 (13-40)	27 (13.75-43)	25 (13-41)
3	29 (15-44)	27 (11-42)	27 (12-46)	28 (13.75-40.25)	28 (13-43)
4	29 (14-44)	27 (14-40.5)	29 (16-51)	24.5 (12.75-38)	29 (14-43)
5	30 (13.25-50.75)	26 (11-42)	29 (13-49)	27 (11-41.25)	28 (13-45)
6	32.5 (14.25-49.75)	23 (11-40)	34 (14-54)	22.5 (12.75-39.25)	27 (13-46)
7	28.5 (12-43.75)	20 (7-38)	33 (12-53)	19 (8.75-35)	23 (9-42)
8	27 (11-40)	19 (5.5-39)	29 (9-48)	17.5 (7-34)	22 (8-40)
Total	215 (141-330.75)	218 (113-310)	232 (126-356)	201.5 (125.75-285.5)	216 (126-319)

Number of App Sessions

Table 2 displays the number of app launches by week across treatment arm. The median number of app sessions was 216 (IQR 126-325) across all apps. There was a significant effect for recommendations, with those receiving app recommendations having a median of 232 (IQR 126-356) app sessions versus those who did not receive recommendations, who had a median of 202 (IQR 126-286) app session (P=.04). There was no significant effect for coaching on number of app sessions (coached median 215 [IQR 141-331]; self-guided median 218 [IQR 113-310]; P=.36).

Number of Apps Downloaded

Participants who received coaching downloaded a median of 11 apps (IQR 10-12), whereas those who did not receive coaching downloaded a median of 7 apps (IQR 4-10), a difference that was statistically significant (Wilcoxon P<.001). The effect of receipt of recommendations on number of apps used did not reach significance (Wilcoxon P=.08).

Use Data During 6-Month Follow-Up

After completion of the trial, 253 (84.0%, 253/301) participants continued using the IntelliCare apps. Among those who continued to use the apps, the median time from the end of treatment to last use was 92 days (IQR 14-178), with a median of 83 (IQR 11-286) sessions. Neither length of use nor number of app sessions varied by treatment group (χ^2_3=1.4, P=.72; χ^2_3=0.3, P=.97, respectively).

XSL•FO
RenderX

Depression and Anxiety Outcomes

Table 3 displays the unadjusted outcomes across treatment groups. Participants showed significant reductions in the PHQ-9 and GAD-7 over time across both treatment arms ($F_{3,844}$=16.8, P<.001; $F_{3,844}$=16.8, P<.001, respectively).

Coached treatment produced significantly larger reductions in the GAD-7 than self-guided treatment ($F_{1,844}$=4.97; P=.03); however, there was no interaction between coaching and time ($F_{3,841}$=0.32; P=.81). The benefits of coaching did not reach significance for PHQ-9 ($F_{1,844}$=3.59; P=.06), and there was no evidence of an interaction with time ($F_{3,841}$=0.81; P=.49).

There was a significant interaction between receiving recommendations and time for the PHQ-9 ($F_{3,841}$=2.73; P=.04), such that those who received recommendations showed stronger improvements ($F_{3,841}$=14.56; P<.001) than those who did not ($F_{3,841}$=5.09; P=.002). Simple effects for recommendations on PHQ-9 are not reported, given the significant interaction effects. There was no significant effect of receiving weekly app recommendations on the GAD-7 ($F_{1,844}$=0.05; P=.82) and no significant interaction with time ($F_{3,841}$=0.65; P=.58).

There was no significant interactive effect of coaching and receiving recommendations over time for either the PHQ-9 or GAD-7 ($F_{3,835}$=0.19, P=.90; and $F_{3,835}$=0.73, P=.53, respectively).

Table 3. Unadjusted means (standard deviation) for the Patient Health Questionnaire-9 and Generalized Anxiety Disorder-7.

Treatment arm	Baseline	Week 4	Week 8	After 3 months	After 6 months
PHQ-9[a]					
Coached	12.95 (4.59)	8.81 (5.30)	7.17 (5.36)	7.03 (5.03)	6.93 (5.58)
Self-guided	13.47 (4.67)	9.56 (5.01)	8.43 (4.75)	8.45 (5.15)	8.32 (5.32)
Recommendations	13.01 (4.49)	8.81 (5.3)	7.51 (5.15)	7.41 (4.99)	7.74 (5.77)
No recommendations	13.40 (4.77)	8.99 (4.95)	8.13 (5.01)	8.12 (5.27)	7.56 (5.19)
GAD-7[b]					
Coached	12.1 (4.28)	7.86 (4.64)	6.76 (4.80)	6.26 (4.77)	6.24 (4.65)
Self-guided	11.86 (3.75)	8.34 (4.6)	7.45 (4.5)	7.19 (4.59)	6.99 (4.85)
Recommendations	11.87 (3.94)	8.3 (4.64)	7.06 (4.56)	6.65 (4.67)	6.66 (4.95)
No recommendations	12.09 (4.1)	7.91 (4.61)	7.17 (4.75)	6.83 (4.73)	6.59 (4.59)

[a]PHQ-9: Patient Health Questionnaire-9.

[b]GAD-7: Generalized Anxiety Disorder-7.

Secondary Analysis Using Multiple Imputation

As there was a small but statistically significant difference in the lost to follow-up rate, we conducted a secondary analysis using the expectation-maximization algorithm to impute 5 distinct datasets, in which 4-week outcomes were imputed for any participant who did not have at least one follow-up assessment. This allowed all participants to be included in our generalized linear mixed models. Parameter estimates and corresponding standard errors from each of the 5 models were combined and included in the SAS MIANALYZE procedure to derive valid inferences for the parameters of interest. Conclusions drawn from those analyses were consistent with those presented in our results above, namely, the interaction of the recommendation system and time for PHQ-9 (P=.04), the effect of coaching on GAD-7 (P=.02), and the changes in PHQ-9 and GAD-7 over time (P s<.001).

Relationship Between App Use and Outcomes

End of treatment PHQ-9, controlling for baseline, was significantly related to the number of app sessions (beta=−.01; P<.001), time to last use (beta=−.09; P=.001), and number of apps downloaded (beta=−.26; P=.001). GAD-7 outcome, controlling for baseline, was significantly related to the number of app downloads (beta=−.16; P=.03), but the effect did not reach significance for number of app sessions (beta=−.003; P=.05) and time to last use (beta=−.04; P=.08) did not reach significance. There were no significant interaction effects for treatment arm and PHQ-9 with number of app sessions (P=0.26), time to last use (P=0.70), or number of downloads (P=0.69). Similarly, there were no signification interaction effects for treatment arm and GAD-7 with number of app sessions (P=0.49), time to last use (P=0.77), or number of downloads(P=0.64).

Discussion

Principal Findings

Participants using the IntelliCare app platform showed substantial reductions in symptoms of depression and anxiety, similar to effects previously observed [18]. Coaching resulted in significantly lower levels of anxiety relative to self-guided treatment; however, the effect of coaching on depression was only marginal (P=.06). Although there was a difference between depression and anxiety in whether the criterion for significance was met, both P values were close to the .05 cutoff, and thus, there was no meaningful difference in the effect of coaching on depression versus anxiety. Receiving weekly recommendations resulted in significantly greater reductions in depression than

XSL•FO
RenderX

not receiving recommendations, but there was no similar effect for anxiety. This difference was large. We speculate that the recommendations are more useful for people with depression, as they address motivational challenges faced by people with depression.

App use was strong, with a median of 216 app sessions per participant (an average of 3.9 sessions per day) and a median last day of use being day 56 of 56 days of treatment, with no substantial change in the rate of use over the 8 weeks. Furthermore, 84.0% (253/301) of the participants continued using the apps for a median of 92 days after the completion of the 8-week treatment. This high level of engagement stands in stark contrast to most digital mental health apps, which tend to show sharp drop-offs in the first weeks [7,37]. This is likely because of 2 factors. First, the novelty of having new apps to use over the course of an 8-week treatment likely increases engagement [38]. Second, most of the apps are brief, requiring less than 30 seconds to use, allowing users to fit them into the context of their lives [18]. This suggests that the strategy of providing a platform of simple apps that patients can integrate into the fabric of their lives elicits stronger engagement than more traditional forms of digital mental health that are based on psychotherapy models and require greater time commitments.

People who received weekly recommendations to try new apps engaged in significantly more app sessions, compared with those not receiving recommendations. There was a similar trend for an effect of weekly recommendations on number of apps downloaded and time to last use, although these did not reach significance. These findings are generally consistent with findings of studies of IntelliCare downloads from the Google Play Store, in which users who had installed the Hub app and received recommendations were more likely to download recommended apps and used them more frequently, compared with those who did not download the Hub app and therefore did not receive recommendations [19,20]. These findings support the idea that providing recommendations for new apps on a regular basis can promote behavioral engagement with an app platform.

Participants in the coached conditions downloaded more apps than did those who were self-guided; however, there were no effects of coaching on any other use metrics. This suggests that coaches can help people stay engaged with the platform by trying new apps. However, we speculate that once an app is on the person's phone, the user's determination as to whether it is of sufficient value to continue using it is less modifiable by coaches, at least with the present coaching model.

Although the effect of coaching on anxiety and depression was significant or marginally significant, the effect sizes were smaller than we had expected. This stands in contrast to a number of meta-analyses that have consistently shown coaching to have a strong effect [8,39]. There are a number of potential reasons for our weaker than expected findings for coaching. First, the uncoached participants did have an initial 10-min call with a coach to ensure the app was properly installed and that the person knew how to engage with the platform. This may have provided some motivation and reduced confusion that could have led to nonengagement. It is also possible that the automated reminders to use the apps that are part of each app's design fulfilled some of the coaches' function in encouragement. However, such reminders are common features of mobile apps, and thus, we expect these automated reminders alone do not fully account for the weak coaching effects. The weaker than expected effects of coaching may be because of the strong app usage observed in this study. The design of the IntelliCare platform emphasized the usability of apps over the application of a theory-based approach, such as modeling the design of an app on CBT. Apps were designed to be simple and quick to use, thereby fitting into the fabric of users' lives. Coaches have often been employed to encourage the use of intervention technologies. It may be that the coach's role of encouraging adherence becomes less important as we improve the usability of the apps. Although coaches will likely continue to be beneficial for some people using the IntelliCare platform, these findings suggest that better design of technologies may limit the need for human support, thereby increasing their scalability.

The relationship between engagement metrics and outcomes was mixed, with number of app use sessions being significantly related to both depression and anxiety outcomes, but time to last use and number of app downloads were only related significantly to depression outcomes and not anxiety. However, consistent with much of the literature, even where relationships were significant, they were not strong [40]. There are a number of potential reasons for this. One is that the engagement metrics were strong with fairly high consistency, and thus, the weaker findings may be an artifact of this low variability. However, it is also likely that the relationship between these behavioral engagement metrics and overall symptom change during the intervention obscures more complex relationships. For example, although engagement may reduce symptoms, higher symptoms may increase engagement in the immediate time frame [35]. Thus, simple associations between overall engagement and symptom reduction over the course of treatment may obscure more complex relationships over shorter time frames. Another problem may be that behavioral engagement metrics do not capture meaningful engagement. Indeed, the field of human-computer interaction has viewed engagement more holistically than psychology, considering not only behavioral metrics but also many other subjective factors related to the user's cognitive and emotional engagement with the apps and intervention [41,42]. This richer conceptualization of engagement may provide a richer understanding of the user's experience and thus may be more strongly related to clinical outcomes relative to metrics that rely solely on app usage data.

Limitations

This study has a number of limitations that should be considered in the interpretation of these results. First, lost to follow-up rates were slightly albeit significantly greater in the coached arms than in the noncoached. It is notable that most of those in the coaching arms who were lost to follow-up never initiated treatment, suggesting a small number of people may prefer uncoached interventions. Nevertheless, this difference in lost to follow-up rates across the treatment cells likely did not impact the findings, as it was very small (9% in arm with the highest rate and 3% overall), and secondary analyses imputing missing values showed no difference in outcomes. Second, as a research

XSL•FO
RenderX

study, participants had to go through the usual consenting and screening procedures and agree to regular follow-up assessments. This likely resulted in a sample that was more motivated to engage in digital mental health treatment than the average person with depression or anxiety. Thus, the robust level of engagement seen in this sample may not be strong in real-world treatment settings. It is possible that strategies to support engagement (ie, coaching and recommendations) may be more important for less motivated groups. Similarly, we did not conduct diagnostic evaluations and therefore cannot determine how this sample may or may not be similar to populations in health care settings. Finally, this trial did not control for receipt of the IntelliCare apps. Thus, although the reductions in depression and anxiety are substantial, we cannot rule out the possibility that these changes are because of factors other than the treatment, such as the natural course of the symptoms, user expectancies, or research procedures such as repeated assessment.

Conclusions

This study explores a new paradigm in digital mental health interventions. IntelliCare moves away from the single app for a mental health problem and recognizes 1 of the basic properties of digital tools—that they are broken down into their smallest, simplest elements possible, thereby allowing people to bundle them as they see fit [43]. The strong engagement in IntelliCare suggests that this principle also applies to digital mental health interventions and tools. Engagement with the platform is enhanced through weekly recommendations to try new apps. There is some support for the use of coaches to enhance anxiety outcomes and recommendations to enhance depression outcomes. This suggests that coaching may not be necessary for all people using modular, platform-based digital mental health treatments such as IntelliCare.

Acknowledgments

This work was supported by the United States National Institute of Mental Health grant R01 MH100482 to DCM. EL was supported by a research grant K08 MH112878 from the National Institute of Mental Health. SS is an investigator with the Implementation Research Institute, Washington University, St. Louis, and received an award from the National Institute of Mental Health (R25-MH08091607) and the Department of Veterans Affairs, Health Services Research and Development Service, Quality Enhancement Research Initiative.

Conflicts of Interest

DM has accepted honoraria from Apple Inc and has an ownership interest in Actualize Therapy, which has a license from Northwestern University for IntelliCare. EL has received consulting fees from Actualize Therapy. SS serves as a scientific advisor to Joyable, Inc, and has received stock options in Joyable. The other authors have no conflicts of interest to declare.

References

1. Baxter AJ, Scott KM, Vos T, Whiteford HA. Global prevalence of anxiety disorders: a systematic review and meta-regression. Psychol Med 2013 May;43(5):897-910. [doi: 10.1017/S003329171200147X] [Medline: 22781489]
2. Ferrari AJ, Somerville AJ, Baxter AJ, Norman R, Patten SB, Vos T, et al. Global variation in the prevalence and incidence of major depressive disorder: a systematic review of the epidemiological literature. Psychol Med 2013 Mar;43(3):471-481. [doi: 10.1017/S0033291712001511] [Medline: 22831756]
3. GBD 2015 Disease and Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990-2015: a systematic analysis for the global burden of disease study 2015. Lancet 2016 Oct 8;388(10053):1545-1602 [FREE Full text] [doi: 10.1016/S0140-6736(16)31678-6] [Medline: 27733282]
4. Greenberg PE, Fournier AA, Sisitsky T, Pike CT, Kessler RC. The economic burden of adults with major depressive disorder in the United States (2005 and 2010). J Clin Psychiatry 2015 Feb;76(2):155-162 [FREE Full text] [doi: 10.4088/JCP.14m09298] [Medline: 25742202]
5. Mohr DC, Ho J, Duffecy J, Baron KG, Lehman KA, Jin L, et al. Perceived barriers to psychological treatments and their relationship to depression. J Clin Psychol 2010 Apr;66(4):394-409 [FREE Full text] [doi: 10.1002/jclp.20659] [Medline: 20127795]
6. Mohr DC, Hart SL, Howard I, Julian L, Vella L, Catledge C, et al. Barriers to psychotherapy among depressed and nondepressed primary care patients. Ann Behav Med 2006 Dec;32(3):254-258. [doi: 10.1207/s15324796abm3203_12] [Medline: 17107299]
7. Andrews G, Basu A, Cuijpers P, Craske MG, McEvoy P, English CL, et al. Computer therapy for the anxiety and depression disorders is effective, acceptable and practical health care: an updated meta-analysis. J Anxiety Disord 2018 Apr;55:70-78 [FREE Full text] [doi: 10.1016/j.janxdis.2018.01.001] [Medline: 29422409]
8. Firth J, Torous J, Nicholas J, Carney R, Pratap A, Rosenbaum S, et al. The efficacy of smartphone-based mental health interventions for depressive symptoms: a meta-analysis of randomized controlled trials. World Psychiatry 2017 Oct;16(3):287-298 [FREE Full text] [doi: 10.1002/wps.20472] [Medline: 28941113]

XSL•FO
RenderX

9. Firth J, Torous J, Nicholas J, Carney R, Rosenbaum S, Sarris J. Can smartphone mental health interventions reduce symptoms of anxiety? A meta-analysis of randomized controlled trials. J Affect Disord 2017 Aug 15;218:15-22 [FREE Full text] [doi: 10.1016/j.jad.2017.04.046] [Medline: 28456072]
10. Poushter J. Pew Research Center. 2017. Smartphones Are Common in Advanced Economies, But Digital Divides Remain URL: http://www.pewresearch.org/fact-tank/2017/04/21/smartphones-are-common-in-advanced-economies-but-digital-divides-remain/
11. Larsen ME, Nicholas J, Christensen H. Quantifying app store dynamics: longitudinal tracking of mental health apps. JMIR Mhealth Uhealth 2016 Aug 9;4(3):e96 [FREE Full text] [doi: 10.2196/mhealth.6020] [Medline: 27507641]
12. Sucala M, Cuijpers P, Muench F, Cardo R, Soflau R, Dobrean A, et al. Anxiety: there is an app for that. A systematic review of anxiety apps. Depress Anxiety 2017 Jun;34(6):518-525. [doi: 10.1002/da.22654] [Medline: 28504859]
13. Huguet A, Rao S, McGrath PJ, Wozney L, Wheaton M, Conrod J, et al. A systematic review of cognitive behavioral therapy and behavioral activation apps for depression. PLoS One 2016;11(5):e0154248 [FREE Full text] [doi: 10.1371/journal.pone.0154248] [Medline: 27135410]
14. Vaish R, Wyngarden K, Chen J, Cheung B, Bernstein MS. Twitch Crowdsourcing: Crowd Contributions in Short Bursts of Time. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2014 Presented at: CHI'14; April 26-May 1, 2014; Toronto, Ontario, Canada p. 3645-3654. [doi: 10.1145/2556288.2556996]
15. Andrews S, Ellis DA, Shaw H, Piwek L. Beyond self-report: tools to compare estimated and real-world smartphone use. PLoS One 2015;10(10):e0139004 [FREE Full text] [doi: 10.1371/journal.pone.0139004] [Medline: 26509895]
16. Reger GM, Browne KC, Campellone TR, Simons C, Kuhn E, Fortney JC, et al. Barriers and facilitators to mobile application use during PTSD treatment: clinician adoption of PE coach. Prof Psychol: Res Pract 2017 Dec;48(6):510-517. [doi: 10.1037/pro0000153]
17. Torous J, Nicholas J, Larsen ME, Firth J, Christensen H. Clinical review of user engagement with mental health smartphone apps: evidence, theory and improvements. Evid Based Ment Health 2018 Aug;21(3):116-119. [doi: 10.1136/eb-2018-102891] [Medline: 29871870]
18. Mohr DC, Tomasino KN, Lattie EG, Palac HL, Kwasny MJ, Weingardt K, et al. IntelliCare: an eclectic, skills-based app suite for the treatment of depression and anxiety. J Med Internet Res 2017 Jan 5;19(1):e10 [FREE Full text] [doi: 10.2196/jmir.6645] [Medline: 28057609]
19. Cheung K, Ling W, Karr CJ, Weingardt K, Schueller SM, Mohr DC. Evaluation of a recommender app for apps for the treatment of depression and anxiety: an analysis of longitudinal user engagement. J Am Med Inform Assoc 2018 Aug 1;25(8):955-962 [FREE Full text] [doi: 10.1093/jamia/ocy023] [Medline: 29659857]
20. Lattie EG, Schueller SM, Sargent E, Stiles-Shields C, Tomasino KN, Corden ME, et al. Uptake and usage of IntelliCare: a publicly available suite of mental health and well-being apps. Internet Interv 2016 May;4(2):152-158 [FREE Full text] [doi: 10.1016/j.invent.2016.06.003] [Medline: 27398319]
21. Anguera JA, Jordan JT, Castaneda D, Gazzaley A, Areán PA. Conducting a fully mobile and randomised clinical trial for depression: access, engagement and expense. BMJ Innov 2016 Jan;2(1):14-21 [FREE Full text] [doi: 10.1136/bmjinnov-2015-000098] [Medline: 27019745]
22. Arean PA, Hallgren KA, Jordan JT, Gazzaley A, Atkins DC, Heagerty PJ, et al. The use and effectiveness of mobile apps for depression: results from a fully remote clinical trial. J Med Internet Res 2016 Dec 20;18(12):e330 [FREE Full text] [doi: 10.2196/jmir.6482] [Medline: 27998876]
23. Kelders SM, Bohlmeijer ET, Pots WT, van Gemert-Pijnen JE. Comparing human and automated support for depression: fractional factorial randomized controlled trial. Behav Res Ther 2015 Sep;72:72-80. [doi: 10.1016/j.brat.2015.06.014] [Medline: 26196078]
24. Richards D, Richardson T, Timulak L, McElvaney J. The efficacy of internet-delivered treatment for generalized anxiety disorder: a systematic review and meta-analysis. Internet Interv 2015 Sep;2(3):272-282. [doi: 10.1016/j.invent.2015.07.003]
25. Palac HL, Alam N, Kaiser SM, Ciolino JD, Lattie EG, Mohr DC. A practical do-it-yourself recruitment framework for concurrent ehealth clinical trials: simple architecture (part 1). J Med Internet Res 2018 Nov 1;20(11):e11049 [FREE Full text] [doi: 10.2196/11049] [Medline: 30389650]
26. Lattie EG, Kaiser SM, Alam N, Tomasino KN, Sargent E, Rubanovich CK, et al. A practical do-it-yourself recruitment framework for concurrent ehealth clinical trials: identification of efficient and cost-effective methods for decision making (part 2). J Med Internet Res 2018 Nov 29;20(11):e11050 [FREE Full text] [doi: 10.2196/11050] [Medline: 30497997]
27. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613 [FREE Full text] [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]
28. 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]
29. Noth KN, Bardsley LR, Lattie EG, Mohr DC. DigitalHub - Northwestern University. 2018. IntelliCare Study Coaching Manual URL: https://digitalhub.northwestern.edu/files/00fa4294-5b9f-4afc-897a-7fffceae8f3f
30. Mohr DC, Cuijpers P, Lehman K. Supportive accountability: a model for providing human support to enhance adherence to eHealth interventions. J Med Internet Res 2011 Mar 10;13(1):e30 [FREE Full text] [doi: 10.2196/jmir.1602] [Medline: 21393123]

XSL•FO
RenderX

31. Schueller SM, Tomasino KN, Mohr DC. Integrating human support into behavioral intervention technologies: the efficiency model of support. Clin Psychol: Sci Pract 2016 Nov 17;24(1):27-45. [doi: 10.1111/cpsp.12173]
32. Spitzer RL, Kroenke K, Williams JB. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. Primary care evaluation of mental disorders. Patient health questionnaire. J Am Med Assoc 1999 Nov 10;282(18):1737-1744. [doi: 10.1001/jama.282.18.1737] [Medline: 10568646]
33. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
34. Pham Q, Graham G, Carrion C, Morita PP, Seto E, Stinson JN, et al. A library of analytic indicators to evaluate effective engagement with consumer mhealth apps for chronic conditions: scoping review. JMIR Mhealth Uhealth 2019 Jan 18;7(1):e11941 [FREE Full text] [doi: 10.2196/11941] [Medline: 30664463]
35. Mohr DC, Duffecy J, Jin L, Ludman EJ, Lewis A, Begale M, et al. Multimodal e-mental health treatment for depression: a feasibility trial. J Med Internet Res 2010 Dec 19;12(5):e48 [FREE Full text] [doi: 10.2196/jmir.1370] [Medline: 21169164]
36. R Project. 2017. The R Project for Statistical Computing URL: https://www.R-project.org/
37. Gilbody S, Brabyn S, Lovell K, Kessler D, Devlin T, Smith L, REEACT Collaborative. Telephone-supported computerised cognitive-behavioural therapy: REEACT-2 large-scale pragmatic randomised controlled trial. Br J Psychiatry 2017 May;210(5):362-367. [doi: 10.1192/bjp.bp.116.192435] [Medline: 28254959]
38. Kwasny MJ, Schueller SM, Lattie EG, Gray EL, Mohr DC. Towards managing a platform of mental health apps: a secondary analysis of the IntelliCare field trial. J Med Internet Res 2019:- (epub ahead of print). [doi: 10.2196/preprints.11572]
39. 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]
40. Fuhr K, Schröder J, Berger T, Moritz S, Meyer B, Lutz W, et al. The association between adherence and outcome in an internet intervention for depression. J Affect Disord 2018 Mar 15;229:443-449. [doi: 10.1016/j.jad.2017.12.028] [Medline: 29331706]
41. Doherty K, Doherty G. Engagement in HCI: conception, theory and measurement. ACM Comput Surv 2019 Jan 23;51(5):1-39. [doi: 10.1145/3234149]
42. Perski O, Blandford A, West R, Michie S. Conceptualising engagement with digital behaviour change interventions: a systematic review using principles from critical interpretive synthesis. Transl Behav Med 2017 Jun;7(2):254-267 [FREE Full text] [doi: 10.1007/s13142-016-0453-1] [Medline: 27966189]
43. Kelly K. The Inevitable. New York: Viking Press; 2016.

Abbreviations

CBT: cognitive behavioral therapy
GAD-7: Generalized Anxiety Disorder-7
IQR: interquartile range
PHQ-9: Patient Health Questionnaire-9

Edited by S Kitsiou; submitted 05.02.19; peer-reviewed by N Zimmerman, L Cadmus-Bertram, S Kelders; comments to author 11.04.19; revised version received 06.06.19; accepted 20.07.19; published 28.08.19.

Please cite as:
Mohr DC, Schueller SM, Tomasino KN, Kaiser SM, Alam N, Karr C, Vergara JL, Gray EL, Kwasny MJ, Lattie EG
Comparison of the Effects of Coaching and Receipt of App Recommendations on Depression, Anxiety, and Engagement in the IntelliCare Platform: Factorial Randomized Controlled Trial
J Med Internet Res 2019;21(8):e13609
URL: http://www.jmir.org/2019/8/e13609/
doi:10.2196/13609
PMID:31464192

©David C Mohr, Stephen M Schueller, Kathryn Noth Tomasino, Susan M Kaiser, Nameyeh Alam, Chris Karr, Jessica L Vergara, Elizabeth L Gray, Mary J Kwasny, Emily G Lattie. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 28.08.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Factors Influencing Patients' Intentions to Use Diabetes Management Apps Based on an Extended Unified Theory of Acceptance and Use of Technology Model: Web-Based Survey

Yiyu Zhang[1,2,3], MMed; Chaoyuan Liu[4], MMed; Shuoming Luo[1,2,3], MD; Yuting Xie[1,2,3], MD; Fang Liu[1], MNurs; Xia Li[1,2,3*], MD; Zhiguang Zhou[1,2,3*], MD

[1]Department of Metabolism and Endocrinology, The Second Xiangya Hospital, Central South University, Changsha, China

[2]Key Laboratory of Diabetes Immunology, Ministry of Education, Changsha, China

[3]National Clinical Research Center for Metabolic Diseases, Changsha, China

[4]Department of Oncology, The Second Xiangya Hospital, Central South University, Changsha, China

[*]these authors contributed equally

Corresponding Author:
Zhiguang Zhou, MD
Department of Metabolism and Endocrinology
The Second Xiangya Hospital
Central South University
No 139, Renmin Road
Changsha, 410011
China
Phone: 86 073185292154
Fax: 86 073185367220
Email: zhouzhiguang@csu.edu.cn

Abstract

Background: Diabetes poses heavy social and economic burdens worldwide. Diabetes management apps show great potential for diabetes self-management. However, the adoption of diabetes management apps by diabetes patients is poor. The factors influencing patients' intention to use these apps are unclear. Understanding the patients' behavioral intention is necessary to support the development and promotion of diabetes app use.

Objective: This study aimed to identify the determinants of patients' intention to use diabetes management apps based on an integrated theoretical model.

Methods: The hypotheses of our research model were developed based on an extended Unified Theory of Acceptance and Use of Technology (UTAUT). From April 20 to May 20, 2019, adult patients with diabetes across China, who were familiar with diabetes management apps, were surveyed using the Web-based survey tool Sojump. Structural equation modeling was used to analyze the data.

Results: A total of 746 participants who met the inclusion criteria completed the survey. The fitness indices suggested that the collected data fit well with the research model. The model explained 62.6% of the variance in performance expectancy and 57.1% of the variance in behavioral intention. Performance expectancy and social influence had the strongest total effects on behavioral intention (β=0.482; *P*=.001). Performance expectancy (β=0.482; *P*=.001), social influence (β=0.223; *P*=.003), facilitating conditions (β=0.17; *P*=.006), perceived disease threat (β=0.073; *P*=.005), and perceived privacy risk (β=–0.073; *P*=.012) had direct effects on behavioral intention. Additionally, social influence, effort expectancy, and facilitating conditions had indirect effects on behavioral intention that were mediated by performance expectancy. Social influence had the highest indirect effects among the three constructs (β=0.259; *P*=.001).

Conclusions: Performance expectancy and social influence are the most important determinants of the intention to use diabetes management apps. Health care technology companies should improve the usefulness of apps and carry out research to provide clinical evidence for the apps' effectiveness, which will benefit the promotion of these apps. Facilitating conditions and perceived privacy risk also have an impact on behavioral intention. Therefore, it is necessary to improve facilitating conditions and provide

XSL•FO
RenderX

solid privacy protection. Our study supports the use of UTAUT in explaining patients' intention to use diabetes management apps. Context-related determinants should also be taken into consideration.

(J Med Internet Res 2019;21(8):e15023) doi:10.2196/15023

KEYWORDS

diabetes mellitus; mobile applications; survey; structural equation modeling; China

Introduction

Background

Diabetes poses heavy social and economic burdens worldwide. The estimated number of adult patients with diabetes in 2017 was 451 million worldwide, and this figure is expected to increase to 693 million by 2045 [1]. Nearly 5 million deaths in the adult population were caused by diabetes in 2017 [1]. According to a national survey in 2013, the prevalence of diabetes in China was estimated to be 10.9%, representing more than 100 million adults in China [2]. Optimal glycemic control can prevent diabetes-related complications [3]. However, in China, less than half of the patients treated for diabetes were found to have appropriate glycemic control [2]. Poor blood sugar control can lead to various life-threatening complications such as blindness, renal failure, stroke, and myocardial infarction [4]. The estimated cost of diabetes worldwide in 2015 was as high as US $1.31 trillion [5].

Diabetes self-management education and support are critical for diabetes management [6]. However, doctors in large hospitals in China are overloaded with work, and the time spent with each patient in outpatient departments is very limited and usually less than 3 min [7]. Diabetes patients receive little diabetes education in such a short time. Most patients with suboptimal glycemic control lack diabetes-related knowledge and self-care practices [8]. Moreover, due to the imbalance of medical resources in China, it is inconvenient for patients from remote rural areas to seek medical care in large hospitals [9]. Therefore, patients with diabetes in rural areas have a higher mortality [10].

Diabetes management apps enable patients to record their blood sugar, receive diabetes-related information, and communicate with health care providers and peers anytime and anywhere [11]. These apps show promising potential for diabetes self-management [12]. Several studies have shown that diabetes management apps have benefits such as glycosylated hemoglobin reduction [11,13-15], reduced feelings of loneliness [16], reduced hypoglycemic fears, and improved behavioral scores [17]. However, surveys have shown that the uptake of diabetes management apps among diabetes patients is poor. In a survey in America in 2014, the use of diabetes management apps was approximately 3.6% among Latino patients with diabetes [18]. In Australia, 8% of patients with type 2 diabetes reported using diabetes management apps [19]. Our previous Web-based survey also showed the same pattern, and only 10.8% of patients with type 2 diabetes used diabetes management apps [20]; these results were similar to those of a survey conducted in New Zealand [21].

The actual use of a technology is often determined by the intention of its use [22]. Understanding the factors that influence patients' use intention will help manufacturers further improve the design of diabetes management apps and promote their use. However, the factors influencing patients' intention to use diabetes management apps are unclear. Several studies have applied umbrella theoretical models to understand the determinants of use intentions for mobile health (mHealth) services [23-27] or health information technology [28]. However, a theoretical model must be identified and tested for different technologies and in different user groups, to provide a context-related understanding of technology adoption [22]. Diabetes management apps have unique functions, and patients with diabetes have unique characteristics. Therefore, it is necessary to analyze the factors influencing the use intention for diabetes management apps based on an integrated theoretical model. To our knowledge, relevant theoretical models have not been applied to the field of diabetes management apps.

Theoretical Background

Venkatesh et al [22] integrated the following eight theories (Table 1) to form the UTAUT: technology acceptance model (TAM), theory of reasoned action, motivational model, theory of planned behavior (TPB), combined TAM and TPB, model of personal computer use, diffusion of innovations theory, and social cognitive theory. They found that the UTAUT outperformed the eight individual models [22]. The UTAUT is the most frequently used theoretical model in information technology and has been applied to a wide range of areas, such as electronic health (eHealth) services [24,29-30], electronic medical record systems [31,32] and other health-related information technologies [33,34]. According to the UTAUT, performance expectancy, effort expectancy and social influence are the core determinants of behavioral intention, and facilitating conditions and behavioral intentions are direct determinants of use behavior. Performance expectancy and effort expectancy are equivalent to relative advantage and complexity of the diffusion of innovation theory [35,36]. Venkatesh proposed the updated UTAUT2 in a consumer information technology context and found a direct association between facilitating conditions and behavioral intentions. The new model incorporates three new constructs: hedonic motivation, price value, and habit [37]. However, patients do not use diabetes management apps for the intention of enjoyment. Moreover, the Tavares and Oliveira study concerning electronic health record patient portals did not find an association between hedonic motivation and behavioral intention [38]. Diabetes management apps are offered to users for free [39] and represent a relatively new technology in China; thus, we did not incorporate the new constructs of the UTAUT2.

XSL•FO
RenderX

Table 1. Summary of technology acceptance theories.

Theory	Application fields	Constructs
Technology acceptance model (TAM) [23,40]	Originally designed to predict the acceptance and use of information technology, TAM has been applied to a wide range of technologies and users	Perceived Usefulness, Perceived Ease of Use, Subjective Norm
Theory of reasoned action [41]	Originating from social psychology, this model has been used widely to predict human behaviors	Attitude Toward Behavior, Subjective Norm
Theory of planned behavior (TPB) [42,43]	Extension of the Theory of Reasoned Action to deal with behaviors over which people have incomplete volitional control	Attitude Toward Behavior, Subjective Norm, Perceived Behavioral Control
Motivational model [44,45]	Widely used in psychology to explain human behavior	Extrinsic Motivation, Intrinsic Motivation
Combined TAM and TPB [46]	A hybrid model of the TPB and TAM	Attitude Toward Behavior, Subjective Norm, Perceived Behavioral Control, Perceived Usefulness
Model of personal computer use [47]	This model was adopted to predict personal computer utilization	Job Fit, Complexity, Long-Term Consequence, Affect Toward Use, Social Factor, Facilitating Conditions
Diffusion of innovations theory [48]	Grounded from sociology, this model has been applied to a wide range of innovations, such as information systems	Relative Advantage, Ease of Use, Image, Visibility, Compatibility, Results Demonstrability, Voluntariness of Use
Social cognitive theory [49]	Widely used in social behaviors, this model was also applied to information technologies	Outcome Expectations - personal, Self-efficacy, Affect, Anxiety

Research Hypotheses

Performance expectancy, which is similar to perceived usefulness in the TAM, is defined as the degree to which use of a specific technology benefits users [37]. Several studies have shown that performance expectancy is a major determinant of the intention to use health information technologies [28,50-52]. Overall, patients tend to use eHealth tools that are beneficial for them [53]. Thus, we propose the following hypothesis:

> *H1: Performance expectancy positively influences the behavioral intention of patients to use diabetes management apps.*

Effort expectancy is defined as the degree of the ease of use of a specific technology [37]. If patients find mHealth technology easy to use, they will have a higher intention to use it. This hypothesis has been tested in many studies [29,36-38], especially among the elderly [24,54]. Therefore, we propose the following hypothesis:

> *H2: Effort expectancy positively influences the behavioral intention of patients to use diabetes management apps.*

The study by Alaiad [34], concerning home health care robots, found that effort expectancy is a strong determinant of performance expectancy, and the investigation of home telehealth services acceptance behavior by Cimperman al [55] also found such an association. Several other studies also revealed that performance expectancy was predicted by effort expectancy [30,56,57]. If patients find a technology easy to use, they may find it useful. Therefore, we pose the following hypothesis:

> *H3: Effort expectancy positively influences performance expectancy.*

Facilitating conditions are defined as the consumers' awareness of the available resources to support the use of a particular technology [37]. Although the original UTAUT model did not show a direct association between facilitating conditions and behavioral intention (showing an association between facilitating conditions and use) [22], the UTAUT2 and several other studies on the consumer environment demonstrated this relationship [24,33,37,58]. The facilitation available to each mobile app consumer can vary significantly across mobile devices and network access levels. Thus, we pose the following hypothesis:

> *H4: Facilitating conditions positively influence the behavioral intention of patients to use diabetes management apps.*

In their study regarding telemedicine for diabetes management, Rho et al [51] showed that facilitating conditions have an indirect effect on behavioral intention, which is mediated by performance expectancy [51]. Other studies also showed that facilitating conditions can affect performance expectancy [59]. Thus, we propose the following hypothesis:

> *H5: Facilitating conditions positively influence performance expectancy.*

Social influence is defined as the extent to which people think that others who are important to them or who can influence their behavior think that they should use a specific technology, and it is similar to the subjective norm in the TAM [22]. Studies regarding health information technologies showed that social influence affects behavioral intention [24,36]. In health care circumstances, patients' intention to adopt a health behavior is often influenced by their doctors, peers with the same disease,

XSL•FO
RenderX

and family members [60]. Thus, we propose the following hypothesis:

H6: Social influence positively influences the behavioral intention of patients to use diabetes management apps.

One study on a Web-based interactive self-management technology revealed that social influence affected behavioral intention indirectly through the mediation of perceived usefulness [61]. Since physicians are perceived as an expert authority, patients' perceived usefulness of a health care tool is often influenced by their physician's opinion. Thus, we propose the following hypothesis:

H7: Social influence positively influences performance expectancy.

Context can be defined as the environment in which a technology is used, and it may affect an individual's behavioral intention [62]. The UTAUT is not derived from the environment of health information technology consumers [22,37]. According to the Health Belief Model, individuals will not take health-related actions unless they feel susceptible to or experience the severity of a disease [63]. The model has been widely employed to predict health behavioral intentions [36,63,64]. Individuals with a higher perceived health threat have greater motivation to adopt mHealth apps [62]. In this study, perceived disease threat (PDT) refers to a patient's awareness of his/her hyperglycemia condition and concern for its potential consequences. Thus, we pose the following hypothesis:

H8: Perceived disease threat positively influences the behavioral intention of patients to use diabetes management apps.

An investigation by Ahadzadeh et al [64] found that perceived health risk and health consciousness influenced perceived usefulness of the health-related internet [64], and a study by Dou et al [65] on a hypertension management mobile app found that the perceived health threat had significant positive effects on perceived usefulness [65]. Thus, we propose the following hypothesis:

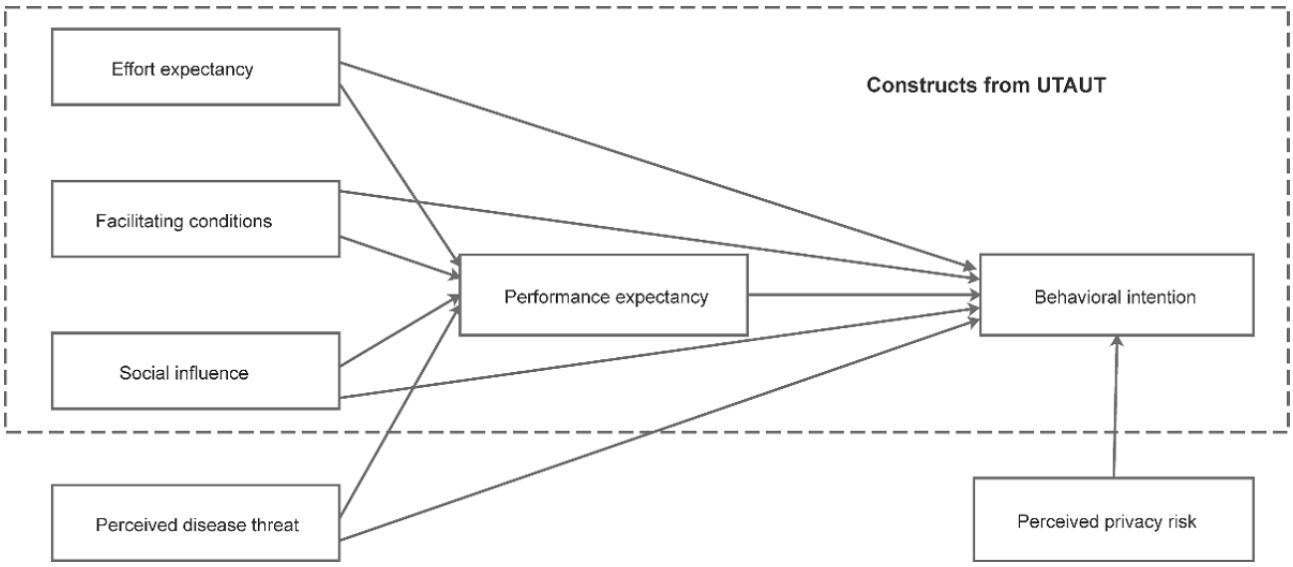
H9: Perceived disease threat positively influences performance expectancy.

Although mHealth services may improve the quality of health care and users' quality of life, they also generate security and privacy issues [66]. The possible risks of mHealth apps include information leakage and theft. Consumers may want to use mHealth services but may not want to disclose their personal information. We define perceived privacy risk as patients' feeling of a lack of control over their personal information after they have adopted mobile apps, and it is not consistent with a real privacy risk. Studies have shown that privacy risks negatively influence patients' intention to use mHealth services [23,67]. Thus, we propose the following hypothesis:

H10: Perceived privacy risks negatively influence the behavioral intention of patients to use diabetes management apps.

The 10 research hypotheses are summarized in the research model (Figure 1).

Figure 1. Research model. UTAUT: Unified Theory of Acceptance and Use of Technology.



Methods

Survey Instrument Design

All survey items were adopted from previous studies related to health information technology. The questionnaire items (Table 2) were translated from English to Chinese by an expert proficient in both English and Chinese, and they were discussed among an expert group selected based on their expertise and our previous explorative studies [20,68]. Items were slightly changed to adapt them to the diabetes management apps. Some items were removed or replaced to ensure face validity, content validity, and construct validity [32]. Back translation was performed from Chinese to English by another qualified translator. The items were measured with a 7-point Likert scale ranging from "strongly disagree" (1) to "strongly agree" (7).

XSL•FO
RenderX

Table 2. Measurement items of the constructs.

Construct	Item
PE[a] [29,55]	
PE1	Diabetes management apps help me to monitor my blood sugar.
PE2	Diabetes management apps educate me in how to deal with my diabetes.
PE3	Overall, diabetes management apps are useful in managing my blood sugar.
EE[b] [24,34,36]	
EE1	My interaction with diabetes management apps is clear and understandable.
EE2	Learning how to use diabetes management apps is easy for me.
EE3	I find diabetes management apps easy to use.
SI[c] [24,29,36,55]	
SI1	People whose opinions that I value (eg, my doctors) think I should use diabetes management apps.
SI2	People who influence my behavior (eg, peers with diabetes) think I should use diabetes management apps.
SI3	People who are important to me (eg, family members) think I should use diabetes management apps.
FC[d] [24,29,34,36]	
FC1	I have the resources (eg, network) necessary to use diabetes management apps.
FC2	I have the knowledge necessary to use diabetes management apps.
FC3	I can get help from others when I have difficulties using diabetes management apps (dropped).
PDT[e] [65]	
PDT1	I am aware that my blood sugar control is not optimal.
PDT2	I am very concerned about my blood sugar.
PDT3	I am very concerned about diabetes-associated complications.
PPR[f] [23]	
PPR1	I think my personal privacy information will be used for other purposes if I use diabetes management apps.
PPR2	Because of security issues, I face the risk of personal information leakage if I use diabetes management apps.
PPR3	I think that when I use diabetes management apps, my personal information will be abused by cyber criminals.
BI[g] [24,29,36]	
BI1	I intend to use or continue to use diabetes management apps.
BI2	I plan to use diabetes management apps frequently.
BI3	Overall, I have a high intention to use diabetes management apps.

[a]PE: performance expectancy.

[b]EE: effort expectancy.

[c]SI: social influence.

[d]FC: facilitating condition.

[e]PDT: perceived disease threat.

[f]PPR: perceived privacy risk.

[g]BI: behavioral intention.

We performed a pilot survey to validate the questionnaire in 10 patients with diabetes who were familiar with diabetes management apps. Context-specific adjustments were made according to the feedback from the pilot survey. On the pilot survey, patients responded that mobile apps were offered to them for free; thus, they had no opinion about the price value. Accordingly, we dropped the perceived value construct. Data on demographic characteristics such as age, sex, type of diabetes, and education level were also collected.

Data Collection

The target population was adult patients with diabetes who were familiar with diabetes management apps. Patients under the age of 18 years and those who were unfamiliar with diabetes management apps were excluded from the survey. Data were collected using the Web-based survey tool Sojump (Changsha ran Xing InfoTech Ltd, China). From April 20 to May 20, 2019, we sent the survey link to diabetologists at hospitals

XSL•FO
RenderX

collaborating in a latent autoimmune diabetes of adults study in 25 major cities in China [69]. The diabetologists shared the survey link through their WeChat contacts network. In addition to facilitating this snowball sampling, we published a survey link on three public diabetes-related WeChat accounts that had nearly 60,000 subscribed followers, and we asked patients with diabetes to complete the questionnaires. Before the survey, we introduced the purpose of the survey and explained the definition of diabetes management apps. After consent was obtained, the survey continued. The questionnaires were completed by the patients themselves. Each WeChat account and mobile IP address could complete the questionnaire only once. A set of electronic diabetes education materials was offered to participants as compensation after completing the questionnaire. The study was approved by the ethics committee of the Second Xiangya Hospital, Central South University.

Data Analysis

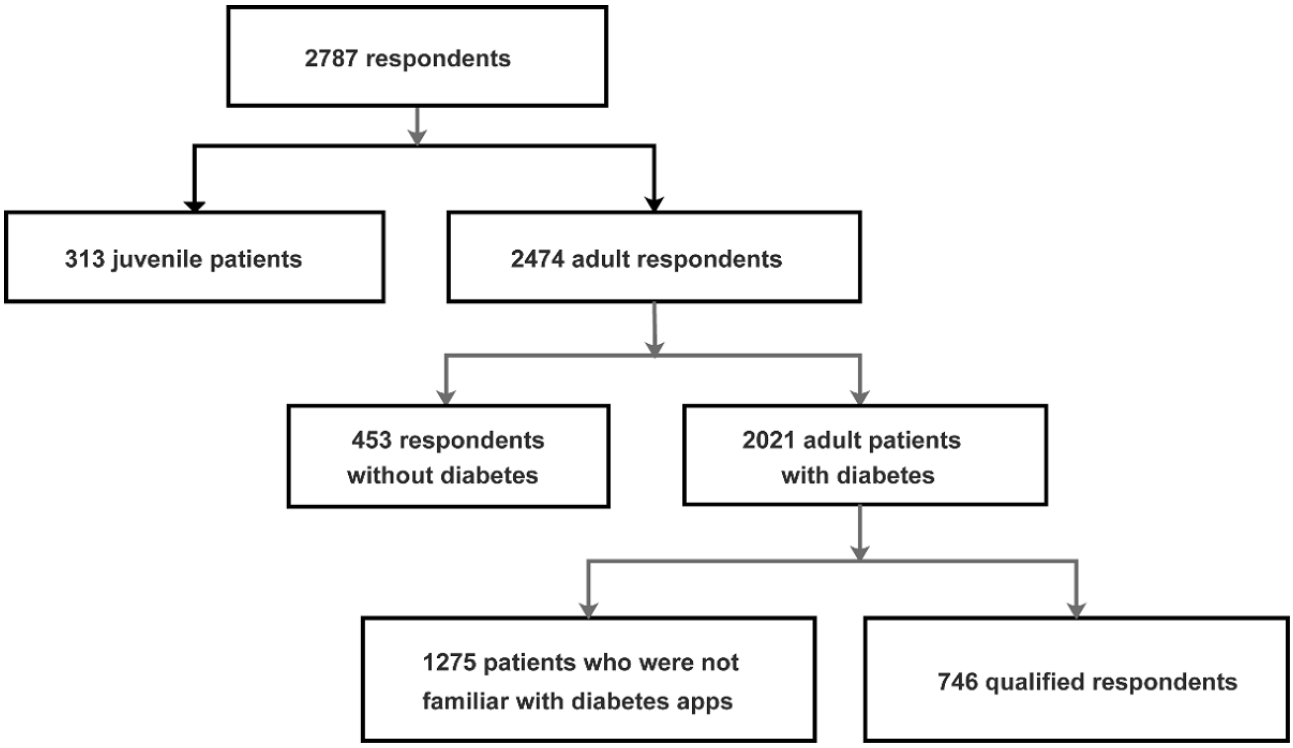
The demographic characteristics of patients were analyzed by descriptive statistics. Patients' acceptance (behavioral intention) of diabetes management apps was measured using three items (Table 2), with higher scores indicating elevated acceptance. An independent *t* test was used to evaluate the differences among acceptance scores between patients with type 1 diabetes and those with type 2 diabetes. Before evaluating the structural model, we assessed the measurement model to evaluate construct reliability, convergent validity, discriminant validity, and data fit indexes. Reliability was measured using the composite reliability and Cronbach alpha. The composite reliability and Cronbach alpha of all constructs should be greater than 0.70 [23,24]. We measured the convergent validity based on the average variance extracted (AVE) of the constructs, and the threshold was higher than 0.50 [24,65]. Discriminant validity is acceptable if the correlation coefficients between any two constructs are smaller than the square root of the corresponding AVE [62]. The model fit was generally considered acceptable when the root mean square error of approximation values was below 0.05; the ratio of χ^2 and df was below 3; and the goodness of fit index, the adjusted goodness of fit index, the comparative fit index, the normed fit index, and the incremental fit index were above 0.90 [23,70]. The data were analyzed using SPSS [computer program] (Version 23.0. Armonk, NY: IBM Corp; 2015), and structural equation modeling analysis was performed using AMOS [computer program] (Version 23.0. Armonk, NY: IBM Corp; 2015) via a maximum likelihood estimation [32,62,71]. We performed a bootstrap analysis with 5000 bootstrap bias-corrected samples to calculate the total, direct, and indirect effects of the variables [70,72]. Values of $P<.05$ (two-tailed) were considered to indicate statistical significance.

Results

Sample Characteristics

Figure 2 shows the sampling procedure and results. A total of 746 participants who met the inclusion criteria completed the survey. The qualified respondent characteristics are shown in Table 3. On an average, the patients' acceptance (behavioral intention) of diabetes management apps (min 1, max 7) was high, with a mean score of 5.65 (SD 0.99), and differences were not observed between patients with type 1 diabetes and those with type 2 diabetes (mean 5.59, SD 1.02 vs mean 5.67, SD 0.97; $P=.097$).

Figure 2. Sampling procedure and results.



XSL•FO
RenderX

Table 3. Demographic characteristics of the qualified respondents (N=746).

Characteristics	Value, n (%)
Gender	
Male	373 (50.0)
Female	373 (50.0)
Age (years)	
18-39	298 (39.9)
40-59	349 (46.8)
≥60	99 (13.3)
Education level	
Primary school or lower	16 (2.1)
Middle school	91 (12.2)
High school	219 (29.4)
University or higher	420 (56.3)
Residence	
Rural	170 (22.8)
Urban	576 (77.2)
Diabetes type	
Type 1	230 (30.8)
Type 2	455 (61.0)
Others	33 (4.4)
Not clearly classified	28 (3.8)
Disease duration (years)	
<1	156 (20.9)
1-4	228 (30.6)
5-10	153 (20.5)
>10	209 (28)

Measurement Model Testing

One indicator (FC3) with a factor loading below 0.50 was removed [62,73]. The results of the measurement model are shown in Table 4. The composite reliability, Cronbach alpha, and AVE of each construct are greater than the recommended values, indicating good reliability and convergent validity. As shown in Table 5, the correlation coefficients between any two constructs are smaller than the square root of the corresponding AVE, indicating acceptable discriminant validity. Table 6 shows the fit indexes of the research model, which indicate that the data collected fit well with the research model.

XSL•FO
RenderX

Table 4. Results of the measurement model.

Constructs and items	Factor loadings	Mean score (SD)	AVE[a]	CR[b]	Cronbach alpha
PE[c]			0.579	0.804	0.794
PE1	0.838	5.83 (1.05)			
PE2	0.74	5.81 (0.87)			
PE3	0.697	5.83 (0.93)			
EE[d]			0.768	0.908	0.892
EE1	0.835	5.79 (0.99)			
EE2	0.898	5.72 (0.99)			
EE3	0.894	5.59 (1.01)			
SI[e]			0.632	0.836	0.866
SI1	0.895	5.21 (1.13)			
SI2	0.797	5.3 (1.13)			
SI3	0.678	5.49 (1.10)			
FC[f]			0.668	0.799	0.79
FC1	0.892	5.99 (0.87)			
FC2	0.735	5.94 (0.88)			
PDT[g]			0.557	0.779	0.743
PDT1	0.531	4.23 (1.68)			
PDT2	0.986	5.12 (1.49)			
PDT3	0.646	5.45 (1.43)			
PPR[h]			0.804	0.925	0.924
PPR1	0.865	4.53 (1.38)			
PPR2	0.948	4.54 (1.41)			
PPR3	0.874	3.57 (1.35)			
BI[i]			0.846	0.943	0.943
BI1	0.904	5.63 (1.03)			
BI2	0.951	5.61 (1.07)			
BI3	0.904	5.72 (1.05)			

[a]AVE: average variance extracted.

[b]CR: composite reliability.

[c]PE: performance expectancy.

[d]EE: effort expectancy.

[e]SI: social influence.

[f]FC: facilitating conditions.

[g]PDT: perceived disease threat.

[h]PPR: perceived privacy risk.

[i]BI: behavioral intention.

XSL•FO
RenderX

Table 5. Square root of average variance extracted of latent variables and correlation coefficient matrix. Italicized values represent square root of the average variance extracted; the values below them indicate the correlation coefficients.

Variable	EE[a]	SI[b]	FC[c]	PDT[d]	PPR[e]	PE[f]	BI[g]
EE	*0.876*						
SI	0.492	*0.795*					
FC	0.581	0.311	*0.817*				
PDT	–0.018	0.01	0.075	*0.746*			
PPR	–0.157	–0.238	–0.065	0.111	*0.897*		
PE	0.567	0.578	0.43	0.043	–0.211	*0.761*	
BI	0.504	0.527	0.451	0.086	–0.221	0.646	*0.92*

[a]EE: effort expectancy.

[b]SI: social influence.

[c]FC: facilitating conditions.

[d]PDT: perceived disease threat.

[e]PPR: perceived privacy risk.

[f]PE: performance expectancy.

[g]BI: behavioral intention.

Table 6. Fit indexes of the research model.

Fit	χ^2/df	GFI[a]	AGFI[b]	NFI[c]	CFI[d]	RMSEA[e]	IFI[f]
Research model	2.63	0.949	0.929	0.96	0.975	0.047	0.975
Recommended value	<3	>0.9	>0.9	>0.9	>0.9	<0.05	>0.9

[a]GFI: goodness of fit index.

[b]AGFI: adjusted goodness of fit index.

[c]NFI: normed fit index.

[d]CFI: comparative fit index.

[e]RMSEA: root mean square error of approximation.

[f]IFI: incremental fit index.

Structural Model Testing

Figure 3 shows that 2 (H2 and H9) of the 10 research hypotheses were rejected. The nonstandardized regression weights for all other links were significant at $P<.05$. Table 7 shows the total, direct, and indirect effects (standardized regression weights) between the model variables.

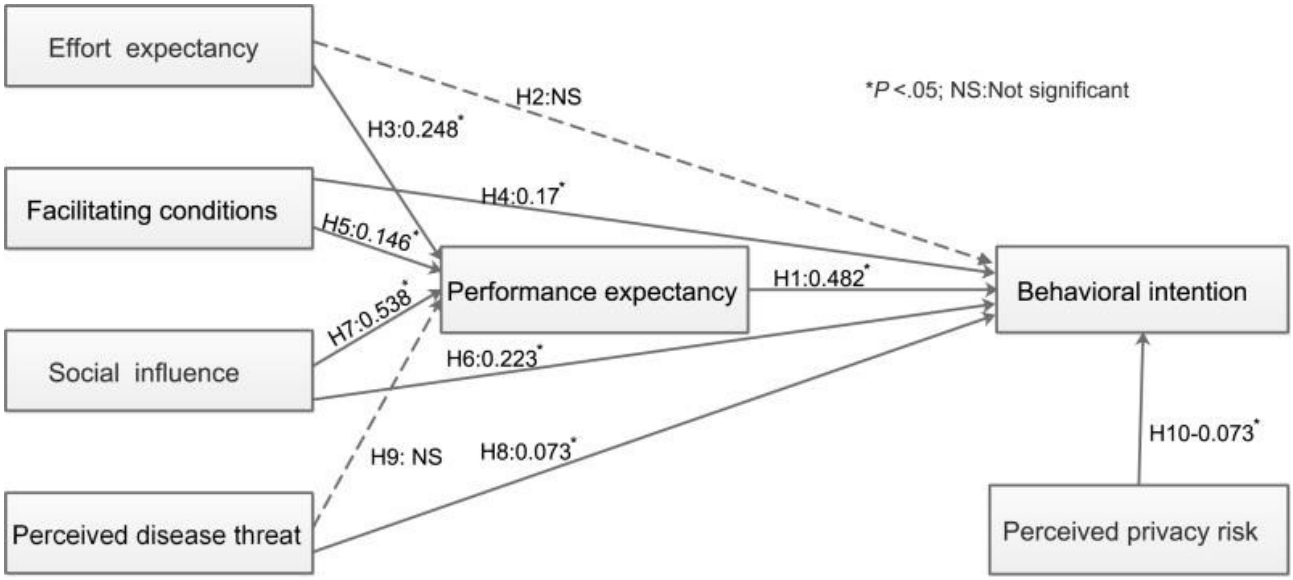
Social influence, effort expectancy, and facilitating conditions explained 62.6% of the variance in performance expectancy. The effect of social influence on performance expectancy was strongest among the three variables (β=0.538, $P=.001$). Effort expectancy and facilitating conditions had moderate effects on performance expectancy (β=0.248, $P=.003$ and β=0.146, $P=.016$, respectively).

Performance expectancy had the strongest direct effect on behavioral intention (β=0.482, $P=.001$). Social influence and facilitating conditions had moderate direct effects on behavioral intention (β=0.223, $P=.003$ and β=0.17, $P=.006$, respectively). Perceived disease threat had a slight positive direct effect on behavioral intention (β=0.073, $P=.005$). Perceived privacy risk had a slight negative direct effect on behavioral intention (β=–0.073, $P=.012$). Additionally, social influence, effort expectancy, and facilitating conditions had indirect effects on behavioral intention, and these effects were mediated by performance expectancy. Social influence had the highest indirect effects among the three constructs (β=0.259, $P=.001$).

Overall, performance expectancy, social influence, disease threat, perceived privacy risk, and facilitating conditions explained 57.1% of the variance in behavioral intention. Performance expectancy and social influence had the strongest total effects on behavioral intention (β=0.482, $P=.001$). Facilitating conditions had a moderate total effect on behavioral intention (β=0.240, $P=.001$). Perceived disease threat had a slight total effect on behavioral intention (β=0.082, $P=.002$). Perceived privacy risk had a slight negative total effect on behavioral intention (β=–0.073, $P=.012$).

XSL•FO
RenderX

Figure 3. Research model explaining performance expectancy and behavioral intention (direct effects). H1: Performance expectancy positively influences the behavioral intention of patients to use diabetes management apps, H2: Effort expectancy positively influences the behavioral intention of patients to use diabetes management apps, H3: Effort expectancy positively influences performance expectancy, H4: Facilitating conditions positively influence the behavioral intention of patients to use diabetes management apps, H5: Facilitating conditions positively influence performance expectancy, H6: Social influence positively influences the behavioral intention of patients to use diabetes management apps, H7: Social influence positively influences performance expectancy, H8: Perceived disease threat positively influences the behavioral intention of patients to use diabetes management apps, H9: Perceived disease threat positively influences performance expectancy.



XSL•FO
RenderX

Table 7. Standardized regression weights between the model variables.

Variable	PE[a] (R^2=62.6%)		BI[b] (R^2=57.1%)	
	β	*P* value	β	*P* value
EE[c]				
Direct	0.248	.003	–0.019	.85[d]
Indirect	—[e]	—	0.119	.002
Total	0.248	.003	0.1	.12[d]
SI[f]				
Direct	0.538	.001	0.223	.003
Indirect	—	—	0.259	.001
Total	0.538	.001	0.482	.001
FC[g]				
Direct	0.146	.02	0.17	.006
Indirect	—	—	0.07	.01
Total	0.146	.02	0.24	.001
PDT[h]				
Direct	–0.032	.49[d]	0.073	.005
Indirect	—	—	0.009	.46[d]
Total	–0.032	.49[d]	0.082	.002
PPR[i]				
Direct	—	—	–0.073	.01
Indirect	—	—	—	—
Total	—	—	–0.073	.01
PE[j]				
Direct	—	—	0.482	.001
Indirect	—	—	—	—
Total	—	—	0.482	.001

[a]PE: performance expectancy.
[b]BI: behavioral intention.
[c]EE: effort expectancy.
[d]Not significant.
[e]Not available.
[f]SI: social influence.
[g]FC: facilitating conditions.
[h]PDT: perceived disease threat.
[i]PPR: perceived privacy risk.
[j]PE: performance expectancy.

Discussion

Principal Findings

Our study found that performance expectancy and social influence were the most important determinants of patients' intention to use diabetes management apps. Several studies on mHealth services also revealed that performance expectancy was the major determinant of behavioral intention [24,65,74]. If patients with diabetes believe they can benefit from diabetes management apps, their willingness to use them will be stronger. Our previous study found that some patients were reluctant to continue using diabetes management apps because they thought the apps were useless, and the experts surveyed believed that one reason for the inefficacy of apps was their lack of comprehensiveness or functionality [20]. Physical activity,

XSL•FO
RenderX

nutrition, blood glucose testing, medication, health feedback, and education are all important for diabetes management; however, few apps integrate all six diabetes management tasks [75]. Information quality is a determinant of people's intention to seek and use health information from internet sources [76]. However, few apps provide information cited from accredited sources [77]. Blood sugar monitoring is the most frequently used feature of diabetes management apps [21]. However, patients think that merely recording their blood sugar is of little use to them [68]. Therefore, the benefits of diabetes management apps to patients are limited to a certain extent, and low performance expectancy affects patients' willingness to use apps.

Although the direct effect of social influence on intention to use diabetes management apps was moderate, it had a significant indirect effect on behavioral intention, and this effect was mediated by performance expectancy. The effect of social influence on behavioral intention is consistent with the findings of another study on multiple sclerosis management mobile apps [74]. A survey by Hennemann et al [33] also found a prominent effect of social influence on the acceptance of Web-based aftercare. Patients' uptake of health-related actions is susceptible to the influence of doctors, family members, and peers with the same disease. Our previous survey found that nearly half of the patients used apps because they were recommended to use them by other patients or doctors [20]. Because of the lack of clinical evidence on apps' effectiveness, doctors do not know which apps are suitable to recommend to their patients [20]. Therefore, high-quality randomized controlled trials are needed to provide evidence-based information for doctors to recommend diabetes management apps, which will benefit the promotion of apps.

Facilitating conditions had a moderate direct effect on behavioral intention to use apps for diabetes management, and it also had a slight indirect impact on behavioral intention; this impact was mediated by performance expectancy. This result was consistent with the study of Rho et al [51] on the acceptance behavior of telemedicine for diabetes management. Despite the rapid development of smartphones in China, the use of smartphones and networks is still limited in some remote rural areas. China is vigorously advocating internet health care [78], which requires improvements to basic network facilities. App manufacturers should also provide continuous assistance services and use guidelines to support patients' use of diabetes management apps.

Perceived disease threat had slight positive effects on patients' intention to use diabetes management apps. The study by [65] Dou et al revealed that perceived health threat predicted patients' intentions to use a hypertension management mobile app [65]. Several other studies concerning health information technology demonstrated the effect of disease threat on behavioral intention [29,64,71]. The awareness rate of diabetes mellitus in Chinese diabetic patients is low [2]. We should improve diabetes awareness among diabetic patients and help them correctly understand the disease. Improving patient awareness of the disease will promote patients' intention to use diabetes management apps to manage their disease.

The negative impact of privacy risk on health information technology acceptance intention is inconsistent across studies. Our study found that perceived privacy risk had a slight negative effect on patients' intention to use diabetes management apps. This finding is consistent with the study on mHealth services acceptance behavior [67]. A survey in America found a moderate negative effect of privacy risk on patients' intention to use home health care robots [34]. However, a study in Bangladesh found that privacy had no effect on the adoption intention of eHealth [57]. This finding might be attributed to the different awareness levels of privacy protection in different regions. With the development of health information technology, patients are increasingly aware of privacy protection. Although our research found that perceived privacy risk has only a weak effect on the intention to use diabetes management apps, solid privacy protection measures are necessary [11].

Our study found that effort expectancy did not affect the intention to use apps for diabetes management. One study on hypertensive patients' intention to use a hypertension management mobile app in China also did not find such an association [65]. Some studies regarding health information technology found no association between effort expectancy and behavioral intention [36,58]. However, several other studies found that effort expectancy had positive effects on behavioral intention [23,29,38], especially among the elderly [54,55]. This difference might be related to the ease of use of different technologies. However, our sample was relatively young and well educated, and some patients had been using apps for a long time, which may have biased the results.

Limitations

First, our study used a Web-based survey. Moreover, some selection bias was unavoidable. Our surveyed patients were relatively young and highly educated; thus, a higher awareness of diabetes management apps was observed among these patients. Previous studies have demonstrated that the use of mHealth apps among younger patients and those with higher education is relatively high [20,79,80]. This bias might have influenced our results to a certain degree. For example, effort expectancy might be a determinant of the use intention among the elderly. Therefore, further offline population-based surveys are necessary, especially among the elderly. Second, our survey did not study the effect of behavioral intention on actual use. Although intention to use is a determinant of use behavior, there is usually a gap between actual use and behavioral intention [81]. However, when people have the intention to use diabetes management apps, they do not necessarily start using the apps right away. Rather, the use behavior may lag behind the intention to use it. Therefore, cross-sectional surveys may not be able to observe the impact of behavioral intention on use behavior, and further longitudinal surveys are needed to observe this impact and other factors that may affect use behavior, such as facilitating conditions. Third, our model explained only 57.1% of the variance in behavioral intention, which indicates that some other factors affecting behavioral intention may have been overlooked. Future studies could include other constructs such as compatibility of the diffusion of innovation theory [82]. Fourth, our model is for diabetes management apps, and it should be applied to other chronic disease management apps

XSL•FO
RenderX

with caution. Finally, although diabetes management apps on the market are all offered for free to patients in China at present, some apps offer in-app purchases, such as diabetes education materials and telemedicine services. Therefore, the impact of perceived value on use intention needs to be further investigated.

Conclusions

Performance expectancy and social influence are the most important determinants of patients' intention to use diabetes management apps. Therefore, manufacturers must improve the usefulness of diabetes management apps and carry out research to provide clinical evidence for the effectiveness of these apps, which will benefit the promotion of apps. Facilitating conditions and perceived privacy risk also have an impact on behavioral intention. Therefore, it is necessary to improve facilitating conditions and provide solid privacy protection. Our study supports the use of the UTAUT in explaining patients' intention to use diabetes management apps. In addition, context-related determinants should be considered to understand patients' behavior intentions.

Acknowledgments

We thank all respondents for participating in the survey; the doctors from LADA Study collaborative hospitals in China for spreading the survey link; and Yang Lijun, Zhu Junping and Li Wenjie for publishing the survey link on the three Wechat public accounts. We also thank American Journal Experts for their language editing. This work was supported by the National Key R&D Program of China (2018YFC1315603, 2017YFC1309604, 2016YFC1305000, 2016YFC1305001).

Conflicts of Interest

None declared.

References

1. Cho NH, Shaw JE, Karuranga S, Huang Y, da Rocha Fernandes JD, Ohlrogge AW, et al. IDF Diabetes Atlas: Global estimates of diabetes prevalence for 2017 and projections for 2045. Diabetes Res Clin Pract 2018 Apr;138:271-281. [doi: 10.1016/j.diabres.2018.02.023] [Medline: 29496507]
2. Wang L, Gao P, Zhang M, Huang Z, Zhang D, Deng Q, et al. Prevalence and Ethnic Pattern of Diabetes and Prediabetes in China in 2013. JAMA 2017 Dec 27;317(24):2515-2523. [doi: 10.1001/jama.2017.7596] [Medline: 28655017]
3. Holman RR, Paul SK, Bethel MA, Matthews DR, Neil HAW. 10-year follow-up of intensive glucose control in type 2 diabetes. N Engl J Med 2008 Oct 9;359(15):1577-1589. [doi: 10.1056/NEJMoa0806470] [Medline: 18784090]
4. Nathan DM. Long-term complications of diabetes mellitus. N Engl J Med 1993 Jun 10;328(23):1676-1685. [doi: 10.1056/NEJM199306103282306] [Medline: 8487827]
5. Bommer C, Heesemann E, Sagalova V, Manne-Goehler J, Atun R, Bärnighausen T, et al. The global economic burden of diabetes in adults aged 20-79 years: a cost-of-illness study. Lancet Diabetes Endocrinol 2017 Dec;5(6):423-430. [doi: 10.1016/S2213-8587(17)30097-9] [Medline: 28456416]
6. Beck J, Greenwood DA, Blanton L, Bollinger ST, Butcher MK, Condon JE, et al. 2017 National Standards for Diabetes Self-Management Education and Support. Diabetes Educ 2019 Dec;45(1):34-49. [doi: 10.1177/0145721718820941] [Medline: 30558523]
7. Li Q, Xie P. Outpatient workload in China. Lancet 2013 Jun 08;381(9882):1983-1984. [doi: 10.1016/S0140-6736(13)61198-8] [Medline: 23746895]
8. Zhou Y, Liao L, Sun M, He G. Self-care practices of Chinese individuals with diabetes. Exp Ther Med 2013 Apr;5(4):1137-1142 [FREE Full text] [doi: 10.3892/etm.2013.945] [Medline: 23599736]
9. Chen Y, Yin Z, Xie Q. Suggestions to ameliorate the inequity in urban/rural allocation of healthcare resources in China. Int J Equity Health 2014;13:34 [FREE Full text] [doi: 10.1186/1475-9276-13-34] [Medline: 24884614]
10. Bragg F, Holmes MV, Iona A, Guo Y, Du H, Chen Y, China Kadoorie Biobank Collaborative Group. Association Between Diabetes and Cause-Specific Mortality in Rural and Urban Areas of China. JAMA 2017 Dec 17;317(3):280-289 [FREE Full text] [doi: 10.1001/jama.2016.19720] [Medline: 28114552]
11. Veazie S, Winchell K, Gilbert J, Paynter R, Ivlev I, Eden KB, et al. Rapid Evidence Review of Mobile Applications for Self-management of Diabetes. J Gen Intern Med 2018 Jul;33(7):1167-1176. [doi: 10.1007/s11606-018-4410-1] [Medline: 29740786]
12. Istepanian RSH. Mobile applications for diabetes management: efficacy issues and regulatory challenges. Lancet Diabetes Endocrinol 2015 Dec;3(12):921-923. [doi: 10.1016/S2213-8587(15)00414-3] [Medline: 26545304]
13. Hou C, Xu Q, Diao S, Hewitt J, Li J, Carter B. Mobile phone applications and self-management of diabetes: A systematic review with meta-analysis, meta-regression of 21 randomized trials and GRADE. Diabetes Obes Metab 2018 Dec;20(8):2009-2013. [doi: 10.1111/dom.13307] [Medline: 29582538]
14. Bonoto BC, de Araújo VE, Godói IP, de Lemos LLP, Godman B, Bennie M, et al. Efficacy of Mobile Apps to Support the Care of Patients With Diabetes Mellitus: A Systematic Review and Meta-Analysis of Randomized Controlled Trials. JMIR Mhealth Uhealth 2017 01;5(3):e4 [FREE Full text] [doi: 10.2196/mhealth.6309] [Medline: 28249834]

XSL•FO
RenderX

15. Sun C, Sun L, Xi S, Zhang H, Wang H, Feng Y, et al. Mobile Phone-Based Telemedicine Practice in Older Chinese Patients with Type 2 Diabetes Mellitus: Randomized Controlled Trial. JMIR Mhealth Uhealth 2019 Jan 04;7(1):e10664 [FREE Full text] [doi: 10.2196/10664] [Medline: 30609983]
16. Husted GR, Weis J, Teilmann G, Castensøe-Seidenfaden P. Exploring the Influence of a Smartphone App (Young with Diabetes) on Young People's Self-Management: Qualitative Study. JMIR Mhealth Uhealth 2018 Feb 28;6(2):e43 [FREE Full text] [doi: 10.2196/mhealth.8876] [Medline: 29490897]
17. Garg SK, Shah VN, Akturk HK, Beatson C, Snell-Bergeon JK. Role of Mobile Technology to Improve Diabetes Care in Adults with Type 1 Diabetes: The Remote-T1D Study iBGStar in Type 1 Diabetes Management. Diabetes Ther 2017 Aug;8(4):811-819 [FREE Full text] [doi: 10.1007/s13300-017-0272-5] [Medline: 28555339]
18. Williams JP, Schroeder D. Popular Glucose Tracking Apps and Use of mHealth by Latinos With Diabetes: Review. JMIR Mhealth Uhealth 2015;3(3):e84 [FREE Full text] [doi: 10.2196/mhealth.3986] [Medline: 26307533]
19. Trawley S, Baptista S, Browne JL, Pouwer F, Speight J. The Use of Mobile Applications Among Adults with Type 1 and Type 2 Diabetes: Results from the Second MILES-Australia (MILES-2) Study. Diabetes Technol Ther 2017 Dec;19(12):730-738. [doi: 10.1089/dia.2017.0235] [Medline: 29028442]
20. Zhang Y, Li X, Luo S, Liu C, Xie Y, Guo J, et al. Use, Perspectives, and Attitudes Regarding Diabetes Management Mobile Apps Among Diabetes Patients and Diabetologists in China: National Web-Based Survey. JMIR Mhealth Uhealth 2019 Feb 08;7(2):e12658 [FREE Full text] [doi: 10.2196/12658] [Medline: 30735147]
21. Boyle L, Grainger R, Hall RM, Krebs JD. Use of and Beliefs About Mobile Phone Apps for Diabetes Self-Management: Surveys of People in a Hospital Diabetes Clinic and Diabetes Health Professionals in New Zealand. JMIR Mhealth Uhealth 2017 Jun 30;5(6):e85 [FREE Full text] [doi: 10.2196/mhealth.7263] [Medline: 28666975]
22. Venkatesh, Morris, Davis, Davis. User Acceptance of Information Technology: Toward a Unified View. MIS Quarterly 2003;27(3):425. [doi: 10.2307/30036540]
23. Deng Z, Hong Z, Ren C, Zhang W, Xiang F. What Predicts Patients' Adoption Intention Toward mHealth Services in China: Empirical Study. JMIR Mhealth Uhealth 2018 Aug 29;6(8):e172 [FREE Full text] [doi: 10.2196/mhealth.9316] [Medline: 30158101]
24. Hoque R, Sorwar G. Understanding factors influencing the adoption of mHealth by the elderly: An extension of the UTAUT model. Int J Med Inform 2017 May;101:75-84. [doi: 10.1016/j.ijmedinf.2017.02.002] [Medline: 28347450]
25. Guo X, Sun Y, Wang N, Peng Z, Yan Z. The dark side of elderly acceptance of preventive mobile health services in China. Electron Markets 2012 Dec 11;23(1):49-61. [doi: 10.1007/s12525-012-0112-4]
26. Lin SP. Determinants of adoption of Mobile Healthcare Service. IJMC 2011;9(3):298. [doi: 10.1504/IJMC.2011.040608]
27. Meng F, Guo X, Peng Z, Lai K, Zhao X. Investigating the Adoption of Mobile Health Services by Elderly Users: Trust Transfer Model and Survey Study. JMIR Mhealth Uhealth 2019 Jan 08;7(1):e12269. [doi: 10.2196/12269] [Medline: 30622092]
28. Kijsanayotin B, Pannarunothai S, Speedie SM. Factors influencing health information technology adoption in Thailand's community health centers: applying the UTAUT model. Int J Med Inform 2009 Jun;78(6):404-416. [doi: 10.1016/j.ijmedinf.2008.12.005] [Medline: 19196548]
29. Koivumäki T, Pekkarinen S, Lappi M, Väisänen J, Juntunen J, Pikkarainen M. Consumer Adoption of Future MyData-Based Preventive eHealth Services: An Acceptance Model and Survey Study. J Med Internet Res 2017 Dec 22;19(12):e429 [FREE Full text] [doi: 10.2196/jmir.7821] [Medline: 29273574]
30. van Houwelingen CT, Ettema RG, Antonietti MG, Kort HS. Understanding Older People's Readiness for Receiving Telehealth: Mixed-Method Study. J Med Internet Res 2018 Apr 06;20(4):e123 [FREE Full text] [doi: 10.2196/jmir.8407] [Medline: 29625950]
31. Kim S, Lee K, Hwang H, Yoo S. Analysis of the factors influencing healthcare professionals' adoption of mobile electronic medical record (EMR) using the unified theory of acceptance and use of technology (UTAUT) in a tertiary hospital. BMC Med Inform Decis Mak 2016 Jan 30;16:12 [FREE Full text] [doi: 10.1186/s12911-016-0249-8] [Medline: 26831123]
32. Maillet , Mathieu L, Sicotte C. Modeling factors explaining the acceptance, actual use and satisfaction of nurses using an Electronic Patient Record in acute care settings: an extension of the UTAUT. Int J Med Inform 2015 Jan;84(1):36-47. [doi: 10.1016/j.ijmedinf.2014.09.004] [Medline: 25288192]
33. 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]
34. Alaiad A, Zhou L. The determinants of home healthcare robots adoption: an empirical investigation. Int J Med Inform 2014 Nov;83(11):825-840. [doi: 10.1016/j.ijmedinf.2014.07.003] [Medline: 25132284]
35. Yi MY, Jackson JD, Park JS, Probst JC. Understanding information technology acceptance by individual professionals: Toward an integrative view. Information & Management 2006 Apr;43(3):350-363. [doi: 10.1016/j.im.2005.08.006]
36. Tavares J, Oliveira T. New Integrated Model Approach to Understand the Factors That Drive Electronic Health Record Portal Adoption: Cross-Sectional National Survey. J Med Internet Res 2018 Nov 19;20(11):e11032 [FREE Full text] [doi: 10.2196/11032] [Medline: 30455169]

XSL•FO
RenderX

37. Venkatesh, Thong, Xu. Consumer Acceptance and Use of Information Technology: Extending the Unified Theory of Acceptance and Use of Technology. MIS Quarterly 2012 Mar;36(1):157-117. [doi: 10.2307/41410412] [Medline: 22217802]
38. Tavares J, Oliveira T. Electronic Health Record Patient Portal Adoption by Health Care Consumers: An Acceptance Model and Survey. J Med Internet Res 2016 Mar 02;18(3):e49 [FREE Full text] [doi: 10.2196/jmir.5069] [Medline: 26935646]
39. Nie L, Xie B, Yang Y, Shan YM. Characteristics of Chinese m-Health Applications for Diabetes Self-Management. Telemed J E Health 2016 Jul;22(7):614-619. [doi: 10.1089/tmj.2015.0184] [Medline: 27171016]
40. Rahimi B, Nadri H, Lotfnezhad Afshar H, Timpka T. A Systematic Review of the Technology Acceptance Model in Health Informatics. Appl Clin Inform 2018 Dec;9(3):604-634 [FREE Full text] [doi: 10.1055/s-0038-1668091] [Medline: 30112741]
41. Goodarzi F, Araban M, Eslami AA, Zamani-Alavijeh F. Development and psychometric evaluation of the diabetic Men's dietary behaviors inventory based on the theory of reasoned action. Arch Public Health 2019;77:3 [FREE Full text] [doi: 10.1186/s13690-018-0328-7] [Medline: 30675349]
42. Ajzen I. The theory of planned behavior. Organizational Behavior and Human Decision Processes 1991 Dec;50(2):179-211. [doi: 10.1016/0749-5978(91)90020-T]
43. Armitage CJ, Conner M. Efficacy of the Theory of Planned Behaviour: a meta-analytic review. Br J Soc Psychol 2001 Dec;40(Pt 4):471-499. [Medline: 11795063]
44. Vallerand R. Toward A Hierarchical Model of Intrinsic and Extrinsic Motivation. Advances in Experimental Social Psychology 1997;29:271-360. [doi: 10.1016/S0065-2601(08)60019-2]
45. Davis FD, Bagozzi RP, Warshaw PR. Extrinsic and Intrinsic Motivation to Use Computers in the Workplace1. J Appl Social Pyschol 1992 Jul;22(14):1111-1132. [doi: 10.1111/j.1559-1816.1992.tb00945.x]
46. Taylor S, Todd PA. Understanding Information Technology Usage: A Test of Competing Models. Information Systems Research 1995 Jun;6(2):144-176. [doi: 10.1287/isre.6.2.144]
47. Thompson RL, Higgins CA, Howell JM. Personal Computing: Toward a Conceptual Model of Utilization. MIS Quarterly 1991 Mar;15(1):125. [doi: 10.2307/249443]
48. Moore GC, Benbasat I. Development of an Instrument to Measure the Perceptions of Adopting an Information Technology Innovation. Information Systems Research 1991 Sep;2(3):192-222. [doi: 10.1287/isre.2.3.192]
49. Compeau DR, Higgins CA. Computer Self-Efficacy: Development of a Measure and Initial Test. MIS Quarterly 1995 Jun;19(2):189 [FREE Full text] [doi: 10.2307/249688]
50. Tsai C. Integrating social capital theory, social cognitive theory, and the technology acceptance model to explore a behavioral model of telehealth systems. Int J Environ Res Public Health 2014 May;11(5):4905-4925 [FREE Full text] [doi: 10.3390/ijerph110504905] [Medline: 24810577]
51. Rho MJ, Kim HS, Chung K, Choi IY. Factors influencing the acceptance of telemedicine for diabetes management. Cluster Comput 2014 Mar 12;18(1):321-331. [doi: 10.1007/s10586-014-0356-1]
52. Wang L, Rau PP, Salvendy G. Older Adults' Acceptance of Information Technology. Educational Gerontology 2011 Dec;37(12):1081-1099. [doi: 10.1080/03601277.2010.500588]
53. Portz J, Bayliss E, Bull S. Using the Technology Acceptance Model to Explore User Experience, Intent to Use, and Use Behavior of a Patient Portal Among Older Adults With Multiple Chronic Conditions: Descriptive Qualitative Study. J Med Internet Res Apr 2019;21(4):8. [Medline: 30958272]
54. de Veer AJE, Peeters JM, Brabers AEM, Schellevis FG, Rademakers JJDJM, Francke AL. Determinants of the intention to use e-Health by community dwelling older people. BMC Health Serv Res 2015 15;15:103 [FREE Full text] [doi: 10.1186/s12913-015-0765-8] [Medline: 25889884]
55. Cimperman M, Makovec BM, Trkman P. Analyzing older users' home telehealth services acceptance behavior-applying an Extended UTAUT model. Int J Med Inform 2016 Jun;90:22-31. [doi: 10.1016/j.ijmedinf.2016.03.002] [Medline: 27103194]
56. Lee S, Choi M, Rho M, Kim D, Choi I. Factors Affecting User Acceptance in Overuse of Smartphones in Mobile Health Services: An Empirical Study Testing a Modified Integrated Model in South Korea. Front Psychiatry 2018;9:A. [Medline: 30631283]
57. Hoque MR, Bao Y, Sorwar G. Investigating factors influencing the adoption of e-Health in developing countries: A patient's perspective. Inform Health Soc Care 2017 Jan;42(1):1-17. [doi: 10.3109/17538157.2015.1075541] [Medline: 26865037]
58. Jewer J. Patients' intention to use online postings of ED wait times: A modified UTAUT model. Int J Med Inform 2018;112:34-39. [Medline: 29500019]
59. Aggelidis V, Chatzoglou P. Using a modified technology acceptance model in hospitals. Int J Med Inform 2009;78(2):115-126. [Medline: 18675583]
60. Or CKL, Karsh B. A systematic review of patient acceptance of consumer health information technology. J Am Med Inform Assoc 2009 Aug;16(4):550-560 [FREE Full text] [doi: 10.1197/jamia.M2888] [Medline: 19390112]
61. Or CKL, Karsh B, Severtson DJ, Burke LJ, Brown RL, Brennan PF. Factors affecting home care patients' acceptance of a web-based interactive self-management technology. J Am Med Inform Assoc 2011 Feb;18(1):51-59 [FREE Full text] [doi: 10.1136/jamia.2010.007336] [Medline: 21131605]
62. Lee E, Han S, Jo SH. Consumer choice of on-demand mHealth app services: Context and contents values using structural equation modeling. Int J Med Inform 2017 Dec;97:229-238. [doi: 10.1016/j.ijmedinf.2016.10.016] [Medline: 27919381]

XSL•FO
RenderX

63. Janz NK, Becker MH. The Health Belief Model: a decade later. Health Educ Q 1984;11(1):1-47. [Medline: 6392204]
64. Ahadzadeh AS, Pahlevan SS, Ong FS, Khong KW. Integrating health belief model and technology acceptance model: an investigation of health-related internet use. J Med Internet Res 2015;17(2):e45 [FREE Full text] [doi: 10.2196/jmir.3564] [Medline: 25700481]
65. Dou K, Yu P, Deng N, Liu F, Guan Y, Li Z, et al. Patients' Acceptance of Smartphone Health Technology for Chronic Disease Management: A Theoretical Model and Empirical Test. JMIR Mhealth Uhealth 2017 Dec 06;5(12):e177 [FREE Full text] [doi: 10.2196/mhealth.7886] [Medline: 29212629]
66. Kotz D, Gunter CA, Kumar S, Weiner JP. Privacy and Security in Mobile Health: A Research Agenda. Computer 2016 Jun;49(6):22-30. [doi: 10.1109/MC.2016.185]
67. Guo X, Zhang X, Sun Y. The privacy–personalization paradox in mHealth services acceptance of different age groups. Electronic Commerce Research and Applications 2016 Mar;16:55-65. [doi: 10.1016/j.elerap.2015.11.001]
68. Zhang Y, Li X, Luo S, Liu C, Liu F, Zhou Z. Exploration of Users' Perspectives and Needs and Design of a Type 1 Diabetes Management Mobile App: Mixed-Methods Study. JMIR Mhealth Uhealth 2018 Sep 21;6(9):e11400 [FREE Full text] [doi: 10.2196/11400] [Medline: 30249580]
69. Zhou Z, Xiang Y, Ji L, Jia W, Ning G, Huang G, LADA China Study Group. Frequency, immunogenetics, and clinical characteristics of latent autoimmune diabetes in China (LADA China study): a nationwide, multicenter, clinic-based cross-sectional study. Diabetes 2013 Feb;62(2):543-550 [FREE Full text] [doi: 10.2337/db12-0207] [Medline: 23086039]
70. Zhao H, Zhang H, Xu Y, He W, Lu J. Why Are People High in Dispositional Awe Happier? The Roles of Meaning in Life and Materialism. Front Psychol 2019;10:1208 [FREE Full text] [doi: 10.3389/fpsyg.2019.01208] [Medline: 31191402]
71. Kim J, Park H. Development of a health information technology acceptance model using consumers' health behavior intention. J Med Internet Res 2012;14(5):e133 [FREE Full text] [doi: 10.2196/jmir.2143] [Medline: 23026508]
72. Orkibi H, Ronen T. Basic Psychological Needs Satisfaction Mediates the Association between Self-Control Skills and Subjective Well-Being. Front Psychol 2017;8:936 [FREE Full text] [doi: 10.3389/fpsyg.2017.00936] [Medline: 28638362]
73. Dwivedi YK, Shareef MA, Simintiras AC, Lal B, Weerakkody V. A generalised adoption model for services: A cross-country comparison of mobile health (m-health). Government Information Quarterly 2016 Jan;33(1):174-187. [doi: 10.1016/j.giq.2015.06.003]
74. Apolinário-Hagen J, Menzel M, Hennemann S, Salewski C. Acceptance of Mobile Health Apps for Disease Management Among People With Multiple Sclerosis: Web-Based Survey Study. JMIR Form Res 2018 Dec 12;2(2):e11977 [FREE Full text] [doi: 10.2196/11977] [Medline: 30684408]
75. Chavez S, Fedele D, Guo Y, Bernier A, Smith M, Warnick J, et al. Mobile Apps for the Management of Diabetes. Diabetes Care 2017 Oct;40(10):e145-e146. [doi: 10.2337/dc17-0853] [Medline: 28774944]
76. Deng Z, Liu S, Hinz O. The health information seeking and usage behavior intention of Chinese consumers through mobile phones. Info Technology & People 2015 Jun;28(2):405-423. [doi: 10.1108/ITP-03-2014-0053]
77. Huang Z, Soljak M, Boehm BO, Car J. Clinical relevance of smartphone apps for diabetes management: A global overview. Diabetes Metab Res Rev 2018 May;34(4):e2990. [doi: 10.1002/dmrr.2990] [Medline: 29431916]
78. National Health and Wellness Commission, National Administration of Traditional Chinese Medcine. Notice on Printing and Distributing 3 Documents Such as the Internet Management and Management Measures (Trial). 2018 Jul 17. URL: http://bgs.satcm.gov.cn/zhengcewenjian/2018-09-17/7909.html [accessed 2019-07-26] [WebCite Cache ID 73MT7S8am]
79. Lu C, Hu Y, Xie J, Fu Q, Leigh I, Governor S, et al. The Use of Mobile Health Applications to Improve Patient Experience: Cross-Sectional Study in Chinese Public Hospitals. JMIR Mhealth Uhealth 2018 May 23;6(5):e126. [doi: 10.2196/mhealth.9145] [Medline: 29792290]
80. Shen C, Wang MP, Chu JT, Wan A, Viswanath K, Chan SSC, et al. Health App Possession Among Smartphone or Tablet Owners in Hong Kong: Population-Based Survey. JMIR Mhealth Uhealth 2017 Jun 05;5(6):e77 [FREE Full text] [doi: 10.2196/mhealth.7628] [Medline: 28583905]
81. Bhattacherjee A, Sanford C. The intention–behaviour gap in technology usage: the moderating role of attitude strength. Behaviour & Information Technology 2009 Jul;28(4):389-401. [doi: 10.1080/01449290802121230]
82. Zhang X, Yu P, Yan J, Ton AMSI. Using diffusion of innovation theory to understand the factors impacting patient acceptance and use of consumer e-health innovations: a case study in a primary care clinic. BMC Health Serv Res 2015 Feb 21;15:71 [FREE Full text] [doi: 10.1186/s12913-015-0726-2] [Medline: 25885110]

Abbreviations

AVE: average variance extracted
BI: behavioral intention
CR: composite reliability
EE: effort expectancy
FC: facility conditions
mHealth: mobile health
PDT: perceived disease threat

XSL•FO
RenderX

PE: performance expectancy
PPR: perceived privacy risk
SI: social influence
TAM: technology acceptance model
TPB: theory of planned behavior
UTAUT: Unified Theory of Acceptance and Use of Technology

Edited by G Eysenbach; submitted 13.06.19; peer-reviewed by J Tavares, J Apolinário-Hagen, W Zhang, C Reis, JA Sim; comments to author 04.07.19; revised version received 21.07.19; accepted 22.07.19; published 13.08.19.

Please cite as:
Zhang Y, Liu C, Luo S, Xie Y, Liu F, Li X, Zhou Z
Factors Influencing Patients' Intentions to Use Diabetes Management Apps Based on an Extended Unified Theory of Acceptance and Use of Technology Model: Web-Based Survey
J Med Internet Res 2019;21(8):e15023
URL: http://www.jmir.org/2019/8/e15023/
doi:10.2196/15023
PMID:31411146

©Yiyu Zhang, Chaoyuan Liu, Shuoming Luo, Yuting Xie, Fang Liu, Xia Li, Zhiguang Zhou. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 13.08.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

A Comparison of Physical Activity Mobile Apps With and Without Existing Web-Based Social Networking Platforms: Systematic Review

Jasmine Maria Petersen[1], BSc; Ivanka Prichard[1], PhD; Eva Kemps[2], PhD

[1]College of Nursing and Health Sciences, Flinders University, Adelaide, Australia

[2]College of Education, Psychology and Social Work, Flinders University, Adelaide, Australia

Corresponding Author:
Jasmine Maria Petersen, BSc
College of Nursing and Health Sciences
Flinders University
GPO Box 2100
Adelaide,
Australia
Phone: 61 8 82013713
Email: jasmine.petersen@flinders.edu.au

Abstract

Background: Physical activity mobile apps present a unique medium to disseminate scalable interventions to increase levels of physical activity. However, the effectiveness of mobile apps has previously been limited by low levels of engagement. Existing Web-based social networking platforms (eg, Facebook and Twitter) afford high levels of popularity, reach, and sustain engagement and, thus, may present an innovative strategy to enhance the engagement, and ultimately the effectiveness of mobile apps.

Objective: This study aimed to comparatively examine the effectiveness of, and engagement with, interventions that incorporate physical activity mobile apps in conjunction with and without existing Web-based social networking platforms (eg, Facebook and Twitter).

Methods: A systematic review was conducted by following the Preferred Reporting Items for Systematic Reviews and Meta-Analysis Guidelines. A systematic search of the following databases was conducted: Medline, PsycINFO, Web of Science, Scopus, CINAHL, ProQuest, SPORTDiscus, EMBASE, and Cochrane. According to the comparative objective of this review, 2 independent literature searches were conducted. The first incorporated terms related to apps and physical activity; the second also incorporated terms related to Web-based social networking. The results of the two searches were synthesized and compared narratively.

Results: A total of 15 studies were identified, 10 incorporated a physical activity app alone and 5 incorporated an app in conjunction with an existing Web-based social networking platform. Overall, 10 of the 15 interventions were effective in improving one or more physical activity behaviors. Specifically, improvements in physical activity behaviors were reported in 7 of the 10 interventions incorporating physical activity apps alone and in 3 of the 5 interventions incorporating physical activity apps in conjunction with existing Web-based social networking platforms. Interventions incorporating physical activity apps alone demonstrated a decline in app engagement. In contrast, the physical activity apps in conjunction with existing Web-based social networking platforms showed increased and sustained intervention engagement.

Conclusions: The interventions incorporating physical activity apps in conjunction with and without existing Web-based social networking platforms demonstrated effectiveness in improving physical activity behaviors. Notably, however, the interventions that incorporated existing Web-based social networking platforms achieved higher levels of engagement than those that did not. This review provides preliminary evidence that existing Web-based social networking platforms may be fundamental to increase engagement with physical activity interventions.

(J Med Internet Res 2019;21(8):e12687) doi:10.2196/12687

KEYWORDS

physical activity; mobile applications; social networking

XSL•FO
RenderX

Introduction

Physical inactivity is a global pandemic. Globally, 1.4 billion adults (28%) are not meeting the physical activity guidelines (150 min of physical activity per week), a figure that is steadily increasing [1]. This is of public health concern given the consistently documented benefits of physical activity, including a reduced risk of cardiovascular disease, hypertension, osteoporosis, diabetes mellitus, obesity, mental illness, and premature mortality [2-4]. Thus, it is important to develop innovative, scalable interventions to increase levels of physical activity.

Advancements in mobile technology, specifically the development of mobile apps, present a unique medium to deliver interventions targeted at improving health behaviors. Mobile apps are software programs developed for mobile phones and tablets that hold potential to influence health behaviors owing to their widespread reach, accessibility, and convenience [5]. Recently, there has been a proliferation of mobile health apps, with estimates of over 318,000 available for download, double the number available 2 years ago [6]. Among mobile health apps, physical activity apps account for the largest proportion (30%) and are expected to increase 87% faster than any other category of health app [7]. Despite the ever-increasing ubiquity of physical activity mobile apps, previous reviews have only demonstrated modest evidence from such apps in terms of the magnitude of their effectiveness to positively influence physical activity behavior [8-11]. This indicates that there is potential to improve the effectiveness of physical activity mobile apps.

The effectiveness of mobile apps is influenced by levels of engagement with the app [8]. Specifically, a dose-response has been identified, such that increasing levels of engagement, and thus greater exposure to intervention content, is associated with improved behavioral outcomes [12]. Unfortunately, commercial research has identified a lack of commitment to sustained engagement with health and physical activity apps, reporting that few individuals (10%) engage with downloaded apps for more than 7 days [13,14]. An initial review of interventions incorporating physical activity apps also revealed rapid declines in app engagement over intervention periods of 3 and 9 months [15]. A more recent review further documented that interventions incorporating apps were effective only in the short term (<3 months), and this was purportedly linked to declining levels of engagement over time [11]. This is concerning given that long-term engagement in physical activity behaviors is important to attain any associated health benefits [16]. It is clear that strategies are needed to enhance engagement with mobile apps targeted at increasing physical activity. This, however, requires a greater understanding of the specific features of mobile apps that may augment engagement, and ultimately enhance their effectiveness.

An important consideration in the endeavor to improve the effectiveness of physical activity mobile apps is the appropriate utilization of behavior change theory. This is fundamental as the existing empirical literature has consistently identified that effective physical activity interventions are informed by theory [17,18]. However, previous research within the realm of physical activity interventions incorporating mobile apps has documented that the utilization of behavior change theory is largely lacking [19-22]. In addition, among the physical activity apps that are informed by theory, a diverse range of theories have been utilized including the Health Belief Model; Transtheoretical Model; Self-determination Theory; and Social Cognitive Theory [19-22]. This has limited the formation of conclusions regarding the most appropriate theoretical foundation(s) to inform the development of apps [23].

Behavior change theories are important in isolating specific features to incorporate into an intervention that will effectively facilitate behavior change. Given this, it is not surprising that an emerging body of research examining the content of physical activity mobile apps has identified that apps are lacking in the inclusion of features underpinned by behavior change theory [19-22]. Nevertheless, the limited theory-driven research to date has identified one particular feature, namely social support, that has been consistently incorporated into physical activity mobile apps and is underpinned by a myriad of behavior change theories [19-22]. Social support is commonly integrated into apps via Web-based social networking, which allows individuals to construct a personal profile and connect with other users [21]. Web-based social networks incorporated into physical activity mobile apps have a range of functionalities, including features that allow users to share physical activity data, receive *likes* and comments on their behavior (facilitating social interactions), and thus foster the provision of social support [21].

Typically, social support has been documented as a fundamental component of health interventions delivered face to face and has been associated with increased intervention engagement [12,24] and sustained behavior change [25]. Although face-to-face interventions may effectively facilitate high levels of support through interpersonal interactions, several limitations including time, cost, and resource intensiveness may hinder the viability of such interventions. Web-based social networks overcome many of the barriers of face-to-face interventions and afford several advantages including greater accessibility of immediate and continuous support, anonymity, and wide reach. Additionally, Web-based social networks incorporated into Web-based interventions targeting weight-related outcomes (eg, body weight and body mass index [BMI]) have demonstrated that the support provided is comparable with that attained in face-to-face interventions [26]. Thus, it has been suggested that the support provided by Web-based social networks may emulate the interpersonal support achieved through face-to-face interventions [27]. Evidently, Web-based social networking may be valuable in facilitating the provision of social support and fundamental in enhancing intervention engagement and thus effectiveness.

Previous research has ascertained 2 types of Web-based social networks incorporated into health interventions: (1) health-focused social networks (ie, networks developed by a researcher or integrated into health apps allowing users to connect with other users), and (2) existing social networking platforms (eg, Facebook and Twitter) [28,29]. In total, 2 systematic reviews have examined interventions (predominately Web-based) targeting health behaviors, including obesity, physical activity, sexual health, and smoking cessation, that

XSL•FO
RenderX

either incorporated or were exclusively delivered via Web-based social networks (health-focused and existing) [28,29]. These reviews demonstrated positive effects of Web-based social networking in modifying health behaviors [28,29]. However, neither review [28,29] was able to identify the differing effectiveness of health-focused and existing Web-based social networks on influencing health outcomes and levels of engagement, as the 2 types of social networks were not evaluated independently. Notably, in both reviews, it was proposed that the inherent nature of existing Web-based social networking platforms may be harnessed to address issues of engagement and reach, ultimately enhancing the effectiveness of health interventions [28,29].

A recent meta-analysis [30] of interventions (eg, Web-based, face-to-face, and text messaging) targeting weight-related behaviors (eg, physical activity) and body weight status (eg, BMI) that either incorporated or were exclusively delivered via existing Web-based social networking platforms reported that these interventions produced significant reductions in body weight, BMI, and waist circumference, and significantly increased the average number of daily steps. This demonstrates that interventions incorporating, or exclusively delivered via existing Web-based social networking platforms, have the capacity to effectively modify a range of health-related outcomes. This may be attributed to the unique nature of existing Web-based social networking platforms, including their enormous popularity and widespread reach, with over 2.46 billion users worldwide, a figure that is continuing to rise [31]. Additionally, existing Web-based social networking platforms achieve high levels of sustained engagement, with estimates that 76% of Facebook users log in daily, 51% engage multiple times per day, and 70% continue to use the platform after 24 months [31]. Therefore, interventions that incorporate existing Web-based social networking platforms may achieve heightened effectiveness in their capacity to reach large audiences and sustain high levels of engagement.

Previously, no review has exclusively examined the effectiveness of interventions that incorporate physical activity mobile apps in conjunction with existing Web-based social networking platforms. The high prevalence of physical activity mobile apps, coupled with the promising capabilities of existing Web-based social networking platforms to augment app effectiveness, highlights an important avenue that warrants examination. Thus, this review examined the influence of existing Web-based social networking platforms on the effectiveness of, and engagement with, mobile apps that target physical activity. To isolate the influence of existing Web-based social networking platforms, this review provides a comparison between interventions that incorporate physical activity mobile apps in conjunction with and without existing Web-based social networking platforms.

Methods

Overview

The systematic review was conducted according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) Guidelines [32] (see Figures 1 and 2) and was registered with the International Prospective Register of Systematic Review (registration number CRD42018106456). An academic health librarian assisted with the development of the search strategy. The search strategy incorporated key terms and thesaurus terms related to mobile apps (eg, application, app, mobile phone, and iPhone), physical activity (eg, exercise, fitness, sports, inactive, and sedentary behavior) and Web-based social networks (eg, social network, social medium, Facebook, Twitter, and Instagram; see Multimedia Appendices 1 and 2 for complete search strategy). However, according to the comparative aims of this review, 2 independent searches were conducted, which differed such that one incorporated the terms related to apps and physical activity (app-alone search) and the other also incorporated the terms related to Web-based social networking (app Web-based social networking search). Both searches were conducted on the July 3, 2018, using the following 9 databases: Medline, PsycINFO, Web of Science, Scopus, CINAHL, ProQuest, SPORTDiscus, EMBASE, and Cochrane. The search results were limited to the English language, peer-reviewed, and year of publication between 2007 (the year smartphones were introduced) and the July 3, 2018.

Inclusion Criteria and Study Selection

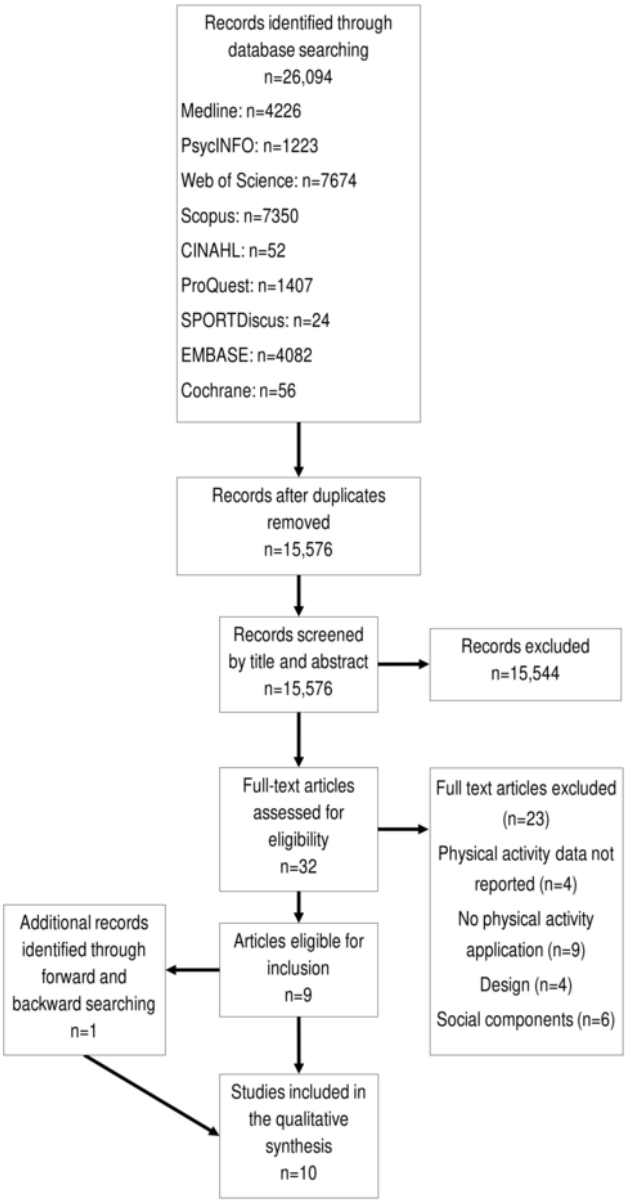
Studies from the 2 independent searches were selected if (1) a mobile app was incorporated as the main component of the intervention; (2) the primary or secondary outcome was to promote physical activity; (3) physical activity outcomes were reported; and (4) baseline and postintervention assessments of physical activity outcomes were included. The inclusion criteria differed slightly between the 2 searches to fulfill the comparative aims of the review. Specifically, the first search, termed app-alone, attempted to exclusively isolate the effect of physical activity apps, such that studies were deemed relevant if they did not incorporate any type of Web-based social network (health-focused or existing) or social component. Conversely, to ascertain the additive effects of an existing Web-based social network over and above that of an app, the second search, termed app Web-based social networking, required studies to specifically incorporate an existing Web-based social networking platform (eg, Facebook and Twitter) into their design. Included studies utilized an experimental or within-subjects pre-post design to determine the effectiveness of the intervention. Studies incorporating populations capable of engaging in physical activity were eligible for inclusion. In total, 2 reviewers independently screened the titles, abstracts, and full-text papers for eligibility and any disagreements were resolved by discussion. Forward (screening the citations of included studies) and backward (screening the reference lists of included studies) searching was conducted to ensure all relevant publications were identified.

Data Extraction

Data extraction was conducted by the first author using a standardized form developed for this review. Extracted information included sample characteristics, study design, features of the mobile app, details of the Web-based social network, physical activity outcomes (time points reported), any additional outcomes reported (eg, engagement and psychosocial outcomes), and behavior change theories reported.

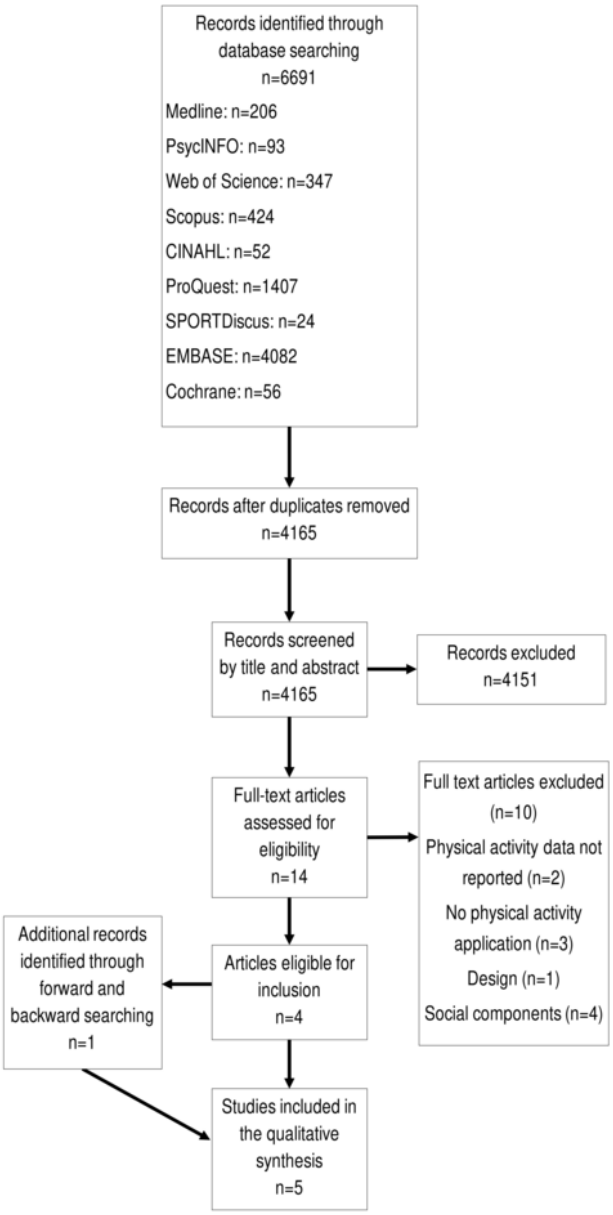
XSL•FO
RenderX

Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flowchart: App-alone search.



XSL•FO
RenderX

Figure 2. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flowchart: App Web-based Social Networking Search.



Reporting of Methodological Characteristics

A 25-item tool devised by Maher et al [28] based on the Consolidated Standards of Reporting Trials (CONSORT) checklist [33] that examines reported methodological characteristics was used to assess methodological risk of bias. The tool was deemed to be relevant for this study as most of the items (20 out of 25) were applicable to both pre-post designs and randomized controlled trials [11,28]. The checklist was scored according to the extent to which each item was (1) fulfilled; (0.5) partially fulfilled; and (0) not fulfilled. A higher score is indicative of a lower risk of bias. In total, 2 independent reviewers assessed all included studies, and any disagreements were discussed and resolved.

Data Synthesis

The primary outcome was physical activity behavior. The secondary outcomes included engagement with the intervention and psychosocial outcomes related to physical activity. In line with the comparative aims of the review, the app-alone and app Web-based social networking studies were compared in relation to both the primary and secondary outcomes. To determine whether the interventions effectively improved physical activity behavior, *P* values were evaluated. Specifically, interventions that were randomized controlled trials were identified to be effective if significant differences between groups across time were reported. Interventions of a pre-post study design were identified to be effective if significant changes across time were reported. Effect sizes were also examined and taken into account when evaluating the effectiveness of the interventions. The benchmark criteria for effect sizes are 0.20 for a small effect, 0.50 for a medium effect, and 0.80 for a large effect [34].

XSL•FO
RenderX

Results

Study Selection

The first database search (app-alone) identified 15,576 studies, following the removal of duplicates. Title and abstract screening deemed 15,544 studies ineligible for inclusion. In total, 32 full-text articles were screened for inclusion, with 23 studies excluded at this point (see Figure 1 for reasons). Forward and backward searching identified 1 additional study that was eligible for inclusion. A total of 10 app-alone studies were deemed relevant according to the predefined criteria and thus were included in this review (Figure 1).

The second database search (app Web-based social networking) identified 4165 studies, after removing duplicates. Title and abstract screening identified 4151 ineligible studies. In total, 14 full-text articles were screened for inclusion, resulting in 10 studies being excluded (see Figure 2 for reasons). Screening of reference lists and forward searching identified 1 additional study that was eligible for inclusion. A total of 5 studies were deemed suitable to be included in this review (Figure 2).

Thus, the following review included a total of 15 studies. Of these, 10 studies used an app alone, and 5 studies incorporated an app in conjunction with an existing Web-based social networking platform. These numbers of studies are similar to those of a recent comparative review [35].

Characteristics of Included Studies

The characteristics of the app-alone studies are tabulated in Multimedia Appendix 3 and those of the app Web-based social networking studies are tabulated in Multimedia Appendix 4. The app-alone and app Web-based social networking studies were comparable in years of publication and the countries where the studies were conducted. However, the study designs differed such that the app-alone studies predominately utilized an experimental design (*n*=7) [36-42], whereas the app Web-based social networking studies predominantly utilized within-subjects pre-post designs (*n*=4) [43-46]. Across the 7 app-alone studies that utilized an experimental design, the control groups received either a no intervention control (*n*=1) [40]; minimal intervention (eg, accelerometer or print materials; *n*=5) [36-39,42]; or an app that differed slightly (fewer features; *n*=1) [41]. In contrast, the 1 app Web-based social networking study that included a control utilized a waitlisted control condition [47]. Among all included studies, 2 app Web-based social networking studies [45,47] aimed to modify physical activity in conjunction with dietary quality. Across the app-alone and app Web-based social networking studies, a greater number of interventions utilized newly designed apps (*n*=10) [40-45,47-50] than commercially available apps (*n*=5) [36-39,46]. The app-alone and app Web-based social networking studies incorporated samples that were similar in size, age, and the predominance of female participants. The samples that were composed of women, were women who were healthy [37,39,43], overweight and obese [47,48,50], insufficiently active [49,50], or nurses [44,45]. Although both the app-alone and app Web-based social networking studies largely recruited from a specific population (*n*=11) [38,40-42,44-50], disparities were noted among the app-alone and app Web-based social networking studies in relation to the populations recruited. Specifically, the app-alone interventions recruited samples that were sedentary (*n*=3) [38,40,49], low active (*n*=3) [41,42,50], obese or overweight (*n*=2) [48,50], in primary care (*n*=1) [36], pregnant (*n*=1) [38], or diagnosed with type 2 diabetes (*n*=1) [49]. Contrastingly, the app Web-based social networking interventions targeted samples that were nurses (*n*=2) [44,45], breast cancer survivors (*n*=1) [46], and obese or overweight (*n*=1) [47]. The average intervention duration for app-alone studies ranged from 1 week [40] to 14 weeks [50], comparable with the intervention durations of the app Web-based social networking studies that ranged from 3 weeks [44] to 3 months [45]. One app-alone study incorporated a 3-month follow-up assessment [42], whereas 2 app Web-based social networking studies incorporated follow-up assessments at 1 week postintervention [46] and 6 months postintervention [45].

Among the app-alone and app Web-based social networking studies, all apps targeted aerobic physical activity including light physical activity (*n*=6) [39,42,45,46,48,49], moderate physical activity (*n*=2) [39,42], moderate-to-vigorous physical activity (MVPA; *n*=6) [41,45-49], vigorous physical activity (*n*=2) [39,42], and daily steps (*n*=9) [36-38,42-46,50]. The apps incorporated a diverse range of features targeted at encouraging physical activity, including monitoring or tracking of behavior (*n*=9) [36,37,41,43,44,46-48,50], feedback (*n*=7) [36-38,40-43], information or education relating to physical activity (*n*=4) [38,40,41,47], goal setting (*n*=5) [41-43,45,50], and reinforcements (*n*=4) [40,41,48,50]. Both the app-alone and app Web-based social networking studies were underpinned by a diverse range of behavior change theories, namely the Social Cognitive Theory [38,40,41,43,45,46,50], Self-determination Theory [39], Control Theory [40,45], Goal-Setting Theory [45], attitude-social influence self-efficacy model [42], the Behavior Change Wheel [37], and the Theory of Reasoned Action [43].

Description of the Existing Web-Based Social Networks

Among the app Web-based social networking studies, all 5 incorporated Facebook as the existing Web-based social networking platform; however, this platform was differentially utilized. In total, 2 studies provided participants with a link to a private Facebook group [45,47]; and 1 study incorporated a public Facebook page that included educational tips related to physical activity and participants were encouraged to comment and generate posts [46]. Alternatively, in 2 studies, the app had the functionality to connect to Facebook, whereby participants could share their physical activity data and receive *likes* and comments [43,44]. The existing Web-based social networks most often utilized features that facilitated social interaction (sharing physical activity posts, liking or commenting on others posts, and communicating with others; *n*=5) [43-47], social comparison (viewing posts of others' physical activity performance; *n*=3) [43,44,46], and competition (ranking table and group averages; *n*=2) [43,44].

Measures of Physical Activity and Additional Outcomes

Both the app-alone and app Web-based social networking studies primarily measured physical activity objectively (*n*=14) [36-38,40-50], specifically by utilizing an accelerometer (*n*=8)

XSL•FO
RenderX

[41-43,45-49], pedometer (*n*=3) [36,37,44], Fitbit (*n*=2) [38,50] or inclinometer (*n*=1) [40]. Among all included studies, 2 app-alone studies measured physical activity by self-report, specifically by using the International Physical Activity Questionnaire (IPAQ)-Long form [42] and IPAQ-Short form [39]. Physical activity outcomes predominantly targeted for modification included light physical activity (*n*=6) [39,42,45,46,48,49], moderate physical activity (*n*=2) [39,42], MVPA (*n*=6) [41,45-49], vigorous physical activity (*n*=2) [39,42], daily steps (*n*=9) [36-38,42-46,50], or sedentary behavior (*n*=5) [40,45,46,48,49]. Across all studies, the underlying psychosocial outcomes related to physical activity (ie, self-efficacy and exercise motivation) were assessed by 4 app-alone studies [38,39,41,42] and 2 app Web-based social networking studies [45,46].

The Effectiveness of the Intervention

Table 1 provides a summary of the intervention effects on physical activity outcomes. Across all included studies, 10 of the 15 interventions effectively improved one or more physical activity behaviors [36,37,40,41,43,44,46,48-50], including 7 of the 10 app-alone interventions [36,37,40,41,48-50] and 3 of the 5 app Web-based social networking interventions [43,44,46]. Improvements were reported in either the intervention conditions relative to a control condition (*n*=3) [36,37,40] or over time (*n*=7) [41,43,44,46,48-50] for one or more physical activity behaviors. Specifically, the physical outcomes reported were increases in daily steps (*n*=6) [36,37,43,44,46,50]; increases in light physical activity (*n*=2) [48,49]; increases in MVPA (*n*=3) [41,46,48]; and decreases in sedentary behavior (*n*=3) [40,46,48]. In total, 5 studies, 3 app-alone studies [38,39,42] and 2 app Web-based social networking studies [45,47], did not find an intervention effect across groups [38,39,42,47] or across time [45] in any of the physical activity behaviors measured. Effect sizes varied widely among both the app-alone and app Web-based social networking studies. Across the app-alone studies, effect sizes were small (*n*=2) [36,37], medium (*n*=2) [40,41], and large (*n*=1) [40]. Similarly, the distribution of effect sizes reported among the app Web-based social networking studies ranged from small (*n*=2) [45,46] to medium (*n*=2) [44,46] to large (*n*=1) [46].

Table 2 provides a summary of the intervention effects on psychosocial outcomes. The app-alone and app Web-based social networking studies overall reported mixed results in relation to psychosocial outcomes associated with physical activity. Specifically, 2 app-alone studies [39,42] and 1 app Web-based social networking study [45] revealed no significant intervention effects on any of the assessed psychosocial outcomes. In total, 2 app-alone studies reported significant decreases in perceptions of barriers to exercising in the intervention condition; however, not in the alternative outcomes assessed (eg, perceived social support and self-efficacy) [38,41]. Contrastingly, 1 app Web-based social networking study reported improvements over time in all psychosocial outcomes assessed (eg, social support, physical activity self-efficacy, and enjoyment) [46].

XSL•FO
RenderX

Table 1. Summary of intervention effects on physical activity outcomes.

Study	Physical activity (PA) outcomes			Engagement
	Daily steps	Light, moderate, moderate-to-vigorous physical activity (MVPA), and vigorous PA	Sedentary behavior	
App-alone studies				
Arrogi et al, 2017 [40]	—[a]	—	[++][b]	—
Bond et al, 2014 [48]	—	[+][c]	[+]	—
Choi et al, 2016 [38]	[-][d]	—	—	x[e]
Cowdery et al, 2015 [39]	—	[-]	—	—
Fanning et al, 2017 [41]	—	[+]	—	x
Glynn et al, 2014 [36]	[++]	—	—	—
Korinek et al, 2018 [50]	[+]	—	—	—
Pellegrini et al, 2015 [49]	—	[+/−][f]	[-]	✓[g]
Simons et al, 2018 [42]	—	[-]	—	x
Walsh et al, 2016 [37]	[++]	—	—	—
App Web-based social networking studies				
Al Ayubi et al, 2014 [43]	[+]	—	—	✓
Foster et al, 2010 [44]	[+]	—	—	✓
Hurkmanns et al, 2018 [47]	—	[-]	—	—
Pope et al, 2018 [46]	[+]	[+]	[+]	✓
Torquati, Kolbe-Alexander et al, 2018 [45]	[-]	[-]	[-]	x

[a]Not applicable.

[b]Significant between-group improvement in outcome.

[c]Significant within-group improvement in outcome.

[d]No improvement in outcome.

[e]Unfavorable (low) engagement.

[f]Mixed results; engagement.

[g]Favorable (high) engagement.

XSL•FO
RenderX

Table 2. Summary of intervention effects on psychosocial outcomes.

Study	Psychosocial outcomes								Behavior change theories
	Social support	PA[a] self-efficacy	PA motivation	Barriers to PA	PA enjoyment	Outcome expectations	Perceived benefits of PA	Perceived PA competency	
App-alone studies									
Arrogi et al, 2017 [40]	—[b]	—	—	—	—	—	—	—	SCT[c], CT[d]
Bond et al, 2014 [48]	—	—	—	—	—	—	—	—	—
Choi et al, 2016 [38]	[-][e]	[-]	—	[++][f]	—	—	—	—	SCT
Cowdery et al, 2015 [39]	—	—	[-]	—	[-]	—	—	[-]	SDT[g]
Fanning et al, 2017 [41]	—	[-]	—	[+][h]	—	[-]	—	—	SCT
Glynn et al, 2014 [36]	—	—	—	—	—	—	—	—	—
Korinek et al, 2018 [50]	—	—	—	—	—	—	—	—	SCT
Pellegrini et al, 2015 [49]	—	—	—	—	—	—	—	—	—
Simons et al, 2018 [42]	[-]	[-]	—	[-]	—	—	[-]	—	ASE[i]
Walsh et al, 2016 [37]	—	—	—	—	—	—	—	—	COM-B[j]
App Web-based social networking studies									
Al Ayubi et al, 2014 [43]	—	—	—	—	—	—	—	—	SCT, TRA[k]
Foster et al, 2010 [44]	—	—	—	—	—	—	—	—	—
Hurkmanns et al, 2018 [47]	—	—	—	—	—	—	—	—	—
Pope et al, 2018 [46]	[+]	[+]	—	—	[+]	—	—	—	SCT
Torquati, Kolbe-Alexander et al, 2018 [45]	[-]	[-]	—	—	—	—	—	—	SCT, GST[l], CT

[a]PA: physical activity.

[b]Not applicable.

[c]SCT: Social Cognitive Theory.

[d]CT: Control Theory.

[e]No improvement in outcome.

[f]Significant between-group improvement in outcome.

[g]SDT: Self-Determination Theory.

[h]Significant within-group improvement in outcome.

[i]ASE: Attitude-social Influence Self-efficacy Model.

[j]COM-B: The Capability, Opportunity, Motivation, Behavior framework.

[k]TRA: The Theory of Reasoned Action.

[l]GST: Goal setting Theory.

Measures of Engagement

Notably, only 4 of the 10 app-alone studies (40%) reported on app usage [38,41,42,49], whereas 80% (*n*=4) of the app Web-based social networking studies assessed engagement with intervention materials (app and Web-based social network) [43-46]. Among the studies that assessed app engagement, objective measures were primarily utilized (*n*=6) [38,41-44,49]. This included the use of Google Analytics to monitor app logins and duration of use (*n*=2) [42,44], the functionality of the app to record logins (*n*=1) [41] or days and minutes of use (*n*=2) [43,49], or monitoring of engagement with app content (eg, reading or responding to automated messages and logging in activity diary; *n*=2) [38,42]. Self-report measures of app engagement were also utilized in 2 app Web-based social networking studies [45,46]. This included questionnaires

XSL•FO
RenderX

whereby participants were asked to report frequency and duration of app use (*n*=1) [46] or engagement with app content (eg, willingness to use app and follow instructions; *n*=1) [45]. All studies that measured app engagement objectively (*n*=6) [38,41-44,49] monitored app usage over the duration of the intervention period. Conversely, among the 2 studies that utilized self-report measures, the questionnaires were completed at 2 time points: at mid and postintervention [46] and at postintervention and 6-month follow-up [45]. Among the app Web-based social networking studies, 2 reported engagement with the existing Web-based social network, such that the number of Facebook posts generated and posts viewed was monitored [45,46].

Engagement With the Intervention

Among the 4 app-alone studies that assessed engagement with the app, 1 reported that, on average, the app was used on 21 days for a cumulative total of 7.6 hours, over a 1-month intervention period [49]. The other 3 studies reported a notable decline in app engagement [38,41,42]. Specifically, decreases were reported in the frequency and duration of app usage [41,42] and engagement with app content (logging physical activity and reading or responding to messages) [38,42] over 9-week [42] and 12-week intervention periods [38,41]. Among the app Web-based social networking studies, a single study reported limited engagement with the intervention materials over a 3-month intervention period, reporting that 68.4% of participants used the app less than once a month or never and 47.5% of participants engaged with the Facebook page on only one occasion per week [45]. Conversely, 2 reported increases in minutes of app usage following the provision of access to the existing Web-based social network [43,44], and 1 reported sustained engagement with intervention materials (app and Facebook page; *n*=1) [46].

Comparison of Effective and Ineffective Interventions

As can be seen in Table 1, across all included studies, 7 of the 10 app-alone interventions (70%) [36,37,40,41,48-50] and 3 of the 5 app Web-based social networking interventions (60%) [43,44,46] were effective in improving one or more physical activity behaviors, as identified by *P* values and/or effect sizes. Among the effective interventions, the intervention durations were relatively short, ranging from 1 week [40] to 14 weeks [50]. In comparison, the ineffective interventions typically incorporated longer intervention durations, ranging from 9 weeks [42] to 3 months [45]. Notably, 6 of the 10 (60%) effective interventions recruited low-active (*n*=2) [41,50] or sedentary participants (*n*=2) [40,49], or documented that participants engaged in low levels of baseline physical activity (*n*=2) [37,48]. By contrast, only 2 of the 5 (40%) ineffective interventions recruited low-active (*n*=1) [42] or sedentary participants (*n*=1) [38]. The effective interventions all exclusively targeted physical activity behaviors. The 2 app Web-based social networking interventions that were not effective [45,47] both targeted the modification of physical activity in conjunction with diet quality. Across all included studies, objective measures of physical activity were predominately utilized (*n*=14) [36-38,40-50], and the type of objective measure used (eg, accelerometer) was comparable among the effective and ineffective interventions. However, 2 of the 5 ineffective interventions utilized self-report measures to assess the physical activity behaviors [39,42]. Both the effective (*n*=6) [37,40,41,43,46,50] and ineffective (*n*=4) [38,39,42,45] interventions were largely underpinned by behavior change theories. Among the 10 effective studies, 7 (70%) used newly designed apps [40,41,43,44,48-50] and 3 (30%) used commercially available apps [36,37,46]. Among the 5 ineffective studies, 3 (60%) used newly designed apps [42,45,47] and 2 (40%) used a commercially designed app [38,39].

In total, 2 of the effective interventions [41,46] assessed psychosocial outcomes, and mixed findings were reported. Specifically, the app-alone study that incorporated a newly designed app reported no changes in physical activity self-efficacy or physical activity outcome expectancies but identified a decrease in perceptions of barriers to exercising [41]. In contrast, the app Web-based social networking study that incorporated a commercially available app reported increases in physical activity self-efficacy, physical activity enjoyment, and social support [46]. In total, 4 of the ineffective studies assessed psychosocial outcomes [38,39,42,45], and although 1 study identified a decrease in the lack of energy as a barrier to exercising [38], no changes were reported in any of the alternative outcomes assessed, including social support [38,42,45], physical activity self-efficacy [38,42,45], physical activity enjoyment [39], physical activity motivation [39], perceived competency for exercising regularly [39], and perceived benefits to exercising [42].

Among the effective studies, 1 app-alone study [41] and all app Web-based social networking studies (*n*=3) [43,44,46] reported on app engagement. The app-alone study reported a decline in app usage over the 12-week intervention period [41]. In contrast, in the app Web-based social networking studies, higher app usage following the provision of access to the Web-based social networking functionalities [43,44] and sustained engagement with intervention materials (app and Facebook page) were reported [46]. Among the ineffective studies, 3 of the 5 studies reported on intervention engagement [38,42,45]. Of these studies, all reported unfavorable intervention engagement, specifically declines in app engagement during a 9-week [42] and 12-week intervention period [38], and low engagement with intervention materials (app and Facebook group) [45]. Additionally, among the effective app Web-based social networking interventions, the existing social networks utilized were a public Facebook page (*n*=1) [46] or a physical activity app that incorporated functionalities to connect with Facebook (*n*=2) [43,44]. Among the 2 ineffective app Web-based social networking interventions, both incorporated a private Facebook group as the existing Web-based social network [45,47].

Reporting of Methodological Characteristics

The reported methodological characteristics were examined to generate a methodological risk of bias score. Scores ranged from 9.5 (out of 20) to 20.5 (out of 25) in the app-alone studies (Multimedia Appendix 5) and from 8.5 (out of 20) to 18 (out of 25) in the app Web-based social networking studies (Multimedia Appendix 6). The app-alone and app Web-based

XSL•FO
RenderX

social networking studies all fulfilled the checklist criteria for scientific background and a detailed description of the intervention. Among the randomized controlled trials (*n*=8), few adequately reported on the allocation concealment mechanisms (*n*=3) [36,41,42] or blinding (*n*=3) [36,41,47]; however, most did report on randomization procedures (*n*=7) [36-39,41,42,47]. Notably, both the app-alone and app Web-based social networking studies rarely fulfilled the criterion detailing how the sample size was calculated (*n*=8) [36,37,39,41-43,47,48] or appropriately reported on the study outcomes (effect sizes; *n*=7) [36-38,40,41,44,48].

Discussion

Principal Findings

This review examined the influence of existing Web-based social networking platforms on the engagement with, and effectiveness of, mobile apps that target physical activity. Specifically, to isolate the influence of existing Web-based social networking platforms, the review provided a comparison between interventions that incorporated physical activity apps in conjunction with and without existing Web-based social networking platforms.

The review identified that physical activity mobile apps show promise in their capacity to improve physical activity behaviors. Of the included studies, 10 of the 15 interventions effectively improved one or more physical activity behaviors [36,37,40,41,43,44,46,48-50]. Specifically, 7 of the 10 app-alone studies [36,37,40,41,48-50] and 3 of the 5 app Web-based social networking studies [43,44,46] reported improvements. At a surface level, these findings indicate that the app Web-based social networking interventions may be no more effective than the app-alone interventions. However, this may be attributed to methodological disparities between the app-alone and app Web-based social networking interventions rather than the presence of Web-based social networking *per se*. Specifically, heterogeneity in the recruited samples may have influenced physical activity outcomes and thus must be considered in the formation of accurate conclusions regarding intervention effectiveness. This is highlighted in the comparison of 2 app-alone [48,50] and an app Web-based social networking intervention [47] that all targeted the modification of physical activity in overweight or obese individuals. The 2 app-alone interventions [48,50] both improved physical activity levels, whereas the app Web-based social networking study did not [47]. However, both app-alone studies [48,50] reported low baseline levels of physical activity, which may have influenced intervention outcomes. Furthermore, the differences in the samples recruited may also be responsible for overall differences in intervention effectiveness between the app-alone and app Web-based social networking studies. Specifically, 80% (*n*=8) of the app-alone interventions recruited low-active (*n*=3) [41,42,50] or sedentary participants (*n*=3) [38,40,49] or reported that participants engaged in low levels of physical activity at baseline (*n*=2) [37,48]. Of these interventions, 75% (*n*=6) [37,40,41,48-50] reported improvements in physical activity behaviors. This is consistent with previous literature documenting that physical activity interventions demonstrate greater effectiveness among low-active individuals, as there is a larger potential for improvement in behavior [51]. In contrast, none of the app Web-based social networking interventions incorporated recruitment criteria regarding sedentary or physical activity behaviors or reported low baseline levels of physical activity. Thus, the disparity among the samples may have influenced intervention outcomes, limiting the formation of appropriate conclusions regarding the influence of existing Web-based social networking platforms on intervention effectiveness. Future research is needed to evaluate the effectiveness of apps in conjunction with Web-based social networks in low-active or sedentary populations.

The comparability of intervention engagement between the app-alone and the app Web-based social networking interventions is also somewhat limited by the lack of reporting on engagement in the app-alone studies. This is consistent with existing reviews that have documented a lack of assessment of engagement in interventions targeting health behaviors [8]. This presents a shortcoming of research to date, such that the previously limited assessment of engagement has hindered the identification of intervention components that may be associated with engagement. This review identified clear differences in the levels of engagement reported among the app-alone and app Web-based social networking studies. The app-alone studies that reported on patterns of engagement identified declines in app engagement over 9-week [42] and 12-week intervention periods [38,41]. Of these studies, 1 reported improvement in physical activity behaviors [41], whereas the other 2 did not [38,42]. Across the app Web-based social networking studies, 1 study reported low engagement with intervention materials (app and Facebook group), and notably no improvement in physical activity outcomes [45]. In contrast, all other app Web-based social networking studies reported increases in engagement following the provision of access to the existing Web-based social networking platform [43,44] and sustained engagement with intervention materials (app and Facebook page) [46]. Among these studies, all reported improvements in physical activity behaviors [43,44,46], in line with previous evidence linking engagement with intervention effectiveness [8,12]. Thus, the app-alone studies demonstrated the typically observed decline in app engagement [38,41,42], whereas the app Web-based social networking studies showed increased and sustained intervention engagement [43,44,46]. This review provides preliminary evidence that existing Web-based social networks may be an important component in increasing engagement with physical activity interventions.

The existing Web-based social networking platform incorporated into all the app Web-based social networking interventions was Facebook, including either a public Facebook page [46], a private Facebook group [45,47], or a physical activity app that had the functionality to connect to Facebook [43,44]. The existing Web-based social networks utilized a diverse range of features that primarily facilitated social interaction, social comparison, and competition. However, the heterogeneity in the features utilized, and the predominance of studies that incorporated several different features, limited the capacity to ascertain the association between specific features of Web-based social networking and app engagement. Interestingly, the

XSL•FO
RenderX

findings indicated that the differential use of the Facebook platform may have influenced intervention effectiveness. The interventions incorporating a private Facebook group did not report improvements in physical activity behaviors [45,47]. Of these interventions, one study [45] reported on intervention engagement and psychosocial constructs, identifying low intervention engagement, and no changes in social support or self-efficacy. Contrastingly, the interventions that incorporated a Facebook page [46], or an app that connected with Facebook [43,44] showed improvements in physical activity behaviors and resulted in increased and sustained engagement. Additionally, increases were reported in social support, self-efficacy, and physical activity enjoyment in one of these studies [46]. Importantly, these are all psychosocial constructs associated with facilitating physical activity behaviors [25], intervention engagement [12,24], and sustained behavior change [25]. Notably, among the interventions that produced favorable outcomes [43,44,46], participants' existing networks were leveraged via apps that connected with Facebook [43,44], or a Facebook page [46]. Contrastingly, the interventions that produced unfavorable outcomes [45,47] incorporated private Facebook groups that generated an artificial Web-based social network, such that participants were required to create connections with unknown others. This indicates that network dynamics may be an important underlying determinant of the influence of Web-based social networks on intervention outcomes.

Implications for Future Research

This review suggests that the way in which Web-based social networking platforms are utilized must be considered in the development of interventions as it has important implications for intervention effectiveness. This highlights a gap in the literature, such that little guidance exists in relation to optimally harnessing Web-based social networking platforms in behavior change interventions. Future research must endeavor to identify specific features of Web-based social networking platforms that are associated with intervention engagement, to ascertain how best to incorporate Web-based social networking into health interventions. However, this will require a greater understanding of the mechanisms (eg, social support) underlying the influence of Web-based social networking on health behaviors, to elucidate how best to leverage specific features of Web-based social networking platforms in health interventions. In addition, Web-based social networking is evolving rapidly, and, thus, an understanding of the underlying mechanisms will be advantageous in identifying how to optimally leverage a diverse range of social networking platforms in future interventions.

The present review further ascertained disparities among the designs and quality of app-alone and app Web-based social networking studies. The app-alone interventions were predominately randomized controlled trials; by contrast, the app Web-based social networking studies were largely pre-post within-subjects designs. Thus, future research must endeavor to utilize study designs of a higher standard (ie, randomized controlled trials) to increase the quality of evidence pertaining to the effectiveness of interventions incorporating physical activity apps in conjunction with Web-based social networking. Furthermore, the app-alone and app Web-based social networking studies incorporated predominately short intervention durations, and across all studies in the review, only 3 included follow-up assessments, at 1-week postintervention [46], 3 months [42], and 6 months postintervention [45]. The dearth of evidence regarding the long-term efficacy of mobile apps is frequently documented as an important shortcoming. Evaluating the long-term effectiveness of mobile apps is imperative, as sustained engagement in physical activity behavior is required to attain the associated health benefits [16].

The review identified several features of the interventions that may be important in guiding the design of future interventions. Specifically, interventions that were effective targeted exclusively the modification of physical activity behaviors. This is consistent with previous research identifying that single behavior change interventions targeting physical activity are more effective than interventions that target multiple behaviors (eg, physical activity and dietary behavior) [52,53]. Although interventions that target multiple health behaviors simultaneously have the potential to maximize health benefits, evidence suggests that the modification of one behavior will enhance intervention outcomes [52,53]. Furthermore, the interventions that were effective incorporated objective measures of physical activity [36,37,40,41,43,44,46,48-50]. Interestingly, the 2 studies that incorporated a self-report measure of physical activity did not report an increase in physical activity over intervention periods of 9 [42] and 12 weeks [39]. It is possible that self-report measures as opposed to objective measures such as accelerometers afford lower sensitivity to detect changes in physical activity behaviors over short intervention periods [54]. Indeed, a previous review has demonstrated that 69% of studies that incorporated self-report measures, as opposed to 20% of studies that measured physical activity objectively, found no effect on physical activity [9]. In addition, in this review, comparatively, there was no difference in the effectiveness of interventions that used a newly designed app as opposed to a commercially available app. Despite this, the interventions largely utilized newly designed apps. This is problematic as commercially available apps are ubiquitous and highly accessible to the general public; however, evidence of their effectiveness is lacking [19,20,22]. Thus, future research should evaluate the effectiveness of commercially available physical activity mobile apps.

Overall, the mobile apps were effective in increasing physical activity in a diverse range of population samples, including inactive [41,50], sedentary [40,49], obese or overweight individuals [48,50], breast cancer survivors [46], and individuals diagnosed with type 2 diabetes [49]. However, all studies exclusively targeted adults, ranging from 20 [37] to 53 years [49]. Thus, future research must endeavor to evaluate the applicability of physical activity mobile apps in conjunction with existing Web-based social networks in alternative age groups, in particular among adolescents, a highly inactive population subgroup [55], and among the highest users of existing Web-based social networking platforms [56]. This will ensure that mobile apps are an appropriate medium to disseminate physical activity interventions that are scalable, owing to their applicability to the population broadly.

XSL•FO
RenderX

This review also has important implications for guiding the development of an appropriate theoretical foundation to inform future physical activity mobile apps. The included interventions incorporated mobile apps predominately underpinned by behavior change theory [37-43,45,46,50]. This suggests that there was no association between mobile app effectiveness and the utilization of any one particular theory. Additionally, across the included studies a diverse range of behavior change theories were utilized, limiting the formation of conclusions regarding the most effective theory to guide the development of physical activity mobile apps. This is consistent with previous research examining the content of physical activity mobile apps that has documented challenges ascertaining the theory or combination of theories associated with physical activity mobile app effectiveness [23].

The physical activity apps examined in this review incorporated a diverse range of features. The most common among these were monitoring or tracking of behavior, feedback, information or education related to physical activity, goal setting, and providing reinforcements (eg, points). Much of the previous research that has examined the content of physical activity apps has utilized a taxonomy developed by Abraham and Michie [57] that functions to isolate the presence of behavior change techniques common to many behavior change theories. This research has identified that feedback, self-monitoring, and goal setting are features frequently integrated into apps, in line with findings by this review [19,58,59]. Notably, Abraham and Michie [60] highlight that these features are also commonly associated with effectively modifying physical activity behavior. This may have underpinned the capacity of the majority of the apps in the current review to improve physical activity behavior. However, the specific number or combination of features that may have a greater influence on the effectiveness of physical activity apps is currently unknown, and thus requires future examination.

Limitations

To our knowledge, this is the first review to isolate the influence of existing Web-based social networking platforms by providing a comparison between interventions that incorporate mobile physical activity apps in conjunction with and without existing Web-based social networking platforms. Despite the novel nature of this review, several limitations must be noted. First, to date, there are only a small number of studies that have incorporated physical activity mobile apps in conjunction with an existing Web-based social networking platform. Additionally, owing to the heterogeneity of the identified studies in relation to the target population, intervention, study design, and outcomes measured, the results could not be validly pooled, precluding the ability to conduct a meta-analysis, and, thus, form definitive conclusions regarding the influence of Web-based social networks. Second, all interventions incorporated apps that targeted aerobic activity, and, thus, the findings may not generalize to apps aimed at other types of physical activity such as strength training. Future research should endeavor to examine apps targeted at all forms of physical activity. Third, among the included studies the methodological risk of bias varied, with some studies receiving low scores, limiting the trust that may be placed in their findings. Finally, there is a possibility of publication bias as the search did not incorporate gray literature or non-English publications.

Conclusions

In conclusion, the unprecedented growth in physical activity mobile apps presents an innovative medium to disseminate scalable interventions to increase levels of physical activity worldwide. However, previous literature has consistently documented that the effectiveness of mobile apps is limited by low levels of engagement. The popularity, reach, and engagement afforded by existing Web-based social networking platforms provides an unparalleled opportunity to serve as an adjunct to mobile apps to augment engagement, and ultimately effectiveness. Thus, this review aimed to provide insight into the influence of existing Web-based social networks by providing a comparison between interventions that incorporated mobile apps in conjunction with and without existing Web-based social networking platforms. Both the interventions incorporating physical activity apps in conjunction with and without existing Web-based social networking platforms demonstrated effectiveness in improving physical activity behaviors. Notably, however, interventions that incorporated existing Web-based social networking platforms achieved higher levels of engagement than those that did not. This provides preliminary evidence that existing Web-based social networking platforms may be fundamental in overcoming the previously documented low engagement associated with physical activity mobile apps. This is of particular importance as greater app engagement is associated with increased exposure to intervention content, and ultimately an enhanced capacity of the app to effectively improve physical activity behavior. Thus, existing Web-based social networks must be further evaluated by conducting rigorously designed randomized controlled trials. Importantly, future research must endeavor to provide a greater understanding of the mechanisms underlying the influence of Web-based social networking on physical activity behaviors, to ascertain how best to leverage specific features of Web-based social networking platforms. This review makes an important contribution to guiding future research, by providing an initial insight into mobile apps and existing Web-based social networking platforms, imperative to improving the development of interventions targeted at increasing physical activity levels.

Acknowledgments

The authors thank Nikki May for providing guidance in the development of the search strategy, and Marcela Radunz, Nepheli Beos, and Chloe Craig for their assistance with the screening process and the study quality assessments.

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1
App-alone search strategy.
[PDF File (Adobe PDF File), 48KB - jmir_v21i8e12687_app1.pdf]

Multimedia Appendix 2
App Web-based social networking search strategy.
[PDF File (Adobe PDF File), 46KB - jmir_v21i8e12687_app2.pdf]

Multimedia Appendix 3
Characteristics of app-alone studies.
[PDF File (Adobe PDF File), 122KB - jmir_v21i8e12687_app3.pdf]

Multimedia Appendix 4
Characteristics of app Web-based social networking studies.
[PDF File (Adobe PDF File), 88KB - jmir_v21i8e12687_app4.pdf]

Multimedia Appendix 5
Risk of bias based on Consolidated Standards of Reporting Trials Checklist; App-alone studies.
[PDF File (Adobe PDF File), 106KB - jmir_v21i8e12687_app5.pdf]

Multimedia Appendix 6
Risk of bias based on Consolidated Standards of Reporting Trials Checklist; App Web-based social networking studies.
[PDF File (Adobe PDF File), 82KB - jmir_v21i8e12687_app6.pdf]

References

1. Guthold R, Stevens GA, Riley LM, Bull FC. Worldwide trends in insufficient physical activity from 2001 to 2016: a pooled analysis of 358 population-based surveys with 1·9 million participants. Lancet Glob Health 2018 Dec;6(10):e1077-e1086 [FREE Full text] [doi: 10.1016/S2214-109X(18)30357-7] [Medline: 30193830]
2. Warburton DE, Nicol CW, Bredin SS. Health benefits of physical activity: the evidence. Can Med Assoc J 2006 Mar 14;174(6):801-809 [FREE Full text] [doi: 10.1503/cmaj.051351] [Medline: 16534088]
3. Barbour KA, Edenfield TM, Blumenthal JA. Exercise as a treatment for depression and other psychiatric disorders: a review. J Cardiopulm Rehabil Prev 2007;27(6):359-367. [doi: 10.1097/01.HCR.0000300262.69645.95] [Medline: 18197069]
4. 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]
5. Payne HE, Lister C, West JH, Bernhardt JM. Behavioral functionality of mobile apps in health interventions: a systematic review of the literature. JMIR Mhealth Uhealth 2015 Feb 26;3(1):e20 [FREE Full text] [doi: 10.2196/mhealth.3335] [Medline: 25803705]
6. Aitken M, Clancy B, Nass D. IQVIA. 2017. The Growing Value of Digital Health: Evidence and Impact on Human Health and the Healthcare System URL: https://www.iqvia.com/institute/reports/the-growing-value-of-digital-health [accessed 2018-08-05] [WebCite Cache ID 73bvSzBtO]
7. Grand View Research. 2017. mHealth Apps Market Size, Share & Trends Analysis Report by Type (Fitness, Lifestyle Management, Nutrition & Diet, Women's Health, Medication Adherence, Healthcare Providers/Payers), and Segment Forecasts, 2019 - 2026 URL: https://www.grandviewresearch.com/industry-analysis/mhealth-app-market [accessed 2018-11-01] [WebCite Cache ID 73bv3rLj7]
8. Schoeppe S, Alley S, van Lippevelde W, Bray NA, Williams SL, Duncan MJ, et al. Efficacy of interventions that use apps to improve diet, physical activity and sedentary behaviour: a systematic review. Int J Behav Nutr Phys Act 2016 Dec 7;13(1):127 [FREE Full text] [doi: 10.1186/s12966-016-0454-y] [Medline: 27927218]
9. Covolo L, Ceretti E, Moneda M, Castaldi S, Gelatti U. Does evidence support the use of mobile phone apps as a driver for promoting healthy lifestyles from a public health perspective? A systematic review of randomized control trials. Patient Educ Couns 2017 Dec;100(12):2231-2243. [doi: 10.1016/j.pec.2017.07.032] [Medline: 28855063]
10. Zhao J, Freeman B, Li M. Can mobile phone apps influence people's health behavior change? An evidence review. J Med Internet Res 2016 Dec 31;18(11):e287 [FREE Full text] [doi: 10.2196/jmir.5692] [Medline: 27806926]

XSL•FO
RenderX

11. Romeo A, Edney S, Plotnikoff R, Curtis R, Ryan J, Sanders I, et al. Can smartphone apps increase physical activity? Systematic review and meta-analysis. J Med Internet Res 2019 Mar 19;21(3):e12053 [FREE Full text] [doi: 10.2196/12053] [Medline: 30888321]
12. Poirier J, Cobb NK. Social influence as a driver of engagement in a web-based health intervention. J Med Internet Res 2012 Feb 22;14(1):e36 [FREE Full text] [doi: 10.2196/jmir.1957] [Medline: 22356829]
13. 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 Apr 18;15(4):e86 [FREE Full text] [doi: 10.2196/jmir.2583] [Medline: 23598614]
14. Grady A, Yoong S, Sutherland R, Lee H, Nathan N, Wolfenden L. Improving the public health impact of ehealth and mhealth interventions. Aust N Z J Public Health 2018 Apr;42(2):118-119. [doi: 10.1111/1753-6405.12771] [Medline: 29384248]
15. Jee H. Review of researches on smartphone applications for physical activity promotion in healthy adults. J Exerc Rehabil 2017 Feb;13(1):3-11 [FREE Full text] [doi: 10.12965/jer.1732928.464] [Medline: 28349027]
16. Jepson RG, Harris FM, Platt S, Tannahill C. The effectiveness of interventions to change six health behaviours: a review of reviews. BMC Public Health 2010 Sep 8;10:538 [FREE Full text] [doi: 10.1186/1471-2458-10-538] [Medline: 20825660]
17. 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]
18. Webb TL, Joseph J, Yardley L, Michie S. Using the internet to promote health behavior change: a systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. J Med Internet Res 2010 Feb 17;12(1):e4 [FREE Full text] [doi: 10.2196/jmir.1376] [Medline: 20164043]
19. Bondaronek P, Alkhaldi G, Slee A, Hamilton FL, Murray E. Quality of publicly available physical activity apps: review and content analysis. JMIR Mhealth Uhealth 2018 Mar 21;6(3):e53 [FREE Full text] [doi: 10.2196/mhealth.9069] [Medline: 29563080]
20. Conroy DE, Yang CH, Maher JP. Behavior change techniques in top-ranked mobile apps for physical activity. Am J Prev Med 2014 Jun;46(6):649-652. [doi: 10.1016/j.amepre.2014.01.010] [Medline: 24842742]
21. Mollee J, Middelweerd A, Kurvers R, Klein M. What technological features are used in smartphone apps that promote physical activity? A review and content analysis. Pers Ubiquit Comput 2017 Jul 5;21(4):633-643. [doi: 10.1007/s00779-017-1023-3]
22. Cowan LT, van Wagenen SA, Brown BA, Hedin RJ, Seino-Stephan Y, Hall PC, et al. Apps of steel: are exercise apps providing consumers with realistic expectations?: a content analysis of exercise apps for presence of behavior change theory. Health Educ Behav 2013 Apr;40(2):133-139. [doi: 10.1177/1090198112452126] [Medline: 22991048]
23. Stuckey MI, Carter SW, Knight E. The role of smartphones in encouraging physical activity in adults. Int J Gen Med 2017;10:293-303 [FREE Full text] [doi: 10.2147/IJGM.S134095] [Medline: 28979157]
24. 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 Dec;7(2):254-267 [FREE Full text] [doi: 10.1007/s13142-016-0453-1] [Medline: 27966189]
25. Courneya KS, Plotnikoff RC, Hotz SB, Birkett NJ. Social support and the theory of planned behavior in the exercise domain. Am J Health Behav 2000 Jul 1;24(4):300-308. [doi: 10.5993/AJHB.24.4.6]
26. Bennett GG, Herring SJ, Puleo E, Stein EK, Emmons KM, Gillman MW. Web-based weight loss in primary care: a randomized controlled trial. Obesity (Silver Spring) 2010 Feb;18(2):308-313 [FREE Full text] [doi: 10.1038/oby.2009.242] [Medline: 19696764]
27. Santarossa S, Kane D, Senn CY, Woodruff SJ. Exploring the role of in-person components for online health behavior change interventions: can a digital person-to-person component suffice? J Med Internet Res 2018 Dec 11;20(4):e144 [FREE Full text] [doi: 10.2196/jmir.8480] [Medline: 29643048]
28. Maher CA, Lewis LK, Ferrar K, Marshall S, de Bourdeaudhuij I, Vandelanotte C. Are health behavior change interventions that use online social networks effective? A systematic review. J Med Internet Res 2014 Feb 14;16(2):e40 [FREE Full text] [doi: 10.2196/jmir.2952] [Medline: 24550083]
29. Laranjo L, Arguel A, Neves AL, Gallagher AM, Kaplan R, Mortimer N, et al. The influence of social networking sites on health behavior change: a systematic review and meta-analysis. J Am Med Inform Assoc 2015 Jan;22(1):243-256 [FREE Full text] [doi: 10.1136/amiajnl-2014-002841] [Medline: 25005606]
30. An R, Ji M, Zhang S. Effectiveness of social media-based interventions on weight-related behaviors and body weight status: review and meta-analysis. Am J Health Behav 2017 Nov 1;41(6):670-682. [doi: 10.5993/AJHB.41.6.1] [Medline: 29025495]
31. Statista. 2019. Number of Monthly Active Facebook Users Worldwide as of 1st Quarter 2019 (in Millions) URL: https://www.statista.com/statistics/264810/number-of-monthly-active-facebook-users-worldwide/ [accessed 2019-04-09] [WebCite Cache ID 73bsWkt05]
32. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009 Jul 21;6(7):e1000097 [FREE Full text] [doi: 10.1371/journal.pmed.1000097] [Medline: 19621072]

XSL•FO
RenderX

33. Moher D, Hopewell S, Schulz KF, Montori V, Gøtzsche PC, Devereaux PJ, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. Br Med J 2010 Mar 23;340:c869 [FREE Full text] [doi: 10.1136/bmj.c869] [Medline: 20332511]
34. Cohen J. A power primer. Psychol Bull 1992 Jul;112(1):155-159. [doi: 10.1037/0033-2909.112.1.155] [Medline: 19565683]
35. Nour M, Yeung SH, Partridge S, Allman-Farinelli M. A narrative review of social media and game-based nutrition interventions targeted at young adults. J Acad Nutr Diet 2017 May;117(5):735-52.e10. [doi: 10.1016/j.jand.2016.12.014] [Medline: 28238894]
36. Glynn LG, Hayes PS, Casey M, Glynn F, Alvarez-Iglesias A, Newell J, et al. Effectiveness of a smartphone application to promote physical activity in primary care: the SMART MOVE randomised controlled trial. Br J Gen Pract 2014 Jul;64(624):e384-e391 [FREE Full text] [doi: 10.3399/bjgp14X680461] [Medline: 24982490]
37. Walsh JC, Corbett T, Hogan M, Duggan J, McNamara A. An mhealth intervention using a smartphone app to increase walking behavior in young adults: a pilot study. JMIR Mhealth Uhealth 2016 Sep 22;4(3):e109 [FREE Full text] [doi: 10.2196/mhealth.5227] [Medline: 27658677]
38. Choi J, Lee JH, Vittinghoff E, Fukuoka Y. mHealth physical activity intervention: a randomized pilot study in physically inactive pregnant women. Matern Child Health J 2016 May;20(5):1091-1101 [FREE Full text] [doi: 10.1007/s10995-015-1895-7] [Medline: 26649879]
39. Cowdery J, Majeske P, Frank R, Brown D. Exergame apps and physical activity: the results of the ZOMBIE trial. Am J Health Educ 2015 Jul 6;46(4):216-222 [FREE Full text] [doi: 10.1080/19325037.2015.1043063]
40. Arrogi A, Bogaerts A, Seghers J, Devloo K, Abeele VV, Geurts L, et al. Evaluation of stAPP: a smartphone-based intervention to reduce prolonged sitting among Belgian adults. Health Promot Int 2019 Feb 1;34(1):16-27. [doi: 10.1093/heapro/dax046] [Medline: 28973149]
41. Fanning J, Roberts S, Hillman CH, Mullen SP, Ritterband L, McAuley E. A smartphone 'app'-delivered randomized factorial trial targeting physical activity in adults. J Behav Med 2017 Oct;40(5):712-729. [doi: 10.1007/s10865-017-9838-y] [Medline: 28255750]
42. Simons D, de Bourdeaudhuij I, Clarys P, de Cocker K, Vandelanotte C, Deforche B. Effect and process evaluation of a smartphone app to promote an active lifestyle in lower educated working young adults: cluster randomized controlled trial. JMIR Mhealth Uhealth 2018 Aug 24;6(8):e10003 [FREE Full text] [doi: 10.2196/10003] [Medline: 30143477]
43. Al Ayubi SU, Parmanto B, Branch R, Ding D. A persuasive and social mhealth application for physical activity: a usability and feasibility study. JMIR Mhealth Uhealth 2014 May 22;2(2):e25 [FREE Full text] [doi: 10.2196/mhealth.2902] [Medline: 25099928]
44. Foster D, Linehan C, Kirman B, Lawson S, James G. Motivating Physical Activity at Work: Using Persuasive Social Media for Competitive Step Counting. In: Proceedings of the 14th International Academic MindTrek Conference: Envisioning Future Media Environments. 2010 Presented at: MindTrek'10; October 6-8, 2010; Tampere, Finland p. 111-116. [doi: 10.1145/1930488.1930510]
45. Torquati L, Kolbe-Alexander T, Pavey T, Leveritt M. Changing diet and physical activity in nurses: a pilot study and process evaluation highlighting challenges in workplace health promotion. J Nutr Educ Behav 2018;50(10):1015-1025. [doi: 10.1016/j.jneb.2017.12.001] [Medline: 29650395]
46. Pope Z, Lee JE, Zeng N, Lee HY, Gao Z. Feasibility of smartphone application and social media intervention on breast cancer survivors' health outcomes. Transl Behav Med 2019 Dec 1;9(1):11-22. [doi: 10.1093/tbm/iby002] [Medline: 29471477]
47. Hurkmans E, Matthys C, Bogaerts A, Scheys L, Devloo K, Seghers J. Face-to-face versus mobile versus blended weight loss program: randomized clinical trial. JMIR Mhealth Uhealth 2018 Jan 11;6(1):e14 [FREE Full text] [doi: 10.2196/mhealth.7713] [Medline: 29326093]
48. Bond DS, Thomas JG, Raynor HA, Moon J, Sieling J, Trautvetter J, et al. B-MOBILE--a smartphone-based intervention to reduce sedentary time in overweight/obese individuals: a within-subjects experimental trial. PLoS One 2014;9(6):e100821 [FREE Full text] [doi: 10.1371/journal.pone.0100821] [Medline: 24964010]
49. Pellegrini CA, Hoffman SA, Daly ER, Murillo M, Iakovlev G, Spring B. Acceptability of smartphone technology to interrupt sedentary time in adults with diabetes. Transl Behav Med 2015 Sep;5(3):307-314 [FREE Full text] [doi: 10.1007/s13142-015-0314-3] [Medline: 26327936]
50. Korinek EV, Phatak SS, Martin CA, Freigoun MT, Rivera DE, Adams MA, et al. Adaptive step goals and rewards: a longitudinal growth model of daily steps for a smartphone-based walking intervention. J Behav Med 2018 Dec;41(1):74-86. [doi: 10.1007/s10865-017-9878-3] [Medline: 28918547]
51. Lubans DR, Morgan PJ, Tudor-Locke C. A systematic review of studies using pedometers to promote physical activity among youth. Prev Med 2009 Apr;48(4):307-315. [doi: 10.1016/j.ypmed.2009.02.014] [Medline: 19249328]
52. Sweet SN, Fortier MS. Improving physical activity and dietary behaviours with single or multiple health behaviour interventions? A synthesis of meta-analyses and reviews. Int J Environ Res Public Health 2010 Dec;7(4):1720-1743 [FREE Full text] [doi: 10.3390/ijerph7041720] [Medline: 20617056]
53. Prochaska JJ, Prochaska JO. A review of multiple health behavior change interventions for primary prevention. Am J Lifestyle Med 2011 May;5(3):- [FREE Full text] [doi: 10.1177/1559827610391883] [Medline: 24358034]

XSL•FO
RenderX

54. Shephard RJ. Limits to the measurement of habitual physical activity by questionnaires. Br J Sports Med 2003 Jun;37(3):197-206; discussion 206 [FREE Full text] [doi: 10.1136/bjsm.37.3.197] [Medline: 12782543]
55. World Health Organisation. 2018. Physical Activity: Key Facts URL: http://www.who.int/news-room/fact-sheets/detail/physical-activity [accessed 2018-07-02] [WebCite Cache ID 73bw6arvq]
56. Anderson M, Jiang J. Pew Research Center. 2018. Teens, Social Media & Technology URL: http://www.pewinternet.org/2018/05/31/teens-social-media-technology-2018/ [accessed 2018-07-05] [WebCite Cache ID 73bvpc0hZ]
57. Abraham C, Michie S. A taxonomy of behavior change techniques used in interventions. Health Psychol 2008 May;27(3):379-387. [doi: 10.1037/0278-6133.27.3.379] [Medline: 18624603]
58. Middelweerd A, Mollee JS, van der Wal CN, Brug J, Te Velde SJ. Apps to promote physical activity among adults: a review and content analysis. Int J Behav Nutr Phys Act 2014 Jul 25;11:97 [FREE Full text] [doi: 10.1186/s12966-014-0097-9] [Medline: 25059981]
59. Simões P, Silva AG, Amaral J, Queirós A, Rocha NP, Rodrigues M. Features, behavioral change techniques, and quality of the most popular mobile apps to measure physical activity: systematic search in app stores. JMIR Mhealth Uhealth 2018 Oct 26;6(10):e11281 [FREE Full text] [doi: 10.2196/11281] [Medline: 30368438]
60. 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]

Abbreviations

BMI: Body Mass Index
CONSORT: Consolidated Standards of Reporting Trials
IPAQ: International Physical Activity Questionnaire
MVPA: moderate-to-vigorous physical activity
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

Edited by G Eysenbach; submitted 04.11.18; peer-reviewed by C Bedard, S Edney, B Chaudhry, L Grepo; comments to author 21.03.19; revised version received 13.05.19; accepted 10.06.19; published 16.08.19.

Please cite as:
Petersen JM, Prichard I, Kemps E
A Comparison of Physical Activity Mobile Apps With and Without Existing Web-Based Social Networking Platforms: Systematic Review
J Med Internet Res 2019;21(8):e12687
URL: https://www.jmir.org/2019/8/e12687/
doi:10.2196/12687
PMID:31420956

©Jasmine Maria Petersen, Ivanka Prichard, Eva Kemps. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 16.08.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Automated Text Messaging With Patients in Department of Veterans Affairs Specialty Clinics: Cluster Randomized Trial

Vera Yakovchenko[1], MPH, MS; Timothy P Hogan[1,2], PhD; Thomas K Houston[3], MPH, MD; Lorilei Richardson[1], PhD; Jessica Lipschitz[4,5], PhD; Beth Ann Petrakis[1], MPA; Chris Gillespie[1], PhD; D Keith McInnes[1,6], ScD

[1]Center for Healthcare Organization and Implementation Research, Bedford Department of Veterans Affairs Medical Center, Department of Veterans Affairs, Bedford, MA, United States

[2]Division of Health Informatics and Implementation Science, Department of Population and Quantitative Health Sciences, University of Massachusetts Medical School, Worcester, MA, United States

[3]Division of Health Informatics and Implementation Science, Department of Population and Quantitative Health Sciences, University of Massachusetts Medical School, Amherst, MA, United States

[4]Brigham and Women's Hospital Department of Psychiatry, Boston, MA, United States

[5]Harvard Medical School, Boston, MA, United States

[6]Department of Health Law Policy and Management, Boston University School of Public Health, Boston, MA, United States

Corresponding Author:
Vera Yakovchenko, MPH, MS
Center for Healthcare Organization and Implementation Research
Bedford Department of Veterans Affairs Medical Center
Department of Veterans Affairs
Bedford, MA,
United States
Phone: 1 781 687 3374
Email: vera.yakovchenko@va.gov

Abstract

Background: Acceptability of mobile phone text messaging as a means of asynchronous communication between health care systems and patients is growing. The US Department of Veterans Affairs (VA) has adopted an automated texting system (aTS) for national rollout. The aTS allows providers to develop clinical texting protocols to promote patient self-management and allows clinical teams to monitor patient progress between in-person visits. Texting-supported hepatitis C virus (HCV) treatment has not been previously tested.

Objective: Guided by the Practical, Robust Implementation and Sustainability Model (PRISM), we developed an aTS HCV protocol and conducted a mixed methods, hybrid type 2 effectiveness implementation study comparing two programs supporting implementation of the aTS HCV protocol for medication adherence in patients with HCV.

Methods: Seven VA HCV specialty clinics were randomized to usual aTS implementation versus an augmented implementation facilitation program. Implementation process measures included facilitation metrics, usability, and usefulness. Implementation outcomes included provider and patient use of the aTS HCV protocol, and effectiveness outcomes included medication adherence, health perceptions and behaviors, and sustained virologic response (SVR).

Results: Across the seven randomized clinics, there were 293 facilitation events using a core set of nine implementation strategies (157 events in augmented implementation facilitation, 136 events in usual implementation). Providers found the aTS appropriate with high potential for scale-up but not without difficulties in startup, patient selection and recruitment, and clinic workflow integration. Patients largely found the aTS easy to use and helpful; however, low perceived need for self-management support contributed to high declination. Reach and use was modest with 197 patients approached, 71 (36%) enrolled, 50 (25%) authenticated, and 32 (16%) using the aTS. In augmented implementation facilitation clinics, more patients actively used the aTS HCV protocol compared with usual clinic patients (20% vs 12%). Patients who texted reported lower distress about failing HCV treatment (13/15, 87%, vs 8/15, 53%; *P*=.05) and better adherence to HCV medication (11/15, 73%, reporting excellent adherence vs 6/15, 40%; *P*=.06), although SVR did not differ by group.

Conclusions: The aTS is a promising intervention for improving patient self-management; however, augmented approaches to implementation may be needed to support clinician buy-in and patient engagement. Considering the behavioral, social,

XSL•FO
RenderX

organizational, and technical scale-up challenges that we documented, successful and sustained implementation of the aTS may require implementation strategies that operate at the clinic, provider, and patient levels.

Trial Registration: Retrospectively registered at ClinicalTrials.gov NCT03898349; https://clinicaltrials.gov/ct2/show/NCT03898349

(J Med Internet Res 2019;21(8):e14750) doi:10.2196/14750

KEYWORDS

implementation facilitation; texting; veterans; eHealth; self-management; digital health; digital medicine

Introduction

Short message service (SMS or texting) is becoming an accepted means of asynchronous communication between health care systems and patients, supporting appointment attendance, medication taking, and medication refill reminders [1]. Texting interventions have been studied across a range of clinical domains and stages of care [2]. Despite the ubiquity of cell phones and the established benefits of texting interventions, there has been limited research on the implementation of patient texting in health care systems [2]. Although texting interventions have yielded improvements in processes of care and health outcomes, results have been achieved through heterogenous approaches, revealing significant implementation knowledge gaps [3-5]. Furthermore, technologies intended to directly engage patients (patient-facing technologies) encounter distinct implementation challenges when clinical staff are needed to promote patients' adoption of the technology.

Implementation facilitation, a kind of meta-strategy composed of multiple implementation strategies, uses experts in clinical, process, and implementation issues to solve problems and offer support that enables others to institute and sustain practice change [6-8]. Although facilitation is versatile, it has not been extensively studied as part of health information technology implementation or sustainability efforts [9,10].

In 2016, the US Department of Veterans Affairs (VA) began piloting an automated text messaging system (aTS) for patient self-management modeled after texting systems used in the UK National Health Service (NHS), Australia, and Canada [11,12]. The aTS, titled Annie after Annie Fox, RN, the first nurse to be awarded a Purple Heart, provides patients with a technology to become more engaged in their own care through condition-specific protocols to accentuate the key points of care plans. The aTS can send one-way messages and interpret patient messages following specified syntax and then reply through rule-based logic with two-way (bidirectional) messages.

The rollout of VA's aTS presented a unique opportunity to examine implementation in the context of a large, integrated health care system. We chose to evaluate the implementation and effectiveness of the aTS across specialty care hepatitis C virus (HCV) clinics and test the utility of implementation facilitation. We focused exclusively on HCV treatment because it involves a time-limited predefined course of daily medication, and high adherence is required to achieve sustained virologic response (SVR) and avoid drug resistance [13,14]. Adherence to follow-up appointments and regular bloodwork are also essential to successful treatment.

While there is ample evidence regarding the effectiveness of texting to support medication taking in other disease contexts, no studies, to our knowledge, have examined HCV treatment support via texting [2]. Given the evidence for similar conditions, we hypothesized that texting for HCV would have comparable effects. As such, we conducted a hybrid type 2 implementation study to simultaneously examine both clinical and implementation outcomes and thus generate evidence in this area [15]. Our aims were to (1) qualitatively and quantitatively assess implementation outcomes for the aTS and (2) assess impact of the aTS on HCV clinical outcomes in a real-world setting.

Methods

Design

This was a multisite, mixed methods, randomized, two-group hybrid type 2 study design comparing the effectiveness of usual implementation (UI) and augmented implementation (AI) facilitation. Matched comparison (ie, no intervention) sites helped to determine effectiveness of the aTS in aiding HCV treatment. The study was reviewed by the institutional review board at the Edith Nourse Rogers Memorial Veterans Hospital in Bedford, Massachusetts, and determined to be a quality improvement study and therefore exempt (VA Handbook 1058.05) [16]. The project was conducted from February 2017 through February 2018. Due to an error, this study was registered retrospectively. The study was registered at ClinicalTrials.gov (NCT03898349).

Setting and Participants

Nine HCV clinics participated in this study. Seven served as intervention sites and two as matched comparisons. The group of clinics selected reflects a purposive sample based on criteria including clinic size and complexity as well as geography. HCV clinics were recruited via a national HCV provider email listserv and monthly HCV provider phone call. The seven HCV intervention clinics selected were randomly assigned to either UI or AI, using set randomization, which, with small sample sizes, helps to achieve balance on a set of relevant characteristics—in this case, urban/suburban setting and HCV patient volume [17]. Two additional comparison clinics were selected purposively because they had similar patient volume, clinic complexity, and geographic locale to the other participating clinics [18]. In total, data were collected from nine VA clinics: four AI clinics, three UI clinics, and two comparison clinics. Care teams within the HCV clinics had different compositions and involved pharmacists, nurse practitioners, registered nurses, and social worker to varying degrees.

XSL•FO
RenderX

Text Messaging Protocol for Hepatitis C Virus Treatment

At the time of implementation, usual care for HCV included starting a patient on daily oral medication for 8 to 16 weeks with follow-up in-person visits, blood lab work, and medication refills at 2- or 4-week intervals. Using the aTS in the context of HCV was intended to improve processes, outcomes of care, and satisfaction with care. The HCV texting protocol included reminder text messages for each modifiable behavior in the HCV treatment process: medication taking, appointment attendance, laboratory completion, and self-efficacy to encourage continued engagement in treatment. Typically, clinic providers and staff would reach out by phone or letter to remind patients of HCV appointments, labs, and refills. To ensure alignment with standard treatment processes and local clinic workflow, each of the seven clinics co-designed and tailored the HCV texting protocol (eg, adjusting messaging logic from 2 to 4 weeks for different treatment intervals) in conjunction with study team members (VY, KM, and national aTS program office technical and clinical specialists). Motivational messages were written to be supportive in nature, promote self-management, and increase feelings of connection to the treatment team. Principles of universal design were also incorporated to ensure that different ranges of abilities, access, and equity were considered and the widest reach and benefit of the HCV texting protocol could be achieved [19]. Sample HCV texting protocol messages are as follows:

- Medication reminder: “Hi, it’s Annie, with a helpful reminder. Did you remember to take your HepC medication today?”
- Appointment reminder: “Hi, Annie here. Don’t forget about your upcoming HepC appointment. If you do not know when your appointment is, please call to find out.”
- Lab reminder: “Annie & VA Liver team here. You are due to have a blood draw this week so we can see how your HepC meds are working. Please call your VA care team if you need help getting your labs.”
- Motivational message: “Don’t forget that the HepC Team is here to support you in your HepC treatment efforts. Call your care team if we can help you. – Annie”

Messages could be tailored for content (eg, adding clinic name, phone number, or appointment date) and timing (eg, adjusting time of day that medication reminder is sent) through patient and provider discussion at the time of aTS enrollment or later to reflect patient preferences. Veterans who did not use the aTS received otherwise standard HCV care.

Usual Implementation and Augmented Implementation Facilitation

Facilitation was employed to support adoption of the aTS at participating clinics, with facilitation delivered by a primary (VY) and a secondary (KM) external facilitator. During the 4-month preimplementation phase, the functions of the external facilitators included engaging local, regional, and national stakeholders to garner support for the aTS. The implementation phase took place over 6 months and differed between UI and AI. The postimplementation (evaluation) phase took place over 3 months.

Usual Implementation Clinics

UI clinics received the start-up experience that VA designed for all new clinics instituting the aTS. This involved a live virtual demonstration of the aTS and access to an aTS resource website that included promotional materials and training guides. UI clinics could receive troubleshooting assistance from the external facilitators by phone or email but only if and when they reached out to the them.

Augmented Implementation Clinics

In addition to the start-up experience for UI sites, AI clinics received an implementation toolkit, support for local champion development, and proactive outreach by the primary external facilitator. The toolkit was developed by our team based on a formative evaluation that involved visits to five VA medical centers around the country that were using a pilot version of the aTS for conditions other than HCV. The toolkit contained sections on evidence of texting in health care, suggestions for gaining leadership and clinic support for technology like the aTS, use of champions to support aTS adoption, tips and tools on how to use the aTS, and aTS promotional materials to encourage clinic and patient participation. Each AI clinic received one in-person visit from the primary external facilitator early in their implementation efforts. Additionally, the primary external facilitator initiated check-ins with AI clinic champions throughout implementation.

Facilitation was delivered via email, phone, and in person. In the preimplementation period, to establish rapport and trust, there was more emphasis on phone calls and in-person meetings, whereas during implementation, those modes were used less while use of emails increased. Facilitation calls lasted from 5 minutes to 90 minutes (40-minute average) and site visits by the external facilitator lasted 2 to 4 hours. The facilitation meta-strategies included assessing for readiness to implement, site visits, identifying and preparing champions, developing and distributing educational materials, building a coalition, local technical assistance, and tailoring implementation to context [20,21].

Conceptual Framework

Our evaluation was guided by the Practical, Robust Implementation and Sustainability Model (PRISM), which defines a set of factors for consideration when designing, implementing, sustaining, and evaluating interventions [22]. PRISM posits that the extent to which an intervention achieves results can be linked to the four PRISM domains: intervention characteristics (via patient and organizational perspectives), intervention recipients (via patient and organizational perspectives), external environment, and implementation and sustainability infrastructure.

Measures and Data Collection

PRISM domains (intervention characteristics, recipients via patient and organizational perspectives, external environment, and implementation and sustainability infrastructure) guided the measures and data collection and are denoted in parentheticals.

XSL•FO
RenderX

Implementation Processes: Facilitation (Implementation and Sustainability Infrastructure)

The primary external facilitator logged facilitation events on a tracking sheet, including facilitation date, length of time, mode of delivery (ie, email, phone call, or in-person visit), purpose, notes, and other observations. If multiple facilitation events occurred in one day, only one event per person per day was counted.

Implementation Outcomes: Texting Use (Intervention)

Providers logged the number of patients who were offered the aTS and noted whether patients enrolled or declined, including the reason for declining. The content of patient text message replies was extracted from the aTS portal. To be eligible for the texting protocol, patients had to be starting HCV medication treatment during the implementation period. There were four steps to initiate a patient on the aTS: (1) providers verbally offered patients the aTS, (2) providers registered interested patients in the aTS portal and assigned them the HCV protocol, (3) once a patient was registered, the aTS would send an automated text message requesting the patient authenticate themselves by replying to this initial text message thus prompting the assigned HCV protocol to begin, and (4) patients actively texted with the aTS.

Clinical Effectiveness Outcomes

Medication adherence was measured via patient text response rate, operationalized as the number of days of text-confirmed medication taking divided by the number of days receiving medication reminder texts. Consistent with other adherence standards, an affirmative text response rate of ≥80% was considered high adherence [23]. Clinical data, including HCV treatment regimen and duration and lab results, were extracted from VA's national HCV dashboard based on data from VA corporate data warehouse. The goal of HCV treatment is to achieve cure or SVR, meaning there is an undetectable HCV lab result 12 weeks after completion of treatment.

Questionnaires (Recipients, Intervention, Implementation)

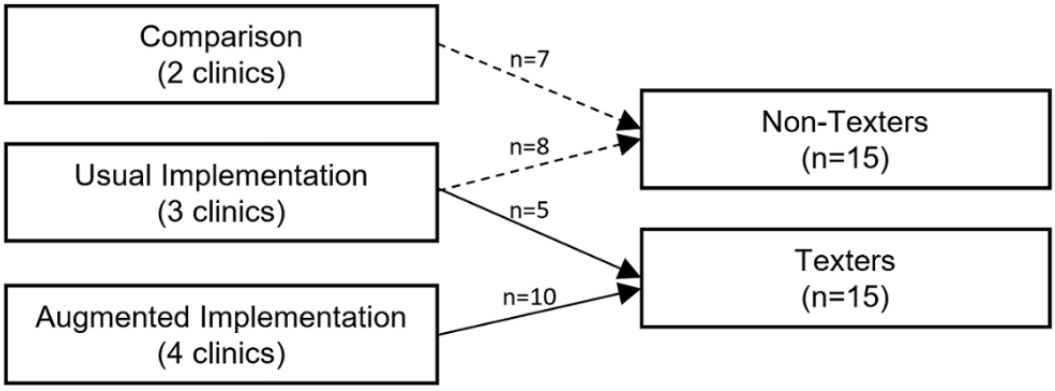
Patients at each clinic completed baseline and follow-up (after 8 to 12 weeks of using the aTS) questionnaires. The comparison clinics followed the same schedule, although without any use of texting. Patient questionnaires covered the topics of self-rated health status, adherence, illness perception, health engagement and activation, technology use, experiences with the aTS (usability, usefulness, working alliance), and demographics [24-29]. Provider questionnaires followed the same schedule and covered topics of technology experience, quality improvement culture, climate and readiness for implementation, satisfaction with current local HCV care processes, experiences with the aTS (usability and usefulness), and demographics [30-33]. Questionnaires were pretested for clarity, redundancy, and relevancy by two patients and two providers and two implementation scientists independent of the study team.

Semistructured Interviews (All Practical, Robust Implementation and Sustainability Model Domains)

Qualitative semistructured telephone interviews were conducted with patients and providers who used or were invited to use the aTS. Interviews were conducted in the follow-up period during September and October 2017. The interview guides were informed by PRISM domains and explored issues regarding barriers and facilitators to aTS uptake and use (intervention, implementation, and sustainability), usability and usefulness of the aTS (intervention), and how the aTS was experienced by patients and providers (recipients) in the course of treatment and daily practice (external environment). Interviews were conducted by members of the study team not involved in facilitation (BP, CG) and lasted about 30 minutes.

Due to the small number of participants who successfully used the aTS, it was decided that comparing texters, regardless of group (AI or UI), against nontexters was necessary. For effectiveness outcomes measures, we combined patients who were using the aTS regardless of whether they were in UI or AI clinics. These were referred to as texters. In contrast, nontexters were defined as patients who agreed to participate in the project but never completed the step of authenticating themselves with the aTS (at either UI or AI clinics), and thus never received any text messages, as well as patients from the two comparison clinics that did not implement the aTS (Figure 1). For each questionnaire or qualitative interview completed, patients received a $10 store gift card to compensate them for their time.

Figure 1. Flow diagram of clinics in the study and participants, identified as texters and nontexters.



XSL•FO
RenderX

Data Analysis

Descriptive and bivariate analyses of facilitation log data were conducted to compare facilitation dose between UI and AI groups. Descriptive and bivariate analyses were conducted on provider and patient questionnaires, text message frequencies, and clinical data to assess differences between implementation groups (UI and AI) in implementation outcomes. We then compared clinical effectiveness outcomes between texters and nontexters. We examined patient progression through the aTS initiation process by calculating the percentage retained from one step to the next by UI and AI group. Chi-square tests were used to assess differences between the two groups. All analyses were conducted in RStudio version 1.0.153 (The R Foundation), and statistical significance was defined as $P<.05$.

Qualitative interviews were audio recorded, transcribed verbatim, and analyzed using NVivo 11 (QSR International Pty Ltd) software. Thematic analysis of all qualitative data (interview transcripts, facilitator meeting notes, text messages) was conducted [34]. PRISM domains provided deductive a priori codes and other codes emerged through inductive coding. The triangulation of quantitative and qualitative data served as the final step of analysis.

Results

Findings are arranged by relevant PRISM domain (intervention characteristics, recipients via both patient and organizational perspectives, external environment, and implementation and sustainability infrastructure).

Recipient Perspectives and Organizational Characteristics

Of the nine HCV clinics, seven were in the northeastern United States (including the two comparison clinics), and two were in the western United States (one each AI and UI). In total, fifteen providers across the intervention clinics (seven in UI and eight in AI) were trained to use the aTS and completed a baseline demographic questionnaire. Ten of these providers eventually enrolled patients and completed a follow-up questionnaire (five providers were unable to enroll any patients). At baseline, clinic and provider characteristics were balanced on age, sex, and technology experience (data not shown) across AI, UI, and comparison clinics. Provider surveys indicated there were no differences between clinics in the two implementation arms on readiness to implement the aTS, including on measures of perceived evidence strength for texting, organizational context, and implementation climate (data not shown). These surveys indicated there were, however, differences in satisfaction with clinic HCV treatment practices: 100% (7/7) of UI compared with 50% (4/8) of AI providers were satisfied with their local HCV treatment processes ($P=.04$).

External Environment

Providers were generally eager to support HCV treatment with the aTS because improving HCV treatment was a national and local VA priority. One UI provider mentioned that their clinic had already been considering creating a texting reminder system:

> *We're actually really excited to use it [the aTS] with our patients. This is something that we had talked about doing or developing something like this to see what the impact could be on improving adherence to appointments, adherence to medications for patients.*

Another provider lamented:

> *I wish this [the aTS] could have come earlier.*

Implementation Facilitation Processes

Figure 2 summarizes email, phone, and in-person facilitation events (n=293) across the seven intervention clinics. Facilitation effort was relatively modest, initially, as new clinics were adopting and learning the aTS, rising to the busiest period in months 4 to 7 and tapering off in months 8 to 10. AI clinics had an average of 39 facilitation events compared with 45 for UI clinics ($P=.17$), or about weekly contact. Only 10% of providers reported (via postimplementation survey) that they could have implemented the aTS without facilitation. The association between facilitation dose and provider-initiated aTS recruitment was positive and linear, although not statistically significant ($r=.71$, $P=.07$).

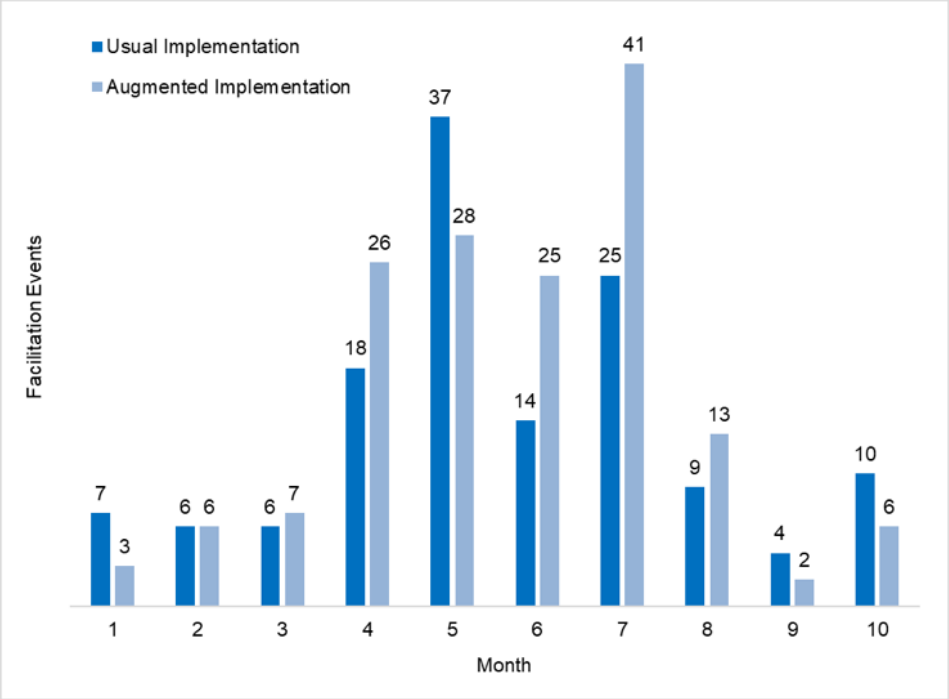
Providers across implementation arms had largely positive feedback about the facilitation received to support aTS implementation. One provider explained the value of an accessible facilitator:

> *...whenever you're using new technology and new approaches with the technology component, it's just good to have somebody that you can, who's very responsive...and can find out the answer for you in a timely fashion.*

In the case of AI clinics, providers highlighted the value of the in-person site visit because "when [the facilitator] came it kind of clinched it," suggesting the one-on-one visit helped providers make the decision to use the aTS and provided an important opportunity to ask clarifying questions and cement more of the technical and logistical aspects of implementation. AI clinic providers had mixed impressions of the toolkit. Providers felt it provided needed information and was easy to understand; however, some felt it was too long and overly dense. Several suggested that an abridged quick start guide would have been more useful.

XSL•FO
RenderX

Figure 2. Facilitation events over time by clinic and implementation arm.



Implementation Outcomes

Implementation outcomes, UI versus AI, are depicted in the retention diagram in Figure 3, which shows the percentage of patients approached to participate in the aTS who progressed from one step to the next. Across the seven aTS intervention clinics, a total of 625 patients started HCV treatment during the implementation period. Providers comparably offered the aTS to 33% (92/278) of patients at UI sites and 30% (105/347) of patients at AI sites who were starting HCV treatment (*P*=.45). Notably, UI sites did better in the registration and authentication steps, and AI sites did better in reaching the final step of aTS message interaction. By implementation group there was a significant difference in patient registration (42/278, 15%, UI vs 29/347, 8%, AI; *P*=.01) and a borderline significant difference in patient authentication (26/278, 9%, UI vs 24/347, 7%, AI; *P*=.06). However, compared to UI site patients, a greater percentage of AI site patients who authenticated their phone numbers went on to actively respond to HCV messages (11/278, 4%, UI vs 21/347, 6%, AI; *P*<.001). Of 35 patients (14 UI vs 21 AI) receiving medication reminder text messages, 91% (32/35) replied to at least one medication reminder. The mean medication reminder response rate was 78% (SD 26%; median 89%), with no difference between groups (77% UI vs 79% AI; *P*=.87).

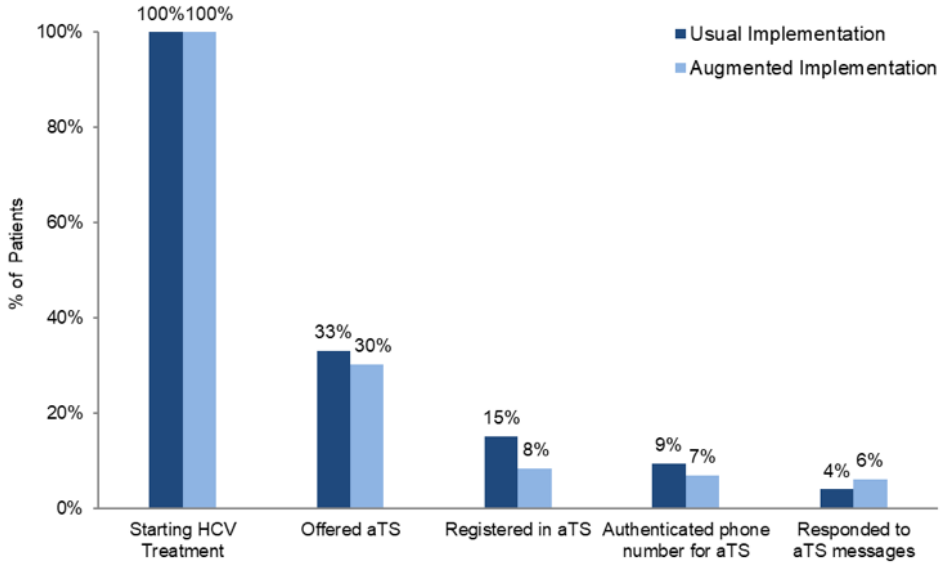
Qualitative interviews with providers indicated that they were often choosing to offer the aTS to younger and clinically less complex candidate patients who they perceived to be more technologically savvy and thus more likely to agree to use the system. In correlation analysis, the number of patients to whom the aTS was offered was inversely associated with providers' baseline satisfaction with HCV care process (*r*=–0.65, *P*=.06) and their length of time working in the VA health care system (*r*=–0.78, *P*=.01). Nevertheless, all providers were surprised that a high percentage of patients declined the aTS offer:

> *I thought a lot of people are going to be able to participate, but I guess when we started offering them [the text messages] some patients don't, I guess they're not used to it, so they, most of the patients that I offer to decline to participate.*

Patient reasons for declining the aTS fell into four categories: general disinterest (48%), texting apprehension, including cost concerns (30%), beliefs of already being good at medication adherence (12%), or beliefs that texting would duplicate other self-management approaches (10%).

XSL•FO
RenderX

Figure 3. Texting engagement by implementation and facilitation arm. HCV: hepatitis C virus; aTS: automated texting system.



Patient Characteristics

There were few demographic differences between texters and nontexters (Table 1), although texters were more likely to be black (13/15, 87%, vs 6/15, 40%; *P*=.02) and less likely to have a self-reported mental health or substance use disorder (9/15, 60%, vs 14/15, 93%; *P*=.03).

Clinical Effectiveness Outcomes

On the follow-up survey (Table 2), more texters versus nontexters reported excellent health (4/15, 27%, vs 0/15, 0%; *P*=.01) and greater ability to prevent or reduce problems associated with their health (9/15, 60%, vs 2/15, 13%; *P*=.03). A greater proportion of texters compared with nontexters had overall less negative HCV illness perception and reported fewer concerns about failing HCV treatment (13/15, 87%, vs 8/15, 53%, respectively; *P*=.05). Texters reported better or about equal HCV treatment adherence than nontexters, including having higher perceptions of their own HCV adherence than did nontexters (11/15, 73%, reporting excellent adherence vs 6/15, 40%; *P*=.06). For patients for whom lab results were available, 96% (27/28) of texters had achieved SVR compared with 94% (153/163) of nontexters (*P*>.99), reflecting the high cure rates possible with current HCV treatments and our small sample size.

XSL•FO
RenderX

Table 1. Patient characteristics by automated test messaging system use at follow-up (n=30).

Characteristic	Nontexters[a] (n=15)	Texters[b] (n=15)	*P* value
Age in years, mean (SD)	62 (11)	62 (5)	.56
Gender (male), n (%)	14 (94)	14 (94)	>.99
Race, n (%)			
African American or black	6 (40)	13 (87)	—
White	8 (53)	2 (13)	—
Other	1 (7)	0 (0)	—
Hispanic/Latino, n (%)	1 (7)	0 (0)	>.99
Marital status, n (%)			.93
Single/never married	5 (33)	4 (27)	—
Married/in a relationship	2 (13)	2 (13)	—
Divorced/separated	7 (47)	7 (47)	—
Widowed	1 (7)	2 (13)	—
Employment, currently working, n (%)	2 (13)	3 (20)	.38
Income, n (%)			.25
>$10,000	5 (33)	2 (13)	—
$10,000 to $20,000	5 (33)	4 (27)	—
$20,000 to $40,000	4 (27)	4 (27)	—
<$40,000	1 (7)	5 (33)	—
Education, n (%)			.27
High school or less	9 (60)	5 (33)	—
Some college or more	6 (40)	10 (67)	—
Housing, n (%)			.10
Own apartment or house	7 (47)	11 (73)	—
Hospital, domiciliary, shelter, street, drug treatment center	6 (40)	1 (7)	—
With friend or relative	2 (13)	3 (20)	—
Self-reported mental health or substance use disorder, n (%)	14 (93)	9 (60)	.03
Social/emotional support			.61
Always/usually	3 (20)	8 (53)	—
Sometime/rarely	12 (80)	7 (47)	—
Texting history, n (%)			—
Unlimited texting plan	10 (67)	11 (73)	.83
Daily texting	8 (53)	12 (80)	.27

[a]Nontexters: patients who never received any text messages (could be from usual or augmented implementation clinics or comparison clinics).

[b]Texters: patients who used the automated texting system regardless of whether they were in usual or augmented implementation clinics.

XSL•FO
RenderX

Table 2. Patient self-reported outcomes by texting group.

Characteristics	Nontexters[a] (n=15)	Texters[b] (n=15)	*P* value
Health status			
In general, how would you rate your health (excellent)?, n (%)	0 (0)	4 (27)	.01
For how many days during the past 30 days..., n			
was your physical health not good?	10	5	.08
was your mental health not good?	13	8	.23
did poor physical or mental health keep you from doing your usual activities?	10	5	.25
Patient activation (strongly agree), n (%)			
I am responsible for taking care of my health.	10 (67)	12 (80)	.61
I am able to prevent or reduce problems associated with my health.	2 (13)	9 (60)	.03
I can follow through on recommended medical treatment.	5 (33)	10 (67)	.11
Illness perception (strongly disagree), n (%)			
Feeling angry, scared and/or depressed when I think about living with HCV[c].	4 (27)	7 (47)	.32
Feeling that HCV controls my life.	5 (33)	8 (53)	.58
Feeling overwhelmed by the demands of living with HCV.	5 (33)	9 (60)	.21
Not feeling motivated to go through HCV treatment.	7 (47)	10 (67)	.51
Feeling that I am failing with my HCV treatment.	8 (53)	13 (87)	.05
HCV treatment behavior (strongly disagree), n (%)			
I forget to take my HCV medicine(s).	11 (73)	12 (80)	>.99
I decide not to take my HCV medicine(s).	11 (73)	14 (93)	.26
I forget to get my HCV prescription(s) filled.	12 (80)	11 (73)	.52
I tend to forget to get my HCV lab and blood work done.	7 (47)	13 (87)	.07
I run out of my HCV medicine(s).	11 (73)	12 (80)	>.99
I tend to miss my doctors' appointments, n (%)	5 (33)	11 (73)	.10
Self-report HCV medication adherence			
Percentage of HCV medication taken correctly in last 4 weeks, mean (SD)	95 (13)	99 (2)	.20
Ability to take HCV medication as prescribed (excellent), n (%)	6 (40)	11 (73)	.06

[a]Nontexters: patients who never received any text messages (could be from usual or augmented implementation clinics or comparison clinics).

[b]Texters: patients who used the automated texting system regardless of whether they were in usual or augmented implementation clinics.

[c]HCV: hepatitis C virus.

Intervention Usability, Workflow, and Value

Usefulness and usability of the aTS was assessed with patients and providers. Among those patients using the aTS (texters), 15 completed a follow-up survey and 13 also participated in a semistructured interview. Another 15 patients who did not use the aTS (nontexters) completed a follow-up survey. Among patients, there were no differences by implementation arm (AI vs UI) on patient measures of usability, usefulness, and degree of working alliance with the aTS (data not shown). Texters reported mostly positive sentiments about interacting with the aTS, saying it was easy, simple, and "not rocket science." Still, some patients struggled with the aTS authentication syntax ("Start") due to capitalization and punctuation errors (eg, "START" and "start."). Most (12/15, 80%) preferred text reminders to appointment reminders delivered by mail or phone. About half (8/15, 53%) reported the cost of text messaging could be a barrier to use, although at least 73% (11/15) had an unlimited texting plan. In patients' opinions, the aTS had benefits in helping stay engaged in care, stay connected to their health care team, and assisting with HCV medication adherence (all 100% positive endorsements). As one patient recognized: "I wouldn't have been as efficient or effective without some assistance" and it "[gave] me encouragement for doing what I was supposed to." Most patients (13/15, 87%) were on multiple medications and viewed the aTS as supporting their overall medication-taking routine: "I use [the aTS] for other medications as well...I just group them all in together." Providers corroborated patient feedback by observing that the aTS "helps relieve a lot of [patient] anxiety about missing a med."

Provider perspectives on usefulness and usability were captured via follow-up questionnaires and semistructured interviews. Most providers (7/10, 70%) logged into the aTS at least weekly

XSL•FO
RenderX

to either assign a protocol to a patient or monitor messages in the aTS dashboard. Some providers tailored protocols according to patient preferences, while other providers opted to retain default protocol settings to streamline the enrollment process. Some indicated that the message history data could be made more usable.

Even with training and ongoing support, providers had startup difficulties with the aTS:

> *I mean it took me a little while to familiarize myself because the training versus actually doing it yourself, you know, there's a learning curve...*

Half of providers (5/10, 50%) felt that enrolling patients was difficult, and half also felt the aTS did not easily integrate into clinical workflow. One third (3/10, 30%) said using the aTS added a lot of work to their workday. Providers described the aTS as a little bulky, a little cumbersome, and labor intensive. To enhance uptake of the aTS, providers recommended the system become more streamlined and more intuitive, particularly at the stage of registering patients. Providers also suggested integrating the aTS with other VA technologies, including registration kiosks, the electronic medical record, and patient portal (MyHealtheVet).

Although providers saw the aTS as a potential benefit to their practice, sometimes they could not accommodate the additional time to educate and enroll patients. A provider commented that they "underappreciated the coaching that the patients require at the time of enrolling." There was a tendency for providers to view the enrollment process as not in their scope of work and as a task more suited for nurses or support staff, suggesting that there be, "someone [nonprovider] assigned to assist with Annie."

Notwithstanding difficulties incorporating the aTS into established clinical practices, providers believed the aTS was welcomed by patients (7/10, 70%), could enable clinics to meet patient-centered care goals (9/10, 90%), and could lead to cost savings for the VA (6/10, 60%). Most providers intended to continue using the aTS for HCV (9/10, 90%) and would use it for other health conditions after study completion (8/10, 80%).

Discussion

Principal Findings

To our knowledge, this is this first randomized evaluation of implementation facilitation strategies intended to increase adoption of VA's aTS. We found that patients and providers largely accepted and deemed the aTS appropriate, easy to use, and useful, albeit with substantial barriers to uptake and sustained use. Greater patient recruitment by providers was associated with more facilitation and lower baseline satisfaction with clinic HCV care processes. We found differences in implementation outcomes. For some of the stages of aTS implementation, patient engagement was higher in AI (vs UI). Overall, 1 out of 6 patients (16%) who were offered the aTS received aTS text messages. Texters, compared with nontexters, felt more connected to their care team, confident in their HCV medication taking, and more activated for self-management. There were no differences, however, in clinical effectiveness outcomes between patients based on use of the aTS system (texters vs nontexters). Our results suggest that the aTS may not be easy to implement but has better chances of success if several PRISM domains are attended to at the adoption, implementation, and sustainability phases.

Augmented Implementation Facilitation

Within the implementation and sustainability infrastructure domain, a need for facilitated implementation was identified. Considerable effort was needed to assist all clinics with aTS adoption. Once implementation began, the aTS called for episodic technical assistance rather than high-intensity, sustained facilitation [35]. More facilitator-provider interaction appeared to be related to higher aTS recruitment, as others have demonstrated [36]. We also found that facilitation may have replaced the need for a toolkit in this study because the facilitator was highly accessible to providers. There is mounting evidence that toolkits and manuals tend to be underused when other implementation strategies are available [37,38]. This may be due to toolkit development being unstandardized and thus highly variable. For this reason, Hempel and colleagues [39] recently provided recommendations for the content, development, and evaluation of quality improvement toolkits. Of the implementation facilitation meta-strategies in this study, it appeared site visits had a strong influence on implementation, suggesting that the interpersonal component of a facilitator-provider dyad is paramount to successful implementation.

Our results also indicate that despite differences in planned facilitation approaches, there were no differences in the dose of facilitation delivered, suggesting facilitation efficiency is enhanced when it is delivered in person or is channeled through a champion (as with augmented implementation sites) [40]. Our findings are consistent with current literature that facilitator effort tapers once clinics have commenced implementation [37,41]. Facilitation, in many ways, is not formulaic and is inherently dependent on local context and need. Notably, in our work, a common implementation strategy was not included—there was no explicit benchmarking or audit and feedback component to the facilitation efforts. More guidelines on how to gauge specific facilitation need (ie, activities, dose, intensity, timing of start, and removal) are necessary, as are ways to track, evaluate, and replicate with fidelity.

Assessing Readiness to Implement

Taking an organizational perspective to understand both the intervention and its recipients illuminates several important factors that can influence aTS uptake. Organizational readiness between clinics differed in only one area: satisfaction with local HCV care practices. Providers who were unsatisfied with their HCV care practices may have perceived their clinics as needing improvement and thus were more likely to embrace the aTS. As such, some clinics were primed for the introduction of a new practice, despite no mandate or clinical practice guideline motivation. Once providers began implementing, however, several believed the aTS did not integrate well into their workflow and was a more fitting task for nonclinicians. Insights from organizational readiness assessments tend to be underused but may help understand facilitation mechanisms of action,

XSL•FO
RenderX

which likewise remain poorly understood and often unmeasured [42].

We also found variation in the assumptions made by providers regarding patient candidacy for aTS recruitment. UI clinics registered and enrolled more patients; however, AI clinic patients authenticated their phone numbers at higher rates thus triggering the text messages to start. Providers at AI clinics may have been targeting patients more selectively, thus making their recruitment more efficient. While potentially successful at their AI clinics, in general, providers should be wary of the validity of their selection heuristics, which studies indicate are often unreliable [43]. It may be more appropriate to offer the aTS universally to patients rather than selectively recruiting those deemed more apt to agree, thereby reducing proficient user bias. Future studies could explore alternative methods for introducing the aTS and consenting and registering patients, such as through patient opt-out and self-enrollment approaches. Since study completion, adjustments to the aTS registration procedure now allow for support staff to register patients in the aTS portal on behalf of a consenting clinician.

Patient Behavior Change

There were several facilitating patient perspective elements within the intervention and recipients domains important to aTS use and sustainability. There was near universal positive feedback from patients about the ease of use and benefits of the aTS, however there were notable drop-offs in engagement during phone number authentication due to syntax errors and possible changes in willingness to use. After interacting with the aTS, texters compared with nontexters reported feeling more activated for self-management and adherence to medication and improved health status, but no significant differences in clinical outcomes were detected. As the transtheoretical model posits and as the aTS begins to show, attitudinal changes precede behavior change and may produce meaningful behavior change [44]. Because polypharmacy among veterans is common, additional work is required to understand the types of patients most likely to benefit from texting interventions [45-47]. There may be other important patient moderating variables, such as age and rurality, that determine whether individuals choose to adopt the aTS [48]. While communication preferences are shifting in favor of texting and other virtual modalities, slower-than-anticipated uptake of technology tools for health care is common [49]. Nonetheless, texting differs from other health technologies in that it reaches patients where they are, at any time, and can accept responses whenever patients are ready to offer them. In depth qualitative work is needed to understand barriers to implementation and sustained use of the aTS and how to redesign the enrollment and engagement process.

Strengths and Limitations

This was the first study of VA's aTS. Strengths of this study are that PRISM was used to guide the evaluation, it was a randomized design, and we simultaneously studied implementation and effectiveness outcomes using mixed methods. We believe this study makes an important contribution to advancing the implementation science of texting interventions.

Our study has several potential limitations. First, our study was conducted within the VA and for the treatment of HCV; therefore, not all settings or health behaviors were represented, which limits our generalizability. Second, we selected clinics that expressed interest in using the aTS, raising the potential for selection bias. Third, the sicker, more socially and psychologically vulnerable patients may not have been invited by providers to participate in the aTS because of concerns that it could be confusing or costly for them. This selection bias may explain some of the apparent beneficial findings for texters. Also, feedback on the aTS may have been subject to social desirability bias. We did not adjust for multiple comparisons due to the explanatory nature of the study.

Conclusions

Increasingly, health care systems are using technology to meet patient expectations for electronic transactions, information exchange, and on-demand access to providers. This was the first study to examine the implementation and effectiveness of an automated text messaging system in the VA generally and for HCV treatment specifically. Despite positive perceptions of the aTS, patient enrollment was challenging; however, augmented facilitation resulted in greater sustained engagement of patients once they enrolled. Importantly, among patients who used the aTS (the texters) there was an indication of improved illness perception, health engagement, and patient activation. Our results suggest that the aTS can serve as an adjunct tool to usual HCV care, provided it is appropriately integrated into clinical workflow. Our study has implications for health care systems making efforts to engage patients beyond episodic in-person visits through patient-facing technologies. Findings suggest that a large pool of potential texting adopters have yet to realize benefits from this technology. While novel technologies such as the aTS have considerable potential, they also present distinct behavioral, social, and technical challenges for implementation and scale-up.

Acknowledgments

We would like to thank Drs Neil Evans and Kathy Frisbee from the Office of Connected Care for their support of this work and Mary Lou Glazer, Ralph Strenglein, and Brian Vetter for their contributions to training on the aTS and development our HCV texting protocol. This study was supported by the US Veterans Health Administration, Office of Connected Care, and Quality Enhancement Research Initiative Program (grant #PEC 15-470). The views expressed in this article are those of the authors and do not necessarily reflect the position and/or policy of the VA or the US Government. The data used for this research are available from the corresponding author on reasonable request and subject to VA guidelines.

XSL•FO
RenderX

Authors' Contributions

TPH, TKH, and KM conceived the study and obtained funding. TPH, TKH, DKM, LR, and VY helped to conceptualize the study design and data collection tools. VY conducted the analyses. All authors performed significant editing of the manuscript and read and approved the final manuscript.

Conflicts of Interest

None declared.

Editorial notice: This randomized study was not prospectively registered. The editor granted an exception of ICMJE rules for prospective registration of randomized trials because the risk of bias appears low. However, readers are advised to carefully assess the validity of any potential explicit or implicit claims related to their primary outcomes or effectiveness, as the lack of registration means that authors could change their outcome measures retrospectively. Best practice is to register trials and/or publish protocols prospectively.

References

1. Fjeldsoe BS, Marshall AL, Miller YD. Behavior change interventions delivered by mobile telephone short-message service. Am J Prev Med 2009 Feb;36(2):165-173. [doi: 10.1016/j.amepre.2008.09.040] [Medline: 19135907]
2. Hall AK, Cole-Lewis H, Bernhardt JM. Mobile text messaging for health: a systematic review of reviews. Annu Rev Public Health 2015 Mar 18;36:393-415 [FREE Full text] [doi: 10.1146/annurev-publhealth-031914-122855] [Medline: 25785892]
3. 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]
4. Krishna S, Boren SA, Balas EA. Healthcare via cell phones: a systematic review. Telemed J E Health 2009 Apr;15(3):231-240. [doi: 10.1089/tmj.2008.0099] [Medline: 19382860]
5. 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]
6. Kirchner JE. Department of Veterans Affairs. The use of facilitation as an implementation strategy URL: https://www.hsrd.research.va.gov/for_researchers/cyber_seminars/archives/1233-notes.pdf [accessed 2019-07-19]
7. Berta W, Cranley L, Dearing JW, Dogherty EJ, Squires JE, Estabrooks CA. Why (we think) facilitation works: insights from organizational learning theory. Implement Sci 2015 Oct 06;10(1):1 [FREE Full text] [doi: 10.1186/s13012-015-0323-0] [Medline: 26443999]
8. Ritchie MJ, Dollar KM, Kearney LK, Kirchner JE. Research and services partnerships: responding to needs of clinical operations partners: transferring implementation facilitation knowledge and skills. Psychiatr Serv 2014 Feb 01;65(2):141-143. [doi: 10.1176/appi.ps.201300468] [Medline: 24492898]
9. Lessard S, Bareil C, Lalonde L, Duhamel F, Hudon E, Goudreau J, et al. External facilitators and interprofessional facilitation teams: a qualitative study of their roles in supporting practice change. Implement Sci 2016 Jul 16;11:97 [FREE Full text] [doi: 10.1186/s13012-016-0458-7] [Medline: 27424171]
10. Stetler CB, Legro MW, Rycroft-Malone J, Bowman C, Curran G, Guihan M, et al. Role of “external facilitation” in implementation of research findings: a qualitative evaluation of facilitation experiences in the Veterans Health Administration. Implement Sci 2006 Oct 18;1:23 [FREE Full text] [doi: 10.1186/1748-5908-1-23] [Medline: 17049080]
11. Cottrell E, Chambers R, O'Connell P. Using simple telehealth in primary care to reduce blood pressure: a service evaluation. BMJ Open 2012;2(6):1 [FREE Full text] [doi: 10.1136/bmjopen-2012-001391] [Medline: 23117563]
12. McInnes DK, Houston TK, Woods SS, Frisbee KL, Evans NC. Disseminating information systems across the Atlantic: collaboration between U.K. National Health Service and U.S. Department of Veterans Affairs. Implementation Sci 2015 Aug 14;10(S1):A62. [doi: 10.1186/1748-5908-10-S1-A62]
13. Mathes T, Antoine S, Pieper D. Factors influencing adherence in Hepatitis-C infected patients: a systematic review. BMC Infect Dis 2014 Apr 15;14:203 [FREE Full text] [doi: 10.1186/1471-2334-14-203] [Medline: 24731285]
14. Pawlotsky J. Hepatitis C virus resistance to direct-acting antiviral drugs in interferon-free regimens. Gastroenterology 2016 Dec;151(1):70-86. [doi: 10.1053/j.gastro.2016.04.003] [Medline: 27080301]
15. Curran GM, Bauer M, Mittman B, Pyne JM, Stetler C. Effectiveness-implementation hybrid designs: combining elements of clinical effectiveness and implementation research to enhance public health impact. Med Care 2012 Mar;50(3):217-226 [FREE Full text] [doi: 10.1097/MLR.0b013e3182408812] [Medline: 22310560]
16. VA Office of Research & Development. 2011. VHA Handbook 1058.05, VHA Operations Activities That May Constitute Research URL: https://www.research.va.gov/resources/policies/oro-120811.cfm [accessed 2019-07-19]
17. Efird J. Blocked randomization with randomly selected block sizes. Int J Environ Res Public Health 2011 Jan;8(1):15-20 [FREE Full text] [doi: 10.3390/ijerph8010015] [Medline: 21318011]
18. Veterans Health Administration. VHA Facility Complexity Model 2017 URL: https://www.vendorportal.ecms.va.gov/FBODocumentServer/DocumentServer.aspx?DocumentId=2793591&FileName=VA118-16-R-1059-A00002002.docx [accessed 2019-07-26]

XSL•FO
RenderX

19. Story M. Maximizing usability: the principles of universal design. Assistive Technol 1998;10(1):4-12. [Medline: 10181150]
20. Powell BJ, Waltz TJ, Chinman MJ, Damschroder LJ, Smith JL, Matthieu MM, et al. A refined compilation of implementation strategies: results from the Expert Recommendations for Implementing Change (ERIC) project. Implement Sci 2015 Feb 12;10:21 [FREE Full text] [doi: 10.1186/s13012-015-0209-1] [Medline: 25889199]
21. Rogal SS, Yakovchenko V, Waltz TJ, Powell BJ, Kirchner JE, Proctor EK, et al. The association between implementation strategy use and the uptake of hepatitis C treatment in a national sample. Implement Sci 2017 Dec 11;12(1):60 [FREE Full text] [doi: 10.1186/s13012-017-0588-6] [Medline: 28494811]
22. Feldstein AC, Glasgow RE. A practical, robust implementation and sustainability model (PRISM) for integrating research findings into practice. Jt Comm J Qual Patient Saf 2008 Apr;34(4):228-243. [Medline: 18468362]
23. Haynes RB. A critical review of determinants of patient compliance with therapeutic regimens. In: Sackett DL, Haynes RB, editors. Compliance with Therapeutic Regimens. Baltimore: Johns Hopkins University Press; 1976:26-39.
24. Fisher L, Glasgow RE, Mullan JT, Skaff MM, Polonsky WH. Development of a brief diabetes distress screening instrument. Ann Fam Med 2008;6(3):246-252 [FREE Full text] [doi: 10.1370/afm.842] [Medline: 18474888]
25. Hays RD, Bjorner JB, Revicki DA, Spritzer KL, Cella D. Development of physical and mental health summary scores from the patient-reported outcomes measurement information system (PROMIS) global items. Qual Life Res 2009 Sep;18(7):873-880 [FREE Full text] [doi: 10.1007/s11136-009-9496-9] [Medline: 19543809]
26. Behavioral Risk Factor Surveillance System Survey Questionnaire 2018. Atlanta: US Department of Health and Human Services, Centers for Disease Control and Prevention URL: https://www.cdc.gov/brfss/questionnaires/pdf-ques/2018_BRFSS_English_Questionnaire.pdf [accessed 2019-07-19]
27. Mayberry LS, Gonzalez JS, Wallston KA, Kripalani S, Osborn CY. The ARMS-D out performs the SDSCA, but both are reliable, valid, and predict glycemic control. Diabetes Res Clin Pract 2013 Nov;102(2):96-104 [FREE Full text] [doi: 10.1016/j.diabres.2013.09.010] [Medline: 24209600]
28. Davis FD. User acceptance of information technology: system characteristics, user perceptions and behavioral impacts. Int J Man-Machine Stud 1993 Mar;38(3):475-487. [doi: 10.1006/imms.1993.1022]
29. Horvath AO, Greenberg LS. Development and validation of the Working Alliance Inventory. J Counseling Psychol 1989;36(2):223-233. [doi: 10.1037/0022-0167.36.2.223]
30. Hannon PA, Helfrich CD, Chan KG, Allen CL, Hammerback K, Kohn MJ, et al. Development and pilot test of the workplace readiness questionnaire, a theory-based instrument to measure small workplaces' readiness to implement wellness programs. Am J Health Promot 2017 Jan;31(1):67-75 [FREE Full text] [doi: 10.4278/ajhp.141204-QUAN-604] [Medline: 26389975]
31. Jacobs SR, Weiner BJ, Bunger AC. Context matters: measuring implementation climate among individuals and groups. Implement Sci 2014 Apr 17;9:46 [FREE Full text] [doi: 10.1186/1748-5908-9-46] [Medline: 24742308]
32. Wen K, Gustafson DH, Hawkins RP, Brennan PF, Dinauer S, Johnson PR, et al. Developing and validating a model to predict the success of an IHCS implementation: the Readiness for Implementation Model. J Am Med Inform Assoc 2010;17(6):707-713 [FREE Full text] [doi: 10.1136/jamia.2010.005546] [Medline: 20962135]
33. Brooke J. SUS—a quick and dirty usability scale. Usability Evaluation in Industry 1996;189(194):4-7.
34. Fereday J, Muir-Cochrane E. Demonstrating rigor using thematic analysis: a hybrid approach of inductive and deductive coding and theme development. Int J Qual Methods 2016 Nov 29;5(1):80-92. [doi: 10.1177/160940690600500107]
35. Harvey G, Loftus-Hills A, Rycroft-Malone J, Titchen A, Kitson A, McCormack B, et al. Getting evidence into practice: the role and function of facilitation. J Adv Nurs 2002 Mar;37(6):577-588. [Medline: 11879422]
36. Baskerville NB, Liddy C, Hogg W. Systematic review and meta-analysis of practice facilitation within primary care settings. Ann Fam Med 2012;10(1):63-74 [FREE Full text] [doi: 10.1370/afm.1312] [Medline: 22230833]
37. Davis MM, Howk S, Spurlock M, McGinnis PB, Cohen DJ, Fagnan LJ. A qualitative study of clinic and community member perspectives on intervention toolkits: unless the toolkit is used it won't help solve the problem. BMC Health Serv Res 2017 Dec 18;17(1):497 [FREE Full text] [doi: 10.1186/s12913-017-2413-y] [Medline: 28720092]
38. Gagliardi AR, Brouwers MC. Integrating guideline development and implementation: analysis of guideline development manual instructions for generating implementation advice. Implement Sci 2012 Jul 23;7:67 [FREE Full text] [doi: 10.1186/1748-5908-7-67] [Medline: 22824094]
39. Hempel S, Miake-Lye I, Brega AG, Buckhold F, Hassell S, Nowalk MP, et al. Quality improvement toolkits: recommendations for development. Am J Med Qual 2019 Jan 24:1062860618822102. [doi: 10.1177/1062860618822102] [Medline: 30675797]
40. Miech EJ, Rattray NA, Flanagan ME, Damschroder L, Schmid AA, Damush TM. Inside help: an integrative review of champions in healthcare-related implementation. SAGE Open Med 2018;6:2050312118773261 [FREE Full text] [doi: 10.1177/2050312118773261] [Medline: 29796266]
41. Kirchner JE, Ritchie MJ, Pitcock JA, Parker LE, Curran GM, Fortney JC. Outcomes of a partnered facilitation strategy to implement primary care-mental health. J Gen Intern Med 2014 Dec;29 Suppl 4:904-912 [FREE Full text] [doi: 10.1007/s11606-014-3027-2] [Medline: 25355087]
42. Helfrich CD, Blevins D, Smith JL, Kelly PA, Hogan TP, Hagedorn H, et al. Predicting implementation from organizational readiness for change: a study protocol. Implement Sci 2011 Jul 22;6:76 [FREE Full text] [doi: 10.1186/1748-5908-6-76] [Medline: 21777479]

XSL•FO
RenderX

43. McHorney CA. The Adherence Estimator: a brief, proximal screener for patient propensity to adhere to prescription medications for chronic disease. Curr Med Res Opin 2009 Jan;25(1):215-238. [doi: 10.1185/03007990802619425] [Medline: 19210154]
44. Prochaska JO, Velicer WF. The transtheoretical model of health behavior change. Am J Health Promot 1997;12(1):38-48. [Medline: 10170434]
45. Witry M, Klein D, Alexander B, Franciscus C, Turvey C. Medication list discrepancies and therapeutic duplications among dual use veterans. Fed Pract 2016 Sep;33(9):14-20 [FREE Full text] [Medline: 27840570]
46. Rossi M, Feske U, Shaffer S, Kreutzer J. Polypharmacy e-consults: a method to assess medications and address polypharmacy in rural patients. Innov Aging 2017 Jun 30;1(suppl_1):1156. [doi: 10.1093/geroni/igx004.4218]
47. Nguyen KA, Haggstrom DA, Ofner S, Perkins SM, French DD, Myers LJ, et al. Medication use among veterans across health care systems. Appl Clin Inform 2017 Dec 08;8(1):235-249 [FREE Full text] [doi: 10.4338/ACI-2016-10-RA-0184] [Medline: 28271121]
48. Connolly SL, Miller CJ, Koenig CJ, Zamora KA, Wright PB, Stanley RL, et al. Veterans' attitudes toward smartphone app use for mental health care: qualitative study of rurality and age differences. JMIR Mhealth Uhealth 2018 Aug 22;6(8):e10748 [FREE Full text] [doi: 10.2196/10748] [Medline: 30135050]
49. Health information technology: HHS should assess the effectiveness of its efforts to enhance patient access to and use of electronic health information (GAO-17-305). Washington: US Government Accountability Office; 2017 Mar 15. URL: https://www.gao.gov/assets/690/683388.pdf [accessed 2019-07-19]

Abbreviations

AI: augmented implementation
aTS: automated text messaging system
HCV: hepatitis C virus
NHS: National Health Service
PRISM: Practical, Robust Implementation and Sustainability Model
SMS: short message service
SVR: sustained virologic response
UI: usual implementation
VA: Department of Veterans Affairs

Edited by G Eysenbach; submitted 21.05.19; peer-reviewed by M Muldoon, L Garvin, E Da Silva; comments to author 18.06.19; revised version received 12.07.19; accepted 19.07.19; published 04.08.19.

Please cite as:
Yakovchenko V, Hogan TP, Houston TK, Richardson L, Lipschitz J, Petrakis BA, Gillespie C, McInnes DK
Automated Text Messaging With Patients in Department of Veterans Affairs Specialty Clinics: Cluster Randomized Trial
J Med Internet Res 2019;21(8):e14750
URL: https://www.jmir.org/2019/8/e14750/
doi:10.2196/14750
PMID:31444872

©Vera Yakovchenko, Timothy P Hogan, Thomas K Houston, Lorilei Richardson, Jessica Lipschitz, Beth Ann Petrakis, Chris Gillespie, D Keith McInnes. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 04.08.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Adherence and Satisfaction of Smartphone- and Smartwatch-Based Remote Active Testing and Passive Monitoring in People With Multiple Sclerosis: Nonrandomized Interventional Feasibility Study

Luciana Midaglia[1,2], MD; Patricia Mulero[1], MD; Xavier Montalban[1,3], MD, PhD; Jennifer Graves[4], MAS, MD, PhD; Stephen L Hauser[5], MD; Laura Julian[6], PhD; Michael Baker[7], MSc; Jan Schadrack[7], MD; Christian Gossens[7], MBA, PhD; Alf Scotland[7], MSc; Florian Lipsmeier[7], PhD; Johan van Beek[7], PhD; Corrado Bernasconi[7], MD, PhD; Shibeshih Belachew[7], MD, PhD; Michael Lindemann[7,8], MBA, PhD

[1]Department of Neurology-Neuroimmunology, Multiple Sclerosis Centre of Catalonia, Vall d'Hebron University Hospital, Barcelona, Spain

[2]Department of Medicine, Autonomous University of Barcelona, Barcelona, Spain

[3]Division of Neurology, University of Toronto, Toronto, ON, Canada

[4]Department of Neurology, University of California, San Diego, San Diego, CA, United States

[5]Department of Neurology, University of California, San Francisco, San Francisco, CA, United States

[6]Genentech Inc, South San Francisco, CA, United States

[7]F Hoffmann–La Roche Ltd, Basel, Switzerland

[8]Department of Economics, Baden-Wuerttemberg Cooperative State University, Loerrach, Germany

Corresponding Author:
Christian Gossens, MBA, PhD
F Hoffmann–La Roche Ltd
124 Grenzacherstrasse
Basel,
Switzerland
Phone: 41 61 687 5113
Fax: 41 61 691 9391
Email: christian.gossens@roche.com

Related Article:

This is a corrected version. See correction statement: https://www.jmir.org/2019/10/e16287

Abstract

Background: Current clinical assessments of people with multiple sclerosis are episodic and may miss critical features of functional fluctuations between visits.

Objective: The goal of the research was to assess the feasibility of remote active testing and passive monitoring using smartphones and smartwatch technology in people with multiple sclerosis with respect to adherence and satisfaction with the FLOODLIGHT test battery.

Methods: People with multiple sclerosis (aged 20 to 57 years; Expanded Disability Status Scale 0-5.5; n=76) and healthy controls (n=25) performed the FLOODLIGHT test battery, comprising active tests (daily, weekly, every two weeks, or on demand) and passive monitoring (sensor-based gait and mobility) for 24 weeks using a smartphone and smartwatch. The primary analysis assessed adherence (proportion of weeks with at least 3 days of completed testing and 4 hours per day passive monitoring) and questionnaire-based satisfaction. In-clinic assessments (clinical and magnetic resonance imaging) were performed.

Results: People with multiple sclerosis showed 70% (16.68/24 weeks) adherence to active tests and 79% (18.89/24 weeks) to passive monitoring; satisfaction score was on average 73.7 out of 100. Neither adherence nor satisfaction was associated with specific population characteristics. Test-battery assessments had an at least acceptable impact on daily activities in over 80% (61/72) of people with multiple sclerosis.

XSL•FO
RenderX

Conclusions: People with multiple sclerosis were engaged and satisfied with the FLOODLIGHT test battery. FLOODLIGHT sensor-based measures may enable continuous assessment of multiple sclerosis disease in clinical trials and real-world settings.

Trial Registration: ClinicalTrials.gov: NCT02952911; https://clinicaltrials.gov/ct2/show/NCT02952911

(J Med Internet Res 2019;21(8):e14863) doi:10.2196/14863

KEYWORDS

multiple sclerosis; patient adherence; patient satisfaction; smartphone; wearable electronic devices; mobile phone

Introduction

Disease progression throughout the clinical course of multiple sclerosis (MS) is measured using clinician-reported outcomes, most commonly the Expanded Disability Status Scale (EDSS) [1], magnetic resonance imaging (MRI), and patient-reported outcomes (PROs). Conventional assessment of the clinical course of MS relies on relapse-associated and periodic in-clinic visits. However, the current intermittently conducted clinic-based outcome measures in MS have limitations in studying the insidiously subtle progression in MS and may fail to comprehensively capture transient symptomatic and performance fluctuations that affect people with multiple sclerosis.

The ability of consumer wearable technology to measure functional impairment associated with various neurological disease symptoms through smartphone-based assessments is an important area of research [2-5]. Smartphones as vehicles for sensor-based technologies offer the potential for enhanced active and passive real-time data capture that may fundamentally shift traditional paradigms of clinical monitoring [6-9]. A recent large-scale study demonstrated that more than 95% of people with multiple sclerosis have access to a mobile device and most use it routinely [10]. Recently published studies have described the use of technologies in developing tools to assess people with multiple sclerosis [7,9,11-13]. In a recent study, it was reported that both healthy participants and people with multiple sclerosis were capable of completing daily tasks on a smartphone for 1 year [7]. Collection of data on a variety of cognitive and motor tests via the smartphone may represent a feasible way to gather highly granular data to accurately describe the MS disease course outside of the clinic [7].

The FLOODLIGHT study [NCT02952911] was a prospective pilot study to assess the feasibility of remote measurements using smartphones and smartwatches in people with multiple sclerosis and healthy controls (HCs). The smartwatch and smartphone contained apps that prompted the user to perform various assessments and protocols, referred to as active tests. The app also passively recorded sensor data during daily life, referred to as passive monitoring. The novel smartphone- and smartwatch-based FLOODLIGHT active tests were developed to be self-administered by people with multiple sclerosis to capture MS symptoms, including hand motor function, gait and posture, mood, and cognitive impairment.

The primary objectives of the FLOODLIGHT study, which began in November 2016, were to evaluate participant adherence to smartphone- and smartwatch-based assessments and collect feedback from people with multiple sclerosis and HCs on the smartphone and smartwatch schedule of assessments and its impact on their daily activities using a patient satisfaction questionnaire. Other objectives of the study, which will be addressed in subsequent publications, are to determine the association between exploratory sensor-based outcomes derived from the respective components of the FLOODLIGHT test battery and conventional MS clinical outcomes and explore whether the FLOODLIGHT test battery can differentiate between participants with and without MS.

Methods

Trial Design and Participants

After providing written informed consent, people with multiple sclerosis and HCs, preferentially partners or cohabitants, were evaluated for eligibility for enrollment in the FLOODLIGHT study. Eligibility criteria for people with multiple sclerosis included the ability to comply with the study protocol, age 18 to 55 years, diagnosis of MS (2010 revised McDonald criteria, treated or untreated) [14], EDSS score 0 to 5.5 (inclusive), and weight from 99 to 243 lbs (45 to 110 kg). An EDSS score of 5.5 as a maximum limit was meant to ensure any patient with relapsing or progressive MS would not have any significant difficulty in participating in the proposed testing as per study protocol. Further details on eligibility criteria are provided in Multimedia Appendix 1.

This study was conducted at two sites in two countries with a total of 76 people with multiple sclerosis and 25 HCs; 60 people with multiple sclerosis and 20 HCs were recruited from the Multiple Sclerosis Centre of Catalonia, Vall d'Hebron University Hospital, Barcelona, Spain, and 16 people with multiple sclerosis and 5 HCs were recruited from the University of California, San Francisco, California.

The protocol, informed consent forms, any information given to the participants, and relevant supporting information were reviewed and approved by the institutional review board/ethics committee before the study was initiated. Confidentiality was maintained by assigning each participant enrolled in the study a unique identification number.

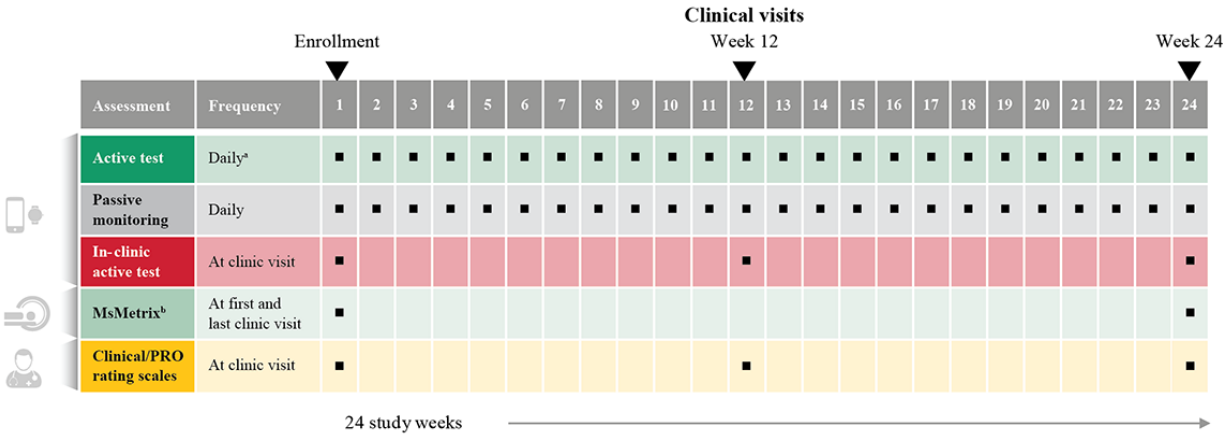
Study Design

The FLOODLIGHT study combines continuous sensor data capture with smartphones and smartwatches and standard clinical outcome measures. Eligible people with multiple sclerosis and HCs were enrolled in the study and assessed clinically at the enrollment visit, week 12, and termination visit (week 24). In addition, participants were asked to perform a set of daily active tests and contribute sensor data via passive

XSL•FO
RenderX

monitoring with smartphone and smartwatches over a period of 24 weeks (Figure 1).

Figure 1. FLOODLIGHT study design. PRO: patient-reported outcome. "a" indicates that active tests were administered weekly or every two weeks (see next figure for schedule). "b" indicates that brain magnetic resonance imaging was performed in people with multiple sclerosis.



In-Clinic Assessments

At each scheduled in-clinic visit (enrollment, week 12, and week 24), the following reference clinical tests were performed for all participants: 9-Hole Peg Test (9HPT), oral version of Symbol Digit Modalities Test (SDMT) [15-17], Timed 25-Foot Walk (T25FW) test, Berg Balance Scale (BBS) [18], Fatigue Scale for Motor and Cognitive Functions (FSMC) [19], and Patient Health Questionnaire–9 (PHQ-9) [20]. For people with multiple sclerosis only, disability was measured by EDSS [1], Patient Determined Disease Steps (PDDS) [21], and Multiple Sclerosis Impact Scale–29 (MSIS-29; version 2) [22-24]. While performing some of the in-clinic tests, people with multiple sclerosis and HCs were asked to carry or wear the smartphone and smartwatch to collect sensor data alongside the in-clinic measures. On the scheduled in-clinic visits, the smartphone and smartwatch FLOODLIGHT active tests were performed under investigator supervision. The satisfaction questionnaire (Multimedia Appendix 2) assessed people with multiple sclerosis' and HCs' experience regarding smartphone and smartwatch use and its impact on their daily activities at the week 12 visit and at the study termination/early discontinuation visit. Brain MRI was performed in people with multiple sclerosis at the enrollment visit and at week 24.

Smartphone and Smartwatch Testing

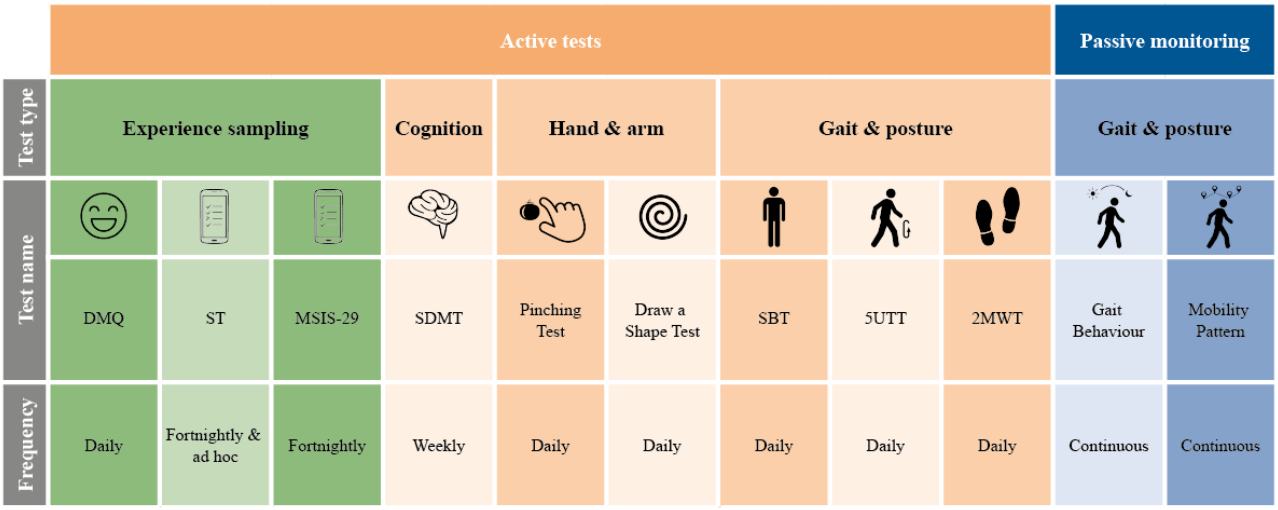
At the enrollment visit, people with multiple sclerosis and HCs were provided with the FLOODLIGHT solution that included a smartphone and smartwatch preconfigured so participants could only run the FLOODLIGHT software. A belt bag was also provided for participants to carry their smartphone in an anterior medial position. The smartphone and smartwatch pair contained preinstalled apps that prompted the user to perform various assessments, referred to as active tests. The apps also passively recorded sensor data, referred to as passive monitoring. At the enrollment visit, participants received training on the use of the smartphone and smartwatch and were provided with supporting content to help them complete the tests successfully. Participants were instructed to complete the active tests at approximately the same time each day and carry the smartphone and smartwatch throughout the day, recharging the devices overnight. Data transfer from the smartphone and smartwatch is described in Multimedia Appendix 1.

FLOODLIGHT Active Tests

People with multiple sclerosis and HCs were asked to perform various active tests (daily, weekly, every two weeks, or on demand) via the smartphone (Figure 2 and Table 1). These novel active tests were developed to be self-administered by people with multiple sclerosis to capture MS symptoms. A range of clinical and sensor-based assessments were chosen to capture the most prominent symptoms of MS from a broad spectrum of symptoms. People with multiple sclerosis and HCs were required to wear the smartwatch throughout the active tests.

XSL•FO
RenderX

Figure 2. FLOODLIGHT active tests and their schedule frequency. DMQ: Daily Mood Question; MSIS-29: Multiple Sclerosis Impact Scale–29; SBT: Static Balance Test; SDMT: Symbol Digit Modalities Test; ST: Symptom Tracker; 2MWT: Two-Minute Walk Test; 5UTT: 5 U-Turn Test.



XSL•FO
RenderX

Table 1. FLOODLIGHT active tests.

Domain and test	Short description
Daily hand motor function tests[a]	
Draw a Shape (DaS) Test	The aim of the DaS Test is to assess fine finger/manual dexterity while the participants are instructed to hold the mobile device in the untested hand and draw on the smartphone touchscreen six prewritten alternating shapes of increasing complexity (linear, rectangular, circular, sinusoidal, and spiral) with the second finger of the tested hand as fast and as accurately as possible within a maximum time (30 seconds for each of the two attempts per shape).
Pinching Test	The aim of the Pinching Test is to assess fine pinching/grasping dexterity while the participants are instructed to hold the mobile device in the untested hand and touch the screen with two fingers from the tested hand (thumb + second or thumb + third finger preferred) to squeeze/pinch as many round shapes (ie, tomatoes) as they can during 30 seconds.
Daily gait tests[b]	
Two-Minute Walk Test (2MWT)	Participants are instructed to walk as fast and as long as they can for 2 minutes but walk safely. The 2MWT is a simple test that is required to be performed on an even ground in a place where participants have identified they could walk straight for as far as ≥200 meters without U-turns. Participants are allowed to wear regular footwear and an assistive device and/or orthotic as needed.
5 U-Turn Test (5UTT)	The aim of this test is to assess difficulties or unusual patterns in performing U-turns while walking on a short distance at comfortable pace. The 5UTT can be performed indoors or outdoors, on an even ground where participants are instructed to walk safely and perform five successive U-turns going back and forward between two points a few meters apart for 1 minute. Participants are allowed to wear regular footwear and an assistive device and/or orthotic as needed.
Static Balance Test (SBT)	Participants are asked to stand still unsupported for 30 seconds with relaxed arms straight alongside the body if possible.
Weekly cognitive test	
Electronic version of the Symbol Digit Modalities Test (SDMT) [15-17]	The aim of SDMT testing is to detect impairment of key neurocognitive functions that underlie many substitution tasks.
Patient-reported outcomes (PROs)	
Daily Mood Question (DMQ)	This test represents an assessment of participants' perceived overall state by responding daily to the question "How do you feel now?" on a 5-item Likert scale, ranging from excellent to horrible.
Electronic version of the Multiple Sclerosis Impact Scale–29, version 2 (MSIS-29) [22-24]; people with multiple sclerosis only	This questionnaire measures the physical and psychological impact of multiple sclerosis.
Multiple Sclerosis Symptom Tracker (MSST); people with multiple sclerosis only	Patients are asked if they experienced new or significantly worsening symptoms during the last 2 weeks. If yes, onset of the symptoms and the patients' perception to whether they think they experienced a relapse (yes, no, or unsure) are recorded.

[a]Tests alternatingly performed with right and left hand; users are instructed on daily alternation.

[b]Recommended position of smartphone in an anterior medial position in the belt bag.

FLOODLIGHT Passive Monitoring

Passive monitoring collected metrics on gait and mobility throughout the day in a continuous and unobtrusive manner. Participants were instructed to carry their smartphone preferably in an anterior medial position in a belt bag or, alternatively, in their pocket and wear the smartwatch all day as they went about their daily routine until the devices ran out of charge.

Statistical Analyses

The analyses of the primary objectives of this study were descriptive. Statistical tests were exploratory and conducted at the two-sided 5% significance level without adjustment for multiple comparisons. The analyses were based on all enrolled patients (full analysis set [FAS]). Patients who prematurely withdrew from the study for any reason were still included in the FAS. Supportive analyses of selected variables were carried out in the per-protocol population, which included people with multiple sclerosis who completed at least 1 week in the study and did not discontinue due to an adverse event or a reason unrelated to the use of the FLOODLIGHT solution (Multimedia Appendix 3).

Adherence was evaluated for the following tests and test groups: all FLOODLIGHT active tests, Two-Minute Walk Test (2MWT), all active tests except 2MWT, smartphone use, smartwatch use, and for the per-protocol population.

Adherence to active tests was measured as the proportion of study weeks with at least 3 days of completed testing (study co-primary endpoint). Adherence to sensor-based passive

XSL•FO
RenderX

monitoring was measured as the proportion of study weeks with at least 3 days of passive monitoring for at least 4 hours per day while the devices were worn by the participant [25] (study co-primary endpoint). Descriptive statistics of calculated adherence were reported for all active tests, all active tests except 2MWT, 2MWT, smartphone use, and smartwatch use. Categorical and numeric variables with fewer than five values were tested for association with adherence using the Kruskal-Wallis test. The association between continuous variables was assessed using the Spearman rank correlation.

Participant complete abandoning of active testing and passive monitoring was also investigated in a time-to-event survival analysis based on the Kaplan-Meier method along the FLOODLIGHT study. The abandoning event was defined as the last week in which the participant was adherent according to the definitions above for active tests and passive monitoring. Active tests performed on days of in-clinic visits were not considered in the adherence calculation to focus the abandoning analysis on the remote use. Participants leaving the study before the terminal visit were considered as censored. The impact of different characteristics on adherence was assessed using Cox regression.

A satisfaction score was developed from the satisfaction questionnaire (Multimedia Appendix 2) that sums the individual answers to questions 1-7 and 10-12 rescaled to 0-100 from their original Likert scale. An interquestion correlation analysis was performed to ensure questions are equally correlated and can be combined. Descriptive statistics of satisfaction score and items are reported, along with covariate analyses of demographics and disease state. Analysis of the change in satisfaction score between week 12 and week 24 are reported using the Wilcoxon signed-rank test.

Patient baseline characteristics incorporated as covariates in the analysis of correlation with FLOODLIGHT adherence and satisfaction outcomes were age, gender, body mass index, time since first MS symptom onset, EDSS, T25FW time, 9HPT time, and the oral SDMT correct responses. A descriptive analysis of safety variables, including adverse events and serious adverse events, was carried out in the FAS.

Results

Baseline Demographics and Characteristics

Participants' baseline demographics for the FAS are described in Table 2. There was an expected greater proportion of females among the people with multiple sclerosis compared with HCs. The majority (69/76, 91%) of people with multiple sclerosis had relapsing-remitting MS (RRMS), with a mild EDSS score (mean 2.4) and presumably "normal" hand/arm function based on an upper limit of normal range defined as the average 9HPT time for HCs plus two standard deviations [26] (Table 2).

Overall, 92% (70/76) of people with multiple sclerosis and 64% (16/25) of HCs who enrolled in the FLOODLIGHT study reached the week 24 visit. Reasons for discontinuation from the study are described in Multimedia Appendix 3.

XSL•FO
RenderX

Table 2. Demographics and characteristics of people with multiple sclerosis and healthy controls (HCs) at baseline.

Parameter	People with multiple sclerosis (n=76)	HCs (n=25)
Age (years), mean (SD)	39.5 (7.9)	34.9 (9.3)
Female, n (%)	53 (70)	7 (28)
Multiple sclerosis (MS) diagnosis, n (%)		
Primary progressive multiple sclerosis	3 (4)	—[a]
Secondary progressive multiple sclerosis	4 (5)	—
Relapsing-remitting multiple sclerosis	69 (91)	—
Time since MS symptom onset (years), mean (SD)	11.3 (7.0)[b]	—
Proportion of people with multiple sclerosis with ≥1 relapse in the past year, n (%)	18 (24)	—
Expanded Disability Status Scale, mean (SD)	2.4 (1.4)	—
Proportion of people with multiple sclerosis with ≥1 T1 Gd[c]-enhancing lesion, n (%)	2 (3)[d]	—
Total FLAIR[e] lesion volume (mL), mean (SD)	6.3 (7.5)[f]	—
9-Hole Peg Test (seconds), mean (SD)		
Dominant hand	22.1 (4.6)[g]	18.9 (2.1)
Nondominant hand	22.8 (4.9)[h]	19.5 (2.0)
Timed 25-Foot Walk (seconds), mean (SD)	6.0 (2.1)[b]	5.0 (1.0)
Symbol Digit Modalities Test (correct responses), mean (SD)	53.8 (11.8)[b]	63.8 (10.0)
Berg Balance Scale, mean (SD)	52.5 (5.7)[i]	56.0 (0)[j]
Patient Determined Disease Steps, mean (SD)	1.5 (1.6)	—
Fatigue Scale for Motor and Cognitive Functions (total score), mean (SD)	59.1 (22.7)[g]	25.5 (6.0)
Patient Health Questionnaire–9, mean (SD)	8.3 (6.1)[k]	2.4 (2.9)[l]
Participants with any previous medications, n (%)	46 (61)	6 (24)
Previous disease-modifying treatment[m], n (%)		
Daclizumab (Zinbryta)	0 (0)	—
Glatiramer acetate (Copaxone)	12 (16)	—
Glatiramer acetate (Glatopa)	1 (1)	—
IFN[n] β-1a IM[o] (Avonex)	4 (5)	—
IFN β-1a SC[p] (Rebif)	5 (7)	—
IFN β-1b SC (Betaseron/Betaferon)	6 (8)	—
IFN β-1b SC (Extavia)	1 (1)	—
Pegylated IFN β-1a (Plegridy)	2 (3)	—
Dimethyl fumarate (Tecfidera)	9 (12)	—
Fingolimod (Gilenya)	9 (12)	—
Teriflunomide (Aubagio)	3 (4)	—
Alemtuzumab (Lemtrada)	2 (3)	—
Mitoxantrone (Novantrone)	1 (1)	—
Natalizumab (Tysabri)	19 (25)	—
Other[q]	5 (7)	—

[a]Not applicable.

[b]n=75.

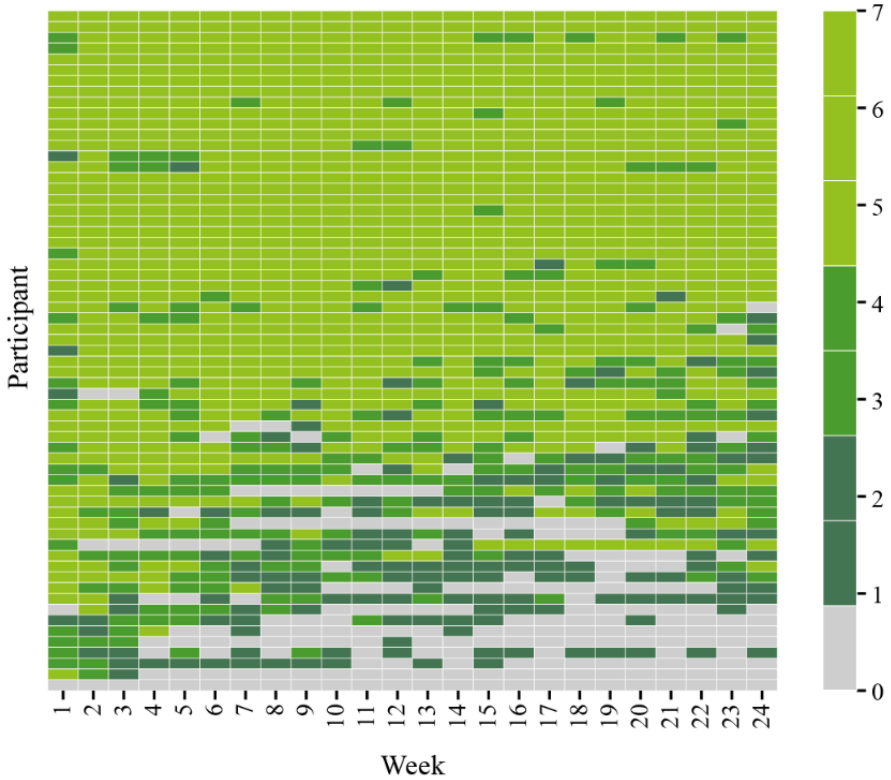
XSL•FO
RenderX

[c]Gd: gadolinium.
[d]n=68.
[e]FLAIR: fluid-attenuated inversion recovery.
[f]n=70.
[g]n=73.
[h]n=74.
[i]n=71.
[j]n=22.
[k]n=60.
[l]n=20.
[m]Total baseline disease-modifying treatment history.
[n]IFN: interferon.
[o]IM: intramuscular.
[p]SC: subcutaneous.
[q]Hidroferol; Radiance study (RPC1063 versus IFN β-1a); Rituximab (Rituxan).

Adherence

Over a period of 18 months (November 2016-April 2018), more than 6 terabytes of raw data were collected from 76 people with multiple sclerosis and 25 HCs. Participants performed 67,544 active tests, of which 9787 were the 2MWT, and recorded 200,171 hours of passive monitoring, of which 113,165 hours were captured with the smartwatch. Over 24 weeks, most participants performed 5 to 7 active tests per week, including the 2MWT (Figure 3). Adherence of people with multiple sclerosis to completing active tests and passive monitoring was good and remained stable over time after week 6 (Figures 4 and 5). Even in the last week of the 24-week study, participants completed all active tests on average 4 out of 7 days per week (Figure 4), and recorded at least 4 hours of data via passive monitoring on average 4 out of 7 days per week (Figure 5). The lowest average adherence over 24 weeks was observed for active tests including the 2MWT and the 2MWT only, with participants showing highest average adherence for passive monitoring (Figure 6). A total of 70% (16.68/24 weeks) of participants were adherent to all active tests, 75% (17.95/24 weeks) to all active tests except 2MWT, 71% (17.13/24 weeks) to 2MWT, 79% (18.89/24 weeks) to smartphone- or smartwatch-based passive monitoring, 66% (15.74/24 weeks) to smartphone-based passive monitoring, and 74% (17.69/24 weeks) to smartwatch-based passive monitoring.

Figure 3. Adherence of people with multiple sclerosis to active tests for individual participants: number of performed active tests per week [level of activity (light green: high; dark green/grey: low] over individual study weeks [columns]).



XSL•FO
RenderX

Figure 4. Adherence of people with multiple sclerosis to active tests. 2MWT: Two-Minute Walk Test.

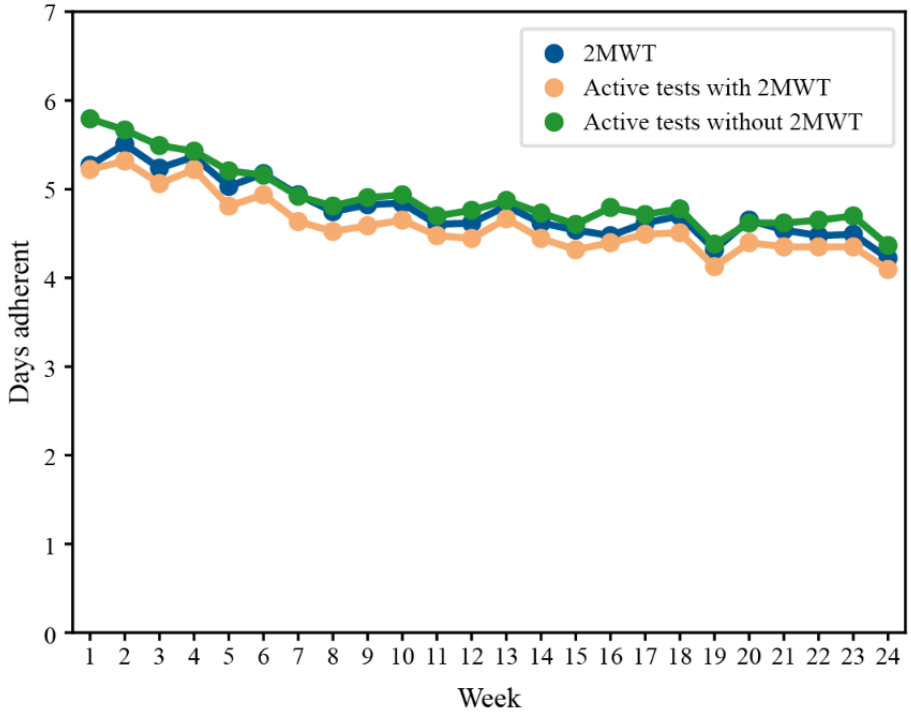
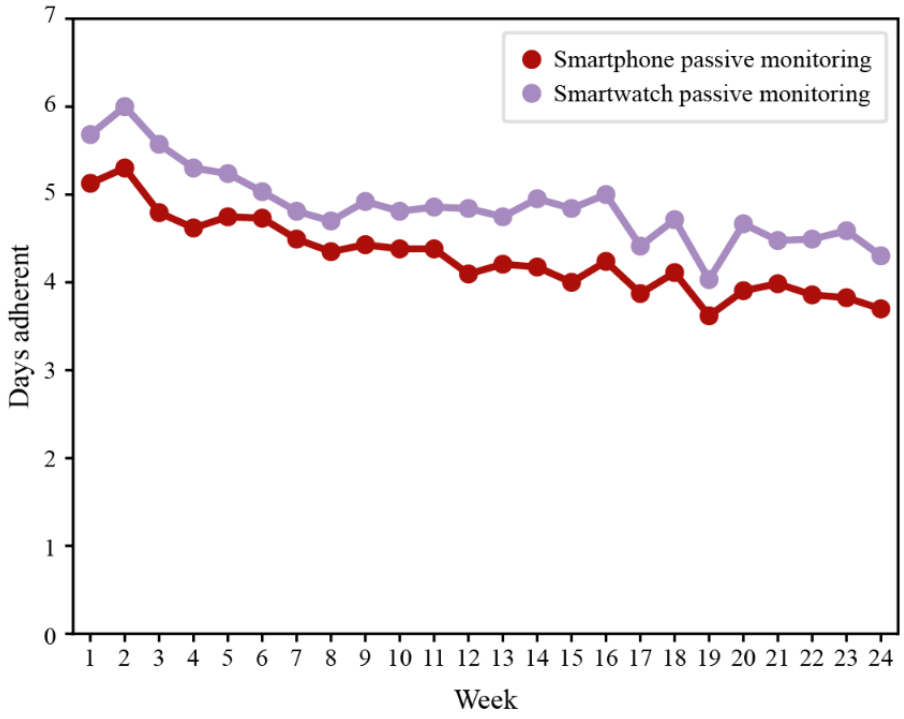
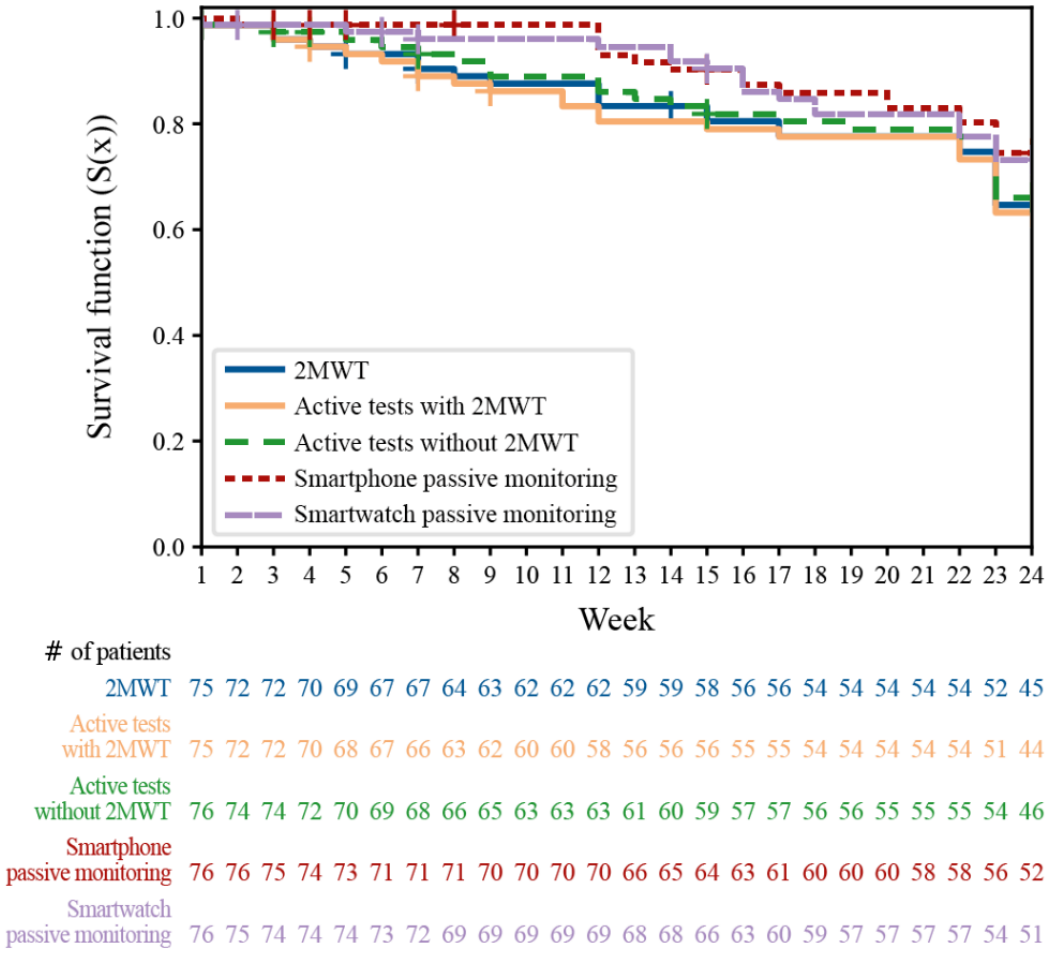


Figure 5. Adherence of people with multiple sclerosis to smartphone and smartwatch passive monitoring. Days with more than 4 hours of passive monitoring on a device are considered as adherent.



XSL•FO
RenderX

Figure 6. Adherence of people with multiple sclerosis to active tests and passive monitoring. The results of the time-to-event survival analysis based on the Kaplan–Meier method along the FLOODLIGHT study. The abandoning event was defined as the last week in which the participant was adherent according to the definitions for active tests and passive monitoring. Active tests performed on days of in-clinic visits were not considered in the adherence calculation, to focus the abandoning analysis on the remote use. Participants leaving the study before the terminal visit were considered as censored. 2MWT: Two-Minute Walk Test.



Correlation was explored between adherence measures and people with multiple sclerosis population characteristics. Only disease duration showed significant small negative correlation with measures of adherence (Spearman rank correlations; 2MWT adherence: –0.42, $P<.001$; smartphone passive monitoring adherence: –0.29, $P=.02$; all active tests except 2MWT adherence: –0.37, $P=.003$; and smartwatch passive monitoring adherence: –0.27, $P=.04$), indicating that disease severity and demographics did not appear to play a significant role in adherence.

Patient Satisfaction

The average overall satisfaction score among people with multiple sclerosis who completed the study at week 12 (n=64) was 74.1 out of a possible 100 and remained stable at week 24 (study termination/early discontinuation visit [n=68]) with 73.7 out of 100 (Wilcoxon signed-rank test $P=.71$). There was one significant association between overall satisfaction score and gender ($P=.04$). Individual questions from the satisfaction questionnaire (Multimedia Appendix 2) were analyzed for their association with people with multiple sclerosis population characteristics, described in Multimedia Appendix 1.

Implications for the use of the FLOODLIGHT test battery in people with multiple sclerosis were assessed from individual questions from the patient satisfaction questionnaire. When asked to rate the impact of the smartphone, smartwatch, and active tests on daily living, more than 80% (61/72) of people with multiple sclerosis perceived the FLOODLIGHT test battery to have at least acceptable impact on daily activities (Figure 7). Nearly 50% (32/71) of participants had no issue with any of the active tests, and only one-third would prefer to avoid the 2MWT, most likely due to increased burden from execution—for example, having to find a place to perform the test or not wanting to go outside in bad weather (Figure 8). Without providing any data feedback to the people with multiple sclerosis throughout the study, more than 60% (46/72) of participants would have liked to continue using FLOODLIGHT “to understand my MS better and improve my disease management” (Figure 9). Approximately 90% (65/72) of people with multiple sclerosis indicated their interest to see the results of the tests, which will be addressed in future Roche-sponsored studies using FLOODLIGHT (CONSONANCE [NCT03523858] and FLOODLIGHT Open [floodlightopen.com]; Figure 10). Analysis of patient responses to the satisfaction questionnaire is described in Multimedia Appendix 1.

XSL•FO
RenderX

Figure 7. Implications of FLOODLIGHT in people with multiple sclerosis for “impact on daily activities” from the patient satisfaction questionnaire.

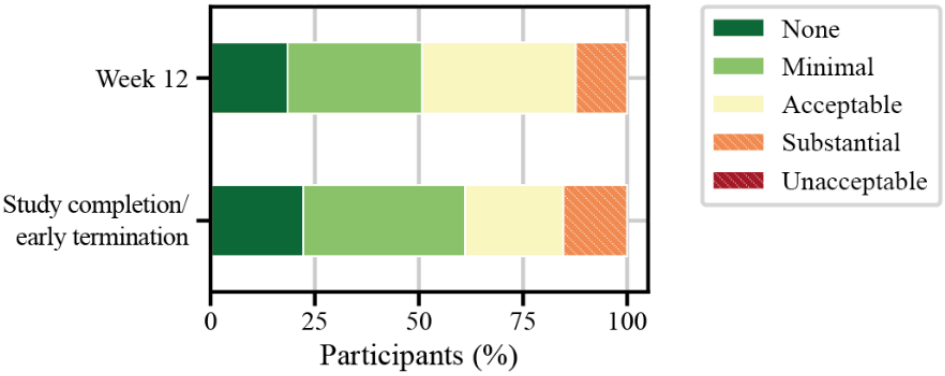


Figure 8. Implications of FLOODLIGHT in people with multiple sclerosis for “avoiding one component of FLOODLIGHT" from the patient satisfaction questionnaire. 2MWT: Two-Minute Walk Test.

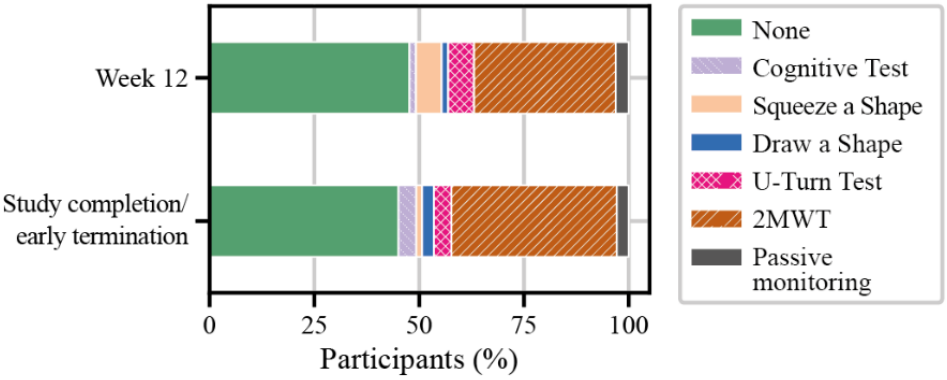


Figure 9. Implications of FLOODLIGHT in people with multiple sclerosis for “desire to continue using the FLOODLIGHT app” from the patient satisfaction questionnaire.

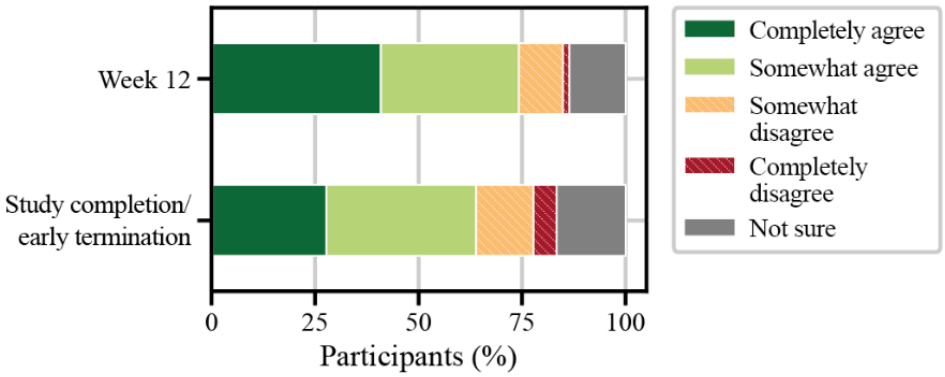
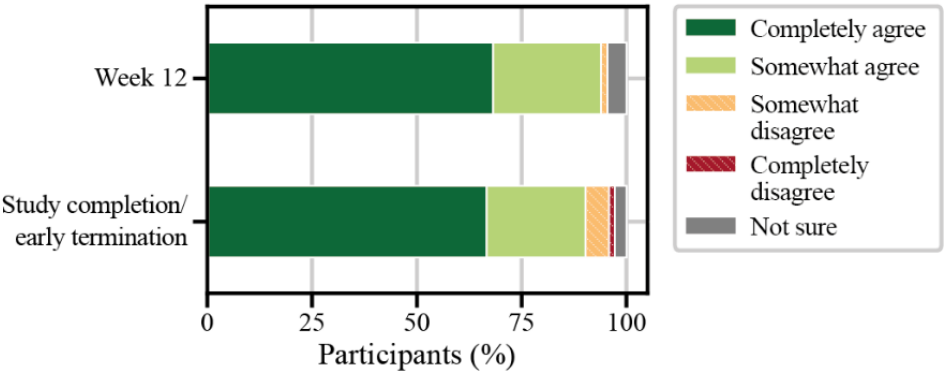


Figure 10. Implications of FLOODLIGHT in people with multiple sclerosis for “prefer to see results immediately to monitor” from the patient satisfaction questionnaire.



XSL•FO
RenderX

Discussion

Principal Findings

This study demonstrates that the use of smartphones and smartwatches for remote daily active testing and continuous passive monitoring is feasible over 6 months and provides further support to earlier studies, which have shown that healthy participants and people with multiple sclerosis were capable of completing daily tasks on a smartphone [7]. This study provides further evidence for the use of digital technology, including smartphones, for data collection. Other studies in MS have used smartphone apps to (1) assess steps when walking on a treadmill [9]; (2) assess pain, fatigue, anxiety, and quality of life [13]; and (3) assess the feasibility of gathering passive and active performance data [7]. Together with the current analyses, these studies document the focus toward developing digital measures to continuously monitor and assess the MS disease.

In this protocol, the FLOODLIGHT solution collected metrics on cognition, mood, upper extremity function, and gait and posture by instructing participants to perform a set of daily active tests, which should take approximately 5 minutes in total to complete and capture activity data via passive monitoring over a period of 24 weeks. A previous study has shown that 51% of participants (22/38 of people with multiple sclerosis and 17/38 of healthy participants) completed 12 months of daily data collection, where participants were prompted to complete one assigned test [7]. In the context of the FLOODLIGHT study, we observed that overall adherence to active tests was 70% (16.68/24 weeks), which appears to be higher than the adherence of participants to 12 months of daily data collection (39/76, 51%) from Bove et al [7]. However, comparisons between the studies are limited, as the study design and burden of testing are different—for example, the app from Bove et al contained 19 different tests, of which participants were prompted to complete one each day. As the FLOODLIGHT app was integrated into standalone devices in this study, deployment of the app on participants' own mobile devices may increase adherence because it removes the need to carry a separate, dedicated device and decreases burden on the individual.

Limitations

As this study remains a pilot investigation designed to collect first experiences from continuous sensor data capture, the main limitation is the small sample size and short duration of follow-up. Future ongoing FLOODLIGHT studies (CONSONANCE and FLOODLIGHT Open) will collect longer term data on smartphone-based sensor data capture in a larger number of participants from a broader disability spectrum. Additionally, whether physical and cognitive limitations in people with secondary progressive MS (SPMS) and people with primary progressive MS (PPMS) differentially impacts adherence compared with people with relapsing MS (RMS) cannot be gleaned from the current data set due to the low numbers of advanced patients enrolled in the study; however, this important question warrants future research exploring remote monitoring in patients with more advanced MS. The importance of continuous monitoring in RMS should also not be overlooked, as the sensitivity of this novel approach aiming at detecting progression in a real-world setting may provide an earlier window into disease progression outside of the clinic.

Comparison With Prior Work

A recent study assessing the feasibility of the MS TeleCoach, a novel intervention offering telemonitoring of fatigue and telecoaching of physical activity in people with multiple sclerosis, showed that participants were highly engaged, with 76% (57/75) of participants completing the study, and 91% (21/23) of a subset of completers showing a median of quite satisfied in the patient satisfaction questionnaire [27]. During the 12-week study period, use of the MS TeleCoach improved fatigue levels in people with multiple sclerosis with moderate to severe fatigue, suggesting that implementation of digital technologies can enhance patient performance. Together with the data presented here, these results indicate that the use of consumer devices by people with multiple sclerosis for sensor data capture fulfills the prerequisites of people with multiple sclerosis satisfaction and acceptable adherence to daily active tests and passive monitoring for potential integration in long-term clinical trials and treatment monitoring. Regarding future studies, for example FLOODLIGHT Open, attempts will be made to improve participant adherence throughout the study by introducing controlled app variations, such as reminders, types of achievements, and fun metrics.

Conclusions

In summary, these analyses showed that people with multiple sclerosis are highly engaged with performing active tests and capturing continuous data via passive monitoring and are satisfied with the FLOODLIGHT test battery. Neither satisfaction nor adherence showed strong correlation with study population characteristics. More than 60% (46/72) of people with multiple sclerosis indicated their interest to continue to use FLOODLIGHT, and approximately 90% (65/72) wanted to see the results of their tests in real time as biofeedback, which was implemented in future studies using the FLOODLIGHT solution. These findings indicate that smartphone-based FLOODLIGHT outcomes may represent a promising avenue to enable a more accurate and continuous assessment of MS disease in clinical trials and real-world practice settings and may eventually also contribute to informing and guiding clinical research and clinical practice in the future.

Acknowledgments

We would like to thank all patients, their families, and the investigators who participated in this trial. This research was funded by F Hoffmann–La Roche Ltd, Basel, Switzerland. We would like to thank the following employees from F Hoffmann–La Roche Ltd who supported and contributed to the study: Atieh Bamdadian, Alessandro Barbato, Jan Beckmann, Sandro Fritz, Nicholas Pierce Heinemeier, Timothy Kilchenmann, Lito Kriara, Bernd Laub, Grégoire Pointeau, Caroline Polakowska, Marcin Puhacz,

XSL•FO
RenderX

Cedric Simillion, Jens Schjodt-Eriksen, Jöerg Sprengel, Ralf Stubner, and Krzysztof Trybus. Writing and editorial assistance for this manuscript was provided by Heather Latimer from Articulate Science, United Kingdom, and funded by F Hoffmann–La Roche Ltd.

Qualified researchers may request access to individual patient-level data through the clinical study data request platform (https://clinicalstudydatarequest.com). Further details on Roche's criteria for eligible studies are available at https://clinicalstudydatarequest.com/Study-Sponsors/Study-Sponsors-Roche.aspx. For further details on Roche's Global Policy on the Sharing of Clinical Information and how to request access to related clinical study documents, see https://www.roche.com/research_and_development/who_we_are_how_we_work/clinical_trials/our_commitment_to_data_sharing.htm.

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: F Hoffmann–La Roche Ltd, Basel, Switzerland, provided financial support for the study and publication of this manuscript.

Conflicts of Interest

XM has received speaker honoraria and travel expense reimbursement for participation in scientific meetings, been a steering committee member of clinical trials, or served on advisory boards of clinical trials for Actelion, Biogen, Celgene, Merck, Novartis, Oryzon, Roche, Sanofi Genzyme, and Teva Pharmaceutical. JG has received grants or research support from Biogen, Genentech Inc, and S3 Group and has received compensation for a nonbranded resident and fellow education seminar supported by Biogen. SLH serves on the scientific advisory boards for Annexon, Symbiotix, Bionure, and Molecular Stethoscope, is on the board of trustees for Neurona Therapeutics, and has received travel reimbursement and writing assistance from F Hoffmann–La Roche Ltd for CD20-related meetings and presentations. LJ is an employee of Genentech Inc and a shareholder of F Hoffmann–La Roche Ltd. MB, JS, and CG are employees and shareholders of F Hoffmann–La Roche Ltd. AS, FL, and JvB are employees of F Hoffmann–La Roche Ltd. CB and ML are contractors for F Hoffmann–La Roche Ltd. SB was an employee of F Hoffmann–La Roche Ltd during the completion of the work related to this manuscript. SB is now an employee of Biogen (Cambridge, MA), which was not in any way associated with this study. LM and PM have nothing to disclose.

Multimedia Appendix 1
Additional trial information.
[PDF File (Adobe PDF File), 278KB - jmir_v21i8e14863_app1.pdf]

Multimedia Appendix 2
Satisfaction questionnaire.
[PDF File (Adobe PDF File), 238KB - jmir_v21i8e14863_app2.pdf]

Multimedia Appendix 3
Participant flow table.
[PDF File (Adobe PDF File), 138KB - jmir_v21i8e14863_app3.pdf]

References

1. Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). Neurology 1983 Nov;33(11):1444-1452. [Medline: 6685237]
2. Bradshaw MJ, Farrow S, Motl RW, Chitnis T. Wearable biosensors to monitor disability in multiple sclerosis. Neurol Clin Pract 2017 Aug;7(4):354-362. [doi: 10.1212/CPJ.0000000000000382] [Medline: 29185551]
3. Lavorgna L, Brigo F, Moccia M, Leocani L, Lanzillo R, Clerico M, et al. e-Health and multiple sclerosis: An update. Mult Scler 2018 Nov;24(13):1657-1664. [doi: 10.1177/1352458518799629] [Medline: 30231004]
4. Busis N. Mobile phones to improve the practice of neurology. Neurol Clin 2010 May;28(2):395-410. [doi: 10.1016/j.ncl.2009.11.001] [Medline: 20202500]
5. Lipsmeier F, Taylor KI, Kilchenmann T, Wolf D, Scotland A, Schjodt-Eriksen J, et al. Evaluation of smartphone-based testing to generate exploratory outcome measures in a phase 1 Parkinson's disease clinical trial. Mov Disord 2018 Dec;33(8):1287-1297 [FREE Full text] [doi: 10.1002/mds.27376] [Medline: 29701258]
6. Prasad S, Ramachandran R, Jennings C. Development of smartphone technology to monitor disease progression in multiple sclerosis. Neurology 2012 Apr 22;78(Meeting Abstracts 1):P01.144. [doi: 10.1212/WNL.78.1_MeetingAbstracts.P01.144]
7. Bove R, White CC, Giovannoni G, Glanz B, Golubchikov V, Hujol J, et al. Evaluating more naturalistic outcome measures: a 1-year smartphone study in multiple sclerosis. Neurol Neuroimmunol Neuroinflamm 2015 Dec;2(6):e162 [FREE Full text] [doi: 10.1212/NXI.0000000000000162] [Medline: 26516627]
8. Boukhvalova AK, Kowalczyk E, Harris T, Kosa P, Wichman A, Sandford MA, et al. Identifying and quantifying neurological disability via smartphone. Front Neurol 2018;9:740 [FREE Full text] [doi: 10.3389/fneur.2018.00740] [Medline: 30233487]

XSL•FO
RenderX

9. Balto JM, Kinnett-Hopkins DL, Motl RW. Accuracy and precision of smartphone applications and commercially available motion sensors in multiple sclerosis. Mult Scler J Exp Transl Clin 2016;2:2055217316634754 [FREE Full text] [doi: 10.1177/2055217316634754] [Medline: 28607720]
10. Haase R, Schultheiss T, Kempcke R, Thomas K, Ziemssen T. Modern communication technology skills of patients with multiple sclerosis. Mult Scler 2013 Aug;19(9):1240-1241. [doi: 10.1177/1352458512471882] [Medline: 23388164]
11. Rudick RA, Miller D, Bethoux F, Rao SM, Lee J, Stough D, et al. The Multiple Sclerosis Performance Test (MSPT): an iPad-based disability assessment tool. J Vis Exp 2014 Jun 30(88):e51318 [FREE Full text] [doi: 10.3791/51318] [Medline: 25046650]
12. Gholami F, Trojan DA, Kovecses J, Haddad WM, Gholami B. A Microsoft Kinect-based point-of-care gait assessment framework for multiple sclerosis patients. IEEE J Biomed Health Inform 2017 Dec;21(5):1376-1385. [doi: 10.1109/JBHI.2016.2593692] [Medline: 27455529]
13. Kos D, Raeymaekers J, Van Remoortel A, D'hooghe MB, Nagels G, D'Haeseleer M, et al. Electronic visual analogue scales for pain, fatigue, anxiety and quality of life in people with multiple sclerosis using smartphone and tablet: a reliability and feasibility study. Clin Rehabil 2017 Sep;31(9):1215-1225. [doi: 10.1177/0269215517692641] [Medline: 28786335]
14. Polman CH, Reingold SC, Banwell B, Clanet M, Cohen JA, Filippi M, et al. Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. Ann Neurol 2011 Feb;69(2):292-302 [FREE Full text] [doi: 10.1002/ana.22366] [Medline: 21387374]
15. Smith A. The symbol-digit modalities test: A neuropsychologic test of learning and other cerebral disorders. Seattle: Helmuth J, editor. Learning disorders. Seattle: Special Child Publications; 1968:83-91.
16. Smith A. Symbol Digits Modalities Test: Manual. Los Angeles: Western Psychological Services; 1982.
17. Rao SM, Losinski G, Mourany L, Schindler D, Mamone B, Reece C, et al. Processing speed test: validation of a self-administered, iPad-based tool for screening cognitive dysfunction in a clinic setting. Mult Scler 2017 Dec;23(14):1929-1937. [doi: 10.1177/1352458516688955] [Medline: 28080262]
18. Berg KO, Wood-Dauphinee SL, Williams JI, Maki B. Measuring balance in the elderly: validation of an instrument. Can J Public Health 1992;83 Suppl 2:S7-S11. [Medline: 1468055]
19. Penner IK, Raselli C, Stöcklin M, Opwis K, Kappos L, Calabrese P. The Fatigue Scale for Motor and Cognitive Functions (FSMC): validation of a new instrument to assess multiple sclerosis-related fatigue. Mult Scler 2009 Dec;15(12):1509-1517. [doi: 10.1177/1352458509348519] [Medline: 19995840]
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. Learmonth YC, Motl RW, Sandroff BM, Pula JH, Cadavid D. Validation of patient determined disease steps (PDDS) scale scores in persons with multiple sclerosis. BMC Neurol 2013 Apr 25;13:37 [FREE Full text] [doi: 10.1186/1471-2377-13-37] [Medline: 23617555]
22. Hobart J, Cano S. Improving the evaluation of therapeutic interventions in multiple sclerosis: the role of new psychometric methods. Health Technol Assess 2009 Feb;13(12):1-177 [FREE Full text] [doi: 10.3310/hta13120] [Medline: 19216837]
23. Hobart J, Lamping D, Fitzpatrick R, Riazi A, Thompson A. The Multiple Sclerosis Impact Scale (MSIS-29): a new patient-based outcome measure. Brain 2001 May;124(Pt 5):962-973. [doi: 10.1093/brain/124.5.962] [Medline: 11335698]
24. Riazi A, Hobart JC, Lamping DL, Fitzpatrick R, Thompson AJ. Multiple Sclerosis Impact Scale (MSIS-29): reliability and validity in hospital based samples. J Neurol Neurosurg Psychiatry 2002 Dec;73(6):701-704 [FREE Full text] [doi: 10.1136/jnnp.73.6.701] [Medline: 12438473]
25. Cheng W, Scotland A, Lipsmeier F, Kilchenmann T, Jin L, Schjodt-Eriksen J, et al. Human activity recognition from sensor-based large-scale continuous monitoring of Parkinson's disease patients. 2017 Presented at: Proceedings of the 2nd IEEE/ACM International Conference on Connected Health; July 17-19, 2017; Philadelphia. [doi: 10.1109/CHASE.2017.87]
26. Wang Y, Bohannon RW, Kapellusch J, Garg A, Gershon RC. Dexterity as measured with the 9-Hole Peg Test (9-HPT) across the age span. J Hand Ther 2015;28(1):53-59. [doi: 10.1016/j.jht.2014.09.002] [Medline: 25449717]
27. D'hooghe M, Van Gassen G, Kos D, Bouquiaux O, Cambron M, Decoo D, et al. Improving fatigue in multiple sclerosis by smartphone-supported energy management: the MS TeleCoach feasibility study. Mult Scler Relat Disord 2018 May;22:90-96 [FREE Full text] [doi: 10.1016/j.msard.2018.03.020] [Medline: 29649789]

Abbreviations

2MWT: Two-Minute Walk Test
5UTT: 5 U-Turn Test
9HPT: 9-Hole Peg Test
BBS: Berg Balance Scale
DMQ: Daily Mood Question
EDSS: Expanded Disability Status Scale
FAS: full analysis set
FLAIR: fluid-attenuated inversion recovery

XSL•FO
RenderX

FSMC: Fatigue Scale for Motor and Cognitive Functions
HCs: healthy controls
MRI: magnetic resonance imaging
MS: multiple sclerosis
MSIS-29: Multiple Sclerosis Impact Scale–29
MSST: Multiple Sclerosis Symptom Tracker
PDDS: Patient Determined Disease Steps
PHQ-9: Patient Health Questionnaire–9
PPMS: primary progressive multiple sclerosis
PRO: patient-reported outcome
RMS: relapsing multiple sclerosis
RRMS: relapsing-remitting multiple sclerosis
SBT: Static Balance Test
SDMT: Symbol Digit Modalities Test
SPMS: secondary progressive multiple sclerosis
T25FW: Timed 25-Foot Walk

Edited by G Eysenbach; submitted 11.06.19; peer-reviewed by A Nguyen, L Visser, L Lavorgna, T Clavier; comments to author 29.06.19; revised version received 11.07.19; accepted 19.07.19; published 30.08.19.

Please cite as:
Midaglia L, Mulero P, Montalban X, Graves J, Hauser SL, Julian L, Baker M, Schadrack J, Gossens C, Scotland A, Lipsmeier F, van Beek J, Bernasconi C, Belachew S, Lindemann M
Adherence and Satisfaction of Smartphone- and Smartwatch-Based Remote Active Testing and Passive Monitoring in People With Multiple Sclerosis: Nonrandomized Interventional Feasibility Study
J Med Internet Res 2019;21(8):e14863
URL: http://www.jmir.org/2019/8/e14863/
doi:10.2196/14863
PMID:31471961

©Luciana Midaglia, Patricia Mulero, Xavier Montalban, Jennifer Graves, Stephen L Hauser, Laura Julian, Michael Baker, Jan Schadrack, Christian Gossens, Alf Scotland, Florian Lipsmeier, Johan van Beek, Corrado Bernasconi, Shibeshih Belachew, Michael Lindemann. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 30.08.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Social Media Surveillance of Multiple Sclerosis Medications Used During Pregnancy and Breastfeeding: Content Analysis

Bita Rezaallah[1,2], DMD, MAS; David John Lewis[2,3], PhD; Carrie Pierce[4], MBA, MPH; Hans-Florian Zeilhofer[5,6], Dr med dent, Dr med; Britt-Isabelle Berg[5,6], Dr med, Dr med dent

[1]Department of Clinical Research, University of Basel, Basel, Switzerland

[2]Patient Safety, Novartis Pharma AG, Basel, Switzerland

[3]School of Health and Human Sciences, University of Hertfordshire, Hatfield, United Kingdom

[4]Booz Allen Hamilton Inc, Boston, MA, United States

[5]Department of Cranio-Maxillofacial Surgery, University Hospital of Basel, Basel, Switzerland

[6]Hightech Research Center of Cranio-Maxillofacial Surgery, University of Basel, Basel, Switzerland

Corresponding Author:
Britt-Isabelle Berg, Dr med, Dr med dent
Department of Cranio-Maxillofacial Surgery
University Hospital of Basel
Spitalstrasse 21
Basel, 4031
Switzerland
Phone: 41 61 2652525
Email: isabelle.berg@usb.ch

Related Article:

This is a corrected version. See correction statement: https://www.jmir.org/2020/2/e18294/

Abstract

Background: Multiple sclerosis (MS) is a chronic neurological disease occurring mostly in women of childbearing age. Pregnant women with MS are usually excluded from clinical trials; as users of the internet, however, they are actively engaged in threads and forums on social media. Social media provides the potential to explore real-world patient experiences and concerns about the use of medicinal products during pregnancy and breastfeeding.

Objective: This study aimed to analyze the content of posts concerning pregnancy and use of medicines in online forums; thus, the study aimed to gain a thorough understanding of patients' experiences with MS medication.

Methods: Using the names of medicinal products as search terms, we collected posts from 21 publicly available pregnancy forums, which were accessed between March 2015 and March 2018. After the identification of relevant posts, we analyzed the content of each post using a content analysis technique and categorized the main topics that users discussed most frequently.

Results: We identified 6 main topics in 70 social media posts. These topics were as follows: (1) expressing personal experiences with MS medication use during the reproductive period (55/70, 80%), (2) seeking and sharing advice about the use of medicines (52/70, 74%), (3) progression of MS during and after pregnancy (35/70, 50%), (4) discussing concerns about MS medications during the reproductive period (35/70, 50%), (5) querying the possibility of breastfeeding while taking MS medications (30/70, 42%), and (6) commenting on communications with physicians (26/70, 37%).

Conclusions: Overall, many pregnant women or women considering pregnancy shared profound uncertainties and specific concerns about taking medicines during the reproductive period. There is a significant need to provide advice and guidance to MS patients concerning the use of medicines in pregnancy and postpartum as well as during breastfeeding. Advice must be tailored to the circumstances of each patient and, of course, to the individual medicine. Information must be provided by a trusted source with relevant expertise and made publicly available.

(J Med Internet Res 2019;21(8):e13003) doi:10.2196/13003

XSL•FO
RenderX

KEYWORDS

pharmacovigilance; machine learning; pregnancy outcome; postpartum; central nervous system agents; risk assessment; text mining

Introduction

Background

Multiple sclerosis (MS) is a chronic disease of the central nervous system [1]. It is more prevalent in females than males, with a ratio of approximately 3:1 [2]. Female MS patients are predominantly of childbearing potential with the average age of disease onset being 29.2 years [1]. The prevalence of MS is more common further from the equator; this maybe because of vitamin D deficiency rather than only genetics [3]. Pregnancy is not contraindicated in MS but remains a concern among female patients for a variety of reasons [2]. Pregnancy appears to have a protective effect in MS such that pregnant women suffer a reduced number of MS relapses, especially during the third trimester (reduction of around 70%). Thereafter, relapse rates tend to increase in the first 3 months postpartum [4,5]. However, this protective effect of pregnancy and the risk of postpartum relapse are both related to each patient's MS history and current disease activity [6].

Pregnant women are usually excluded from clinical trials because of ethical issues [7]; thus, safety information about human drug exposure during pregnancy is very limited at the time a marketing authorization is granted [8]. Pregnancy registries have been developed to address this gap in the safety profile of newly authorized medicines. Despite the evident advantages, such registries often suffer from low enrollment, resulting in delayed findings, selection bias, heterogeneity in data collection methods, and high costs [9]. As a result, prescribing information and patient information leaflets contain limited safety information for pregnant and breastfeeding patients [8]. Despite the evident need, to the best of our knowledge, there are no globally accepted guidelines by regulatory agencies for the medical management of MS during pregnancy and breastfeeding.

The rapid expansion of the internet and the availability of various social media platforms in recent years has increased the frequency with which patients use the internet [10]. In the United States, 90% of adults use the internet regularly, and 72% have searched for health information online [11]. Pregnant women in particular often access the internet to seek health information [12]. A cohort of pregnant women has been identified on Twitter using text mining and machine learning [13]. The availability of data for this cohort of pregnant women in social media provides an opportunity to explore and gain further insights into patient experiences related to MS medications. Therefore, by increasing health care professionals' (HCPs') awareness of patient concerns, carers can better advise patients during clinical visits.

Objective

The objective of this study was to analyze data qualitatively and describe the content of posts in online pregnancy forums to understand better patient experiences resulting from the use of MS medications during pregnancy, postpartum, and breastfeeding.

Methods

Data Acquisition and Classification

We obtained data from publicly available online pregnancy forums. An existing digital monitoring platform called MedWatcher Social (now called Epidemico, Booz Allen Hamilton) was utilized; this system has been described elsewhere [10,13,14]. MedWatcher Social comprises the natural language processing component that acquires public data from the internet, applies classification algorithms, and extracts adverse event–related posts. The aggregated frequency of product-event pairs identified by MedWatcher was concordant with data from the public US Food and Drug Administration (FDA) Adverse Event Reporting System by System Organ Class [9].

The classifier was designed to automatically collect, classify, and analyze social media discussions and threads pertaining to medicinal products [10,13,14]. The system collected online forum posts both retrospectively and prospectively via authorized third-party data vendors using the names of medicinal products as search terms. After data ingestion, a naïve Bayes classifier scored and filtered each post according to its relevance. Using statistical machine learning and a training set of over 360,000 hand-labeled social media posts, the classifier was trained to recognize the following:

- Descriptions of adverse drug reactions
- Medication errors
- Product quality issues
- Other patient experiences with medicinal products

The classifier was also used to exclude "noise" (eg, nonvalid product posts and spam). After filtering the data, natural language processors were applied to recognize and extract product and symptom terms through tokenization and proprietary taxonomies. References to products were standardized and consolidated, and vernacular descriptions of medical concepts were translated into the best matched term within the Medical Dictionary for Regulatory Activities terminology [15]. For this analysis, we identified and extracted a dataset from the system comprising posts acquired from 21 publicly accessible pregnancy social media forums listed in Textbox 1, published between March 2015 and March 2018. The forum data were acquired using third-party data from vendors, namely, Socialgist (SocialGist) and Datasift (DataSift Inc), and thus, was dependent on availability from those vendors. Data were not randomly sampled; rather, we selected any forums that were both available from Socialgist or Datasift and were dedicated to discussions around pregnancy or breastfeeding. The classifier used for the analysis was trained only on English language data, so we only used English language posts for this analysis.

XSL•FO
RenderX

In addition, we identified a list of products authorized for the treatment of MS and filtered the data accordingly. The products were alemtuzumab, teriflunomide, interferon beta-1a, interferon beta-1b, glatiramer acetate, daclizumab, dimethyl fumarate, fingolimod, and natalizumab. Posts mentioning either the active substance or brand name of each medicinal product were collected as shown in Multimedia Appendix 1.

Textbox 1. List of publicly available pregnancy forums.

- Babiesbase.com
- Babycenter.com
- Babycenter.com.au
- Babycentre.co.uk
- Cafemom.com
- Dcurbanmom.com
- Fertility.org
- Magrossesse.com
- Mumsnet.com
- Whattoexpect.com
- Swissmomforum.ch
- Baby-cafe.cz
- Babycenter.ca
- Babycenter.in
- Circleofmoms.com
- Essentialbaby.com.au
- Justmommies.com
- Netmums.com
- Thebump.com
- Fertilethoughts.com
- Scarymommy.com

Content Analysis

After automated classification, reports were manually divided into 2 groups: discussions related to pregnancy or breastfeeding and posts containing no thread relevant to pregnancy or lactation. In this study, we focused only on posts where an individual wrote about an experience related to a current or previous pregnancy, breastfeeding related to medicinal treatment of MS medication, a complication of MS or treatment of this disease.

A human expert reviewed the posts to characterize the experiences described in each post. First, we collected any medical information that a user shared in a post, such as time since their diagnosis of MS, planned or unplanned pregnancy, gestational age, outcome of pregnancy (or multiple pregnancies), number of pregnancies, current or previous pregnancy, concomitant medications, and John Cunningham (JC) virus serology results. Second, for the questions and concerns written in posts, we applied the content analysis method [16,17]. The aim was to use this categorization to identify common themes (threads) and to assess their frequency. To start with, we used open coding for obtaining the sense of the content. The coding team was composed of a physician (BR), a pharmacovigilance expert (DL), a statistician (AZ), and a machine-learning expert (CP). We created a codebook based on features that individual users shared (eg, what were their concerns and what action was taken with the medications). Subsequently the initial codes formed higher order headings of main topics. The entire dataset was reviewed, and posts were assigned to each topic. In addition, we quantified the content by measuring the frequency of each topic, which we cautiously proposed may stand as a proxy for significance [17].

The unit of analysis was the number of posts. It should be noted that in each individual post, the author might have provided comments on more than 1 main topic.

Ethics Statement

All human subject data used in this analysis were publicly available and have been presented in a deidentified format; in no case was any personally identifiable information (PII) reviewed. In fact, the classifier was set up to deidentify individual posts by removing any text relating to PII. We did not contact any individual on social media for follow-up as we felt that this posed unacceptable ethical and potential data privacy concerns. Thus, all of the posts were evaluated without knowledge of the identity of the patients involved.

XSL•FO
RenderX

Results

Data Processing Results

Our initial dataset comprised 376,691 posts that had been shared publicly on the pregnancy forums during the 4-year period of observation. This dataset was reduced to 168 (0.04% of total) posts relevant to pregnancy or breastfeeding and MS after filtering for posts mentioning the specified products. Finally, 16 posts containing spam-like language, non-English text, and nonvalid mentions of the product were automatically identified as irrelevant and were filtered out, leaving 152 posts for analysis as shown in Table 1.

Among the 152 posts, 70 unique posts discussed a current or previous pregnancy and breastfeeding experiences related to MS medications. The remaining 82 posts were noninformative concerning pregnancy and breastfeeding. As a result, we focused on the 70 posts that provided pertinent and substantive information. Table 2 provides illustrative examples of medically relevant information shared by the post authors. We could not identify the gender of individual users in each post but based on the content and the way that the text related personal sentiments and explanations, we assumed that it was predominantly pregnant women who authored the content.

Patients indicated that their newborn children were healthy, with no reports of congenital anomalies, in 18 of 70 posts (25%). MS patients shared in 22 of 70 (31%) posts their gestational age, and in 21 of 70 (30%) posts the year of the first diagnosis of MS was mentioned.

Table 1. Result of data processing.

Number of posts extracted from database via automation	Before spam removal, n	After spam removal, n
Posts mentioning any product	376,691	359,306
Posts mentioning multiple sclerosis products	168	152
Manual selection of unique posts where previous or current pregnancy was mentioned	152	70

Table 2. Medically relevant information shared by multiple sclerosis patients on the online posts.

Information shared in posts	Number of posts (N=70), n (%)	Illustrative text extracted from post
Gestational age	22 (31)	“I am 30 weeks pregnant”
First trimester[a]	8 (11)	“I was 6 weeks when I found I was pregnant...”
Second trimester[a]	7 (10)	“I am 27 weeks pregnant...”
Third trimester[a]	7 (10)	“I am 33 weeks pregnant...”
Time diagnosed for MS[b]	21 (30)	“I got diagnosed in 2009...”
Unplanned pregnancy	22 (31)	“...found out I was pregnant at 8 weeks and immediately stopped Gilenya...”
Planned pregnancy	8 (11)	“...stopped the medication in July to get pregnant...”
Outcome in newborns	18 (25)	“My daughter is [a] healthy one-year old...”
Previous pregnancy	10 (14)	“It’s my second baby...”
First pregnancy	7 (10)	“It’s my first pregnancy...”
Concomitant medication	5 (7)	“...Taking Methadone and Percocet as well...”
JC[c] virus result	3 (4)	“...I am JC positive...”

[a]The first trimester (1-12 weeks), second trimester (13-28 weeks), and third trimester (29-40 weeks) according to definition available in the US Department of Health and Human Services.

[b]MS: multiple sclerosis.

[c]JC: John Cunningham.

Content Analysis Results

Upon detailed review of the content of each post, we identified 6 main topics, which are presented in Tables 3 and 4. Patients used the pregnancy forums as an outlet for the following:

1. Describing in detail personal experiences with medicines, including changes in therapy, stopping medication, taking medication during pregnancy, and breastfeeding
2. Sharing and seeking information about MS medication in pregnancy and postpartum
3. Reporting MS progression (disease status) in this period
4. Expressing uncertainty or fears related to MS medication
5. Discussing or commenting on breastfeeding and MS medication
6. Sharing details or comments on communications with HCPs involved in the care of the pregnant mother or offspring

XSL•FO
RenderX

Table 3. Main topics posted by individuals on social media related to multiple sclerosis, pregnancy, and breastfeeding.

Topic	n (%)[a]
1. Discussion about personal experiences with MS[b] medication in reproductive period	56 (80)
Switched, switching, or will change medication during pregnancy or breastfeeding	28 (40)
Stopped, stopping, or will stop medication during pregnancy or breastfeeding	26 (37)
Took, taking, or will take medication during pregnancy or breastfeeding	22 (31)
2. Reporting MS disease status during and after pregnancy	35 (50)
Reported no relapse and healthy pregnancy	16 (22)
Reported relapse during pregnancy	15 (21)
Reported relapse postpartum	12 (17)
3. Seeking and giving advice	52 (74)
Seeking advice about MS, pregnancy, and postpartum	36 (51)
Giving advice about MS, pregnancy, and postpartum	16 (22)
4. Communication with the HCP[c]	26 (37)
Good communication, and patient express trust in the HCP	8 (11)
Poor communication	18 (25)
5. Discussion related to breastfeeding and MS medication	30 (42)
6. Express uncertainty and fear about MS medication in reproductive period	35 (50)

[a]Percentages are calculated using N=70 total individual posts about pregnancy and breastfeeding. The unit used was topic posted. One post may contain several pieces of information or an individual might have written about more than one pregnancy experience.

[b]MS: multiple sclerosis.

[c]HCP: health care professional.

XSL•FO
RenderX

Table 4. Illustrative example of posts related to each main topic and subtopic.

Topic	Illustrative text extracted from individual posts
Sharing experiences on MS[a] medications	
Stopping medication	"...I don't plan on taking anything [during] this pregnancy either."
Switching treatment	"...I took Copaxone throughout my pregnancy and breastfeeding and then started Tecfidera..."
Taking medication	"...I took Copaxone throughout my pregnancy and breastfeeding under the direction of my neuro [sic]...."
MS disease status	
No relapse	"...I had no issues with my MS during my pregnancy..."
Relapse in pregnancy	"...I have very active MS had 2 relapses in 29 weeks journey. Have been on copaxone [sic] throughout and short steroids course twice..."
Postpartum relapse	"...I didn't start flaring up until my son was over 6 months old. I've been in [sic] Tysabri since...!"
Seeking and giving advice	
Seeking advice	"...I am 8 weeks pregnant and was taking my gilenya [sic] during those 8 weeks meaning the baby will be exposed to it for 2 additional months Has anyone dealt with a pregnancy like this? The doctors have such limited information."
Giving advice	"...MS patients are advicesd [sic] to come off their meds when trying for a baby. my understanding is that Copaxone and the interferons are perfectly ok to take until a positive pregnancy test. I'm a little bitter because I got the same advice and suffered a disabling relapse as a result. Copaxone especially is probably fine to take even during pregnancy (though now that I have finally found luck, I have chosen to stay off during pregnancy and restart after birth and yes I will be breast-feeding). Good luck."
Breastfeeding	"...My neuro [sic] recommended a 3-day steroid infusion treatment. I had to pump and dump [sic] the whole time and for 24 hours following the last infusion..."
Express uncertainty or concerns	"...I just found out I am unexpectedly pregnant and conceived while in gilenya [sic]. Everything everyone has been telling us has made us to start thinking about terminating the pregnancy, which I really badly do not want to do. But if this child is any kind of danger I don't want to risk that. I just want someone to tell me it will be okay. I just don't know if that's realistic..."
Communication with HCPs[b]	
Good communication	"...I have been on Tysabri! I talked to neuro [sic] and she completely calmed my nerves! She just had me stop all meds for now then we'll switch to Copaxone after birth..."
Bad communication	"...My neurologist never mentioned anything, and said I can just start taking Gilenya after I give birth. She said attacks are more common after birth but didn't suggest anything to prevent them..."

[a]MS: multiple sclerosis.

[b]HCP: health care professional.

Discussion

Principal Findings

In this study, we performed text mining and characterization of posts acquired from pregnancy-related online forums where patients discussed MS medications. The aim of this study was to gain a better understanding of information sought by, or provided to, pregnant and breastfeeding MS patients who are active on social media. Our data show that the main topics of concern were switching, stopping, or taking medication during and after pregnancy; there was clear evidence of information seeking related to the risk of MS relapse during pregnancy or postpartum; and finally, questions were raised about breastfeeding while on medication. The most frequently observed content (approximately 80% of all relevant posts) was personal experiences with MS medications. Individuals shared their reasons for personal decisions regarding treatment; described how they felt after changes in therapy: switched, started, or stopped medication; and whether this was because of an HCP's recommendation or because of the patient's personal beliefs.

Patients used online forums to seek information from, and provide advice to, others (the latter occurred in 52/70, ie, 74% of posts). In 36 (51%) posts individuals asked their peers about decisions and outcomes or about experiences when taking a specific medication, queried the safety profile of certain medications, asked about the risk of MS relapses, and enquired about when to restart medical treatment postpartum. Our findings concur with the hypothesis that maternal medicine use is 1 of the 4 topics pregnant women care about most [18]. We had hoped that all of the topics would have been openly discussed with HCPs, but this was not invariably the case. In a number of posts, the patient expressed concerns that they had received medical advice from an HCP and either actively disagreed or least significantly doubted what they had been told. For example:

XSL•FO
RenderX

> *...I went to the infusion center for my first Tysabri treatment, the nurse said my neurologist requested a pregnancy test to rule it out before we got started. Long story short, it came back positive! My treatment was canceled. Here I am 3 years later, and pregnant with our third baby. Coincidentally, I missed my last two months of treatment (I only get it once a month) so it should be well out of my system and there shouldn't be any issues...*

In comparison to our results, a Swedish study found that, when speaking with their midwives, most pregnant women (70%) did not discuss information that they had retrieved from internet despite perceiving this information to be reliable [19]. Interestingly, more than half of the study subjects searched online for topics first raised by a midwife [19]. We were not in a position to explore the reason why patients went online and searched for information about their medicines; however, a Web-based survey among women who used the internet to seek pregnancy information showed that 48.6% of respondents were not satisfied with the information provided by their respective HCPs. The majority of these respondents (46.5%) stated that they primarily turned to the internet because they did not have time during appointments to discuss their concerns [20]. Moreover, pregnant women used the internet because the information given to them by their HCPs was neither clear nor sufficient [20].

In the breastfeeding category, 30 out of 70 (42%) posts described refusal or delay in commencing MS treatment for the sake of breastfeeding, described foregoing breastfeeding to restart treatment, requested evidence of which medication might be safer to take while breastfeeding, and others commented on discarding breast milk, which was suspected to contain medication while receiving treatment (so-called *pump and dump*) [21]. Several individuals shared confusion about the risks and benefits of breastfeeding and expressed anxiety about the dilemma of caring for their own health while not doing any harm to the baby.

There is very limited information about the safety of MS medication during breastfeeding. The in vivo model for drug exposure to breast milk is suboptimal, and human milk biobanks suffer from a paucity of human breast milk samples [22]. This is paradoxical, particularly when one considers the posts concerning the *pump and dump* phenomenon. A small adjustment in behavior, based on medical advice or guidance from a midwife, could yield a range of useful samples for retention and assay within existing biobanks. In addition, it is known that pregnancy registries often have low enrollment rates [23]. In our study, just 2 of 70 posts (2%) mentioned contacting pregnancy registries. One possible solution to increase the enrollment rate of pregnancy registries and human milk biobank centers could be improving the communication to pregnant MS patients about participation both at the point of care and in online forums. A simple scripted explanation about the existence of registries, the purpose of their research, and the impact that they can have on the MS population might yield better recruitment for altruistic reasons. Encouraging individuals to participate in the available biobanks, with all exhibiting a *pump and save rather than pump and dump* philosophy after treatment, could yield valuable evidence to aid decision making.

Another important finding was the rate of unplanned pregnancies with 22 out of 70 (31%) posts describing such events and only 8 of 70 posts (11%) describing planned pregnancies. Nonetheless, in some patient information leaflets for MS medicines, both contraception and careful planning of pregnancy is clearly recommended [24]. A Danish study surveyed 590 MS patients about family planning and reported that 42% of female and 74% of male partners did not know if their MS medication was teratogenic or not. This study also reported that 10% of pregnancies during MS treatment were unplanned; 49% of these pregnancies were terminated [25].

Generally, there are gaps in current methods for collecting and analyzing data pertaining to the safety of medicines during pregnancy and lactation [26,27]. The safety of medicinal products administered during pregnancy and lactation is a complex topic that needs coordinated communication across many disciplines to obtain, analyze, and present information in a harmonized approach. Harmonized methods and metrics among different pregnancy specialties should be developed to allow better analysis of outcomes and end points [26]. In this regard, we are aware of an Innovative Medicine Initiative (IMI) project called *Continuum of Evidence from Pregnancy Exposures, Reproductive Toxicology and Breastfeeding to Improve Outcomes Now* (ConcePTION) [28]. The IMI ConcePTION project is a collaboration between public-private partners and the pharmaceutical industry to address this problem. The aim of ConcePTION is “Building an ecosystem for better monitoring and communicating safety of medicines use in pregnancy and breastfeeding: validated and regulatory endorsed workflows for fast, optimized evidence generation” [28]. Participants in this project and the authors of this paper believe that there is an important societal obligation to reduce uncertainty about the effects of medicines used during pregnancy and breastfeeding.

Furthermore, even when safety data are available, it is often not effectively communicated to patients and HCPs. On September 26, 2017, the Pharmacovigilance Risk Assessment Committee (PRAC) and the European Medicines Agency (EMA) held their first public hearing about safety concerns with the use of medications containing sodium valproate during pregnancy [29]. Patients and carers participated in the public hearing, and both mothers and affected children expressed concern about the lack of effective risk minimization communication for safety of valproate during pregnancy, despite the drug having been authorized for more than 50 years [30]. After the public hearing, the PRAC and the EMA provided new measures for comprehensive risk minimization, including the following [29]:

- A pregnancy prevention program
- Visual warning about the risk in pregnancy on the box (outer packaging)
- A patient reminder card attached to outer package for pharmacists to discuss with patients each time the medicine is dispensed
- Updated educational materials for patients and HCPs

XSL•FO
RenderX

In the valproate pregnancy prevention program, HCPs are instructed to assess patients' potential for becoming pregnant by evaluating their individual circumstances and then assist their patients in making informed decisions. HCPs are responsible for informing their patients about the use of effective contraception methods throughout valproate treatment and to review such treatment annually. Interestingly, as an adjunct to all of these measures, a new risk acknowledgment form has been designed and implemented for patients and their HCPs to document that sufficient advice has been provided and understood [29]. Such comprehensive guidelines and the risk minimization methods adopted for valproate could serve as an example for improving and strengthening the warnings for MS medication in pregnancy.

In 2005, the EMA published guidance for assessing medicinal product risks on human reproduction and lactation [31]. In the United States, the FDA issued the Pregnancy and Lactation Labeling Rule for industry. This document provides a detailed framework for clearly communicating information to prescribers to aid improved decision making [21,32]. It is worth noting a study that reviewed medication risks during pregnancy for 172 drugs approved by the FDA between 2000 and 2010. Among these, in 97.7% of drugs, teratogenic risk in human pregnancy was *undetermined*, and the amount of data for 73.3% of these drugs was described as *none* [33]. For 468 drugs approved by the FDA between 1980 and 2000, the average time required for a drug's risk category to be changed from *undetermined* to a more precise risk was estimated to be 27 years [33]. A Web-based survey reported that patient leaflets were not comprehensive enough to answer pregnant women's questions and did not facilitate decision making [20]. In addition, inconsistencies have been found between the safety information concerning use during pregnancy provided in the US prescribing information and the UK summary of product characteristics for the same medicinal product [8].

Recommendations

Evidently, there is a need to improve regulatory policy and guidance by involving not only health authorities but also HCPs, patients, and other stakeholders including the national Teratology Information Services. We suggest 2 recommendations: (1) to conduct active postmarketing surveillance and (2) to provide globally harmonized evidence-based information for the prescriber, patients, and carers in a timely manner. Inevitably, with the internet and the wide variety of social media available, information is rapidly disseminated, and patients have access to and appear to trust nontraditional sources of medical information. We anticipate that in future, it will not be permissible to take 3 decades to vary existing labelling once sufficient evidence has been generated to provide useful information to patients and prescribers.

Conclusions

Social media can provide insight into patients' real-life experiences with medicinal products during pregnancy as well as their struggle in comprehending the benefits and risks associated with these products. Our study showed that MS patients expressed uncertainty and concerns around reproductive health; however, social media could be utilized as a platform to engage and encourage patients to enroll in pregnancy registries and to donate samples to milk biobank research centers. The adoption of these simple methods would support the generation of essential missing safety data and would support the communication of risk minimization strategies to pregnant patients and women of childbearing potential [34].

The role of HCPs involved in supporting pregnant patients, or during early child development, should not be underestimated. HCPs could provide comprehensive information for MS patients throughout different stages of pregnancy and postpartum as well as during breastfeeding. In addition, improving safety data collection and analysis as well as implementing efficient policies in regard to practical guidelines for MS populations of childbearing age would prove advantageous. Future guidelines should address the impact of MS on pregnancy and the effect of pregnancy on MS, the risks of the occurrence of birth defects, recommendations concerning the most effective contraceptive methods, and planning pregnancy as far as possible, to allow optimal wash-out time of medication, disease control during and after pregnancy, approved medication to use in reproductive periods, and lactation guidelines following the treatment [2,35]. Further research is needed to explore the effectiveness of risk minimization methods and to improve communication between HCPs and patients to the extent that it enables and informs shared decision making.

Limitations of the Study

Social media surveillance for medicinal product insight poses multiple challenges, which have been addressed in the literature [10,11,13]. In summary, there are technical, regulatory, privacy, and ethical considerations that need to be addressed when leveraging social media for this purpose [11]. In this study, the classifier was specifically selected to conduct research focusing on the exposure to MS medicines, not the effects of MS disease on the outcomes of pregnancy. In addition, these searches were only performed in pregnancy forums where posts related to MS medications were published. Hence, we recommend that further research be conducted in both MS and other disease-specific forums including *multiple sclerosis* term.

Acknowledgments

The authors would like to thank Amin Azmon for the statistical discussions.

Conflicts of Interest

BR is an employee of Novartis. DL is an employee of Novartis and holds shares in Novartis and GlaxoSmithKline. CP was an employee of Booz Allen Hamilton when this research was conducted. BIB and H-FZ have no conflicts of interest.

XSL•FO
RenderX

Multimedia Appendix 1

Search terms used to filter multiple sclerosis product-relevant data.

[PDF File (Adobe PDF File), 13KB - jmir_v21i8e13003_app1.pdf]

References

1. Huebner EA, Strittmatter S. Axon regeneration in the peripheral and central nervous systems. Results Probl Cell Differ 2009;48:339-351 [FREE Full text] [doi: 10.1007/400_2009_19] [Medline: 19582408]
2. Coyle PK. Management of women with multiple sclerosis through pregnancy and after childbirth. Ther Adv Neurol Disord 2016 May;9(3):198-210 [FREE Full text] [doi: 10.1177/1756285616631897] [Medline: 27134675]
3. Koch-Henriksen N, Sørensen PS. The changing demographic pattern of multiple sclerosis epidemiology. Lancet Neurol 2010 May;9(5):520-532. [doi: 10.1016/S1474-4422(10)70064-8] [Medline: 20398859]
4. Finkelsztejn A, Brooks J, Paschoal F, Fragoso Y. What can we really tell women with multiple sclerosis regarding pregnancy? A systematic review and meta-analysis of the literature. BJOG 2011 Jun;118(7):790-797 [FREE Full text] [doi: 10.1111/j.1471-0528.2011.02931.x] [Medline: 21401856]
5. Alwan S, Sadovnick AD. Multiple sclerosis and pregnancy: maternal considerations. Womens Health (Lond) 2012 Jul;8(4):399-414. [doi: 10.2217/whe.12.33] [Medline: 22757731]
6. Portaccio E, Ghezzi A, Hakiki B, Martinelli V, Moiola L, Patti F, MS Study Group of the Italian Neurological Society. Breastfeeding is not related to postpartum relapses in multiple sclerosis. Neurology 2011 Jul 12;77(2):145-150. [doi: 10.1212/WNL.0b013e318224afc9] [Medline: 21734184]
7. Foulkes MA, Grady C, Spong CY, Bates A, Clayton JA. Clinical research enrolling pregnant women: a workshop summary. J Womens Health (Larchmt) 2011 Oct;20(10):1429-1432 [FREE Full text] [doi: 10.1089/jwh.2011.3118] [Medline: 21819233]
8. Rezaallah B, Lewis DJ, Zeilhofer H, Berg B. Risk of cleft lip and/or palate associated with antiepileptic drugs: postmarketing safety signal detection and evaluation of information presented to prescribers and patients. Ther Innov Regul Sci 2019 Dec;53(1):110-119. [doi: 10.1177/2168479018761638] [Medline: 29714593]
9. Sarker A, Chandrashekar P, Magge A, Cai H, Klein A, Gonzalez G. Discovering cohorts of pregnant women from social media for safety surveillance and analysis. J Med Internet Res 2017 Dec 30;19(10):e361 [FREE Full text] [doi: 10.2196/jmir.8164] [Medline: 29084707]
10. Pierce CE, Bouri K, Pamer C, Proestel S, Rodriguez HW, van Le H, et al. Evaluation of Facebook and Twitter monitoring to detect safety signals for medical products: an analysis of recent FDA safety alerts. Drug Saf 2017 Dec;40(4):317-331 [FREE Full text] [doi: 10.1007/s40264-016-0491-0] [Medline: 28044249]
11. Sloane R, Osanlou O, Lewis D, Bollegala D, Maskell S, Pirmohamed M. Social media and pharmacovigilance: a review of the opportunities and challenges. Br J Clin Pharmacol 2015 Oct;80(4):910-920 [FREE Full text] [doi: 10.1111/bcp.12717] [Medline: 26147850]
12. Hämeen-Anttila K, Nordeng H, Kokki E, Jyrkkä J, Lupattelli A, Vainio K, et al. Multiple information sources and consequences of conflicting information about medicine use during pregnancy: a multinational Internet-based survey. J Med Internet Res 2014 Feb 20;16(2):e60 [FREE Full text] [doi: 10.2196/jmir.2939] [Medline: 24565696]
13. Powell GE, Seifert HA, Reblin T, Burstein PJ, Blowers J, Menius JA, et al. Social media listening for routine post-marketing safety surveillance. Drug Saf 2016 May;39(5):443-454. [doi: 10.1007/s40264-015-0385-6] [Medline: 26798054]
14. Freifeld CC, Brownstein JS, Menone CM, Bao W, Filice R, Kass-Hout T, et al. Digital drug safety surveillance: monitoring pharmaceutical products in Twitter. Drug Saf 2014 May;37(5):343-350 [FREE Full text] [doi: 10.1007/s40264-014-0155-x] [Medline: 24777653]
15. MedDRA. Medical Dictionary for Regulatory Activities URL: https://www.meddra.org/
16. Keller MS, Mosadeghi S, Cohen ER, Kwan J, Spiegel BM. Reproductive health and medication concerns for patients with inflammatory bowel disease: thematic and quantitative analysis using social listening. J Med Internet Res 2018 Dec 11;20(6):e206 [FREE Full text] [doi: 10.2196/jmir.9870] [Medline: 29891471]
17. Vaismoradi M, Turunen H, Bondas T. Content analysis and thematic analysis: implications for conducting a qualitative descriptive study. Nurs Health Sci 2013 Sep;15(3):398-405. [doi: 10.1111/nhs.12048] [Medline: 23480423]
18. Hämeen-Anttila K, Jyrkkä J, Enlund H, Nordeng H, Lupattelli A, Kokki E. Medicines information needs during pregnancy: a multinational comparison. BMJ Open 2013 Apr;3(4):e002594 [FREE Full text] [doi: 10.1136/bmjopen-2013-002594] [Medline: 23624989]
19. Larsson M. A descriptive study of the use of the internet by women seeking pregnancy-related information. Midwifery 2009 Feb;25(1):14-20. [doi: 10.1016/j.midw.2007.01.010] [Medline: 17408822]
20. Lagan BM, Sinclair M, Kernohan WG. Internet use in pregnancy informs women's decision making: a web-based survey. Birth 2010 Jun;37(2):106-115. [doi: 10.1111/j.1523-536X.2010.00390.x] [Medline: 20557533]
21. US Food and Drug Administration. 2014. Pregnancy, Lactation, and Reproductive Potential: Labeling for Human Prescription Drug and Biological Products — Content and Format, Guidance for Industry (Draft) URL: https://www.fda.gov/media/90160/download

XSL•FO
RenderX

22. Wang J, Johnson T, Sahin L, Tassinari MS, Anderson PO, Baker TE, et al. Evaluation of the safety of drugs and biological products used during lactation: workshop summary. Clin Pharmacol Ther 2017 Jun;101(6):736-744 [FREE Full text] [doi: 10.1002/cpt.676] [Medline: 28510297]
23. Shere M, Zhao XY, Koren G. The role of social media in recruiting for clinical trials in pregnancy. PLoS One 2014;9(3):e92744. [doi: 10.1371/journal.pone.0092744] [Medline: 24671210]
24. US Food and Drug Administration. Highlights of Prescribing Information URL: https://www.accessdata.fda.gov/drugsatfda_docs/label/2018/022527s024lbl.pdf
25. Rasmussen PV, Magyari M, Moberg JY, Bøgelund M, Jensen UF, Madsen KG. Patient awareness about family planning represents a major knowledge gap in multiple sclerosis. Mult Scler Relat Disord 2018 Aug;24:129-134 [FREE Full text] [doi: 10.1016/j.msard.2018.06.006] [Medline: 30005355]
26. Zuber PL, Moran AC, Chou D, Renaud F, Halleux C, Peña-Rosas JP, et al. Mapping the landscape of global programmes to evaluate health interventions in pregnancy: the need for harmonised approaches, standards and tools. BMJ Glob Health 2018 Oct;3(5):e001053 [FREE Full text] [doi: 10.1136/bmjgh-2018-001053] [Medline: 30364289]
27. Thomas SH, Yates LM. Prescribing without evidence - pregnancy. Br J Clin Pharmacol 2012 Oct;74(4):691-697 [FREE Full text] [doi: 10.1111/j.1365-2125.2012.04332.x] [Medline: 22607226]
28. IMI: Innovative Medicines Initiative. 2017. Topic: ConcePTION – Continuum of Evidence from Pregnancy Exposures, Reproductive Toxicology and Breastfeeding to Improve Outcomes Now URL: https://tinyurl.com/yxdty257
29. European Medicines Agency. 2019. PRAC recommends new measures to avoid valproate exposure in pregnancy URL: https://www.ema.europa.eu/en/news/prac-recommends-new-measures-avoid-valproate-exposure-pregnancy
30. Loiseau P. Clinical experience with new antiepileptic drugs: antiepileptic drugs in Europe. Epilepsia 1999;40(Suppl 6):S3-8; discussion S73 [FREE Full text] [Medline: 10530675]
31. European Medicines Agency. 2005 Nov. Guideline on the Exposure to Medicinal Products During Pregnancy Need for Post-Authorisation Data URL: https://tinyurl.com/y5l8x8a5
32. US Food and Drug Administration. 2005 Apr. Evaluating the Risks of Drug Exposure in Human Pregnancies URL: https://www.fda.gov/downloads/Drugs/Guidances/ucm071645.pdf
33. Adam MP, Polifka JE, Friedman JM. Evolving knowledge of the teratogenicity of medications in human pregnancy. Am J Med Genet C Semin Med Genet 2011 Aug 15;157C(3):175-182. [doi: 10.1002/ajmg.c.30313] [Medline: 21766440]
34. Convertino I, Ferraro S, Blandizzi C, Tuccori M. The usefulness of listening social media for pharmacovigilance purposes: a systematic review. Expert Opin Drug Saf 2018 Nov;17(11):1081-1093. [doi: 10.1080/14740338.2018.1531847] [Medline: 30285501]
35. Smailhodzic E, Hooijsma W, Boonstra A, Langley DJ. Social media use in healthcare: a systematic review of effects on patients and on their relationship with healthcare professionals. BMC Health Serv Res 2016 Aug 26;16:442 [FREE Full text] [doi: 10.1186/s12913-016-1691-0] [Medline: 27562728]

Abbreviations

ConcePTION: Continuum of Evidence from Pregnancy Exposures, Reproductive Toxicology and Breastfeeding to Improve Outcomes Now
EMA: European Medicines Agency
FDA: US Food and Drug Administration
HCPs: health care professionals
IMI: Innovative Medicine Initiative
MS: multiple sclerosis
PII: personally identifiable information
PRAC: Pharmacovigilance Risk Assessment Committee

Edited by G Eysenbach; submitted 05.12.18; peer-reviewed by K Skelton, A Lupattelli; comments to author 08.04.19; revised version received 02.06.19; accepted 29.06.19; published 07.08.19.

Please cite as:
Rezaallah B, Lewis DJ, Pierce C, Zeilhofer HF, Berg BI
Social Media Surveillance of Multiple Sclerosis Medications Used During Pregnancy and Breastfeeding: Content Analysis
J Med Internet Res 2019;21(8):e13003
URL: https://www.jmir.org/2019/8/e13003/
doi:10.2196/13003
PMID:31392963

XSL•FO
RenderX

©Bita Rezaallah, David John Lewis, Carrie Pierce, Hans-Florian Zeilhofer, Britt-Isabelle Berg. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 07.08.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Using Participatory Design Methodologies to Co-Design and Culturally Adapt the Spanish Version of the Mental Health eClinic: Qualitative Study

Laura Ospina-Pinillos[1,2], MD; Tracey Davenport[1], BA (Hons), EMBA; Antonio Mendoza Diaz[3], BPsych, PhD; Alvaro Navarro-Mancilla[4], MIPH, MD; Elizabeth M Scott[5], MBBS, FRANZCP; Ian B Hickie[1], MD, FRANZCP

[1]Brain and Mind Centre, The University of Sydney, Sydney, NSW, Australia

[2]Department of Psychiatry and Mental Health, Pontifical Javeriana University, Bogota, Colombia

[3]School of Psychiatry, Department of Medicine, University of New South Wales, Sydney, NSW, Australia

[4]Neuropsychiatry Research Group, Autonomous University of Bucaramanga, Bucaramanga, Colombia

[5]School of Medicine, University of Notre Dame Australia, Sydney, NSW, Australia

Corresponding Author:
Laura Ospina-Pinillos, MD
Brain and Mind Centre
The University of Sydney
Shops 1-3, 66-70 Parramatta Road
Sydney, NSW, 2051
Australia
Phone: 61 028 627 6946
Email: laura.ospinapinillos@sydney.edu.au

Related Article:

This is a corrected version. See correction statement: http://mhealth.jmir.org/2022/3/e37679/

Abstract

Background: The Mental Health eClinic (MHeC) aims to deliver best-practice clinical services to young people experiencing mental health problems by making clinical care accessible, affordable, and available to young people whenever and wherever they need it most. The original MHeC consists of home page with a visible triage system for those requiring urgent help; a online physical and mental health self-report assessment; a results dashboard; a booking and videoconferencing system; and the generation of a personalized well-being plan. Populations who do not speak English and reside in English-speaking countries are less likely to receive mental health care. In Australia, international students have been identified as disadvantaged compared with their peers; have weaker social support networks; and have higher rates of psychological distress. This scenario is acquiring significant relevance as Spanish-speaking migration is rapidly growing in Australia, and the mental health services for culturally and linguistically diverse populations are limited. Having a Spanish version (MHeC-S) of the Mental Health eClinic would greatly benefit these students.

Objective: We used participatory design methodologies with users (young people aged 16-30 years, supportive others, and health professionals) to (1) conduct workshops with users to co-design and culturally adapt the MHeC; (2) inform the development of the MHeC-S alpha prototype; (3) test the usability of the MHeC-S alpha prototype; (4) translate, culturally adapt, and face-validate the MHeC-S self-report assessment; and (5) collect information to inform its beta prototype.

Methods: A research and development cycle included several participatory design phases: co-design workshops; knowledge translation; language translation and cultural adaptation; and rapid prototyping and user testing of the MHeC-S alpha prototype.

Results: We held 2 co-design workshops with 17 users (10 young people, 7 health professionals). A total of 15 participated in the one-on-one user testing sessions (7 young people, 5 health professionals, 3 supportive others). We collected 225 source documents, and thematic analysis resulted in 5 main themes (help-seeking barriers, technology platform, functionality, content, and user interface). A random sample of 106 source documents analyzed by 2 independent raters revealed almost perfect agreement for functionality (kappa=.86; P<.001) and content (kappa=.92; P<.001) and substantial agreement for the user interface (kappa=.785;

XSL•FO
RenderX

P<.001). In this random sample, no annotations were coded for help-seeking barriers or the technology platform. Language was identified as the main barrier to getting medical or psychological services, and smartphones were the most-used device to access the internet. Acceptability was adequate for the prototype's 5 main elements: home page and triage system, self-report assessment, dashboard of results, booking and video visit system, and personalized well-being plan. The data also revealed gaps in the alpha prototype, such as the need for tailored assessment tools and a greater integration with Spanish-speaking services and communities. Spanish-language apps and e-tools, as well as online mental health information, were lacking.

Conclusions: Through a research and development process, we co-designed and culturally adapted, developed and user tested, and evaluated the MHeC-S. By translating and culturally adapting the MHeC to Spanish, we aimed to increase accessibility and availability of e-mental health care in the developing world, and assist vulnerable populations that have migrated to English-speaking countries.

(J Med Internet Res 2019;21(8):e14127) doi:10.2196/14127

KEYWORDS

telemedicine; medical informatics; eHealth; mental health; cultural characteristics; cultural competency; ethnic groups; transients and migrants; quality of health care; international students; Hispanics; Latinos; community-based participatory research; primary health care; patient participation; patient preference; patient satisfaction; consumer health information

Introduction

Background

The need for mental health services far outweighs the capacity of service providers all over the world [1]. Access to adequate-quality mental health care is also limited for many populations but is particularly limited for vulnerable groups such as the elderly and youth populations, racial and ethnic minorities, the socioeconomically disadvantaged, and rural populations [2]. Limited access to services is of particular concern for young people, as it is well established that 75% of the serious mental diseases and substance use problems emerge before 25 years of age [3]. When young people do seek and receive help, timely and evidence-based treatments are encountered by only a small proportion; in some low- and middle-income countries, the treatment gap can be as high as 90% [4].

Populations who do not speak English in English-speaking countries are less likely to receive mental health care [5]. In Latino populations with mental health problems, the lack of English proficiency is one of the biggest barriers when accessing services [6]. In Australia, non–English-speaking migrant populations struggle to access and understand the local health care system [7]. Language proficiency has been identified as a true barrier for migrant men when using services [8].

International Students

Australia is a popular study destination for students around the world. Most of Australia's international students are enrolled in the higher education sector (44%), followed by the vocational education and training sector (27%), and English-language intensive courses for overseas students (19%) [9]. Studying abroad can be one of the most remarkable and rewarding experiences, but it can also be a source of great distress. The way migration is experienced by each individual highly depends on the push and pull factors that precipitated the migration [10]. In the case of international students, a high motivation to study in a different country can act as a protective factor, but the cultural distance of the host country, the lack of social support, and academic pressure can be powerful stressors. Consequently, several studies have shown increased rates of mental health problems in this population [10-12].

In Australia, international students have been identified to be disadvantaged compared with their peers; have weaker social support networks; have higher rates of psychological distress [13]; and are at higher risk of experiencing an adjustment disorder or other mental health problems [14]. The "International Student Welfare in Australia" report suggested that Australia does not adequately protect international students' human rights and highlights mental health as a key area of concern [15]. Recently, awareness of these issues has been covered by Australian mass media due to 27 suicides in the international student population between 2009 and 2015; sadly, all were reportedly associated with low help-seeking behaviors (22%) [16]. As international students are less likely to receive mental health care, the previously mentioned report encourages institutions to provide information, including available services and increased research in this area. However, most of the research has been focused on tertiary education students and none or very little has been dedicated to language students.

This scenario is acquiring a significant relevance in Australia, where the Spanish-speaking (including Latino) international student migration is rapidly growing. In 2016-2017, language student visa grants (subclass 570) increased by 16.8%, where 3 Spanish-speaking countries (Colombia, Spain, and Chile) were situated in the top 10 countries of applications logged outside Australia [17].

Health Information Technologies

The internet and new and emerging technologies hold enormous promise for significantly expanding the reach of adequate-quality mental health care by addressing several barriers [18]. Interventions delivered through these technologies have the potential to reach a wide geographic area via remote delivery of care [19]; decrease costs in delivering self-help and social networking interventions; and allow for relatively rapid, centralized scaling up of interventions to a public health dissemination level [20]. English-language, Web-based mental health interventions have proven to be effective for self-screening and referral [6], reducing symptoms and

XSL•FO
RenderX

delivering effective treatment for major mental health disorders [21]. A large number of studies, including randomized controlled trials, have also demonstrated the effectiveness of various internet-delivered interventions, such as psychotherapy and psychoeducation [22], treating problematic health behaviors [23], and delivering prevention and treatment programs [24]. Other population-based studies have reported that Web-based tools can enhance the delivery of mental health care in primary care settings [25] and support training and supervision for providers [26]. The number of programs available is growing rapidly [27]. Although positive results are seen from the use of self-directed electronic health interventions, increased effectiveness has been reported if they are used as part of a stepped-care model [28], with the support of a trained health professional [29] or as an adjunct to face-to-face treatment [30].

Despite the growth of such technologies in high-income countries, these technologies are still lacking in low- and middle-income countries and, more specifically, in the Spanish language [31]. Traditional telemedicine has supported the cooperation between developed and developing countries to deliver care across borders by linking professionals rather than providing direct connection between professionals and patients [32]. Telepsychiatry has been used to deliver mental health care to individuals requiring attention, not only locally [33-35], but also internationally, as a means to deliver care to Spanish-speaking individuals residing in a different country [32,36,37]. This type of care is a more efficient alternative, as it doesn't require the use of interpreters and is culturally sensitive [38]. Successful Spanish-language health information technology (HIT) interventions have been applied in several fields, such as cancer; diabetes; and child, infant, or maternal health [39]. Despite this, the HITs available for mental health are scarce. Initial reports have demonstrated their potential utility in the screening of mental health problems [40], as well as in the treatment of depression [31,41,42], anxiety [43], and substance use disorders [44].

Although the development of HITs in Spanish is recent, their usability and retainability among users is of concern [31]. Experience in other languages (mostly English) has demonstrated that participatory design research methodologies that involve stakeholders and end users in the design and development of these systems at all stages could finally increase user engagement and system usability [45-48]. A close collaboration with end users ensures the appropriateness of these systems for culturally and linguistically diverse populations [49]. Therefore, incorporating participatory design research methodologies that puts end users at the center of the design and development process is greatly needed for Spanish-language–based HITs.

Objectives

The University of Sydney's Brain and Mind Centre (Sydney, Australia) is a leader in the development of youth-specific mental health services [50,51] and evidence-based electronic health technologies to engage young people in their own care [52]. The Mental Health eClinic (MHeC) [48,53] was a prototypic Web-based tool designed and developed through a partnership between the Young and Well Cooperative Research Centre and the Brain and Mind Centre. The MHeC aimed to deliver best-practice clinical services to young people experiencing mental health problems by making clinical care accessible, affordable, and available to young people whenever and wherever they need it most. The original MHeC had 5 main elements: a home page with a visible triage system for those requiring urgent help; a comprehensive online physical and mental health self-report assessment; a detailed dashboard of results; a booking and videoconferencing system to enable video visits; and the generation of a personalized well-being plan that included links to evidence-based, young person–suggested, health professional–recommended apps and e-tools [53]. We hypothesized that having a Spanish version of the MHeC (MHeC-S) could greatly benefit young people who are native Spanish speakers living in Australia and who are actively seeking help.

Using a research and development cycle (including several participatory design phases) with end users (young people aged 16 to 30 years, supportive others [such as family, friends, caregivers, coaches, teachers, or community members], and health professionals) as a framework, in this study we aimed to (1) conduct co-design workshops with end users to co-design and culturally adapt the MHeC for Spanish-speaking young people based in Australia; (2) inform the development of the alpha prototype of the MHeC-S; (3) test the usability of the alpha prototype of the MHeC-S; (4) translate, culturally adapt, and face-validate the self-report assessment to a Spanish-speaking population based in Australia; and (5) collect information to inform the beta prototype of the MHeC-S.

Methods

Participants

Participants included community-based young people aged 16 to 30 years; native Spanish speakers living in Australia; and native Spanish-speaking young people attending headspace Camperdown and headspace Campbelltown (headspace Australia's National Youth Mental Health Foundation provides early intervention mental health services and assistance in enhancing young peoples' [aged 12-25 years] well-being; Camperdown and Campbelltown are 2 sociodemographic areas of Sydney, Australia). Additionally, native Spanish-speaking health professionals and supportive others participated. Participants were required to have regular access to a smartphone (with the iOS or Android operating system) and the internet.

The University of Sydney's Human Research Ethics Committee approved the study (Protocol No. 2014/689 for the co-design workshops and Protocol No. 2016/487 for the user testing sessions). Participants were provided with the relevant information about the study (participant information statement) before providing their consent and participating in the study. We also obtained parental consent for participants under 18 years of age. Young people received gift vouchers to thank them for their time and expertise when they attended the co-design workshops and the user testing sessions.

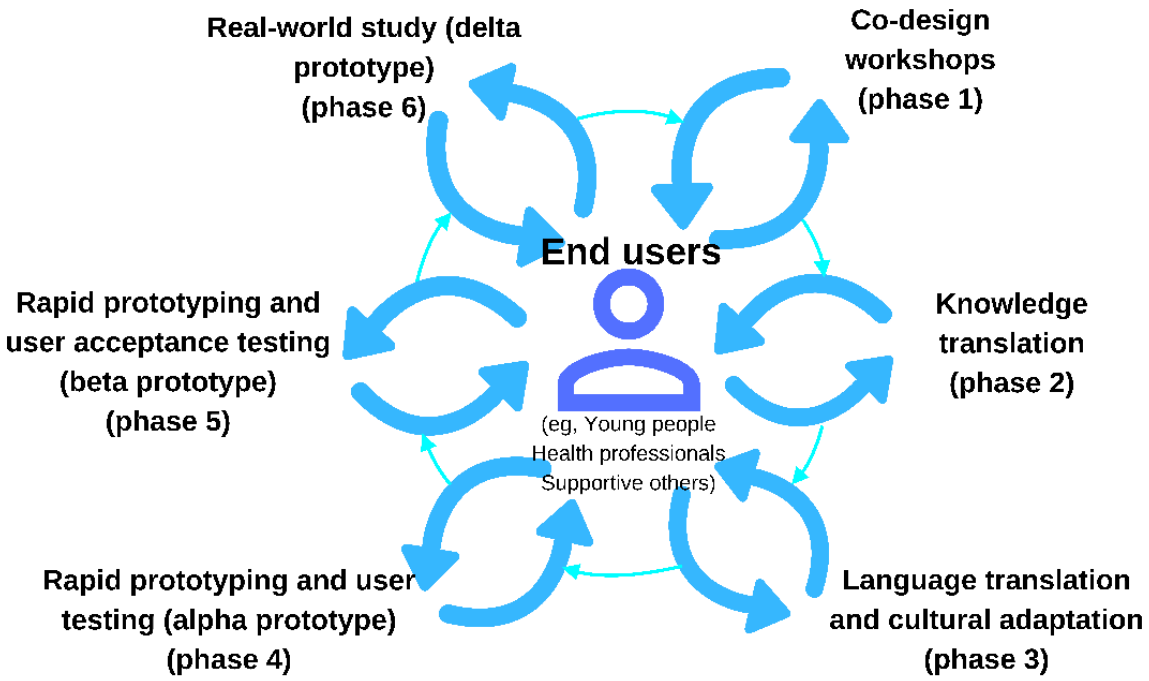
XSL•FO
RenderX

The recruitment strategy included the identification of potential participants through headspace Camperdown and headspace Campbelltown; poster and postcard advertisements displayed in community organizations; Facebook advertisements and a study-specific Facebook page; use of organizational social media channels; universities, institutes of technical and further education, language schools, and vocational and training institutes; and cooperation with Spanish-speaking consulates in Australia.

Procedure

We based the participatory design research methodology on the Young and Well Cooperative Research Centre's guide *Participatory Design of Evidence-Based Online Youth Mental Health Promotion, Intervention and Treatment* [54]. The research and development cycle used our previously established phases for co-design and build of the original version of the prototypic MHeC [53]. The process encompasses several participatory design phases: co-design workshops (phase 1); knowledge translation (phase 2); language translation and cultural adaptation (phase 3); rapid prototyping of the alpha prototype and user testing (phase 4); rapid prototyping and user (acceptance) testing of the beta prototype (phase 5); and real-world trialing of the final prototype (phase 6). This paper reports the initial 4 phases; we will report phase 5 and phase 6 separately (Figure 1).

Figure 1. Research and development cycle of the Spanish version of the Mental Health eClinic.



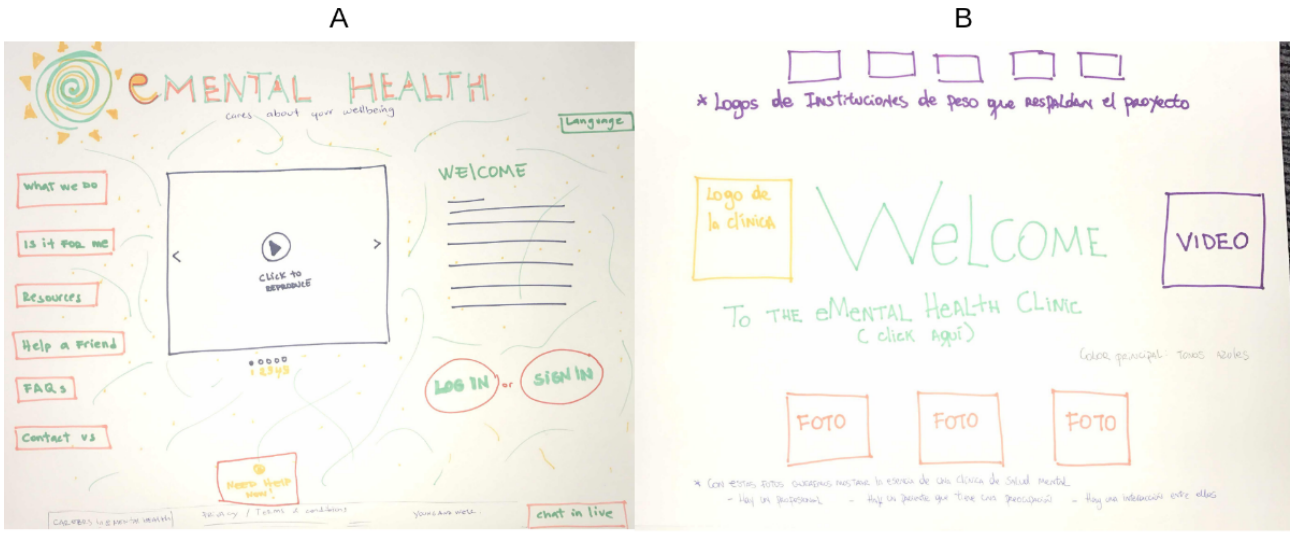
Phase 1: Co-Design Workshops

We held co-design workshops in 2 stages with a maximum of 10 participants per workshop; we ran these with young people and health professionals separately. We did not use technology in the workshops; instead, we conducted the following design activities to facilitate the process: we used design testing using mockups and end user sketching (Figure 2) to obtain information for the content, functionality, and the look and feel of the prototype.

The topics covered in each workshop included defining the advantages and disadvantages of having a Spanish version of the MHeC; defining the barriers of having an MHeC-S; assessing the 5 main elements of the MHeC; and defining the functionality and the user interface. At the end of each workshop, the knowledge translation team analyzed and synthetized the information.

XSL•FO
RenderX

Figure 2. Samples of end user sketches made during a co-design workshop. (A) Hand-drawn sketch by a young person. (B) Hand-drawn sketch by a health professional.



Phase 2: Knowledge Translation

The knowledge translation team independently analyzed the diagrams and notes taken from the previous phase (or workshop), then compared and discussed observations until they reached an agreement. They then synthesized the information by creating wireframes that would be used in the subsequent phase.

Phase 3: Language Translation and Cultural Adaptation

Language of the Prototype

For all of the prototype's language, a native Spanish-born psychiatrist (LOP) performed a simple translation. Then, for general-content items, 2 Spanish-born psychologists (not involved in the publication of this paper) reviewed the translations. A second Spanish-born psychiatrist (ANM) also reviewed specific mental health content or sensitive content such as the dashboard of results and psychoeducational factsheets. Discrepancies between the translations were resolved by consensus in the group.

Translation and Cultural Adaptation of the Self-Report Assessment

Understanding the great need for health professionals and researchers to have available, reliable, and valid measures across different languages and populations, we aimed to translate, back-translate, culturally adapt, and face-validate the Australian self-report assessment using a modified version of the "user-friendly guideline for the translation, adaptation and validation of instruments or scales for cross-cultural health care research" developed by Sousa and Rojjanasrirat [55].

Two Spanish bilingual health professionals (LOP and ANM) independently translated all (health-related) items from English to colloquial Spanish, with the exception of standardized surveys already available in Spanish. A third native Spanish-born psychologist (AMD) reviewed the translated versions and the original English versions. Then, in group discussions, we ensured that all items were linguistically and culturally appropriate by assessing the face-validity of each item in the self-report assessment, as well as assessing the readability and grammatical consistency of the entire assessment. All items were then back-translated to English by one Australian adult (not involved in the publication of this paper) who is fluent in colloquial Spanish and is based in Colombia, has extensive research experience and tertiary qualifications in health, and is accredited to teach English to adults. Discrepancies between the original and back-translated versions were resolved in group sessions between the translators and back-translator.

A literature review was undertaken (by LOP) to identify relevant measures for this population, as well as those instruments already translated and validated into Spanish. The review included both published (identified via PubMed, Google Scholar, SciELO, and LILACS) and gray literature (identified via Google Advanced search) in both English and Spanish. Understanding that some questionnaires might have several translations or versions, we established the following process to select the instruments: first, we selected official translations; if these were unavailable, we selected versions of the published translation and psychometric processes. When more than 1 version or source was available, the 2 previously mentioned health professionals (LOP and ANM) by consensus selected the most appropriate to be included.

Phase 4: Rapid Prototyping and Usability Testing of the Alpha Prototype

Phase 4 involved user testing with end users: young people, health professionals, and supportive others. Sessions used laptops, tablets, and smartphones (with the iOS or Android operating system), where participants had access to the alpha prototype of the MHeC-S. In each 90-minute one-on-one user testing session, a researcher was paired with an end user. Using a think-aloud protocol [56], participants were observed as they navigated the prototype and responses were recorded to questions posed by the researcher about the main elements of the MHeC-S. A total of 4 usability tasks were completed in the session: (1) create an account and log in, (2) find the Need Help Now button, (3) explore the dashboard of results, and (4) book

XSL•FO
RenderX

an appointment. Task completion (yes/no) and task difficulty were measured using the Single Ease Question (SEQ; responses ranged from "very difficult" to "very easy," rated from 1 to 7) [57]. User testing also explored the prototype's utility and the users' inclination to use the MHeC-S, overall comments, and naming of the prototype. Interviews and observations were transcribed. No instructions or clues were provided, and all responses and observations (eg, nonverbal cues) were transcribed.

Data Analysis

We simultaneously collected and analyzed data at the end of each phase in order to facilitate the process. Hence, we explored preliminary findings in the following phase. We determined interrater reliability and analyzed task difficulty using IBM SPSS Statistics for Mac 22.0 (IBM Corporation). We uploaded and analyzed source documents (workshop discussion notes, artifacts [mockups and end user sketches] and user testing notes) using thematic analysis techniques [58] in NVivo 11 for Mac (QSR International) [59]. The thematic analysis framework involved both inductive and deductive coding. Acknowledging that one of the biggest challenges in this project was the translation and cultural adaptation of the MHeC-S, in the deductive code framework we considered the adaptation of the prototype in 4 dimensions: technology platform, functionality, content, and user interface. As defined by Valdez et al in their culturally informed design framework [60], "technology platform" refers to the different types of hardware, "functionality" indicates the actions that can be performed, "content" refers to the message that is transmitted, and "user interface (design)" refers to the visual presentation of the content and functionality. We also enriched this type of coding with our previously established [53] codes (general elements of the MHeC; general look and feel; privacy and data sharing; and interaction of the MHeC with social networks).

Data collection and qualitative analysis were done in Spanish. To facilitate reporting, we provide quotes translated from the original data. Multimedia Appendix 1 shows the original quotes in Spanish.

One researcher coded all the material (coder A: LOP) and a second coder (coder B: not involved in the publication of this paper) reviewed half of the collected documents in order to assess the reliability, assess consistency, and reduce potential bias [61]. We calculated interrater agreement for each theme using the Cohen kappa statistic on a binomial distribution (category present vs category not present) for each of the themes [62] and interpreted the obtained values using Viera and Garrett's criteria: kappa range .01 to .20 indicates slight agreement, kappa range .21 to .40 indicates fair agreement, kappa range .41 to .60 indicates moderate agreement, kappa range .61 to .80 indicates substantial agreement, and kappa range .81 to .99 indicates almost perfect agreement [63].

Results

Workshops and Sessions

In May 2015, we conducted 1 full-day co-design workshop with Spanish-speaking young people based in Australia and 1 full-day co-design workshop with Spanish-speaking health professionals based in Australia. In total, we conducted 3 knowledge translation sessions immediately after each workshop and at the end of the usability testing. The general-content translation process started in June 2015, and the self-report assessment literature review and translation process started in January 2016 and lasted until the end of the same year. We conducted 15 one-on-one user testing sessions between March and November 2017.

Participant Characteristics

A total of 10 young people participated in the co-design workshops; 8 were female and their ages ranged from 17 to 29 years (median age 24 years). Of the young participants, 8 were Colombian and 2 were Chilean. A total of 7 health professionals participated in the workshops; 6 were female and their ages ranged from 22 to 34 years (median age 28 years). Of the health professionals, 3 were from Colombia, 2 were from Chile, and 2 were from Spain.

A total of 15 participants participated in the one-on-one user testing sessions: 7 young people with ages ranging from 19 to 30 years (median age 26 years); 5 health professionals with ages ranging from 27 to 74 years (median age 35 years); and 3 supportive others with ages ranging from 30 to 57 years (median age 30 years). Of these participants, 10 were female and 12 were Colombian, while the rest were from Argentina, Spain, and Venezuela.

Thematic Analysis

We collected and analyzed a total of 225 source documents (2 workshop discussion notes and 208 artifacts produced by participants were collected in the co-design workshops plus 15 user testing notes) during the entire process.

Coding Interrater Reliability

Using inductive coding, 1 new main theme emerged (help-seeking barriers) and, from the deductive coding framework, 4 main themes were reiterated (technology platform, functionality, content, and user interface). Of the 225 source documents, 106 (47.1%) were analyzed by both raters. A total of 378 annotations were recoded from both coders (coder A and coder B). Interrater agreement of functionality theme between coder A and coder B was "almost perfect" (kappa=.86; $P<.001$), with concordance in a total of 93.7% (354/378) of the annotations. Similarly, we obtained an "almost perfect" agreement (kappa=.92; $P<.001$) between raters in relation to the content theme, with 97.6% (369/378) of total concordance. In relation to interface, interrater agreement between coders was "substantial" (kappa=.785; $P<.001$), with concordance in a total of 90.0% (340/378) of the annotations. In this random sample, no annotations were coded for help-seeking barriers or technology platform themes.

Help-Seeking Barriers

Within this domain, participant perceptions of the help-seeking barriers fell into 3 categories: the language barrier, problems recognizing symptoms or poor mental health literacy, and the availability and accessibility of sources of help.

XSL•FO
RenderX

Language Barrier

All participants (32/32, 100%) highlighted language as the main barrier to getting medical or psychological services (Multimedia Appendix 1):

> *...even if I needed to call 000, I wouldn't be sure if they understand what I'm saying...* [Young person, quote A]
>
> *I don't think that I would be able to explain my feelings to someone in English...* [Young person, quote B]

As the aim for most of these students was to learn English (or improve their English level), their communication skills were, in general, limited. This was a source of distress, as they felt limited in their day-to-day living:

> *...it's very hard to arrive here (Australia) and not understand what is happening...* [Young person, quote C]
>
> *...understanding simple instructions—like where is the train stop—is very difficult...* [Young person, quote D]

For some, the language barrier could have a very negative impact on their confidence:

> *...it's like in English I'm a different person; sometimes I feel people think I'm dumb...* [Young person, quote E]
>
> *...the impact on their [international students] confidence is huge. Sometimes I have to remind him [international student] what he is capable of...* [Health professional, quote F]

Problems Recognizing Symptoms or Poor Mental Health Literacy

International students face different issues during migration that could have an impact on their well-being. Participants felt concerned for those who have recently arrived in Australia, as they are perceived as being more vulnerable. According to these participants, a great majority experienced some degree of cultural shock upon their arrival; getting used to regular things such as food, climate, and transport can be relevant stressors among students.

> *...it is hard to try to fit, and try to understand how things work here...* [Young person, quote G]

As Australia's cost of living is high, all participants reported economic concerns (32/32, 100%), whereas some (17/32, 53%) experienced difficulties with housing, getting a job (or a job with fair work conditions), or establishing relationships with peers. All these factors put the students at a higher risk of adaptational problems, which are often unnoticed.

Additionally, the conditions of migration greatly affect individuals' experience in a new country. Some common negative factors were visiting another country for the first time, travelling alone, and not having family members or friends already residing in that country. Most international students need to work to pay their expenses; however, the jobs they find to support their stay are often not related to their already acquired skills, as a young person explained:

> *...the majority [of] us have Bachelor degrees in our home countries...so when we arrive in Australia, the jobs we find are very different from what we have been trained in—most of us have to work cleaning, or as a waitress or in construction* [Young person, quote H]

For many young people, reconciling this discrepancy is challenging.

Availability and Accessibility of Sources of Help

All 12 health professionals and all 3 supportive others believed that international students have a great need for Spanish-language–based mental health services in Australia. They perceived that the cases of young people requiring help are increasing, as a supportive other explained:

> *...possibly one international student dies by suicide every year here, and more frequently we have to provide assistance to students that are hospitalized for a mental health concern...* [Supportive other, quote I]

Young people believed that having an MHeC-S would be of great utility, as they struggle to understand Australia's health system and are not aware of their Overseas Student Health Cover benefits. All young participants (17/17, 100%) knew Australia's national emergency phone number (000). However, just a few (7/17, 41%) of them understood where to go if they needed nonurgent medical care, and all of them stated that they didn't know where to get psychological assistance.

All 17 young people said they would use a system like the MHeC-S, as they felt it would be a tool to increase mental health awareness and access to sources of help. Additionally, students believed they would be more inclined to use the MHeC-S if they knew about it beforehand, perhaps in the information they receive before arriving in Australia. All 12 health professionals imagined the prototype acting as a bridge between established services and centers in Australia such as the Transcultural Mental Health Centre; New South Wales (NSW) Service for the Treatment and Rehabilitation of Torture and Trauma Survivors; Translating and Interpreting Service; NSW Spanish and Latin American Association for Social Assistance; other relevant cultural associations; and diplomatic missions.

In relation to online sources of help, participants stated that Google was their main source for getting information about their health symptoms. However, they did not necessarily trust all the information they obtained. Participants agreed that there is a shortage of Spanish-language online information (from reputable sources such as universities and organizations) and, more specifically, trustworthy apps and e-tools.

Technology Platform

All participants (32/32, 100%) reported that they had constant access to the internet via mobile data plans or Wi-Fi networks. The most commonly used device to access the internet was a smartphone (32/32, 100%), followed by a laptop. All participants agreed that the MHeC-S needs to be accessible via a mobile

XSL•FO
RenderX

device in order to really respond to this population's needs, as some of the students did not have a desktop, laptop, or tablet. All 17 young people reported that mobile phones and the internet were necessary tools in this period of their life, as they used them to communicate with English-speaking people and keep in contact with family and friends overseas. Additionally, they used them in their daily activities (eg, a global positioning system feature), or as a way to find a job or accommodation. As a consequence, they highlighted the importance of the MHeC-S having a responsive Web design, where the prototype needs to work on mobile devices; otherwise, access would be jeopardized.

Functionality

There was adequate acceptability of the 5 main elements of the MHeC-S: a home page with a visible triage system for those requiring urgent help; a comprehensive online physical and mental health self-report assessment; a detailed dashboard of results; a booking and videoconferencing system to enable video visits; and the generation of a personalized well-being plan that includes links to evidence-based, young person–suggested, health professional–recommended apps and e-tools.

Element 1: Home Page and Triage System

When shown the home page, participants (15/32, 47%) suggested that the MHeC-S webpage's domain should be ".com" or ".org," as this would increase the website's credibility. At the same time, they wanted the home page to display all relevant logos of affiliated organizations such as the logo of the University of Sydney and relevant Latin American or Spanish universities associated with the MHeC-S. In this space, they wanted to find a simple description of "...what the MHeC-S has to offer..." (young person, quote J) and perhaps a series of short videos that explain more about the MHeC-S and also contained testimonials. As language might be a concern, participants suggested adding a settings cog on the home page so they could choose their language and, consequently, relevant content would also be prompted. The triage system was widely accepted, as all users understood the need for screening for urgent services and for rapid referral of users (Figure 3).

Figure 3. Home page and triage system.

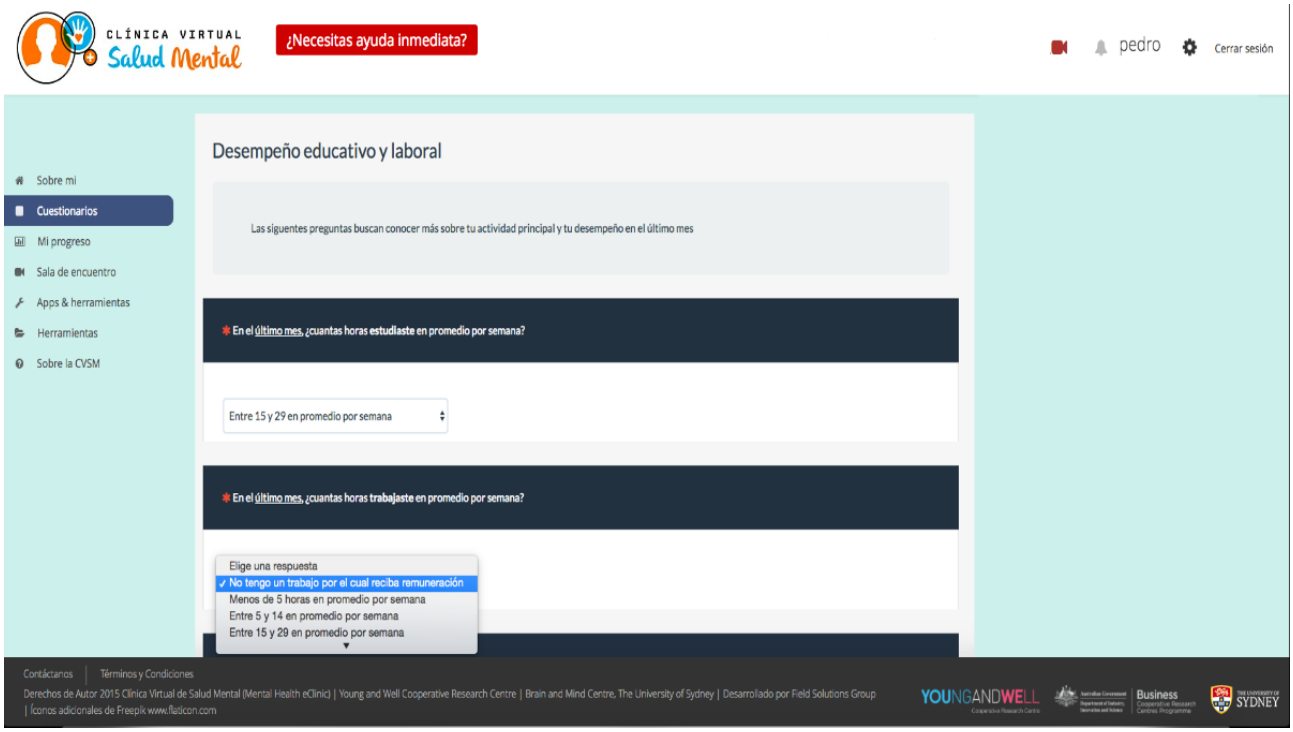
Element 2: Online Physical and Mental Health Self-Report Assessment

When shown the assessment (via questionnaire) (Figure 4), participants liked that the online physical and mental health assessment used rule-based decision algorithms that enable personalization of the assessment to the young person's needs (eg, sex-specific questions or in-depth assessments according to positive screening responses) and inform the dashboard of results. They also accepted the established features of pausing and resuming later, as they would give participants more flexibility to complete the assessment where and when they prefer. Additionally, participants approved the type of questions (eg, Likert-type scale questions and 2-way closed-ended questions) contained in this assessment. However, health professionals (12/12, 100%) suggested including 1 open-text question at the beginning of the assessment with the aim of assessing the reason for accessing the MHeC-S that day, as one clinician explained:

> *I would like to know the reason [why the young person was] visiting the MHeC-S...as we do in practice assessing the presenting or chief complaint...* [Clinician, quote K]

XSL•FO
RenderX

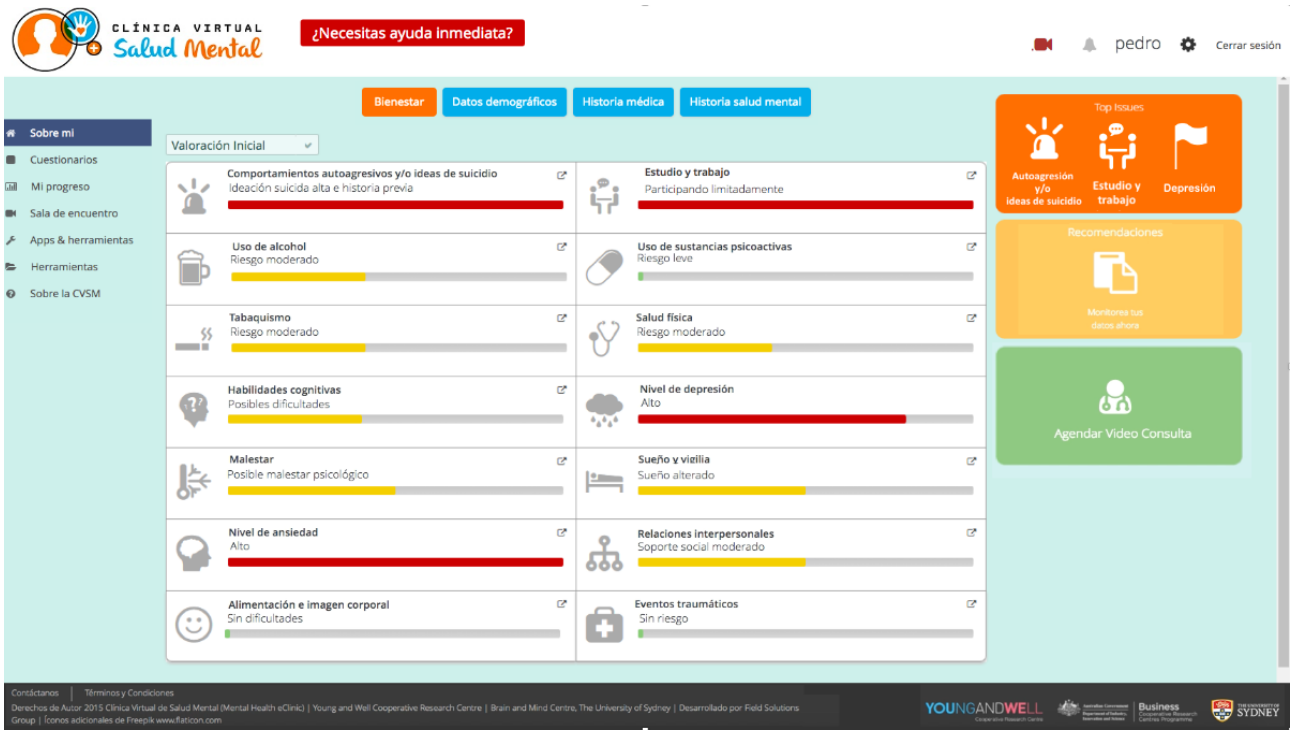
Figure 4. Online physical and mental health self-report assessment.



Element 3: Dashboard of Results and Progress

All 32 participants agreed that after completion of the online self-report assessment a dashboard of results should be displayed immediately (Figure 5). Participants accepted the simple bar and line graphs, colored icons, and traffic light representations, and reported that they were easy to understand. Health professionals (12/12, 100%) believed that the assessment and the dashboard of results were useful tools to inform their practice, not only in their first assessment but also as an ongoing form of care. In relation to the dashboard's language, participants preferred the use of lay terms instead of medical terminology. When medical jargon is needed (eg, psychosis or hypomania), participants proposed that the prototype should display a simple explanation of the term when they click on the word or hover over it.

Figure 5. Dashboard of results and progress.



XSL•FO
RenderX

Element 4: Booking System and Video Visit System

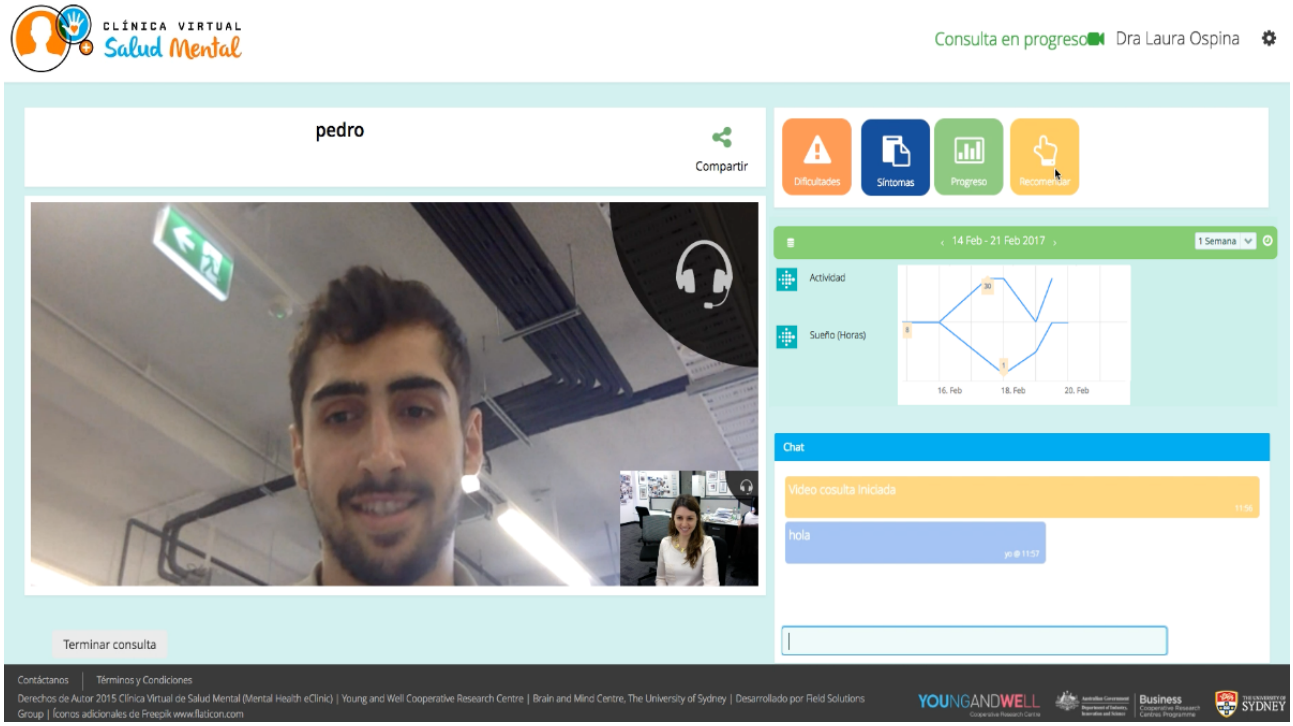
All participants agreed that the booking and video visit system embedded in the MHeC-S (Figure 6) was a secure way of protecting privacy. Due to the limited numbers of Spanish-speaking health professionals in Australia, participants believed that having a video visit with a Spanish-speaking health professional would be an effective way of screening and assessment, as well as providing (and receiving) advice, treatment, and therapy. Importantly, they acknowledged that video visits would be more efficient, as this would save them time and money, as a young person explained:

> *...we will know exactly where to go and not to waste time going from one place to another, searching for someone that understands me...* [Young person, quote L]

Furthermore, the prototype provided them with security, as a health professional explained:

> *...they can always know where to go, like a secure base...* [Health professional, quote M]

Figure 6. Booking system and video visit system.



Element 5: Personalized Well-Being Plan That Includes Links to Evidence-Based, Young Person–Suggested, Health Professional–Recommended Apps and E-Tools

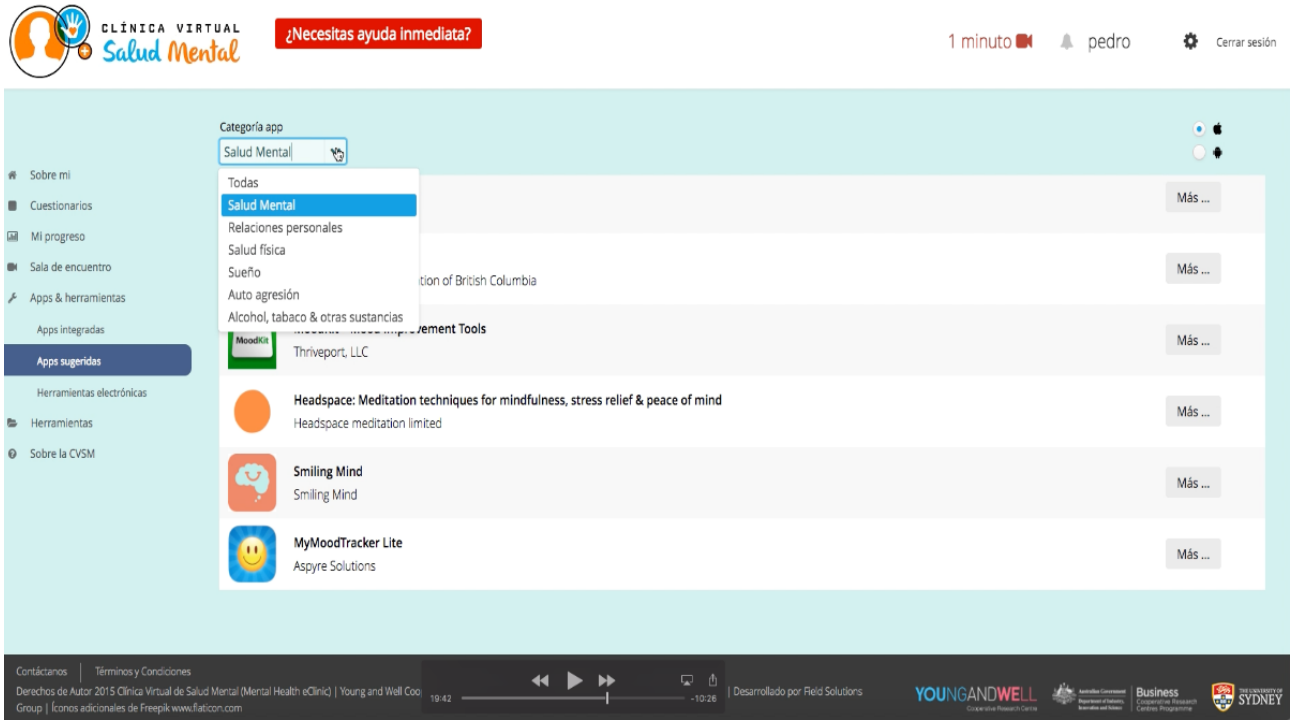
The idea of having a tailored plan and recommendations immediately after the completion of the self-report assessment was widely accepted by participants (32/32, 100%) (Figure 7). Young people (17/17, 100%) said that they would be likely to download and use the recommended apps and e-tools if those matched with their needs. However, most of the participants (29/32, 91%) highlighted a lack of Spanish-language apps and e-tools . Young people (17/17, 100%) said that they would try to use an app in English, but they also recognized that their experience and the benefit would be limited, especially for those apps that have audio resources, as one young person explained:

> *I would try to use it as much as I can, but I think there are going to be many things I don't understand—for example, the mindfulness audios...* [Young person, quote N.]

As potential solutions, participants proposed the creation of videos that contain general information, as well as relaxation and breathing exercises; a detailed directory that describes available English apps and e-tools; and, ideally, the development of several Spanish-language apps, e-tools, and Spanish adaptations of the best evidence-based resources.

XSL•FO
RenderX

Figure 7. Personalized well-being plan that includes links to evidence-based, young person–suggested, health professional–recommended apps and e-tools.



Content: Translation and Cultural Adaptation of the Self-Report Assessment

The self-report assessment included 16 modules (Table 1 [64-89]) with smart skips built in so that it was tailored to each individual and took a minimum amount of time to complete (approximately 45 minutes).

The self-report assessment translation process started in early 2016, with the literature review. We found 8 Spanish-translated versions of measures from the original source: the 2-step method to measure transgender identity [90], 10-item Kessler Psychological Distress Scale [91], Quick Inventory of Depressive Symptomatology [92,93], Community Assessment of Psychic Experiences [94,95], Alcohol Use Disorders Identification Test [96], Alcohol, Smoking and Substance Involvement Screening Test [97], International Physical Activity Questionnaire [98,99], and the Spanish version of the World Mental Health Composite International Diagnostic Interview used in the National Comorbidity Survey Replication Adolescent Supplement [100,101]. We selected 5 because we found their translation and psychometric properties in the academic literature: the Primary Care Posttraumatic Stress Disorder (PTSD) Screen [102], PTSD Checklist-Civilian Version [103], Altman Self-Rating Mania Scale [104], the Cutting down, Annoyance by criticism, Guilty feeling, and Eye-openers questionnaire (which has been widely used in several Spanish-speaking studies and several versions are available online [105,106]; we used the Colombian version for its methods [107]), and the Fagerström Test for Nicotine Dependence [108]. Although we didn't find any versions of the empathy quotient scale in the academic literature, we found a Spanish version provided by the Autism Research Centre of the University of Cambridge [109].

XSL•FO
RenderX

Table 1. Self-report assessments in each of the 19 modules.

Module		Questionnaires
1.	Main reason for visiting the MHeC-S[a]	Short open-text question
2.	General demographics	Items derived from the Second Australian Young and Well National Survey [64] and the 2-step method to measure transgender identity [65]
3.	Social and occupational function	Modified versions of the Brief Disability Questionnaire [66] and the self-report version of the Social and Occupational Functioning Assessment Scale [67]
4.	Psychological distress	10-item Kessler Psychological Distress Scale [68]
5.	Depressed mood	Quick Inventory of Depressive Symptomatology (QIDS-SR-16) [69]
6.	Anxiety	Overall Anxiety Severity and Impairment Scale [70]
7.	Mania-like experiences	Items derived from the Altman Self-Rating Mania Scale [71]
8.	Psychosis-like experiences	Items derived from the Community Assessment of Psychic Experiences-Positive Symptoms Scale [72]
9.	Traumatic experiences	Primary Care PTSD[b] Screen [73] and the PTSD Checklist-Civilian Version [74]
10.	Self-harm behaviors and suicidal ideation	Suicidal Ideation Attributes Scale [75]
11.	Tobacco, alcohol, and substance use	Items adapted from the Alcohol Use Disorders Identification Test [76], the Alcohol, Smoking and Substance Involvement Screening Test [77], the Cutting down, Annoyance by criticism, Guilty feeling, and Eye-openers questionnaire [78], the Drinking Motives Questionnaire [79], the Fagerström Test for Nicotine Dependence [80], and selected items from the National Drug Strategy Household Survey [81]
12.	Physical activity	International Physical Activity Questionnaire [82]
13.	Sleep behaviors	Sleep-related items from the QIDS-SR-16
14.	General mental health conditions	National Comorbidity Survey Replication Adolescent Supplement [83]
15.	Overall heath and somatic distress	Somatic and Psychological Health Report [84], self-perceived health status, and general body measurements
16.	Medical, mental health, and family history	Multiple-choice questions
17.	Cognitive concerns and empathy	Derived from the Subjective Scale to Investigate Cognition in Schizophrenia [85] and the empathy quotient [86]
18.	Eating behaviors and body image	Derived from the Eating Disorder Examination [87]
19.	Social connectedness and support	Derived from the Perceived Social Support/Conflict Measure [88] plus 5 items measuring relationships with peers [89]

[a]MHeC-S: Spanish version of the Mental Health eClinic.

[b]PTSD: posttraumatic stress disorder.

We didn't find any Spanish versions of the items assessing disability, suicide ideation, and anxiety rating. Considering their relevance in overall assessment and potential medicolegal repercussions, we decided to find Spanish-speaking analogs to these measures. We replaced the Brief Disability Questionnaire with the World Health Organization Disability Assessment Schedule 2.0, which has an official translation available [110]. We replaced the Suicidal Ideation Attributes Scale with the Suicide Behaviors Questionnaire-Revised [111] and we replaced the Overall Anxiety Severity and Impairment Scale with the 7-item Generalized Anxiety Disorder scale [112], both of which have their translation process and psychometric properties published.

In July 2016, the rest of the items were carefully translated into 3 individual sessions of colloquial Spanish, and we conducted 1 round of translation and cultural adaptation for Spanish-speaking populations living in Australia. At this stage, we adapted 3 questions in the demographics sections: country of origin; language spoken at home, enriched with relevant dialects from Spanish-speaking regions, such as Quichuan and Catalan; and the ethnicity question, modified to the indigenous populations in Latin America.

To reach agreement, 2 individual sessions of back-translations were performed, followed by 1 round of discussion between the translators and back-translators.

User Interface

All participants accepted the Spanish version of the original MHeC's logo (Figure 8). However, young people preferred a name that they could associate more with general well-being than with mental health, as some of them believed that this would have a wider reach among students.

In relation to language, participants expressed different preferences for the linguistic form in which to address the users (interlocutors); Colombian participants (20/32, 63%) favored the use of formal pronouns (usted), as they considered that the

XSL•FO
RenderX

delivery of online health services should follow the same conventions as face-to-face services:

> *In Colombia, the doctor-patient relationship is always treated in a formal way...* [Health professional, quote O]

Participants of other nationalities (12/32, 38%) preferred the prototype to use the colloquial or familiar pronouns (tú, vos), as the formal pronoun seemed excessively formal in an online context. Considering this discrepancy, all participants agreed that the prototype would use the colloquial or familiar form of the second person singular pronoun (tú), as the target of the MHeC-S is young people from different nationalities. Additionally, participants suggested the possibility of a customizing option to choose to see the prototype (1) completely in Spanish (including menus, links, call-to-action buttons, instructions, videos, apps, and e-tools), (2) in a bilingual version (which would display menus, call-to-action buttons, and instructions in English, but the most relevant content in Spanish, such as the physical and mental health self-report assessment and video visit; or a mix of Spanish and English apps, e-tools, and resources), or (3) completely in English (which would look more like the original MHeC but with relevant information for this population).

Participants in the one-on-one user testing sessions (n=15) assessed the interface in the prototype. These participants approved the MHeC-S's font, color palette (light blue, orange, and green), and the tile-shaped buttons (15/15, 100%). Despite this, young people (7/15, 47%) thought that the Get Started call-to-action button needed to be different (bolder, bigger, brighter, or in a different shape) to get the participants to sign up. In relation to the menus, horizontal or hamburger displays were preferred over the current vertical presentation.

Figure 8. Original Mental Health eClinic logo and its Spanish adaptation. Created by Mandarin Creative

Usability

A total of 15 participants completed 4 usability tasks: (1) create an account and log in, (2) find the Need Help Now button, (3) explore the dashboard of results, and (4) book an appointment. Mean SEQ scores for the tasks were 7 (SD 0); 6.93 (SD 0.26); 5.07 (SD 1.49); 5.80 (1.66) respectively, range 1 to 7. All 15 participants did not report problems completing tasks 1 and 2. In relation to task 3, participants said that exploring all of the dashboard tabs was slightly complicated, as they were not evident at first glance. They had slight difficulty in completing an appointment booking, as the action button was located on the bottom right corner of the site, so this task wasn't intuitive for some participants.

Discussion

Principal Findings

Our study used a comprehensive research and development approach to co-design and culturally adapt a prototypic Web-based mental health clinic (MHeC) for Spanish-speaking young people based in Australia (MHeC-S). Thematic analysis resulted in adequate acceptability of the 5 main elements of the alpha prototype (a home page and triage system; a comprehensive online physical and mental health self-report assessment; a dashboard of results and progress report; a booking and videoconferencing system to enable video visits; and the generation of a personalized well-being plan that includes links to evidence-based, young person–suggested, health professional–recommended apps and e-tools). The data also revealed gaps in the alpha prototype, such as the need for tailored assessment tools and a greater integration with Spanish-speaking services and communities; a lack of Spanish-language apps and e-tools, and of online mental health information was noted. As a consequence, the development of new features included the addition of cultural adjustment items in the online self-report assessment, creation of specific algorithms, and development of several videos and factsheets (Multimedia Appendix 2). In the future, the beta prototype should additionally include refinements of language;

XSL•FO
RenderX

explanations of specific medical terminology; and minor changes in layout and navigation.

Migrants and newly arrived residents have been identified as populations who are difficult to recruit, and then to involve and maintain in research [113], yet this is a population in critical need of support. The research and development cycle that we employed in this study is an optimal methodology to engage, retain, and work more efficiently with hard-to-reach populations. We selected various participatory design methodologies to enhance the generation of new ideas and improve the feedback process. The nature of the research and development cycle and the use of diverse methodologies enabled the research to be conducted and completed in a time-efficient manner.

Previous research has highlighted the need to tailor HIT interventions beyond content and language, by including culture [39]. One of the strengths of this study was the incorporation of the cultural framework as a cornerstone of the research and development cycle. As a consequence, we obtained information about the participants' cultural preferences for the prototype's interface and functionality, as well as the development of culturally appropriate content and features. Performing data collection and analysis in the original language reduced the risk of losing relevant information (or meaning), and decreased research time and costs [114]. Other advantages of this study were the variety of origin of participants (Argentina, Chile, Colombia, Spain, and Venezuela) and the research team (Australia, Chile, Colombia, and Venezuela). Furthermore, this research united all relevant stakeholders (young people, supportive others, and health professionals) in a common goal of adapting this prototype to a population in need.

Although Spanish is the second most common language spoken worldwide and HIT is a growing field, Spanish-speaking populations (including migrants and those residing in low- and middle-income countries) are at risk of experiencing not only physical but also technological social health inequalities [115]. This body of research aims to breach this gap by creating a widely available MHeC-S that works across devices. The participation of end users in the design process ensured that the prototype was accessible to individuals of varying literacy levels with a range of cultural differences. Furthermore, the MHeC-S has the potential to be configured and adapted for use in Spanish-speaking countries and in other multicultural countries with Spanish-speaking migrant populations.

Implications

International education in Australia has grown dramatically and is its third largest export industry, contributing Aus $32.4 billion to the Australian economy [116]. It highlights a significant bilateral exchange (Aus $755 million in 2012) between Latin America and Australia, which is increasingly recognized as an important destination for the English education of Latin Americans [117]. In 2017, the Latin American Spanish-speaking international student population had reportedly increased to more than 21,000 [17]. Our study highlighted a critical concern in the community in relation to a shortage of mental health services targeting the well-being of these students. This is vital, as these students are at higher risk of developing adaptational problems and being socially and linguistically isolated [118]. Participants generally expressed a lack of understanding of the Australian health system, particularly service providers and insurance agencies, resulting in an important barrier for students' help-seeking process. Even for those who do know how to navigate the health system, a reduced English competence could impair the care they do obtain. Protecting, caring for, and enhancing positive experiences for international students is Australia's best strategy to protect and grow this industry.

New and emerging technologies present a solution, as they have changed the way young people communicate, connect, and engage with each other and with society. With the introduction of smartphones, information, services, and resources provided online or via mobile apps can be accessed privately and at any time. This can be empowering for individuals who are marginalized or geographically or socially isolated. It could also help to address the need for Spanish-speaking mental health professionals and interpreters. Having an MHeC-S could greatly benefit young people who are native Spanish speakers living in Australia and who are actively seeking help. This study is the first step toward providing a technology-enabled solution to improve this population's mental health and well-being in Australia. To the best of our knowledge, there has been no research to date in this field.

Conclusion

Further research is needed to understand the psychometric properties of the online self-report assessment (eg, criterion validity) and the integration of the MHeC-S with other apps or e-tools. Importantly, additional steps are needed to evaluate the engagement, efficacy, and effectiveness of the MHeC-S in real-world settings. The MHeC-S shares the same elements and functionality of the original version of the MHeC. Its main difference relies on interaction with face-to-face services. The original MHeC was designed to work with primary care mental health services; however, in the case of the MHeC-S, in Australia it could be used additionally by language schools and Overseas Student Health Cover providers. To the best of our knowledge, this study is the first to explore mental health care barriers and facilitators, and potential technology solutions in a language student population; additional research is needed to expand the knowledge on this topic.

Acknowledgments

The authors would like to thank all young people, supportive others, and health professionals who participated in this study. Additionally, we thank headspace Campbelltown, headspace Camperdown, the Consulate General of Colombia in Sydney, Ms Nancy Benitez Paez (Consul General of Colombia in Sydney), Dr Andres Rangel Martinez-Villalba, Dr German Rueda-Jaimes, and Dr Carlos Filizzola Donado for their collaboration in this study. We would also like to thank Ms Javiera Gálvez and Ms Camila Bravo for their work translating the content of the prototype; Ms Lauren Cannell for her work in the translation and

XSL•FO
RenderX

back-translation process; and Ms Abigail Escobar for coding the data. This project was funded by the Young and Well Cooperative Research Centre (Western Sydney University, Penrith, Australia; 2014-2016), which was led by Professor Jane Burns.

Conflicts of Interest

IBH was an inaugural Commissioner on Australia's National Mental Health Commission (2012-18). He is the Co-Director, Health and Policy at the Brain and Mind Centre (BMC) University of Sydney. The BMC operates early-intervention youth services at Camperdown under contract to headspace. IBH is the Chief Scientific Advisor to, and a 5% equity shareholder in, InnoWell Pty Ltd. InnoWell was formed by the University of Sydney (45% equity) and PwC (Australia; 45% equity) to deliver the $30 M Australian Government Department of Health-funded Project Synergy (2017-20); a three-year program for the transformation of mental health services) and to lead transformation of mental health services internationally through the use of innovative technologies

Multimedia Appendix 1

Original quotes in Spanish.

[PDF File (Adobe PDF File), 47KB - jmir_v21i8e14127_app1.pdf]

Multimedia Appendix 2

Development of new features during the rapid prototyping phase.

[PDF File (Adobe PDF File), 30KB - jmir_v21i8e14127_app2.pdf]

References

1. World Health Organization. Mental Health Atlas 2017. Geneva, Switzerland: WHO; 2018. URL: https://apps.who.int/iris/bitstream/handle/10665/272735/9789241514019-eng.pdf [accessed 2019-06-20]
2. Clarke G, Yarborough BJ. Evaluating the promise of health IT to enhance/expand the reach of mental health services. Gen Hosp Psychiatry 2013;35(4):339-344 [FREE Full text] [doi: 10.1016/j.genhosppsych.2013.03.013] [Medline: 23701698]
3. Gore FM, Bloem PJN, Patton GC, Ferguson J, Joseph V, Coffey C, et al. Global burden of disease in young people aged 10-24 years: a systematic analysis. Lancet 2011 Jun 18;377(9783):2093-2102. [doi: 10.1016/S0140-6736(11)60512-6] [Medline: 21652063]
4. Patel V, Maj M, Flisher AJ, De Silva MJ, Koschorke M, Prince M, WPA Zonal and Member Society Representatives. Reducing the treatment gap for mental disorders: a WPA survey. World Psychiatry 2010 Oct;9(3):169-176 [FREE Full text] [Medline: 20975864]
5. Ohtani A, Suzuki T, Takeuchi H, Uchida H. Language barriers and access to psychiatric care: a systematic review. Psychiatr Serv 2015 Aug 01;66(8):798-805. [doi: 10.1176/appi.ps.201400351] [Medline: 25930043]
6. Kim E, Coumar A, Lober WB, Kim Y. Addressing mental health epidemic among university students via web-based, self-screening, and referral system: a preliminary study. IEEE Trans Inf Technol Biomed 2011 Mar;15(2):301-307. [doi: 10.1109/TITB.2011.2107561] [Medline: 21257386]
7. Blignault I, Ponzio V, Rong Y, Eisenbruch M. A qualitative study of barriers to mental health services utilisation among migrants from mainland China in south-east Sydney. Int J Soc Psychiatry 2008 Mar;54(2):180-190. [doi: 10.1177/0020764007085872] [Medline: 18488410]
8. Kelaher M, Williams G, Manderson L. Towards evidence-based health promotion and service provision for new migrants to Australia. Ethn Health 1999 Nov;4(4):305-313. [doi: 10.1080/13557859998074] [Medline: 10705566]
9. Studies in Australia. International students in Australia. Melbourne, Australia: Good Education Group Pty Ltd; 2019. URL: https://www.studiesinaustralia.com/studying-in-australia/why-study-in-australia/international-students-in-australia [WebCite Cache ID 76KbPEmLW]
10. Bhugra D. Migration and mental health. Acta Psychiatr Scand 2004 Apr;109(4):243-258. [Medline: 15008797]
11. Suh H, van Nuenen M, Rice KG. The CES-D as a measure of psychological distress among international students: measurement and structural invariance across gender. Assessment 2017 Oct;24(7):896-906. [doi: 10.1177/1073191116632337] [Medline: 26887810]
12. Lee RB, Maria MS, Estanislao S, Rodriguez C. Factors associated with higher levels of depressive symptoms among international university students in the Philippines. Southeast Asian J Trop Med Public Health 2013 Nov;44(6):1098-1107. [Medline: 24450248]
13. Khawaja NG, Dempsey J. A comparison of international and domestic tertiary students in Australia. J Psychol Couns Sch 2008;18(1):30-46.
14. Forbes-Mewett H, Sawyer AM. Mental Health Issues Amongst International Students in Australia: Perspectives From Professionals at the Coal-Face. Callaghan, Australia: University of Newcastle; 2011.

XSL•FO
RenderX

15. Pejic D. International Student Welfare in Australia. Melbourne, Australia: International Social Service Australia; 2012. URL: http://www.iss.org.au/wp-content/uploads/2012/07/International-Student-Welfare-in-Australia.pdf [WebCite Cache ID 706gGEzJe]
16. Kwan B. Mental health stigma biggest barrier for international students seeking help. Crows Nest, Australia: SBSNews; 2019. URL: https://www.sbs.com.au/news/mental-health-stigma-biggest-barrier-for-international-students-seeking-help [WebCite Cache ID 76hIjeSD9]
17. Australian Government Department of Home Affairs. Student visa and Temporary Graduate visa program bi-annual report. Canberra, Australia: The Department; 2017. URL: https://www.homeaffairs.gov.au/research-and-stats/files/student-temp-grad-program-report-dec-2018.pdf [WebCite Cache ID 76kRHhjHy]
18. McHugh RK, Barlow DH. The dissemination and implementation of evidence-based psychological treatments. A review of current efforts. Am Psychol 2010;65(2):73-84. [doi: 10.1037/a0018121] [Medline: 20141263]
19. Judd FK, Jackson H, Davis J, Cockram A, Komiti A, Allen N, et al. Improving access for rural Australians to treatment for anxiety and depression: the University of Melbourne Depression and Anxiety Research and Treatment Group-Bendigo Health Care Group initiative. Aust J Rural Health 2001 Apr;9(2):91-96. [Medline: 11259963]
20. Blaya JA, Fraser HSF, Holt B. E-health technologies show promise in developing countries. Health Aff (Millwood) 2010 Feb;29(2):244-251 [FREE Full text] [doi: 10.1377/hlthaff.2009.0894] [Medline: 20348068]
21. Hoek W, Marko M, Fogel J, Schuurmans J, Gladstone T, Bradford N, et al. Randomized controlled trial of primary care physician motivational interviewing versus brief advice to engage adolescents with an Internet-based depression prevention intervention: 6-month outcomes and predictors of improvement. Transl Res 2011 Dec;158(6):315-325. [doi: 10.1016/j.trsl.2011.07.006] [Medline: 22061038]
22. van der Zanden R, Kramer J, Gerrits R, Cuijpers P. Effectiveness of an online group course for depression in adolescents and young adults: a randomized trial. J Med Internet Res 2012;14(3):e86 [FREE Full text] [doi: 10.2196/jmir.2033] [Medline: 22677437]
23. Tait RJ, Christensen H. Internet-based interventions for young people with problematic substance use: a systematic review. Med J Aust 2010 Jun 7;192(11 Suppl):S15-S21. [Medline: 20528701]
24. 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]
25. Hunkeler EM, Meresman JF, Hargreaves WA, Fireman B, Berman WH, Kirsch AJ, et al. Efficacy of nurse telehealth care and peer support in augmenting treatment of depression in primary care. Arch Fam Med 2000 Aug;9(8):700-708. [Medline: 10927707]
26. Epstein JN, Langberg JM, Lichtenstein PK, Kolb R, Altaye M, Simon JO. Use of an internet portal to improve community-based pediatric ADHD care: a cluster randomized trial. Pediatrics 2011 Nov;128(5):e1201-e1208 [FREE Full text] [doi: 10.1542/peds.2011-0872] [Medline: 22007005]
27. Mitchell J, Vella-Brodrick D, Klein B. Positive psychology and the internet: a mental health opportunity. E J Appl Psychol 2010 Dec 30;6(2):30-41. [doi: 10.7790/ejap.v6i2.230]
28. van Straten SA, Seekles W, van't Veer-Tazelaar NJ, Beekman ATF, Cuijpers P. Stepped care for depression in primary care: what should be offered and how? Med J Aust 2010 Jun 7;192(11 Suppl):S36-S39. [Medline: 20528706]
29. Perini S, Titov N, Andrews G. Clinician-assisted Internet-based treatment is effective for depression: randomized controlled trial. Aust N Z J Psychiatry 2009 Jun;43(6):571-578. [doi: 10.1080/00048670902873722] [Medline: 19440890]
30. Saulsberry A, Marko-Holguin M, Blomeke K, Hinkle C, Fogel J, Gladstone T, et al. Randomized clinical trial of a primary care internet-based intervention to prevent adolescent depression: one-year outcomes. J Can Acad Child Adolesc Psychiatry 2013 May;22(2):106-117 [FREE Full text] [Medline: 23667356]
31. Martínez P, Rojas G, Martínez V, Lara MA, Pérez JC. Internet-based interventions for the prevention and treatment of depression in people living in developing countries: a systematic review. J Affect Disord 2018 Dec;234:193-200. [doi: 10.1016/j.jad.2018.02.079] [Medline: 29529553]
32. Saliba V, Legido-Quigley H, Hallik R, Aaviksoo A, Car J, McKee M. Telemedicine across borders: a systematic review of factors that hinder or support implementation. Int J Med Inform 2012 Dec;81(12):793-809. [doi: 10.1016/j.ijmedinf.2012.08.003] [Medline: 22975018]
33. Barrera-Valencia C, Benito-Devia AV, Vélez-Álvarez C, Figueroa-Barrera M, Franco-Idárraga SM. [Cost-effectiveness of synchronous vs. asynchronous telepsychiatry in prison inmates with depression]. Rev Colomb Psiquiatr 2017;46(2):65-73. [doi: 10.1016/j.rcp.2016.04.008] [Medline: 28483175]
34. Castro A, Larraín A, Fritsch R, Rojas G. [Feasibility of implementing tele-psychiatry in Chile]. Rev Med Chil 2012 Jun;140(6):789-796 [FREE Full text] [doi: 10.4067/S0034-98872012000600015] [Medline: 23282619]
35. Moreno FA, Chong J, Dumbauld J, Humke M, Byreddy S. Use of standard webcam and internet equipment for telepsychiatry treatment of depression among underserved Hispanics. Psychiatr Serv 2012 Dec;63(12):1213-1217. [doi: 10.1176/appi.ps.201100274] [Medline: 23026854]
36. Baker-Ericzén MJ, Connelly CD, Hazen AL, Dueñas C, Landsverk JA, Horwitz SM. A collaborative care telemedicine intervention to overcome treatment barriers for Latina women with depression during the perinatal period. Fam Syst Health 2012 Sep;30(3):224-240 [FREE Full text] [doi: 10.1037/a0028750] [Medline: 22709321]

XSL•FO
RenderX

37. Mucic D. International telepsychiatry: a study of patient acceptability. J Telemed Telecare 2008;14(5):241-243. [doi: 10.1258/jtt.2008.080301] [Medline: 18632998]
38. Yellowlees PM, Odor A, Iosif A, Parish MB, Nafiz N, Patrice K, et al. Transcultural psychiatry made simple--asynchronous telepsychiatry as an approach to providing culturally relevant care. Telemed J E Health 2013 Apr;19(4):259-264 [FREE Full text] [doi: 10.1089/tmj.2012.0077] [Medline: 23451812]
39. Chaet AV, Morshedi B, Wells KJ, Barnes LE, Valdez R. Spanish-language consumer health information technology interventions: a systematic review. J Med Internet Res 2016 Dec 10;18(8):e214 [FREE Full text] [doi: 10.2196/jmir.5794] [Medline: 27511437]
40. Ferrari M, Ahmad F, Shakya Y, Ledwos C, McKenzie K. Computer-assisted client assessment survey for mental health: patient and health provider perspectives. BMC Health Serv Res 2016 Sep 23;16(1):516 [FREE Full text] [doi: 10.1186/s12913-016-1756-0] [Medline: 27663508]
41. Aguilera A, Bruehlman-Senecal E, Demasi O, Avila P. Automated text messaging as an adjunct to cognitive behavioral therapy for depression: a clinical trial. J Med Internet Res 2017 May 08;19(5):e148 [FREE Full text] [doi: 10.2196/jmir.6914] [Medline: 28483742]
42. Romero-Sanchiz P, Nogueira-Arjona R, García-Ruiz A, Luciano JV, García CJ, Gili M, et al. Economic evaluation of a guided and unguided internet-based CBT intervention for major depression: results from a multi-center, three-armed randomized controlled trial conducted in primary care. PLoS One 2017;12(2):e0172741 [FREE Full text] [doi: 10.1371/journal.pone.0172741] [Medline: 28241025]
43. Ramos K, Cortes J, Wilson N, Kunik ME, Stanley MA. Vida Calma: CBT for anxiety with a Spanish-speaking Hispanic adult. Clin Gerontol 2017;40(3):213-219. [doi: 10.1080/07317115.2017.1292978] [Medline: 28452668]
44. Muroff J, Robinson W, Chassler D, López LM, Gaitan E, Lundgren L, et al. Use of a smartphone recovery tool for Latinos with co-occurring alcohol and other drug disorders and mental disorders. J Dual Diagn 2017;13(4):280-290. [doi: 10.1080/15504263.2017.1348649] [Medline: 28692420]
45. Bjerkan J, Hedlund M, Hellesø R. Patients' contribution to the development of a web-based plan for integrated care - a participatory design study. Inform Health Soc Care 2015 Mar;40(2):167-184. [doi: 10.3109/17538157.2014.907803] [Medline: 24786524]
46. Peters D, Davis S, Calvo RA, Sawyer SM, Smith L, Foster JM. Young people's preferences for an asthma self-management app highlight psychological needs: a participatory study. J Med Internet Res 2017 Apr 11;19(4):e113 [FREE Full text] [doi: 10.2196/jmir.6994] [Medline: 28400353]
47. Petersen C, Adams SA, DeMuro PR. mHealth: don't forget all the stakeholders in the business case. Med 2.0 2015 Dec 31;4(2):e4 [FREE Full text] [doi: 10.2196/med20.4349] [Medline: 26720310]
48. Ospina-Pinillos L, Davenport T, Iorfino F, Tickell A, Cross S, Scott EM, et al. Using new and innovative technologies to assess clinical stage in early intervention youth mental health services: evaluation study. J Med Internet Res 2018 Sep 10;20(9):e259 [FREE Full text] [doi: 10.2196/jmir.9966] [Medline: 30201602]
49. Ginossar T, Nelson S. Reducing the health and digital divides: a model for using community-based participatory research approach to e-health interventions in low-income Hispanic communities. J Comput Mediat Commun 2010;15(4):a. [doi: 10.1111/j.1083-6101.2009.01513.x]
50. Hickie IB, Luscombe GM, Davenport TA, Burns JM, Highet NJ. Perspectives of young people on depression: awareness, experiences, attitudes and treatment preferences. Early Interv Psychiatry 2007 Nov;1(4):333-339. [doi: 10.1111/j.1751-7893.2007.00042.x] [Medline: 21352121]
51. Hickie IB, McGorry PD. Guidelines for youth depression: time to incorporate new perspectives. Med J Aust 2010 Aug 02;193(3):133-134. [Medline: 20678036]
52. Christensen H, Petrie K. State of the e-mental health field in Australia: where are we now? Aust N Z J Psychiatry 2013 Feb;47(2):117-120. [doi: 10.1177/0004867412471439] [Medline: 23297367]
53. Ospina-Pinillos L, Davenport TA, Ricci CS, Milton AC, Scott EM, Hickie IB. Developing a Mental Health eClinic to improve access to and quality of mental health care for young people: using participatory design as research methodologies. J Med Internet Res 2018 May 28;20(5):e188 [FREE Full text] [doi: 10.2196/jmir.9716] [Medline: 29807878]
54. Hagen P, Collin P, Metcalf A, Nicholas M, Rahilly K, Swainston N. Participatory Design of Evidence-Based Online Youth Mental Health Promotion, Intervention and Treatment. Melbourne, Australia: Young and Well Cooperative Research Centre; 2012:978.
55. Sousa VD, Rojjanasrirat W. Translation, adaptation and validation of instruments or scales for use in cross-cultural health care research: a clear and user-friendly guideline. J Eval Clin Pract 2011 Apr;17(2):268-274. [doi: 10.1111/j.1365-2753.2010.01434.x] [Medline: 20874835]
56. Boren T, Ramey J. Thinking aloud: reconciling theory and practice. IEEE Trans Profess Commun 2000;43(3):261-278. [doi: 10.1109/47.867942]
57. Sauro J. 10 things to know about the Single Ease Question (SEQ). Denver, CO: Measuring U; 2012. URL: https://measuringu.com/seq10/ [WebCite Cache ID 6vcXt6iSr]
58. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]

XSL•FO
RenderX

59. Hoover RS, Koerber AL. Using NVivo to answer the challenges of qualitative research in professional communication: benefits and best practices tutorial. IEEE Trans Profess Commun 2011 Mar;54(1):68-82. [doi: 10.1109/TPC.2009.2036896]
60. Valdez RS, Gibbons MC, Siegel ER, Kukafka R, Brennan PF. Designing consumer health IT to enhance usability among different racial and ethnic groups within the United States. Health Technol 2012 Jul 13;2(4):225-233. [doi: 10.1007/s12553-012-0031-6]
61. Pope C, Ziebland S, Mays N. Qualitative research in health care. Analysing qualitative data. BMJ 2000 Jan 8;320(7227):114-116 [FREE Full text] [Medline: 10625273]
62. Davey JW, Gugiu PC, Coryn CL. Quantitative methods for estimating the reliability of qualitative data. J Multidiscipl Eval 2010;6(13):140-162.
63. Viera AJ, Garrett JM. Understanding interobserver agreement: the kappa statistic. Fam Med 2005 May;37(5):360-363 [FREE Full text] [Medline: 15883903]
64. Milton AC, Ellis LA, Davenport TA, Burns JM, Hickie IB. Comparison of self-reported telephone interviewing and web-based survey responses: findings from the Second Australian Young and Well National Survey. JMIR Ment Health 2017 Sep 26;4(3):e37. [doi: 10.2196/mental.8222] [Medline: 28951382]
65. Tate CC, Ledbetter JN, Youssef CP. A two-question method for assessing gender categories in the social and medical sciences. J Sex Res 2013;50(8):767-776. [doi: 10.1080/00224499.2012.690110] [Medline: 22989000]
66. Ormel J, VonKorff M, Ustun TB, Pini S, Korten A, Oldehinkel T. Common mental disorders and disability across cultures. Results from the WHO Collaborative Study on Psychological Problems in General Health Care. JAMA 1994 Dec 14;272(22):1741-1748. [Medline: 7966922]
67. Goldman HH, Skodol AE, Lave TR. Revising axis V for DSM-IV: a review of measures of social functioning. Am J Psychiatry 1992 Sep;149(9):1148-1156. [Medline: 1386964]
68. Kessler RC, Andrews G, Colpe LJ, Hiripi E, Mroczek DK, Normand SLT, et al. Short screening scales to monitor population prevalences and trends in non-specific psychological distress. Psychol Med 2002 Aug;32(6):959-976. [Medline: 12214795]
69. Rush AJ, Trivedi MH, Ibrahim HM, Carmody TJ, Arnow B, Klein DN, et al. The 16-Item Quick Inventory of Depressive Symptomatology (QIDS), clinician rating (QIDS-C), and self-report (QIDS-SR): a psychometric evaluation in patients with chronic major depression. Biol Psychiatry 2003 Sep 01;54(5):573-583. [Medline: 12946886]
70. Norman SB, Cissell SH, Means-Christensen AJ, Stein MB. Development and validation of an Overall Anxiety Severity And Impairment Scale (OASIS). Depress Anxiety 2006;23(4):245-249. [doi: 10.1002/da.20182] [Medline: 16688739]
71. Altman EG, Hedeker D, Peterson JL, Davis JM. The Altman Self-Rating Mania Scale. Biol Psychiatry 1997 Nov 15;42(10):948-955. [doi: 10.1016/S0006-3223(96)00548-3] [Medline: 9359982]
72. Stefanis NC, Hanssen M, Smirnis NK, Avramopoulos DA, Evdokimidis IK, Stefanis CN, et al. Evidence that three dimensions of psychosis have a distribution in the general population. Psychol Med 2002 Feb;32(2):347-358. [Medline: 11866327]
73. Prins A, Ouimette P, Kimerling R, Camerond RP, Hugelshofer DS, Shaw-Hegwer J, et al. The primary care PTSD screen (PC–PTSD): development and operating characteristics. Prim Care Psychiatry 2004 Jan 01;9(1):9-14. [doi: 10.1185/135525703125002360]
74. Blanchard EB, Jones-Alexander J, Buckley TC, Forneris CA. Psychometric properties of the PTSD Checklist (PCL). Behav Res Ther 1996 Aug;34(8):669-673. [Medline: 8870294]
75. van Spijker BAJ, Batterham PJ, Calear AL, Farrer L, Christensen H, Reynolds J, et al. The suicidal ideation attributes scale (SIDAS): community-based validation study of a new scale for the measurement of suicidal ideation. Suicide Life Threat Behav 2014 Aug;44(4):408-419. [doi: 10.1111/sltb.12084] [Medline: 24612048]
76. Saunders JB, Aasland OG, Babor TF, de la Fuente JR, Grant M. Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO Collaborative Project on Early Detection of Persons with Harmful Alcohol Consumption--II. Addiction 1993 Jun;88(6):791-804. [Medline: 8329970]
77. WHO ASSIST Working Group. The Alcohol, Smoking and Substance Involvement Screening Test (ASSIST): development, reliability and feasibility. Addiction 2002 Sep;97(9):1183-1194. [Medline: 12199834]
78. Ewing JA. Detecting alcoholism. The CAGE questionnaire. JAMA 1984 Oct 12;252(14):1905-1907. [Medline: 6471323]
79. Cooper ML. Motivations for alcohol use among adolescents: development and validation of a four-factor model. Psychol Assess 1994;6(2):117-128. [doi: 10.1037/1040-3590.6.2.117]
80. Heatherton TF, Kozlowski LT, Frecker RC, Fagerström KO. The Fagerström Test for Nicotine Dependence: a revision of the Fagerström Tolerance Questionnaire. Br J Addict 1991 Sep;86(9):1119-1127. [Medline: 1932883]
81. Australian Institute of Health and Welfare. National Drug Strategy Household Survey: detailed report 2013. Canberra, Australia: AIHW; 2014. URL: https://www.aihw.gov.au/getmedia/c2e94ca2-7ce8-496f-a765-94c55c774d2b/16835_1.pdf.aspx?inline=true
82. Craig CL, Marshall AL, Sjöström M, Bauman AE, Booth ML, Ainsworth BE, et al. International physical activity questionnaire: 12-country reliability and validity. Med Sci Sports Exerc 2003 Aug;35(8):1381-1395. [doi: 10.1249/01.MSS.0000078924.61453.FB] [Medline: 12900694]
83. Green JG, Avenevoli S, Gruber MJ, Kessler RC, Lakoma MD, Merikangas KR, et al. Validation of diagnoses of distress disorders in the US National Comorbidity Survey Replication Adolescent Supplement (NCS-A). Int J Methods Psychiatr Res 2012 Mar;21(1):41-51 [FREE Full text] [doi: 10.1002/mpr.357] [Medline: 22086845]

XSL•FO
RenderX

84. Hickie IB, Davenport TA, Hadzi-Pavlovic D, Koschera A, Naismith SL, Scott EM, et al. Development of a simple screening tool for common mental disorders in general practice. Med J Aust 2001 Jul 16;175 Suppl:S10-S17. [Medline: 11556430]
85. Stip E, Caron J, Renaud S, Pampoulova T, Lecomte Y. Exploring cognitive complaints in schizophrenia: the subjective scale to investigate cognition in schizophrenia. Compr Psychiatry 2003;44(4):331-340. [doi: 10.1016/S0010-440X(03)00086-5] [Medline: 12923712]
86. Muncer SJ, Ling J. Psychometric analysis of the empathy quotient (EQ) scale. Pers Indiv Diff 2006 Apr;40(6):1111-1119. [doi: 10.1016/j.paid.2005.09.020]
87. Hay P. The epidemiology of eating disorder behaviors: an Australian community-based survey. Int J Eat Disord 1998 May;23(4):371-382. [Medline: 9561427]
88. Schuster TL, Kessler RC, Aseltine RH. Supportive interactions, negative interactions, and depressed mood. Am J Community Psychol 1990 Jun;18(3):423-438. [Medline: 2264558]
89. Catholic Education Commission, New South Wales. Guidelines for Pastoral Care in Catholic Schools. Canberra, Australia: Commonwealth Department of Education, Science and Training; 2003. URL: http://www.cecnsw.catholic.edu.au/images/Guidelines%20for%20Pastoral%20Care%20In%20Catholic%20Schools.pdf [WebCite Cache ID 6woyDiVoh]
90. Reisner SL, Biello K, Rosenberger JG, Austin SB, Haneuse S, Perez-Brumer A, et al. Using a two-step method to measure transgender identity in Latin America/the Caribbean, Portugal, and Spain. Arch Sex Behav 2014 Nov;43(8):1503-1514 [FREE Full text] [doi: 10.1007/s10508-014-0314-2] [Medline: 25030120]
91. [National Comorbidity Survey. K10 and K6 scales].: World Health Organization; 2006. URL: https://www.hcp.med.harvard.edu/ncs/ftpdir/k6/Spanish_K10.pdf [WebCite Cache ID 6v9fXhI8g]
92. Trujols J, de Diego-Adeliño J, Feliu-Soler A, Iraurgi I, Puigdemont D, Alvarez E, et al. The Spanish version of the Quick Inventory of Depressive Symptomatology-Self-Report (QIDS-SR16): a psychometric analysis in a clinical sample. J Affect Disord 2014 Dec;169:189-196. [doi: 10.1016/j.jad.2014.08.011] [Medline: 25212994]
93. Gili M, Lopez-Navarro E, Homar C, Castro A, García-Toro M, Llobera J, et al. Psychometric properties of Spanish version of QIDS-SR16 in depressive patients. Actas Esp Psiquiatr 2014;42(6):292-299 [FREE Full text] [Medline: 25388771]
94. van Os J, Verdoux H, Hanssen M. CAPE 42. 2001. URL: http://cape42.homestead.com/ [WebCite Cache ID 6v9hNwcwd]
95. Fonseca-Pedrero E, Paino M, Lemos-Giráldez S, Muñiz J. Validation of the Community Assessment Psychic Experiences -42 (CAPE-42) in Spanish college students and patients with psychosis. Actas Esp Psiquiatr 2012;40(4):169-176 [FREE Full text] [Medline: 22851477]
96. Babor T, Higgins-Biddle J. Intervención breve para el consumo de riesgo y perjudicial de alcohol. Geneva, Switzerland: World Health Organization; 2001. URL: http://www.who.int/substance_abuse/activities/en/BImanualSpanish.pdf [WebCite Cache ID 6vArKNVap]
97. World Health Organization. La prueba de detección de consumo de alcohol, tabaco y sustancias (ASSIST). Manual para uso en la atención primaria. Geneva, Switzerland: WHO; 2011.
98. The IPAQ Group. International Physical Activity Questionnaire. 2016. URL: https://sites.google.com/site/theipaq/home [WebCite Cache ID 6vBDlmCVX]
99. The IPAQ Group. IPAQ: formato corto autoadministrado de los últimos 7 días. 2002 Aug. URL: https://docs.google.com/viewer?a=v&pid=sites&srcid=ZGVmYXVsdGRvbWFpbnx0aGVpcGFxfGd4OjJhOTZhOGFhMWNjNzI0NTI [WebCite Cache ID 6vBDTBbaA]
100. Medina-Mora ME, Borges G, Lara C, Benjet C, Blanco J, Fleiz C, et al. Prevalence, service use, and demographic correlates of 12-month DSM-IV psychiatric disorders in Mexico: results from the Mexican National Comorbidity Survey. Psychol Med 2005 Dec;35(12):1773-1783. [doi: 10.1017/S0033291705005672] [Medline: 16300691]
101. Benjet C, Borges G, Medina-Mora ME, Zambrano J, Aguilar-Gaxiola S. Youth mental health in a populous city of the developing world: results from the Mexican Adolescent Mental Health Survey. J Child Psychol Psychiatry 2009 Apr;50(4):386-395. [doi: 10.1111/j.1469-7610.2008.01962.x] [Medline: 19040499]
102. Rendon MJ. The Cultural Adaptation of the Clinician-Administered PTSD Scale for Spanish-Speaking Latinos With Limited English Proficiency in the United States [doctoral thesis]. Coral Gables, FL: University of Miami; 2015. URL: https://scholarlyrepository.miami.edu/cgi/viewcontent.cgi?article=2540&context=oa_dissertations [WebCite Cache ID 777PL8n5q]
103. Miles JNV, Marshall GN, Schell TL. Spanish and English versions of the PTSD Checklist-Civilian version (PCL-C): testing for differential item functioning. J Trauma Stress 2008 Aug;21(4):369-376 [FREE Full text] [doi: 10.1002/jts.20349] [Medline: 18720394]
104. Alvarez Mas P, Benavent Rodríguez P, García Valls JM, Livianos Aldana L, Rojo Moreno L. [Translation and adaptation in the Spanish environment of the Altman Self-Rating Mania Scale]. Actas Esp Psiquiatr 2005;33(3):180-187. [Medline: 15918086]
105. Ferreira Gonzalez L. Cuestionario CAGE: screening de alcoholismo. URL: http://meiga.info/escalas/cuestionariocage.pdf [WebCite Cache ID 6vAtlYltq]
106. Confederación Española de Organizaciones Empresariales. Test Audit y CAGE. 2011. URL: http://contenidos.ceoe.es/PRL/var/pool/pdf/cms_content_documents-file-767-test-audit-y-cage.pdf [WebCite Cache ID 6v0KkwZaV]
107. Campo-Arias A, Barros-Bermúdez J, Rueda-Jaimes G. Propiedades psicométricas del cuestionario CAGE para consumo abusivo de alcohol: resultados de tres análisis. Rev Colombiana Psiquiatr 2009;38(2):294-303.

XSL•FO
RenderX

108. Becoña E, Vázquez FL. The Fagerström Test for Nicotine Dependence in a Spanish sample. Psychol Rep 1998 Dec;83(3 Pt 2):1455-1458. [doi: 10.2466/pr0.1998.83.3f.1455] [Medline: 10079737]
109. Autism Research Centre. Downloadable tests. Cambridge, UK: University of Cambridge; 2017. URL: https://www.autismresearchcentre.com/arc_tests [WebCite Cache ID 6vBJtDZyb]
110. World Health Organization. Medición de la Salud y la Discapacidad, Manual para el Cuestionario de Evaluación de la Discapacidad de la OMS. WHODAS 2.0. Geneva, Switzerland: WHO; 2015:89.
111. Rueda-Jaimes G, Castro-Rueda V, Rangel-Martínez-Villalba A, Corzo-Casasadiego J, Moreno-Quijano C, Camacho P. Validity of the Suicide Behaviors Questionnaire-Revised in patients with short-term suicide risk. Eur J Psychiatry 2017;31(4):145-150. [doi: 10.1016/j.ejpsy.2017.09.002]
112. García-Campayo J, Zamorano E, Ruiz MA, Pardo A, Pérez-Páramo M, López-Gómez V, et al. Cultural adaptation into Spanish of the generalized anxiety disorder-7 (GAD-7) scale as a screening tool. Health Qual Life Outcomes 2010 Jan 20;8:8 [FREE Full text] [doi: 10.1186/1477-7525-8-8] [Medline: 20089179]
113. Shaghaghi A, Bhopal RS, Sheikh A. Approaches to recruiting 'hard-to-reach' populations into research: a review of the literature. Health Promot Perspect 2011;1(2):86-94 [FREE Full text] [doi: 10.5681/hpp.2011.009] [Medline: 24688904]
114. Smith HJ, Chen J, Liu X. Language and rigour in qualitative research: problems and principles in analyzing data collected in Mandarin. BMC Med Res Methodol 2008 Jul 10;8:44 [FREE Full text] [doi: 10.1186/1471-2288-8-44] [Medline: 18616812]
115. 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]
116. Australian Government Department of Education and Training. Research snapshot: 2017-18 export income by state and territory. 2018 Dec. URL: https://internationaleducation.gov.au/research/Research-Snapshots/Documents/Education%20infographic%20Australia%202017%E2%80%9318.pdf [WebCite Cache ID 76kO4kvRa]
117. Carr B, Minns J. Australia and Latin America: Challenges and Opportunities in the New Millennium. Canberra, Australia: ANU Press; 2014.
118. Demes KA, Geeraert N. Measures matter: scales for adaptation, cultural distance, and acculturation orientation revisited. J Cross Cult Psychol 2013 Jun 11;45(1):91-109. [doi: 10.1177/0022022113487590]

Abbreviations

HIT: health information technology
MHeC: Mental Health eClinic
MHeC-S: Spanish version of the Mental Health eClinic
NSW: New South Wales
PTSD: posttraumatic stress disorder
SEQ: Single Ease Question

Edited by G Eysenbach; submitted 27.03.19; peer-reviewed by N Miyoshi, A Nguyen; comments to author 30.05.19; revised version received 11.06.19; accepted 12.06.19; published 02.08.19.

Please cite as:
Ospina-Pinillos L, Davenport T, Mendoza Diaz A, Navarro-Mancilla A, Scott EM, Hickie IB
Using Participatory Design Methodologies to Co-Design and Culturally Adapt the Spanish Version of the Mental Health eClinic: Qualitative Study
J Med Internet Res 2019;21(8):e14127
URL: https://www.jmir.org/2019/8/e14127/
doi:10.2196/14127
PMID:31376271

©Laura Ospina-Pinillos, Tracey Davenport, Antonio Mendoza Diaz, Alvaro Navarro-Mancilla, Elizabeth M Scott, Ian B Hickie. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 02.08.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Internet-Based Cognitive Behavioral Therapy Among Psychologists in a Medical Setting: A Survey on Implementation

Renée V H IJzerman[1], MSc; Rosalie van der Vaart[1], PhD; Andrea W M Evers[1], Prof Dr

Unit of Health, Medical and Neuropsychology, Faculty of Social and Behavioural Sciences, Leiden University, Leiden, Netherlands

Corresponding Author:
Renée V H IJzerman, MSc
Unit of Health, Medical and Neuropsychology
Faculty of Social and Behavioural Sciences
Leiden University
Wassenaarseweg 52
Leiden,
Netherlands
Phone: 31 715275083
Email: r.v.h.ijzerman@fsw.leidenuniv.nl

Abstract

Background: Internet-based cognitive behavioral therapy (iCBT) is an effective treatment for patients with a chronic somatic illness to improve self-management skills and to learn to adjust to their chronic disease and its impact on daily life. However, the implementation of iCBT in clinical practice is challenging.

Objective: This study aimed to examine the current degree of implementation of iCBT among psychologists in a medical setting and discover determinants influencing the implementation of iCBT among nonusers.

Methods: A Web-based survey, based on the Unified Theory of Acceptance and Use of Technology (UTAUT), was distributed among psychologists in a medical setting. The survey included questions regarding the current use of iCBT, intention to use iCBT in the future, and operationalized concepts of the UTAUT constructs, that is, performance expectancy (PE), effort expectancy (EE), social influence (SI), and facilitating conditions (FC).

Results: In total, 107 (24.8%) psychologists completed the survey. Of them, 16.8% have access to iCBT, 15.9% currently use iCBT, and 21.5% are expected to use iCBT within the next year. The constructs PE, EE, and SI together significantly influenced behavioral intention (BI; mean 3.9 [SD=0.8]) among nonusers (R^2=0.490; $F_{4.85}$=20.405; P<.001).

Conclusions: In spite of an average to high BI, the current implementation of iCBT is rather low among psychologists in a medical setting. Further research should focus on reducing the gap between intention to use and actual use by focusing on influencing the predictive UTAUT constructs.

(J Med Internet Res 2019;21(8):e13432) doi:10.2196/13432

KEYWORDS

eHealth; chronic care; self-management; implementation; psychologists; UTAUT

Introduction

Background

Patients with a chronic somatic illness often struggle with a complex interaction of physical, psychological, and social demands related to their disease. These demands significantly influence their disease perception and quality of life [1-3]. Supporting patients in increasing their disease self-management can effectively increase illness adjustment and adherence to treatment, and decrease problems related to illness behaviors and comorbid mental illnesses such as depression and anxiety [4-6]. Psychologists can offer this support as part of outpatient treatment in a hospital clinic. In the Netherlands, these psychologists are called medical psychologists, and patients are referred to them by their medical specialist. Regarding the content of this support, cognitive behavioral therapy (CBT)–based techniques are regularly used. However, CBT is generally only available to a small number of patients in this setting because it is expensive, time consuming [7,8], and sometimes inaccessible because of attendance barriers [9,10].

XSL•FO
RenderX

Internet-Based Cognitive Behavioral Therapy

With the development of internet-based treatment [11], CBT has become much more accessible to a wide range of patients [4]. In the Netherlands, internet-based CBT (iCBT) is usually offered as a guided program, containing modules regarding goal setting, psychoeducation, (behavioral) assignments, relaxation exercises, and diary registrations, which are supported by asynchronous contact with a therapist through messages in a secured email box [12]. The iCBT service can be used in several ways, depending on the severity of the problems of the patient and the patient preferences. It can be provided as a stand-alone service, in which it is a replacement of a regular therapy, or it can be used as an addition to regular treatment, in which iCBT is used blended, combined with face-to-face sessions. In all scenarios, patients use the Web portal from their home environment. Use of iCBT potentially produces similar overall effects compared with face-to-face CBT [13-15] and is significantly efficacious in improving disease-specific symptoms, disease control [9,16-18], and disease-related physical outcomes [13]. Moreover, treatment by traditional forms of psychological therapy can be associated with stigmatizing beliefs about mental illness and negative prejudices about therapists [19], whereas iCBT is usually experienced as easy to access and the most private way to seek help [20]. Moreover, iCBT provides patients the ability to administer treatment and access treatment-related material at any time and any place [16]. Finally, regarding economical perspectives, offering iCBT enables a potential reduction in health care costs [21], the number of needed contact hours, waiting lists, and traveling time for patients [22].

Current Implementation of Internet-Based Cognitive Behavioral Therapy

However, implementation of iCBT in clinical practice remains a challenge [23,24], and different health care settings show different usages of iCBT. Results of a naturalistic study among therapists of a Dutch mental health center showed that after training, over a period of 3 years, only 3.6% of the patients were offered the possibility to participate in iCBT, initiated by 18% of the therapists qualified to provide iCBT [25]. On the contrary, a study among Dutch primary care psychologists (PCPs) and mental health counselors (MHCs) in general practitioner practices showed more optimistic numbers, with 29% of the PCPs and 60% of the MHCs having used Web-based psychological self-management interventions (based on CBT) in their treatments [26]. Regarding iCBT in the hospital setting, research is lacking. Therefore, the aim of this study was to gain insight into use, barriers, and facilitators regarding implementation of iCBT among psychologists in medical settings.

Research into possible factors influencing the degree of implementation of iCBT and iCBT-related interventions among psychologists is relatively young. Dissemination and implementation of digital health interventions often face many barriers on differing levels, for example, related to perceived effectiveness, expected usability, usefulness in the patient population, facilitating conditions (FC) at work, and personal productivity [26,27]. A suitable model to measure acceptance of technology and its associated influences on those multiple levels is the Unified Theory of Acceptance and Use of Technology (UTAUT) [28]. On the basis of this model, determinants influencing behavioral intention (BI) of iCBT are mapped, related to future users' performance expectancy (PE), effort expectancy (EE), social influence (SI), and available FC. [28].

Use of the Unified Theory of Acceptance and Use of Technology

According to the UTAUT model, 4 constructs are crucial in explaining the use of technology, namely *PE*, *EE*, *SI*, and *FC*. *PE* indicates "the degree to which an individual believes that using the system will help him or her to attain gains in job performance" [28]. For psychologists in medical settings, this would mean that the use of iCBT may or may not contribute to their quality of work in a positive manner, for instance, the use of iCBT reduces time that is required to fulfill major responsibilities (eg, providing help to patients). *EE* can be defined as "the degree of ease associated with the use of the system" [28]. In practice, this would mean that the use of iCBT may or may not be clear, understandable, and easy to learn for psychologists. *SI* indicates "the degree to which an individual perceives that important others believe he or she should use the new system" [28]. In practice, this would mean that psychologists may or may not believe that colleagues and management at the workplace think they should use iCBT and support their use of iCBT. *FC* indicates "the degree to which an individual believes that an organizational and technical infrastructure exists to support the use of the system" [28]. For psychologists in medical settings, this would indicate that the use of iCBT may or may not be stimulated by the organization or manager. Moreover, the construct *BI* is a crucial determinant for use behavior and can be defined as "a person's subjective probability to perform a specified behavior" [29]. In practice, this would mean the likelihood of use of iCBT among psychologists in medical settings.

The purpose of this study was to examine the fit and current degree of implementation of iCBT among psychologists in a medical setting and explore the extent to which the determinants of the UTAUT model influence the degree of implementation among nonusers. This was done using a nationally spread survey in the Netherlands to gain broad and generalizable insights into barriers and facilitators regarding implementation of iCBT in hospitals.

Methods

Recruitment

Respondents were recruited by (personal) email in March 2017. Dissemination of the survey was done by the Dutch Association of Medical Psychologists, and email addresses were collected by systematically searching hospitals per province (using Zorgkaart Nederland and Google), after which these hospitals were approached. In case no information on names or email address of psychologists was found on the website, the secretariat of the hospital was contacted by phone to ask whether they would agree to share their psychologists' contact

XSL•FO
RenderX

information. This resulted in a total initial spreading among 432 addresses. However, a snowball effect might have occurred because psychologists were asked to forward our invitation to colleagues to whom it might be relevant as well.

The email included an information letter and a link to the Web-based survey in Qualtrics (2015 Qualtrics, LLC). The information letter explained the purpose of the study and its voluntary nature, use and anonymization of the data, estimated time needed to participate (10 min), and the reward for participating in the study (5 gift certificates of €50 were raffled among the respondents). All email addresses were stored separately and were only used for the raffle. The Web-based survey started with an informed consent form. Not responding to the survey or opting out in the informed consent form was considered as choosing not to participate in the study. Psychologists who were employed at multiple hospitals were asked to choose 1 hospital and fill in the survey based on their chosen hospital. If the survey was not fully completed, data were not included in further statistical analysis. Furthermore, 1 and 2 weeks after sending the invitation emails, reminder emails were sent to all email addresses. The study was approved by the Psychology Research Ethics Committee of Leiden University.

Survey

The survey consisted of 3 parts, namely (1) background information about the psychologist and the hospital, (2) general and work-related use of the internet of the respondent, and (3) an operationalization of the UTAUT constructs. The background information included gender, age, professional background, number of working hours per week, and years of employment. Background information about the hospital included number of colleagues working as a psychologist, number of newly referred patients per month, reasons for referral of patients, consult duration per visit, number of contacts per patient per treatment trajectory, and type of primarily offered help to patients. In part 2, general and work-related use of the internet was measured using questions about quantity of internet use in general, self-perceived skills to use the internet, use of internet for work-related activities, and current available Web-based applications in the hospital.

To operationalize the UTAUT, subscales were created for each construct in the model (PE, EE, SI, FC, and BI). Regarding BI, respondents first answered whether using iCBT was part of their current work activities, after which they had to answer whether they had the intention to start using iCBT in case they currently did not use it. By measuring actual use of iCBT, a distinction between users and nonusers was possible. Subscales of the UTAUT constructs were measured with 3 to 8 items formulated as statements, for instance, “I expect/experience that guided iCBT is effective for my patient population” (PE), “I expect/experience that guided iCBT is time intensive to use” (EE), “I expect/experience that iCBT is seen as a positive development among my colleagues” (SI), “I expect/experience that iCBT fits within the technological circumstances of my practice” (FC), and “I want to use/continue to use guided iCBT” (BI). Respondents answered statements on a 5-point Likert scale, ranging from completely disagree (1) to completely agree (5). A complete overview of the survey is provided in Multimedia Appendix 1.

Statistical Analysis

Statistical analyses were performed using the Statistical Package for the Social Sciences (IBM SPSS Statistics 24). Descriptive statistics were used to describe the study sample, their internet experience, and their scores on the UTAUT constructs. Scale scores and Cronbach alphas were calculated for each UTAUT construct to check internal reliability. In the sample of nonusers, Pearson correlations were calculated to determine the relationship between the UTAUT variables PE, EE, SI, and FC and the dependent variable BI. On the basis of the statistical significance of these relations, multiple regression analysis was carried out to examine the further nature of these relations.

Results

Respondents

A total of 127 psychologists in medical settings filled in the survey. Among them, 107 completed the survey, whereas 20 stopped early. These data were not used for analyses and contained 11.8% of the data. Among the respondents, 58 (58/107, 54.2%) stated that they are members of the Dutch Association of Medical Psychologists, whereas 49 (49/107, 45.8%) stated that they are not members. The estimated response rate is 24.8%, as 432 people were contacted for participation via their email addresses. However, because of the method of data collection, it is not possible to calculate an exact response rate. Table 1 provides an overview of the respondents’ characteristics. Most respondents were female (92/107, 86.0%), with a mean age of 40.9 years (SD=11.3). A large part of the respondents had a professional background as a health care psychologist (45/107, 42.1%) or clinical psychologist (27/107, 25.2%; these are both protected titles in the Netherlands, based on postmaster educational tracks). On average, respondents have been active as psychologists in a medical setting for 10.0 (SD=8.1) years and received up to 20 (69/107, 64.5%) new patients each month (data not shown in Table 1). Regarding consultations, most patients were seen for 5 to 10 times (60/107, 56.1%), and therapy sessions lasted an average of 53 min per session. With regard to characteristics of the care provided, respondents most often reported treating problems with dealing with chronic physical complaints and limitations (93/107, 86.9%), anxiety- and/or mood-related problems (68/107, 63.6%), and pain (39/107, 36.4%). Finally, the type of care primarily provided consisted of interventions aimed at improving mental functioning (93/107, 86.9%), problem clarification and diagnostics (90/107, 84.1%), psychoeducation (80/107, 75.7%), and guiding and supporting self-management (65/107, 60.7%).

XSL•FO
RenderX

Table 1. Background information of the respondents and their hospital (N=107).

Characteristics	Statistics
Gender, n (%)	
Men	15 (14.0)
Women	92 (86.0)
Age (years), mean (SD)	40.5 (11.5)
Professional background, n (%)	
Psychologist, MSc	13 (12.1)
Health care psychologist	45 (42.1)
Clinical psychologist	27 (25.2)
Clinical neuropsychologist	2 (1.9)
Other (eg, in training to become health care psychologist)	20 (18.7)
Number of consultations during 1 complete treatment, n (%)	
<5	39 (36.4)
5-10	60 (56.1)
>10	8 (7.5)
Average number of minutes spent per therapy session, mean (SD)	53.9 (29.2)

Internet Experience

Most respondents used the internet on a daily basis (102/107, 95.3%) and rated their internet skills as (very) good (92/107, 86%). Concerning work functionalities, respondents mainly used the internet to search for medical information (103/107, 96.3%) or have contact with clients through email (71/107, 66.4%). Regarding availability of Web-based health applications in the hospital, respondents primarily reported availability of electronic medical records (104/107, 97.2%), a website with patient information (73/107, 68.2%), and a Web portal with patient records (46/107, 43%). Electronic or Web-based screening (21/107, 19.6%), Web-based self-management modules (18/107, 16.8%), and use of teleconsultation (10/107, 9.3%) were less common.

Use of Guided Internet-Based Cognitive Behavioral Therapy

Regarding the type of problems for which the respondents considered iCBT as an appropriate treatment, anxiety- and/or mood-related problems were most often reported (81/107, 75.7%), followed by problems dealing with chronic physical complaints and limitations (78/107, 72.9%), sleep problems (72/107, 67.3%), and fatigue problems (70/107, 65.4%).

Table 2 shows that 15.9% (17/107) of the respondents used guided iCBT at the time of filling in the survey. More than half of the respondents (68/107, 63.6%) had seen guided iCBT programs before. Among this subgroup, almost three-fourths (49/68, 72.1%) considered iCBT a suitable treatment for problems related to dealing with chronic physical complaints and limitations, and more than three-fourths (54/68, 79.4%) considered iCBT a suitable treatment for anxiety- and/or mood-related problems. In addition, almost one-third of the respondents (31/107, 29%) had received training in the use of guided iCBT and more than a quarter used it as a part of their work activities in the past (31/107, 29%). Within these subgroups, more than half of the respondents (18/31, 58.1% in both subgroups) considered iCBT a suitable treatment for problems related to dealing with chronic physical complaints and limitations, and more than three-fourths (26/31, 83.9% and 25/31, 80.6%, respectively) considered iCBT a suitable treatment for anxiety- and/or mood-related problems. Finally, of the (former) users, a majority (20/31, 64.5%) had used guided iCBT in less than 10 treatments. Regarding future use of iCBT, more than one-fourth of all nonusers (23/90, 25.6%) expected to use guided iCBT within the next year, and more than half (57/90, 63.3%) expected to use guided iCBT within 2 to 5 years.

Facilitators and Barriers

Distribution of the scores on the UTAUT constructs was normal. Internal consistency was acceptable to excellent (see Table 3). On average, responses varied between the answer categories *neutral* and *partly agree*.

Influence of Unified Theory of Acceptance and Use of Technology Constructs on Behavioral Intention Among Nonusers

Pearson correlations showed statistically significant moderate to high positive associations between the UTAUT constructs PE, EE, SI, and FC and the dependent variable of BI, respectively (PE r=0.656; EE r=0.479; SI r=0.251; and FC r=0.443; P<.001 for PE, EE, and FC, and P=.02 for SI). Table 4 shows the results of the multiple regression analysis using the enter method. The multiple regression model with the dependent variable BI explained 49% of variance in BI (R^2=0.490; $F_{4,85}$=20.405; P<.001). The constructs PE, EE, and FC had a significant positive effect on BI regarding use of iCBT, whereas SI did not.

XSL•FO
RenderX

Table 2. Use of guided internet-based cognitive behavioral therapy (N=107).

Question	Statistics, n (%)
Current use of guided iCBT[a]	17 (15.9)
Previous experience with guided iCBT	
Has seen a program of guided iCBT	68 (63.6)
Has received a training in the use of guided iCBT	31 (29.0)
Has applied guided iCBT in previous treatments	31 (29.0)
Completed guided iCBT treatment programs by (former) users (n=31)	
<10 treatments	20 (64.5)
10-20 treatments	4 (12.9)
>20 treatments	7 (22.6)
Expected time frame of iCBT usage in the future among nonusers (n=90)	
Within the next year	23 (25.6)
Within 2-5 years	57 (63.3)
Not within the next 5 years	9 (10.0)
Never	1 (1.1)

[a]iCBT: internet-based cognitive behavioral therapy.

Table 3. Scores on the Unified Theory of Acceptance and Use of Technology constructs (total N=107; users N=17; nonusers N=90).

Construct	Cronbach alpha	Statistics, mean (SD)[a]
Behavioral intention	.90	3.9 (.8)
Users	—[b]	4.5 (.7)
Nonusers	—	3.8 (.8)
Performance expectancy	.80	3.7 (.4)
Users	—	3.7 (.3)
Nonusers	—	3.7 (.4)
Effort expectancy	.71	3.5 (.4)
Users	—	3.5 (.3)
Nonusers	—	3.5 (.4)
Social influence	.78	3.0 (.7)
Users	—	3.2 (.7)
Nonusers	—	2.9 (.6)
Facilitating conditions	.68	3.4 (.6)
Users	—	3.5 (.6)
Nonusers	—	3.3 (.6)

[a]Responses answered all statements of all constructs by a 5-point Likert scale, namely 1: completely disagree, 2: partly disagree, 3: neutral, 4: partly agree, and 5: completely agree.

[b]Not applicable.

XSL•FO
RenderX

Table 4. Results multiple regression analysis of Unified Theory of Acceptance and Use of Technology constructs on behavioral intention (N=90).

Model	B[a]	Beta[b]	SE	*P* value
Constant	−1.823	__[c]	0.669	.01
Performance expectancy	0.803	0.438	0.176	<.001
Effort expectancy	0.625	0.288	0.187	.001
Social influence	−0.222	−0.174	0.139	.11
Facilitating conditions	0.355	0.270	0.154	.02

[a]B: partial regression coefficient.

[b]Beta: standardized regression coefficient.

[c]Not applicable.

Discussion

Principal Findings

Guided iCBT is a suitable and effective treatment for patients with a chronic somatic illness. However, its implementation in clinical practice remains challenging. The main purpose of this study was to examine the fit and the current degree of implementation of iCBT among psychologists in medical settings and determine its facilitators and barriers among nonusers. Most respondents in our sample were female (92/107, 86.0%) and had an official registration as psychologist, which is representative for the Dutch situation [30].

The results of the study showed that the use of technology is highly integrated among psychologists, both on a personal level and work-related level. Moreover, respondents generally considered iCBT as an appropriate treatment for the type of problems they treat their patients for, regardless of their (previous) experience with iCBT. The actual current use of iCBT, however, was limited. Our results show that only 16% (17/107) of the respondents is currently using iCBT. In addition, 29% (31/107) had received training in the use of guided iCBT, and 29% (31/107) had used it as a part of their work activities in the past. This raises the question why only such a small portion of psychologists actually use these Web-based treatment tools nowadays and what happened when they stopped using them. Presumable reasons for ending the use of iCBT could be that people have changed organizations and were not satisfied with the Web-based programs they used or, for instance, a change of policy from their management.

Our results concerning outcomes on the UTAUT model provide some insight into these questions and into what is further needed to increase implementation success. BI regarding the use of iCBT was found to be high among nonusers, with 98.9% (89/90) planning to use it within the upcoming 5 years; therefore, there is a high willingness to use Web-based tools in treatment. We also found that PE, EE, and FC appear to be significant predictors for this BI. These results indicate that aspects such as effectiveness for the client population, easiness of use, and managerial focus impact psychologists' choice in the use of iCBT. These outcomes are in line with recent research on iCBT use in primary and routine care [26,27]. Still, the answers on most constructs had a small dispersion (spreading of the answers; responses were located around the answer categories *partly agree* or *neutral*), indicating that the respondents did not have very outspoken expectations on these determinants. This might indicate an implementation barrier in itself because the needs and beliefs about iCBT might not be strong enough to actually create behavioral change. To create actual change, it seems of major importance that the availability of Web-based self-management tools increases in hospital settings because, currently, only 17% (18/107) have access to iCBT. As a result, a larger part of psychologists in these settings will actually be able to use iCBT when willing to.

In the Netherlands, there is quite a large number of commercially provided iCBT portals, in which hosting and updating is outsourced. Psychologists are able to make use of these programs using a paid license. Although we now know that practicing psychologists would be intended to use these iCBT services, the actual access is probably often depending on the willingness of their management to invest in iCBT services. Future implementation efforts should therefore initially focus on these types of stakeholders, using implementation models that take a broad range of stakeholders into account. An example of such a model is the Consolidated Framework for Implementation Research [31], addressing 5 domains to include in implementation research. These are intervention characteristics (eg, adaptability and relative advantage), the setting within the organization (eg, culture, leadership engagement, and communication), the environment outside the organization (eg, external policies and incentives, peer pressure, needs, and resources), characteristics of the individuals involved (eg, self-efficacy, personal attributes, and individual stages of change), and the process of implementation (eg, planning, execution, and internal implementation leaders).

Future Directions

Based on our results using the UTAUT model, we can further reflect on what follow-up steps in improving implementation success seem essential, partly on an individual level. Concerning PE and EE, this study showed that fewer than one-third of the respondents have received a training in the use of iCBT. Therefore, expectations about whether iCBT asks a lot of new skills, affects work productivity, and increases the quality of care provision appear to affect the implementation process of iCBT. Moreover, these findings are consistent with previous research stating that lack of knowledge of the program and expectations regarding effectiveness and fully mastering the required protocol influence acceptance of iCBT [8,27,32,33].

XSL•FO
RenderX

Therefore, attention should be paid to providing employees more education in iCBT by offering knowledge with regard to availability of existing iCBT programs, their effectiveness, usefulness, and performance productivity and by offering them proper skills training [27,34]. Regarding FC, stimulation of iCBT use from a management level is a key prerequisite for use in daily practice. Moreover, attention should be paid to the degree of customization of iCBT to the daily workflow of psychologists.

Limitations

With this survey, we aimed to reach the most representative sample possible by trying to contact each psychologist in a medical setting in the Netherlands. The possibility exists that people interested in electronic health (eHealth) or actual eHealth users were more inclined to respond to the survey. This could result in fewer psychologists actually using iCBT in the complete population than the results from this survey show. In addition, social desirability may have influenced the provided answers of the respondents because the government encourages the use of eHealth among health professionals, which might have pressured our respondents toward positive answers. Furthermore, it is not possible to exactly estimate how many psychologists actually received our survey because invited psychologists were asked to forward our invitation to colleagues to whom it might be relevant as well. However, our estimated response rate of 24.8% is in line with the average response rate of large surveys [35]. Moreover, by sending out 2 reminders each 7 days after the first invitation email and by making use of personalization, appropriate measures were used to increase the response rate as much as possible [36]. Delay between reminders was based on research stating no significant difference in response rate for follow-up emails sent after 1 or 2 weeks and recommending time lags of 1 week [37].

Conclusions

Despite an average to high BI, current implementation of iCBT is low among psychologists in medical settings. Increasing the availability of iCBT and bringing change in other areas that could influence the gap between intention and actual use is needed. The UTAUT constructs PE, EE, and FC significantly affect BI regarding use of iCBT. Therefore, further research and implementation trajectories should focus on offering a wide range of iCBT programs, education on iCBT facilities and required skills, and customization of iCBT to the daily workflow of psychologists.

Acknowledgments

The authors thank all the health professionals who participated in the study. They also thank the Dutch Society of Medical Psychologists (Landelijke Vereniging Medische Psychologie) for their participation in spreading the survey among their members and Maaike Alvares, Simone Daamen, Nadine Haasnoot, Giovanna Haneveld, Annemarie van der Kaaij, Lenneke Peijs, Pien van Putte, Nadia van Silfhout, and Demi de Vries for their contribution to the data collection.

Conflicts of Interest

None declared.

Multimedia Appendix 1
English version of the survey as distributed among medical psychologists in the Netherlands.
[PDF File (Adobe PDF File), 144KB - jmir_v21i8e13432_app1.pdf]

References

1. 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]
2. Sensky T. Patients' reactions to illness. Br Med J 1990 Mar 10;300(6725):622-623 [FREE Full text] [doi: 10.1136/bmj.300.6725.622] [Medline: 2182151]
3. Strine TW, Chapman DP, Balluz LS, Moriarty DG, Mokdad AH. The associations between life satisfaction and health-related quality of life, chronic illness, and health behaviors among US community-dwelling adults. J Community Health 2008 Feb;33(1):40-50. [doi: 10.1007/s10900-007-9066-4] [Medline: 18080207]
4. Barlow J, Wright C, Sheasby J, Turner A, Hainsworth J. Self-management approaches for people with chronic conditions: a review. Patient Educ Couns 2002;48(2):177-187. [doi: 10.1016/S0738-3991(02)00032-0] [Medline: 12401421]
5. Halford J, Brown T. Cognitive–behavioural therapy as an adjunctive treatment in chronic physical illness. Adv Psychiatr Treat 2018 Jan 2;15(4):306-317. [doi: 10.1192/apt.bp.107.003731]
6. Hofmann SG, Asnaani A, Vonk IJ, Sawyer AT, Fang A. The efficacy of cognitive behavioral therapy: a review of meta-analyses. Cognit Ther Res 2012 Oct 1;36(5):427-440 [FREE Full text] [doi: 10.1007/s10608-012-9476-1] [Medline: 23459093]
7. Hans E, Hiller W. Effectiveness of and dropout from outpatient cognitive behavioral therapy for adult unipolar depression: a meta-analysis of nonrandomized effectiveness studies. J Consult Clin Psychol 2013 Feb;81(1):75-88. [doi: 10.1037/a0031080] [Medline: 23379264]

XSL•FO
RenderX

8. Wilhelmsen M, Høifødt RS, Kolstrup N, Waterloo K, Eisemann M, Chenhall R, et al. Norwegian general practitioners' perspectives on implementation of a guided web-based cognitive behavioral therapy for depression: a qualitative study. J Med Internet Res 2014 Sep 10;16(9):e208 [FREE Full text] [doi: 10.2196/jmir.3556] [Medline: 25208886]
9. Beatty L, Lambert S. A systematic review of internet-based self-help therapeutic interventions to improve distress and disease-control among adults with chronic health conditions. Clin Psychol Rev 2013 Jun;33(4):609-622. [doi: 10.1016/j.cpr.2013.03.004] [Medline: 23603521]
10. Mohr DC, Ho J, Duffecy J, Baron KG, Lehman KA, Jin L, et al. Perceived barriers to psychological treatments and their relationship to depression. J Clin Psychol 2010 Apr;66(4):394-409 [FREE Full text] [doi: 10.1002/jclp.20659] [Medline: 20127795]
11. Kontos E, Blake KD, Chou WS, Prestin A. Predictors of ehealth usage: insights on the digital divide from the Health Information National Trends Survey 2012. J Med Internet Res 2014 Jul 16;16(7):e172 [FREE Full text] [doi: 10.2196/jmir.3117] [Medline: 25048379]
12. Andersson G, Hedman E. Effectiveness of guided internet-based cognitive behavior therapy in regular clinical settings. Verhaltenstherapie 2013;23(3):140-148. [doi: 10.1159/000354779]
13. van Beugen S, Ferwerda M, Hoeve D, Rovers MM, Spillekom-van Koulil S, van Middendorp H, et al. Internet-based cognitive behavioral therapy for patients with chronic somatic conditions: a meta-analytic review. J Med Internet Res 2014 Mar 27;16(3):e88 [FREE Full text] [doi: 10.2196/jmir.2777] [Medline: 24675372]
14. 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]
15. Cuijpers P, Donker T, van Straten A, Li J, Andersson G. Is guided self-help as effective as face-to-face psychotherapy for depression and anxiety disorders? A systematic review and meta-analysis of comparative outcome studies. Psychol Med 2010 Dec;40(12):1943-1957. [doi: 10.1017/S0033291710000772] [Medline: 20406528]
16. Macea DD, Gajos K, Calil YAD, Fregni F. The efficacy of web-based cognitive behavioral interventions for chronic pain: a systematic review and meta-analysis. J Pain 2010 Oct;11(10):917-929. [doi: 10.1016/j.jpain.2010.06.005] [Medline: 20650691]
17. Newman MG, Szkodny LE, Llera SJ, Przeworski A. A review of technology-assisted self-help and minimal contact therapies for anxiety and depression: is human contact necessary for therapeutic efficacy? Clin Psychol Rev 2011 Feb;31(1):89-103. [doi: 10.1016/j.cpr.2010.09.008] [Medline: 21130939]
18. 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]
19. Sibold M, Mittag O, Kulick B, Müller E, Opitz U, Jäckel WH. [Predictors of participation in medical rehabilitation follow-up in working patients with chronic back pain]. Rehabilitation (Stuttg) 2011 Dec;50(6):363-371. [doi: 10.1055/s-0031-1271815] [Medline: 21647850]
20. Kauer SD, Mangan C, Sanci L. Do online mental health services improve help-seeking for young people? A systematic review. J Med Internet Res 2014 Mar 4;16(3):e66 [FREE Full text] [doi: 10.2196/jmir.3103] [Medline: 24594922]
21. 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]
22. Cuijpers P, van Straten A, Andersson G. Internet-administered cognitive behavior therapy for health problems: a systematic review. J Behav Med 2008 Apr;31(2):169-177 [FREE Full text] [doi: 10.1007/s10865-007-9144-1] [Medline: 18165893]
23. Fleming TM, de Beurs D, Khazaal Y, Gaggioli A, Riva G, Botella C, et al. Maximizing the impact of e-therapy and serious gaming: time for a paradigm shift. Front Psychiatry 2016;7:65 [FREE Full text] [doi: 10.3389/fpsyt.2016.00065] [Medline: 27148094]
24. Mohr DC, Riper H, Schueller SM. A solution-focused research approach to achieve an implementable revolution in digital mental health. JAMA Psychiatry 2018 Feb 1;75(2):113-114. [doi: 10.1001/jamapsychiatry.2017.3838] [Medline: 29238805]
25. Kenter RMF, van de Ven PM, Cuijpers P, Koole G, Niamat S, Gerrits RS, et al. Costs and effects of internet cognitive behavioral treatment blended with face-to-face treatment: results from a naturalistic study. Internet Interv 2015 Mar;2(1):77-83. [doi: 10.1016/j.invent.2015.01.001]
26. van der Vaart R, Atema V, Evers AWM. Guided online self-management interventions in primary care: a survey on use, facilitators, and barriers. BMC Fam Pract 2016 Mar 9;17:27 [FREE Full text] [doi: 10.1186/s12875-016-0424-0] [Medline: 26961547]
27. Folker AP, Mathiasen K, Lauridsen SM, Stenderup E, Dozeman E, Folker MP. Implementing internet-delivered cognitive behavior therapy for common mental health disorders: a comparative case study of implementation challenges perceived by therapists and managers in five European internet services. Internet Interv 2018 Mar;11:60-70 [FREE Full text] [doi: 10.1016/j.invent.2018.02.001] [Medline: 30135761]
28. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Q 2003;27(3):435-478 [FREE Full text] [doi: 10.2307/30036540]

XSL•FO
RenderX

29. Yi MY, Jackson JD, Park JS, Probst JC. Understanding information technology acceptance by individual professionals: toward an integrative view. Inform Manage 2006 Apr;43(3):350-363. [doi: 10.1016/j.im.2005.08.006]
30. BIG-Register. 2019. [Basic Occupation and Specialty] URL: https://www.bigregister.nl/over-het-big-register/cijfers/basisberoep-en-specialisme [accessed 2019-04-15] [WebCite Cache ID 77eVg8Taa]
31. Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. Implement Sci 2009 Aug 7;4:50 [FREE Full text] [doi: 10.1186/1748-5908-4-50] [Medline: 19664226]
32. Kivi M, Eriksson MCM, Hange D, Petersson EL, 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 Interv 2015 Sep;2(3):248-256. [doi: 10.1016/j.invent.2015.06.001]
33. Yang Z, Wang X, Su C. A review of research methodologies in international business. Int Bus Rev 2006 Dec;15(6):601-617. [doi: 10.1016/j.ibusrev.2006.08.003]
34. Gagnon MP, Desmartis M, Labrecque M, Car J, Pagliari C, Pluye P, et al. Systematic review of factors influencing the adoption of information and communication technologies by healthcare professionals. J Med Syst 2012 Feb;36(1):241-277 [FREE Full text] [doi: 10.1007/s10916-010-9473-4] [Medline: 20703721]
35. Sánchez-Fernández J, Muñoz-Leiva F, Montoro-Ríos FJ. Improving retention rate and response quality in web-based surveys. Comput Hum Behav 2012 Mar;28(2):507-514. [doi: 10.1016/j.chb.2011.10.023]
36. Yarbrough AK, Smith TB. Technology acceptance among physicians: a new take on TAM. Med Care Res Rev 2007 Dec;64(6):650-672. [doi: 10.1177/1077558707305942] [Medline: 17717378]
37. Sauermann H, Roach M. Increasing web survey response rates in innovation research: an experimental study of static and dynamic contact design features. Res Policy 2013 Feb;42(1):273-286. [doi: 10.1016/j.respol.2012.05.003]

Abbreviations

BI: behavioral intention
CBT: cognitive behavioral therapy
eHealth: electronic health
EE: effort expectancy
FC: facilitating conditions
iCBT: internet-based cognitive behavioral therapy
MHCs: mental health counselors
PCPs: primary care psychologists
PE: performance expectancy
SI: social influence
UTAUT: Unified Theory of Acceptance and Use of Technology

Edited by G Eysenbach; submitted 17.01.19; peer-reviewed by M Kivi, K Mathiasen, K Stasiak; comments to author 28.02.19; revised version received 24.04.19; accepted 20.05.19; published 09.08.19.

Please cite as:
IJzerman RVH, van der Vaart R, Evers AWM
Internet-Based Cognitive Behavioral Therapy Among Psychologists in a Medical Setting: A Survey on Implementation
J Med Internet Res 2019;21(8):e13432
URL: https://www.jmir.org/2019/8/e13432/
doi:10.2196/13432
PMID:31400101

©Renée V H IJzerman, Rosalie van der Vaart, Andrea W M Evers. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 09.08.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Safety of Electronic Cigarette Use During Breastfeeding: Qualitative Study Using Online Forum Discussions

Emily Jade Johnston[1], BSc, MSc; Katarzyna Campbell[1], MSc, PhD; Tim Coleman[1], MD, FRCGP; Sarah Lewis[2], MSc, PhD; Sophie Orton[1], MSc, PhD; Sue Cooper[1], MSc, PhD

[1]Division of Primary Care, School of Medicine, University of Nottingham, Nottingham, United Kingdom

[2]Division of Epidemiology and Public Health, School of Medicine, University of Nottingham, Nottingham, United Kingdom

Corresponding Author:
Emily Jade Johnston, BSc, MSc
Division of Primary Care
School of Medicine
University of Nottingham
Towers Building, 14th Floor
University Park
Nottingham, NG7 2RD
United Kingdom
Phone: 44 1157484622
Email: msxejj@nottingham.ac.uk

Abstract

Background: Electronic cigarettes (e-cigs) are an increasingly popular alternative to smoking, helping to prevent relapse in those trying to quit and with the potential to reduce harm as they are likely to be safer than standard cigarettes. Many women return to smoking in the postpartum period having stopped during pregnancy, and while this can affect their decisions about breastfeeding, little is known about women's opinions on using e-cigs during this period.

Objective: The aim of this study is to explore online forum users' current attitudes, motivations, and barriers to postpartum e-cig use, particularly as a breastfeeding mother.

Methods: Data were collected via publicly accessible (identified by Google search) online forum discussions, and *a priori* codes identified. All transcripts were entered into NVivo for analysis, with a template approach to thematic analysis being used to code all transcripts from which themes were derived.

Results: Four themes were identified: use, perceived risk, social support and evidence, with a number of subthemes identified within these. Women were using e-cigs to prevent postpartum return to smoking, but opinions on their safety were conflicting. They were concerned about possible transfer of harmful products from e-cigs via breastmilk and secondhand exposure, so they were actively seeking and sharing information on e-cigs from a variety of sources. Although some women were supportive of e-cig use, others provided harsh judgement for mothers who used them.

Conclusions: E-cigs have the potential to reduce the number of women who return to smoking in the postpartum period and potentially improve breastfeeding rates, if breastfeeding mothers have access to relevant and reliable information. Health care providers should consider discussing e-cigs with mothers at risk of returning to smoking in the postpartum period.

(J Med Internet Res 2019;21(8):e11506) doi:10.2196/11506

KEYWORDS

e-cigarette; online forum; postpartum relapse; smoking; breastfeeding; forum data

Introduction

Maternal smoking and low breastfeeding rates are both major public health concerns relating to the postpartum period, with health implications for both the mother and her child [1-8]. While many mothers are able to quit smoking during pregnancy, a substantial proportion will return to smoking by six months postpartum [9]. The latest UK statistics show that, despite 81% of mothers initiating breastfeeding at birth, by six months only 1% of UK infants are still breastfed [10].

Studies have consistently reported associations between smoking behavior (abstinence) and breastfeeding patterns [11-15], with

XSL•FO
RenderX

the intention to breastfeed acting as a precipitating factor for reducing postpartum return to smoking, and then the initiation and continuation of breastfeeding being positively associated with smoking abstinence postpartum [16,17]. The intention to return to smoking is one of the strongest predictors of the intention not to breastfeed and the early cessation (<3 months postpartum) of breastfeeding [13,18]. These associations may be partly explained by confounders like sociodemographic factors [19,20], but they may also be attributable to concerns regarding safety of smoking while breastfeeding. This is despite online information from both the American Pediatrics Association [21,22] and the National Health Service (NHS) [23] that promotes the continuation of breastfeeding even if the mother smokes.

A relatively new product that may be useful for preventing return to smoking or supporting cessation of smoking is the electronic cigarette (e-cig) [24]. E-cigs are handheld devices that produce an aerosolized mixture from a solution typically containing concentrated nicotine, flavoring chemicals and propylene glycol [25]. The user draws a deep breath and inhales the vaporized liquid (a process known as vaping), which creates a similar experience to smoking combustible tobacco [26]. Although we accept that the long-term health effects of e-cigs cannot be fully assessed given the relatively short time they have been available for use, the current research suggests that e-cigs are likely safer than smoking traditional cigarettes and their use is proposed as a harm reduction tool for smokers [27,28].

Online forums are a popular and accessible community for mothers to discuss, debate and share opinions on anything relating to motherhood [29]. A method of research known as infodemiology analyzes online discussions to inform public health research [30,31], and previous studies of this type have studied e-cigs [32,33], breastfeeding [34] and parental health-seeking behavior [35]. With the popularity of parenting forums and increased use of social media for health-seeking advice, this study analyzed parenting forum discussions to determine opinions on e-cig use.

Very little research thus far has examined the use of e-cigs amongst postpartum women, with the current research available focusing on the use of e-cigs in pregnancy. Therefore, we aimed to explore online forum users' current attitudes, motivators and barriers to using e-cigs as a breastfeeding mother through the analysis of data extracted from online parenting forums.

Methods

Overview

This was an infodemiological study [30] using online forum data. Qualitative analysis of discussions on online parenting forums has previously been used to explore a wide range of context-specific behaviors, attitudes and beliefs [36,37], including the use of e-cigs during pregnancy [38]. Online support groups provide specific benefits of a virtual group membership compared to a physical group membership, such as being accessible 24-hours a day, 7 days a week, being free to join and participate in, lacking geographical barriers and offering anonymity [39-43]. The use of online support groups or forums can have potentially empowering effects on those who use them by, for example, providing health information, information sharing and input for individuals to make health-based decisions [39]. Online forums can also offer a safe place to discuss sensitive topics or topics in which a person feels they may be judged [44] and have been used for discussions regarding smoking [45]. Therefore, the use of parenting forums is a valuable source of data on what women think about health-related risks and their health-related decision making [46-48].

Inclusion Criteria

Due to varying guidelines between countries on e-cig use, only UK-based forums were used as the United Kingdom already has guidance recommending the use of e-cigs rather than traditional cigarettes during pregnancy [28]. The eligibility for inclusion of a thread (a continuous discussion on a forum) in the final analysis were: (1) it was posted to a forum that is open to public use, without the need to sign up or log in to read posts; (2) it was posted to parenting forums not affiliated with vaping or tobacco companies; (3) it contained a minimum of four unique contributors to the discussion; and (4) the discussion needed to include the mention of either e-cig use or vaping as well as breastfeeding.

Search Strategy

The key words for e-cigs (Textbox 1) were combined using the operator *AND* with key words for breastfeeding, and then searched using the search operator (a Google-based command to filter results) *site: sampleforum.co.uk* via Google search engine.

XSL•FO
RenderX

Textbox 1. Key words (search terms).

Key words: electronic cigarette

- E-Cig(s)
- E-cig(s)
- Electronic cigarette(s)
- Vaping
- Vape(s)

AND

Key words: breastfeeding

- Breastfeeding
- BF(ing)
- Nurse/Nursing
- Breastmilk
- Feeding

The use of search operators was the most effective and thorough way to ensure relevant discussions were obtained while ignoring forums owned by specific groups who may have had competing interests, such as e-cig manufacturers or tobacco companies. These sites were identified by screening the URL name and home page.

A total of 597 Google results were returned using the above search terms, and searches were then adapted to exclude *pregnancy* and *trying to conceive*, in line with the aims of the research. The threads used in analysis were then transcribed for NVivo11.

In the analysis, abbreviations within quotes were expanded in squared brackets, and the data source was identified by the thread (T) and the numbered data set.

Ethical Considerations and Data Collection

Informed individual consent was not obtained as the data were publicly posted on a large forum [49-52]. The British Psychological Society [53] recognizes that although informed consent might not be achievable in this context, certain steps can be taken to protect the participants. Therefore, only data from publicly accessible forums, where users are made aware during the initial sign-up process that all posts are open to public access, were used [54]. Furthermore, as comments on large public forums are less identifiable than those on smaller, private online communities [55], data was only obtained from large forums (for the purposes of this research, a large forum was defined as forums with over 1000 members) [55]. All contributing users were randomly assigned a new name to protect their identity, then names of people, places and institutions were removed from quotes and finally quotations were corrected for spelling and kept brief to reduce the possibility of them being traced back to the original poster.

Ethical approval was obtained from the University of Nottingham Medical School Research Ethics Committee.

Analysis

Template analysis (a template approach to thematic analysis), following the guidelines outlined by King [51,56], was used to analyze the data. Due to the use of *a priori* codes, this permits the analysis of the textual data that had been produced for "a different purpose in a different context" [57]. When analyzing large online support group datasets, template analysis is useful for comparing the perspectives of different contributors. In the current study, the initial template of *a priori* codes was used to code each transcript with codes being continually modified or expanded. After the last transcript was coded, a final version of the template was used to recode all transcripts (Multimedia Appendix 1). A mind map was also created to show integrative relationships and prevalence (Multimedia Appendix 2).

Results

Overview

Of the eight parenting forums identified, two met the inclusion criteria. Using the search operator *site: sampleforum.co.uk* for the two forums, a total of 95 discussion threads were identified and screened for inclusion. From those threads, 39 of them were duplicate results and 46 did not meet the inclusion criteria, leaving a total of 10 results to be analyzed (Table 1).

XSL•FO
RenderX

Table 1. Threads selected for analysis.

Thread number	Opening post title	Website	Sub-group heading	Comments
T[a]1	Vaping whilst Breast feeding?	Babycentre	Vapers Lounge	13
T2	Ecigarette and breastfeeding :(	Babycentre	June 2016 Birth Club	6
T3	Smoking while breastfeeding?	Babycentre	September 2015 Birth Club	19
T4	Does anyone vape?	Babycentre	October 2015 Birth Club	10
T5	Today I am..	Babycentre	February 2015 Birth Club	48
T6	(AIBU[b]?) To smoke an electronic cigarette whilst breastfeeding?	Mumsnet	Am I Being Unreasonable? (AIBU)	23
T7	(AIBU) To use the vape? For friend?	Mumsnet	Am I Being Unreasonable? (AIBU)	6
T8	(AIBU) To use electronic cigarettes even though I'm BF[c]?	Mumsnet	Am I Being Unreasonable? (AIBU)	55
T9	(AIBU) To ask DH[d] to stop vaping?	Mumsnet	Am I Being Unreasonable? (AIBU)	129
T10	(AIBU) To not give up smoking just yet?	Mumsnet	Am I Being Unreasonable? (AIBU)	39

[a]T: thread.

[b]AIBU: am I being unreasonable.

[c]BF: breastfeeding.

[d]DH: dear husband.

Four main themes were identified within the transcripts: use, perceived risk, social support, and evidence, with each of these having a number of subthemes.

Use

First, three subthemes were identified for the main theme of use, which included preventing returning to smoking, quitting smoking and motivation for use.

Based on the results, women were using e-cigs postpartum in a variety of ways. Some reported using them to prevent returning to smoking and described having cravings postpartum that were often associated with specific triggers such as the demands of motherhood, mental health issues or relationship problems. Motivation for use was a separate subtheme, as this applied to those who had already returned to smoking postpartum as well as those who were still abstinent, with women identifying e-cigs as preferable to smoking. Some women reported quitting suddenly and completely throughout pregnancy, but then they experienced cravings postpartum and found these could be alleviated by e-cig use:

> *Before pregnancy I used to smoke roll ups but quit when I found out I was pregnant! But after giving birth I started craving badly so decided that rather than smoking again I would try e-cig.* [Cressida, T2]

Some women had used an e-cig to quit during pregnancy, but then continued use of the e-cig postpartum had prevented them from returning to smoking. Others, however, did not manage to achieve abstinence during pregnancy or had already returned to smoking postpartum, while some had planned to return to smoking postpartum as they enjoyed it. The following quote highlights the experience of one woman who had already identified that she enjoyed smoking and didn’t want to lose that experience, but had found an e-cig to be a suitable alternative:

> *I didn't want to quit, I liked smoking. Bought an e-cig and did 24hrs on it and thought well I can't go back to smoking now. That was 9 months ago and I haven't smoked at all.* [Sakina, T10]

As identified above, women were able to identify what motivated them to seek out an alternative method of nicotine delivery. These reasons were mainly related to the context-specific issues attributed to new motherhood, such as lack of sleep, stress, loss of identity and relationship difficulties, as discussed below:

> *She is going through a massively stressful time right now and struggling to cope. She borrowed her mum's vape and loved it, felt totally better and less stressed straight away.* [Leah, T5]

Perceived Risk and Strategies to Mitigate Risk

The theme of perceived risk had four subthemes identified, including behavioral strategies, psychological strategies, physiological effects and environmental risks.

Although women were using e-cigs postpartum they still had concerns, with the perceived risks of vaping often compared to the risks of smoking:

> *They are not unregulated, we know what's in them and they are at least 95% safer than tobacco.* [Talitha, T6]

Sometimes the e-cigs were compared favorably to regular cigarettes and information on them was used to make assumptions on the safety of e-cigs, such as the guidance on smoking and breastfeeding being used to argue the safety of vaping and breastfeeding:

> *They say it's better for a smoker to smoke and breastfeed than not to breastfeed at all, so I should think the same applies to e-cigs.* [Delilah, T4]

XSL•FO
RenderX

However, there were also unfavorable comparisons, such as the health detriments of smoking being projected onto vaping:

> *I'm not an anxious or risk averse person, really. It's just the link between smoking and SIDS [sudden infant death syndrome] is so strong. What if you do continue to breathe out something for hours after vaping? If in twenty years they turn round and say vaping and co-sleeping causes x, and our baby had x?* [Acacia, T7]

Many of these comparisons were related to nicotine content in e-cigs, with one forum user advising another that it was safe for her husband to vape as he was not the one breastfeeding:

> *The most harmful thing in e cigarettes is the nicotine. Unless your DH [dear husband] plans on doing the breastfeeding, it isnt going to harm your baby.* [Daffodil, T7]

There were also perceived risks associated with the physiological effect of vaping on breastmilk, with discussions on what was likely to transfer to the infant if the mother vaped. There were comments about unknown substances that may be harmful if transferred to an infant; however, the most commonly discussed concerns were about nicotine and the perceived health risks associated with passing nicotine to the baby. The concern was often mixed with judgement, with the emphasis being that a good mother would not smoke or vape:

> *You're basically asking, "AIBU [Am I Being Unreasonable] to feed my baby small amounts of nicotine"? What do you think OP [opening poster]?* [Tirzah , T6]

Mothers were also informed their infant would develop an addiction to nicotine, in that the infant would *'feel like they want a fag [cigarette]'* and would suffer *'withdrawal'* from nicotine when breastfeeding ceased.

The concept of risk came with a variety of strategies to manage these perceived risks. Behavioral strategies were one example, and they involved altering behaviors to reduce exposure to vapor for infants. Alterations included only vaping outdoors or in a separate room, choosing low nicotine juice, or timing vaping around the infant's feeds to allow the maximum time to pass between vaping and the infant receiving inhaled components via breastmilk.

Psychological strategies were also used, which involved justifying any perceived risk in a way that presented a woman's choice to vape in a more favorable light, such as explaining that without vaping they would be stressed, which would be worse for their baby. They also justified the perceived risks of vaping by comparing it to more accepted health behaviors such as drinking coffee, as illustrated in the following quote:

> *Nicotine is in the same drug classification as caffeine so its only as bad as anyone that drinks coffee and breastfeeds.* [Helena, T2]

As well as specific risks to infants via the breast milk, there were wider concerns for risks from the environmental exposure to vapor. This was mainly founded on the basis of the harm from passive smoking, with women more concerned about the exposure to secondhand vapor based on the known harm from secondhand smoke.

Social Support

The third main theme had three subthemes identified within it, including informational, emotional, and instrumental social support.

While discussing risk on the forums, women were also seeking and giving support to one another about vaping, with the social support they received varying in nature. In many ways, the forum users offered positive social support to women who were vaping or considering an e-cig. There was also informational support, such as giving advice on which products to use or how best to use an e-cig. Informational support was often guided by the woman's own experience of vaping and included positive messages to support women, especially those who were trying to quit smoking.

Emotional support came from supportive comments about posters' own experiences as well as the experiences of others who had quit smoking and the health benefits they experienced, or by reassuring a vaper that they would not be judged for vaping. The following forum user discussed her partner's experience of vaping and how she viewed it positively:

> *I would much rather see him vape than smoke and he no longer wheezes when lying down, he is much fitter and it's the first time he has gone longer than a week without smoking. I think he is coming onto three years now.* [Bryony, T7]

However, not all posts were positive and supportive. There were instances of harsh judgement of vaping mothers, or indeed a mother's harsh judgement of herself. There were accusations of not putting their infant's needs before their own and the insinuation that by vaping they were somehow encouraging their child to learn unhealthy coping techniques:

> *Why would anyone condone it? Very strange. She is teaching her son an unhealthy way of coping with normal life stress.* [Tabitha, T6]

The varying forms of support often led to a polarized divide among forum users, with strong views expressed among both those who were provaping and antivaping.

Instrumental support was also identified, which included directing women to the best places to buy products or other forums to use for more information. This was also evident from those opposed to smoking, as they would direct women to alternative products or behaviors to remain smoke-free, including traditional nicotine replacement theory (NRT) use, self-help materials or sometimes more comical ways of both parents remaining smoke free:

> *Reward him with...I dunno. Doughnuts or something. I'd suggest BJs [oral sex] but then I remember how pregnant you are.* [Xanthe, T7]

Evidence

Finally, five subthemes were identified for the main theme of evidence, which included professional, non-professional, anecdotal, lack of evidence and mistrust or uncertainty of

XSL•FO
RenderX

evidence. This theme showed that women accessed a wide variety of sources of information to inform their arguments and opinions, and then interpreted and communicated their understanding of this evidence on forums (sometimes inaccurately).

Professional evidence came from academic articles or via professional websites such as the National Health Service (NHS) and Public Health England (PHE). This evidence was often misinterpreted, particularly by those who were opposed to vaping. One example was an article available on the NHS website about popcorn lung that was incorrectly cited several times across transcripts as evidence of e-cigs being harmful. A further example of using professional evidence is the following poster, who linked a paper by Farsalinos and Polosa [58]:

> *If you want decent info on the risks and benefits of vaping this is a good place to start. You can access the whole paper for free if you create an account.* [Jael, T4]

The most frequently quoted evidence was from nonprofessional sources. This included media articles such as blog posts and newspaper articles, but also links to social media profiles and discussions. There were also examples of websites, like Wikipedia, being cited as sources of evidence against vaping, which was met with ridicule by some provape forum users. The nonprofessional evidence was mostly quoted by those opposed to vaping, whereas professional evidence was equally shared by those both for and against vaping.

Anecdotal evidence was also shared by both sides and appeared to be the most substantial form of evidence accepted by women. The women were often more responsive to the experiences and stories from other forum users than they were to other forms of evidence available, and these forms of evidence often appeared to be more persuasive:

> *Anecdotally I can tell you that when my ExH [Ex-husband] vaped, our cats fled from the vapour and I hated the idea of him vaping inside near the cats. I'd be even more concerned about a baby.* [Grace, T4]

> *I do, I feel so much better too, no coughs or colds. I am positive that e cigs are much much less dangerous than cigarettes and think maybe you're being a bit over anxious.* [Jonquil, T7]

As well as sharing, quoting and interpreting evidence, there was also a general discussion on the lack of evidence available about the safety of e-cigs. This was most often attributed to a lack of empirical evidence of the long-term effects of vaping. Women were anxious to read information that related specifically to their situation and talked about lack of evidence on vaping and breastfeeding or vaping around young children.

This lack of evidence specific to new mothers was also displayed in the final subtheme of mistrust and uncertainty. In the following quote a forum user highlights the use of thalidomide in pregnancy, and how this was perceived as safe:

> *95% safer, Not 100% safe then? Not that long ago the NHS also said Thalidomide was safe. Look how that ended.* [Camelia, T7]

This is evidence of women looking for evidence that relates to their specific circumstances by using the comparison of a professional recommendation that resulted in infant harm. It wasn't just mistrust at the science itself, but also the institutions that make the recommendations:

> *And PHE [Public Health England] have been criticised for their supportive stance on e-cigs. They are very keen to get tobacco smoking down so I can see why they would be supportive.* [Joy, T6]

Discussion

An infodemiologiocal approach was taken for this research, which is the analysis of data on internet forum sites for public health research [30,31]. This research is the first to describe how women are accessing information about e-cig use during the postpartum period and is the first to provide evidence of women proactively using e-cigs to prevent returning to cigarette smoking and to aid smoking cessation, particularly as breastfeeding mothers. Women have concerns regarding the potential risks of using an e-cig and utilize online forums to discuss these risks with other women. This type of forum provides both positive and negative social support.

The themes show that women are accessing both lay and professional information on e-cig safety and their use via multiple sources, but this information is not necessarily being interpreted correctly, or it is being met with a degree of mistrust and uncertainty. There are conflicting opinions on the use of e-cigs while breastfeeding, mainly due to health concerns regarding what may be transferred from e-cig to the breastmilk, and then to the infant, as well as concerns about harm from secondhand vapor exposure.

There are limitations to this form of infodemiological research, with the use of online forum data forgoing the possibility of following up with individual users or seeking clarification on the meaning of their words, which increases the risk of bias during coding. It is also impossible to establish the validity of posts, like being completely confident that a user who identifies as a breastfeeding mother is, in fact, a breastfeeding mother. The transferability of these themes to the general postpartum population is limited due to both the exclusive participation of forum users and also that all the transcripts came from only two parenting forums. There is also no way to completely establish the authenticity of the users on the forum or whether they have connections within the e-cig or tobacco industry. On the other hand, the use of online data has several strengths, like that discussions are free from the response bias that may be present within interviews, and that forum data provides discourse that has been written with the intention of expressing and debating opinion for the purposes of discussion rather than research. The use of forum data provides in-depth qualitative data, and in other research, the use of a discussion analysis tool found that online interactions involving conflicting viewpoints promoted more discussion and critical thinking [59]. Our research is novel in both subject matter and approach, and therefore should be

XSL•FO
RenderX

treated as an exploratory qualitative piece upon which further research can be built. Thus far, this is the only piece of work that considers the motivators, barriers and opinions of breastfeeding mothers using e-cigs postpartum.

This work has improved our understanding of how and why women use e-cigs in the postpartum. For the first time we are able to understand and explore what evidence women are accessing to inform themselves about e-cigs and how this information is then interpreted. It also provides the first evidence of women perceiving their use of e-cigs postpartum to be preventative of a return to smoking. We are also better able to understand the concerns women have around the impact of e-cigs on infant health, in particular that a misattribution of nicotine as the most harmful substance transferred to infants via a smoking mother's milk is sometimes an obstacle to the use of e-cigs postpartum.

Previous research has highlighted that some women perceive that smoking affects the quality of their breastmilk in a way that is detrimental to their infant's health [15,60] despite the previous decade of recommendations from professional health bodies encouraging women who do smoke to continue breastfeeding [23]. This current research helps us better understand the fears women have in relation to e-cig use, most notably in relation to the perceived lack of consistent, evidence-based information about e-cig safety and the effects of nicotine transferring through breastmilk. By addressing these concerns, we could improve the acceptability of using a vape alternative for women who are breastfeeding and also smoke, while minimizing harm to the mother and infant and reducing their fears regarding their child's exposure to vape constituents via breastmilk.

Prior work using forum data considering e-cig use during pregnancy [31] identified three distinct themes explaining the ways in which forum users debated the use of e-cigs while pregnant: (1) quitting (nicotine) abruptly and completely is unsafe; (2) vaping is the lesser of two evils; and (3) vaping is not worth the risk. The authors concluded that women perceived their addiction to cigarettes as more than just a nicotine addiction, and that the behavioral aspects of smoking were also important, hence the potential for e-cigs. In our forum transcripts, women reported using e-cigs either to prevent returning to smoking or to quit smoking after having already relapsed. Returning to smoking, or perceived likelihood of returning to smoking, was often triggered by the demands of motherhood, perceived stress or the feeling of needing some personal time. It is interesting that some women are choosing an e-cig for similar reasons that have previously been identified for why women return to smoking traditional cigarettes, including smoking for relief and nostalgia for their former self [61]. Although evidence is limited, it is suggestive that women who use e-cigs during pregnancy are still likely to return to cigarette smoking, with one qualitative study partly attributing this to a lack of professional consensus within healthcare on the safety of e-cigs [62]. Lack of consistent and transparent information from professional health sources is a significant barrier to e-cig use postpartum, an issue that needs to be addressed given the success some women have reported on using e-cigs to prevent a return to smoking.

In our study, women also displayed mixed views on e-cig safety. The majority of users accepted that e-cigs were probably safer than cigarettes; however, there was a lot of skepticism and mistrust of the evidence for this. Health bodies such as PHE and the NHS were classed as biased due to their targets of reducing cigarette smoking, and comparisons were made regarding previous health recommendations that have since proven to be detrimental. Women were accessing scientific journals to learn more about e-cigs; however, it was often mistranslated. News media stories are often shared among online groups if a headline is particularly provocative, and even when these stories were discredited users felt that these fears must be based on something. Lack of evidence or mistrust of the current evidence appears to be a barrier for the use of e-cigs as a harm reduction tool during the postpartum period, which often gives rise to the thought process that it is better to deal with the familiar than risk the unknown. This skepticism is not confined to e-cigs, as previous research has identified that some women believe NRT patches to be harmful and that smoking is preferable to using them [63].

Uncertainty regarding e-cig safety often led to women discussing the concept of risk either in terms of comparison to smoking or in justifying the perceived risk. This individual assessment of risk is not unique to e-cig use and is attributed to a knowledge deficit between professionals and the lay public [64]. The risk assessment formed by lay people is complex, situationally influenced and reflects their personal values [65], and all of this is particularly relevant when considering the morality of motherhood and the negative attitudes some women hold towards vaping while breastfeeding. For example, the negative attitude towards government-backed advice on risk is assumed to be due to perceived exclusion from science-led and political decision-making [66]. With this in mind, involving women in discussions about e-cig use and safety within usual postnatal visits could help them make an informed choice on e-cig use.

This knowledge deficit could explain the reliance on unverified evidence from social media, news publications or web content found within this study. These sources of information are written to inform a general population but are also written to be read with ease, so this may be why women are engaging more with this type of evidence. There is also a reliance on others for information, with women seeking support and advice from health care professionals but also from other mothers. It is unsurprising that new mothers would seek information that is easy to access and easy to read, but the use of online forums also provides anonymity, which provides some form of protection of self while receiving or giving information [67]. Therefore, despite judgement from other mums, the ability to remain anonymous allows a mother to still have some perceived control over how those around her perceive her morally and ethically.

The concept of risk is a subjective one, and while there was much discussion of it, this risk was not defined apart from discussions on nicotine. There were suggestions of harm drawn from media conclusions, but risk itself was often discussed as a general term. Other lifestyle behaviors such as alcohol and caffeine consumption were often used as a comparison to justify this risk concept, with forum users suggesting that if these

XSL•FO
RenderX

behaviors were acceptable for mothers then vaping was also acceptable. However, the risks of nicotine exposure to the baby were one of the defined examples of risk. There were unsubstantiated attributions that nicotine caused SIDS, fears of infants becoming addicted to nicotine and being forced to experience withdrawal once breastfeeding ceased, and concerns that using e-cigs to manage the mother's mental health needs (such as stress) would lead to children who grew up exposed to unhealthy coping mechanisms. The exposure of infants to nicotine was also the subject of judgement. Some women would argue that a mother asking if it was acceptable to vape was actually asking if it was acceptable to feed her baby nicotine, suggesting that the nicotine became a deliberate exposure rather than a byproduct of the breastmilk. This concern regarding nicotine acted as a barrier to the use of e-cigs postpartum, and although some women acknowledged the known harmful substances in cigarettes, their primary concern seemed to be nicotine. Harmful effects from nicotine are not fully understood but are likely to be minimal compared to the effect of other compounds. Although it is accepted as an extremely addictive substance [68], there are far more worrying compounds within cigarette smoke, which research suggests are either not present in e-cigs or are present at significantly lower levels [69]. There is still limited empirical data on the safety and composition of e-cig vapor; however, there has been some promising toxicity testing that has evaluated the chemical nature of the vapor generated from e–cigarettes [69]. Despite the identification of certain toxicants within e-cig vapor, these levels are <1% of the levels present in cigarette smoke. E-cigs therefore have potential as a harm reduction tool, as confirmed by the PHE report [70].

This research closely relates to the "good mother" social construct [71], as shown with the various justifications of perceived risk, or the moralized stances against the use of any nicotine-containing products by a breastfeeding mother. The role of the mother is one that is subject to historical and cultural experiences, and social networks provide a framework to help make sense of culturally defined experiences and responsibilities [72]. The use of an online forum varies slightly from this by bringing together women from various socioeconomic backgrounds, ages, experiences and cultures to discuss breastfeeding. Therefore, the social constructs of a "good mother" are more explicit, particularly for infant feeding [73], whereby a "good mother" is synonymous with a breastfeeding mother without considering any cultural or environmental context [71]. The justification of risk here is similar to that of mothers who justify smoking by claiming that it is for their baby's sake [74], that is, it is better for the baby to have a mother who isn't stressed or is more alert.

In conclusion, this study has shown women hold a mixture of views on the acceptability of vaping as a mother, but some women are using (or are interested in using) e-cigs in the postpartum period. They are seeking, and need, more reliable information to facilitate their use, especially when breastfeeding. Therefore, we need further research that considers how women could have opportunities to ask and receive advice, perhaps by opening a dialogue on e-cigs between mothers and health care providers which could potentially reduce rates of maternal smoking and increase breastfeeding rates.

Acknowledgments

This paper presents independent research funded by the National Institute for Health Research School for Primary Care Research (NIHR SPCR). The views expressed are those of the authors and not necessarily those of the NIHR, the NHS or the Department of Health and Social Care.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Final coding template.

[PDF File (Adobe PDF File), 587KB - jmir_v21i8e11506_app1.pdf]

Multimedia Appendix 2

Mind Map. Integrative themes - mapping.

[PDF File (Adobe PDF File), 553KB - jmir_v21i8e11506_app2.pdf]

References

1. Marseglia L, Manti S, D'Angelo G, Cuppari C, Salpietro V, Filippelli M, et al. Obesity and breastfeeding: The strength of association. Women Birth 2015 Jun;28(2):81-86. [doi: 10.1016/j.wombi.2014.12.007] [Medline: 25595034]
2. Binns C, Lee M, Low WY. The Long-Term Public Health Benefits of Breastfeeding. Asia Pac J Public Health 2016 Jan;28(1):7-14. [doi: 10.1177/1010539515624964] [Medline: 26792873]
3. Gerhart KD, Stern DA, Guerra S, Morgan WJ, Martinez FD, Wright AL. Protective effect of breastfeeding on recurrent cough in adulthood. Thorax 2018 May 21:1-44. [doi: 10.1136/thoraxjnl-2017-210841] [Medline: 29786547]

XSL•FO
RenderX

4. Oddy WH. Breastfeeding protects against illness and infection in infants and children: a review of the evidence. Breastfeed Rev 2001 Jul;9(2):11-18. [Medline: 11550600]
5. Lanari M, Prinelli F, Adorni F, Di Santo S, Faldella G, Silvestri M, Italian Neonatology Study Group on RSV Infections. Maternal milk protects infants against bronchiolitis during the first year of life. Results from an Italian cohort of newborns. Early Hum Dev 2013 Jun;89 Suppl 1:S51-S57. [doi: 10.1016/S0378-3782(13)70016-1] [Medline: 23809352]
6. Maayan-Metzger A, Schushan-Eisen I, Lubin D, Moran O, Kuint J, Mazkereth R. Delivery room breastfeeding for prevention of hypoglycaemia in infants of diabetic mothers. Fetal Pediatr Pathol 2014 Feb;33(1):23-28. [doi: 10.3109/15513815.2013.842271] [Medline: 24099464]
7. Netzer-Tomkins H, Rubin L, Ephros M. Breastfeeding Is Associated with Decreased Hospitalization for Neonatal Fever. Breastfeed Med 2016 Dec;11:218-221. [doi: 10.1089/bfm.2016.0007] [Medline: 27070106]
8. Penn AH, Carver LJ, Herbert CA, Lai TS, McIntire MJ, Howard JT, et al. Breast Milk Protects Against Gastrointestinal Symptoms in Infants at High Risk for Autism During Early Development. J Pediatr Gastroenterol Nutr 2016 Feb;62(2):317-327 [FREE Full text] [doi: 10.1097/MPG.0000000000000907] [Medline: 26230900]
9. Cooper S, Orton S, Leonardi-Bee J, Brotherton E, Vanderbloemen L, Bowker K, et al. Smoking and quit attempts during pregnancy and postpartum: a longitudinal UK cohort. BMJ Open 2017 Nov 15;7(11):e018746 [FREE Full text] [doi: 10.1136/bmjopen-2017-018746] [Medline: 29146659]
10. McAndrew F, Thompson J, Fellows L, Large A, Speed M, Renfrew M. Infant Feeding Survey 2010. Infant feeding survey 2012:1-136 [FREE Full text] [doi: 10.5255/UKDA-SN-7281-1]
11. Orton S, Coleman T, Coleman-Haynes T, Ussher M. Predictors of Postpartum Return to Smoking: A Systematic Review. Nicotine Tob Res 2018 May 03;20(6):665-673. [doi: 10.1093/ntr/ntx163] [Medline: 29065203]
12. Tong VT, Dietz PM, Morrow B, D'Angelo DV, Farr SL, Rockhill KM, Centers for Disease Control and Prevention (CDC). Trends in smoking before, during, and after pregnancy--Pregnancy Risk Assessment Monitoring System, United States, 40 sites, 2000-2010. MMWR Surveill Summ 2013 Nov 08;62(6):1-19 [FREE Full text] [Medline: 24196750]
13. Horta BL, Kramer MS, Platt RW. Maternal smoking and the risk of early weaning: a meta-analysis. Am J Public Health 2001 Feb;91(2):304-307. [Medline: 11211645]
14. Ratner P, Johnson J, Bottorff J. Smoking relapse and early weaning among postpartum women: is there an association? Birth 1999 Jun;26(2):76-82. [Medline: 10687570]
15. Liu J, Rosenberg KD, Sandoval AP. Breastfeeding duration and perinatal cigarette smoking in a population-based cohort. Am J Public Health 2006 Feb;96(2):309-314. [doi: 10.2105/AJPH.2004.060798] [Medline: 16380564]
16. DiSantis KI, Collins BN, McCoy ACS. Associations among breastfeeding, smoking relapse, and prenatal factors in a brief postpartum smoking intervention. Acta Obstet Gynecol Scand 2010 Apr;89(4):582-586 [FREE Full text] [doi: 10.3109/00016341003678435] [Medline: 20235894]
17. Harmer C, Memon A. Factors associated with smoking relapse in the postpartum period: an analysis of the child health surveillance system data in Southeast England. Nicotine Tob Res 2013 May;15(5):904-909. [doi: 10.1093/ntr/nts221] [Medline: 23045522]
18. Bailey BA, Wright HN. Breastfeeding initiation in a rural sample: predictive factors and the role of smoking. J Hum Lact 2011 Feb;27(1):33-40. [doi: 10.1177/0890334410386955] [Medline: 21177987]
19. Onah S, Osuorah DIC, Ebenebe J, Ezechukwu C, Ekwochi U, Ndukwu I. Infant feeding practices and maternal socio-demographic factors that influence practice of exclusive breastfeeding among mothers in Nnewi South-East Nigeria: a cross-sectional and analytical study. Int Breastfeed J 2014;9:6 [FREE Full text] [doi: 10.1186/1746-4358-9-6] [Medline: 24860612]
20. Mohsin M, Bauman AE. Socio-demographic factors associated with smoking and smoking cessation among 426,344 pregnant women in New South Wales, Australia. BMC Public Health 2005 Dec 21;5:138 [FREE Full text] [doi: 10.1186/1471-2458-5-138] [Medline: 16371166]
21. American Academy of Pediatrics Committee on Drugs. Transfer of drugs and other chemicals into human milk. Pediatrics 2001 Sep;108(3):776-789. [doi: 10.1542/peds.108.3.776] [Medline: 11533352]
22. Minchin MK. Smoking and breastfeeding: an overview. J Hum Lact 1991 Dec;7(4):183-188. [doi: 10.1177/089033449100700415] [Medline: 1818572]
23. NHS: National Health Service. 2018 Dec. Breastfeeding and smoking - Your pregnancy and baby guide URL: http://www.nhs.uk/Conditions/pregnancy-and-baby/Pages/breastfeeding-and-smoking.aspx
24. Hajek P, Etter J, Benowitz N, Eissenberg T, McRobbie H. Electronic cigarettes: review of use, content, safety, effects on smokers and potential for harm and benefit. Addiction 2014 Nov;109(11):1801-1810 [FREE Full text] [doi: 10.1111/add.12659] [Medline: 25078252]
25. Cheng T. Chemical evaluation of electronic cigarettes. Tob Control 2014 May;23 Suppl 2:ii11-ii17 [FREE Full text] [doi: 10.1136/tobaccocontrol-2013-051482] [Medline: 24732157]
26. Barbeau AM, Burda J, Siegel M. Perceived efficacy of e-cigarettes versus nicotine replacement therapy among successful e-cigarette users: a qualitative approach. Addict Sci Clin Pract 2013 Mar 05;8:5 [FREE Full text] [doi: 10.1186/1940-0640-8-5] [Medline: 23497603]

XSL•FO
RenderX

27. McNeill A, Brose L, Calder R, Bauld L, Robson D. A report commissioned by Public Health England. 2018. Evidence review of e-cigarettes and heated tobacco products URL: https://tinyurl.com/yyj34zvd [accessed 2018-07-03] [WebCite Cache ID 70dN8Fxwj]
28. McEwen A, McRobbie H. National Centre for Smoking Cessation and Training. 2016. Electronic cigarettes: A briefing for stop smoking services URL: http://www.ncsct.co.uk/usr/pub/Electronic_cigarettes._A_briefing_for_stop_smoking_services.pdf [accessed 2018-07-03] [WebCite Cache ID 70dTZc6Ws]
29. Taiwo R. Membership and activity in an online parenting community. In: Handbook Of Research On Discourse Behavior And Digital Communication: Language Structures And Social Interaction. Nigeria: Information Science Reference; 2010.
30. Eysenbach G. Infodemiology and infoveillance: framework for an emerging set of public health informatics methods to analyze search, communication and publication behavior on the Internet. J Med Internet Res 2009;11(1):e11 [FREE Full text] [doi: 10.2196/jmir.1157] [Medline: 19329408]
31. Eysenbach G. Infodemiology and infoveillance tracking online health information and cyberbehavior for public health. Am J Prev Med 2011 May;40(5 Suppl 2):S154-S158. [doi: 10.1016/j.amepre.2011.02.006] [Medline: 21521589]
32. Kim AE, Hopper T, Simpson S, Nonnemaker J, Lieberman AJ, Hansen H, et al. Using Twitter Data to Gain Insights into E-cigarette Marketing and Locations of Use: An Infoveillance Study. J Med Internet Res 2015;17(10):e251 [FREE Full text] [doi: 10.2196/jmir.4466] [Medline: 26545927]
33. Chen AT, Zhu S, Conway M. What Online Communities Can Tell Us About Electronic Cigarettes and Hookah Use: A Study Using Text Mining and Visualization Techniques. J Med Internet Res 2015;17(9):e220 [FREE Full text] [doi: 10.2196/jmir.4517] [Medline: 26420469]
34. Wennberg A, Jonsson S, Zadik Janke J, Hörnsten Å. Online Perceptions of Mothers About Breastfeeding and Introducing Formula: Qualitative Study. JMIR Public Health Surveill 2017 Nov 15;3(4):e88 [FREE Full text] [doi: 10.2196/publichealth.8197] [Medline: 29141838]
35. Iftikhar R, Abaalkhail B. Health-Seeking Influence Reflected by Online Health-Related Messages Received on Social Media: Cross-Sectional Survey. J Med Internet Res 2017 Dec 16;19(11):e382 [FREE Full text] [doi: 10.2196/jmir.5989] [Medline: 29146568]
36. Simmons RK, Singh G, Maconochie N, Doyle P, Green J. Experience of miscarriage in the UK: qualitative findings from the National Women's Health Study. Soc Sci Med 2006 Oct;63(7):1934-1946. [doi: 10.1016/j.socscimed.2006.04.024] [Medline: 16781809]
37. Bert F, Gualano MR, Brusaferro S, De Vito E, de Waure C, La TG, et al. Pregnancy e-health: a multicenter Italian cross-sectional study on Internet use and decision-making among pregnant women. J Epidemiol Community Health 2013 Dec 1;67(12):1013-1018. [doi: 10.1136/jech-2013-202584] [Medline: 24072743]
38. Wigginton B, Gartner C, Rowlands IJ. Is It Safe to Vape? Analyzing Online Forums Discussing E-Cigarette Use during Pregnancy. Womens Health Issues 2017;27(1):93-99. [doi: 10.1016/j.whi.2016.09.008] [Medline: 27773530]
39. Chung JE. Social networking in online support groups for health: how online social networking benefits patients. J Health Commun 2014;19(6):639-659. [doi: 10.1080/10810730.2012.757396] [Medline: 23557148]
40. Diefenbeck CA, Klemm PR, Hayes ER. Emergence of Yalom's therapeutic factors in a peer-led, asynchronous, online support group for family caregivers. Issues Ment Health Nurs 2014 Jan;35(1):21-32. [doi: 10.3109/01612840.2013.836260] [Medline: 24350748]
41. Mo PKH, Coulson NS. Online support group use and psychological health for individuals living with HIV/AIDS. Patient Educ Couns 2013 Dec;93(3):426-432. [doi: 10.1016/j.pec.2013.04.004] [Medline: 23669152]
42. Steadman J, Pretorius C. The impact of an online Facebook support group for people with multiple sclerosis on non-active users. Afr J Disabil 2014;3(1):132 [FREE Full text] [doi: 10.4102/ajod.v3i1.132] [Medline: 28730005]
43. Finfgeld DL. Therapeutic groups online: the good, the bad, and the unknown. Issues Ment Health Nurs 2000;21(3):241-255. [Medline: 11075065]
44. Jowett A. A Case for Using Online Discussion Forums in Critical Psychological Research. Qualitative Research in Psychology 2015 Feb 03;12(3):287-297. [doi: 10.1080/14780887.2015.1008906]
45. Burri M, Baujard V, Etter J. A qualitative analysis of an internet discussion forum for recent ex-smokers. Nicotine Tob Res 2006 Dec;8 Suppl 1:S13-S19. [Medline: 17491166]
46. Pedersen S, Smithson J. Mothers with attitude — How the Mumsnet parenting forum offers space for new forms of femininity to emerge online. Women's Studies International Forum 2013 May;38:97-106. [doi: 10.1016/j.wsif.2013.03.004]
47. Walsh AM, Hamilton K, White KM, Hyde MK. Use of online health information to manage children's health care: a prospective study investigating parental decisions. BMC Health Serv Res 2015 Apr 02;15:131 [FREE Full text] [doi: 10.1186/s12913-015-0793-4] [Medline: 25889493]
48. Kuehn BM. Patients go online seeking support, practical advice on health conditions. JAMA 2011 Apr 27;305(16):1644-1645. [doi: 10.1001/jama.2011.509] [Medline: 21521841]
49. Betts D, Dahlen HG, Smith CA. A search for hope and understanding: an analysis of threatened miscarriage internet forums. Midwifery 2014 Jun;30(6):650-656. [doi: 10.1016/j.midw.2013.12.011] [Medline: 24439850]
50. Dahlen HG, Homer CSE. Web-based news reports on midwives compared with obstetricians: a prospective analysis. Birth 2012 Mar;39(1):48-56. [doi: 10.1111/j.1523-536X.2011.00512.x] [Medline: 22369605]

XSL•FO
RenderX

51. Dahlen HG, Homer CSE. 'Motherbirth or childbirth'? A prospective analysis of vaginal birth after caesarean blogs. Midwifery 2013 Feb;29(2):167-173. [doi: 10.1016/j.midw.2011.11.007] [Medline: 22169525]
52. Eastham LA. Research using blogs for data: public documents or private musings? Res Nurs Health 2011 Aug;34(4):353-361. [doi: 10.1002/nur.20443] [Medline: 21626517]
53. British Psychological Society. 2003. Ethics Guidelines for Internet-mediated Research URL: https://www.bps.org.uk/news-and-policy/ethics-guidelines-internet-mediated-research-2017 [accessed 2018-07-03] [WebCite Cache ID 70dNMrHlD]
54. Bradley SK, Carter B. Reflections on the ethics of internet newsgroup research. Int J Nurs Stud 2012 May;49(5):625-630. [doi: 10.1016/j.ijnurstu.2011.10.018] [Medline: 22098925]
55. Eysenbach G, Till JE. Ethical issues in qualitative research on internet communities. BMJ 2001 Nov 10;323(7321):1103-1105 [FREE Full text] [Medline: 11701577]
56. Cassell C, Symon G. Qualitative Methods and Analysis in Organizational Research: A Practical Guide. In: Qualitative methods and analysis in organizational research. London: SAGE; 1998.
57. King N. Doing Template Analysis. In: Qualitative Organizational Research: Core Methods and Current Challenges. Thousand Oaks: SAGE Publications; 2012.
58. Farsalinos KE, Polosa R. Safety evaluation and risk assessment of electronic cigarettes as tobacco cigarette substitutes: a systematic review. Ther Adv Drug Saf 2014 Apr;5(2):67-86 [FREE Full text] [doi: 10.1177/2042098614524430] [Medline: 25083263]
59. Jeong AC. The Sequential Analysis of Group Interaction and Critical Thinking in Online. American Journal of Distance Education 2003 Mar;17(1):25-43. [doi: 10.1207/S15389286AJDE1701_3]
60. Goldade K, Nichter M, Nichter M, Adrian S, Tesler L, Muramoto M. Breastfeeding and smoking among low-income women: results of a longitudinal qualitative study. Birth 2008 Sep;35(3):230-240 [FREE Full text] [doi: 10.1111/j.1523-536X.2008.00244.x] [Medline: 18844649]
61. Bottorff JL, Johnson JL, Irwin LG, Ratner PA. Narratives of smoking relapse: the stories of postpartum women. Res Nurs Health 2000 Apr;23(2):126-134. [Medline: 10782871]
62. Fallin A, Miller A, Assef S, Ashford K. Perceptions of Electronic Cigarettes Among Medicaid-Eligible Pregnant and Postpartum Women. J Obstet Gynecol Neonatal Nurs 2016;45(3):320-325. [doi: 10.1016/j.jogn.2016.02.009] [Medline: 27083420]
63. Hotham ED, Atkinson ER, Gilbert AL. Focus groups with pregnant smokers: barriers to cessation, attitudes to nicotine patch use and perceptions of cessation counselling by care providers. Drug Alcohol Rev 2002 Jun;21(2):163-168. [doi: 10.1080/09595230220139064] [Medline: 12188995]
64. Evans G, Durant J. The relationship between knowledge and attitudes in the public understanding of science in Britain. Public Underst Sci 2016 Dec 21;4(1):57-74 [FREE Full text] [doi: 10.1088/0963-6625/4/1/004]
65. Hansen J, Holm L, Frewer L, Robinson P, Sandøe P. Beyond the knowledge deficit: recent research into lay and expert attitudes to food risks. Appetite 2003 Oct;41(2):111-121. [Medline: 14550309]
66. Frewer LJ, Miles S, Brennan M, Kuznesof S, Ness M, Ritson C. Public preferences for informed choice under conditions of risk uncertainty. Public Underst Sci 2016 Dec 21;11(4):363-372. [doi: 10.1088/0963-6625/11/4/304]
67. Qian H, Scott C. Anonymity and self-disclosure on weblogs. Journal of Computer Mediated Communication 2007;12(4):1428-1451. [doi: 10.1111/j.1083-6101.2007.00380.x]
68. Stolerman IP, Jarvis MJ. The scientific case that nicotine is addictive. Psychopharmacology (Berl) 1995 Jan;117(1):2-10; discussion 14. [Medline: 7724697]
69. Goniewicz ML, Knysak J, Gawron M, Kosmider L, Sobczak A, Kurek J, et al. Levels of selected carcinogens and toxicants in vapour from electronic cigarettes. Tob Control 2014 Mar;23(2):133-139 [FREE Full text] [doi: 10.1136/tobaccocontrol-2012-050859] [Medline: 23467656]
70. McNeil A, Brose L, Calder R, Hitchman S, Hajek P, McRobbie H. Public Health England. 2014. E-cigarettes: an evidence update URL: https://tinyurl.com/yylmf8rr [accessed 2018-07-03] [WebCite Cache ID 70dNWsNpE]
71. Marshall JL, Godfrey M, Renfrew MJ. Being a 'good mother': managing breastfeeding and merging identities. Soc Sci Med 2007 Nov;65(10):2147-2159. [doi: 10.1016/j.socscimed.2007.06.015] [Medline: 17681409]
72. Miller T. Making Sense Of Motherhood. Cambridge: Cambridge University Press; 2019.
73. Murphy E, Parker S, Phipps C. Competing agendas in infant feeding. British Food Journal 1998 Apr 13;100(3):128-132 [FREE Full text] [doi: 10.1108/00070709810207487] [Medline: 29433498]
74. Irwin LG, Johnson JL, Bottorff JL. Mothers who smoke: confessions and justifications. Health Care Women Int 2005 Aug;26(7):577-590. [doi: 10.1080/07399330591004890] [Medline: 16126601]

Abbreviations

AIBU: am I being unreasonable?
DH: dear husband
e-cig: electronic cigarette
ExH: ex-husband

XSL•FO
RenderX

PHE: Public Health England
NHS: National Health Service
NIHR SPCR: National Institute for Health Research School for Primary Care Research
NRT: nicotine replacement theory
SIDS: sudden infant death syndrome
T: thread

Edited by G Eysenbach; submitted 01.08.18; peer-reviewed by J Constantin, S Prior; comments to author 15.11.18; revised version received 15.01.19; accepted 05.03.19; published 12.08.19.

Please cite as:
Johnston EJ, Campbell K, Coleman T, Lewis S, Orton S, Cooper S
Safety of Electronic Cigarette Use During Breastfeeding: Qualitative Study Using Online Forum Discussions
J Med Internet Res 2019;21(8):e11506
URL: https://www.jmir.org/2019/8/e11506/
doi:10.2196/11506
PMID:31407672

©Emily Jade Johnston, Katarzyna Campbell, Tim Coleman, Sarah Lewis, Sophie Orton, Sue Cooper. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 12.08.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Using Twitter to Understand the Human Bowel Disease Community: Exploratory Analysis of Key Topics

Martín Pérez-Pérez[1,2,3], PhD; Gael Pérez-Rodríguez[1,2,3], PhD; Florentino Fdez-Riverola[1,2,3], PhD; Anália Lourenço[1,2,3,4], PhD

[1]Department of Computer Science, University of Vigo, Escuela Superior de Ingeniería Informática, Ourense, Spain

[2]Biomedical Research Centre, Campus Universitario Lagoas-Marcosende, Vigo, Spain

[3]Next Generation Computer Systems Group, School of Computer Engineering, Galicia Sur Health Research Institute, Galician Health Service - University of Vigo, Vigo, Spain

[4]Centre of Biological Engineering, Campus de Gualtar, University of Minho, Braga, Portugal

Corresponding Author:
Anália Lourenço, PhD
Department of Computer Science
University of Vigo
Escuela Superior de Ingeniería Informática
Edificio Politécnico
Campus Universitario As Lagoas s/n
Ourense, 32004
Spain
Phone: 34 988 387 013
Fax: 34 988 387 001
Email: analia@uvigo.es

Abstract

Background: Nowadays, the use of social media is part of daily life, with more and more people, including governments and health organizations, using at least one platform regularly. Social media enables users to interact among large groups of people that share the same interests and suffer the same afflictions. Notably, these channels promote the ability to find and share information about health and medical conditions.

Objective: This study aimed to characterize the bowel disease (BD) community on Twitter, in particular how patients understand, discuss, feel, and react to the condition. The main questions were as follows: Which are the main communities and most influential users?; Where are the main content providers from?; What are the key biomedical and scientific topics under discussion? How are topics interrelated in patient communications?; How do external events influence user activity?; What kind of external sources of information are being promoted?

Methods: To answer these questions, a dataset of tweets containing terms related to BD conditions was collected from February to August 2018, accounting for a total of 24,634 tweets from 13,295 different users. Tweet preprocessing entailed the extraction of textual contents, hyperlinks, hashtags, time, location, and user information. Missing and incomplete information about the user profiles was completed using different analysis techniques. Semantic tweet topic analysis was supported by a lexicon-based entity recognizer. Furthermore, sentiment analysis enabled a closer look into the opinions expressed in the tweets, namely, gaining a deeper understanding of patients' feelings and experiences.

Results: Health organizations received most of the communication, whereas BD patients and experts in bowel conditions and nutrition were among those tweeting the most. In general, the BD community was mainly discussing symptoms, BD-related diseases, and diet-based treatments. Diarrhea and constipation were the most commonly mentioned symptoms, and cancer, anxiety disorder, depression, and chronic inflammations were frequently part of BD-related tweets. Most patient tweets discussed the bad side of BD conditions and other related conditions, namely, depression, diarrhea, and fibromyalgia. In turn, gluten-free diets and probiotic supplements were often mentioned in patient tweets expressing positive emotions. However, for the most part, tweets containing mentions to foods and diets showed a similar distribution of negative and positive sentiments because the effects of certain food components (eg, fiber, iron, and magnesium) were perceived differently, depending on the state of the disease and other personal conditions of the patients. The benefits of medical cannabis for the treatment of different chronic diseases were also highlighted.

XSL•FO
RenderX

Conclusions: This study evidences that Twitter is becoming an influential space for conversation about bowel conditions, namely, patient opinions about associated symptoms and treatments. So, further qualitative and quantitative content analyses hold the potential to support decision making among health-related stakeholders, including the planning of awareness campaigns.

(J Med Internet Res 2019;21(8):e12610) doi:10.2196/12610

KEYWORDS

inflammatory bowel diseases; irritable bowel syndrome; social media; communication; data mining; natural language processing; infodemiology

Introduction

Mass access and diffusion of health-related news and information have dramatically changed in recent years. In addition to traditional media, the internet has become a pivotal instrument for sharing knowledge [1]. In this context, the use of social media has increased exponentially over the last years. More and more people and institutions, including governments and health organizations, use social networks, blogs, content-sharing sites, and wikis on a regular basis [2-4]. Therefore, social media have become a major source of information [5]. Social media creates the opportunity for users to interact among large groups of people that share the same interests and suffer the same afflictions. In particular, these channels promote the ability to find and share information about health and medical conditions [6,7]. Notably, many people who suffer from chronic diseases resort to groups or communities in social networks to share experiences and stand by for news about their afflictions [8-10]. One recurrent example is inflammatory bowel disease (IBD), a chronic, relapsing, and remitting autoimmune disorder with 2 main conditions, Crohn disease and ulcerative colitis [11]. The worldwide incidence of these conditions has been increasing over the last few decades [12]. IBD is frequently diagnosed in the second to fourth decades of life, with a high incidence during the peak female reproductive years [13]. Communities all around the world elected May 19 as the World IBD Day, a reference date to raise awareness of this chronic disease and its associated symptoms [14].

Nowadays, there is an increasing interest in analyzing and studying several factors related to health topics in social media, especially on Twitter [15-17]. Namely, text mining (TM) and natural language processing methods and techniques are being applied to systematically identify the topics under discussion as well as study the relationships among individuals and topics. Indeed, over recent years, an increasing number of publications reported health-related studies in social media. For example, how promotional health information about Lynch syndrome impacts laypeople's discussions [18]; diabetes-related participation on Twitter by describing the frequency and timing of diabetes-related tweets, the geography of tweets, and the types of participants [19]; understanding the use of social media by patients with different types of cancer [17]; and the emergence of health online communities of practice (ie, group of people who share experiences) [16].

Regarding IBD, only a few studies exist, and these are focused on the analysis of particular user communities or particular user details. For example, the study by Guo et al studied the use and quality of social media in patients with IBD [20], and the study by Keller et al discussed how individuals taking IBD medication during reproductive periods made decisions about their medication use [21]. So, the aim of this study was to gain a better understanding of the communities involved in the dissemination of information about bowel disease (BD). A collection of over 24,000 tweets related to BD enabled the identification and geolocation of active users and the analysis of external events that determine tweet content and discussion, as well as the relations between the biomedical and scientific topics being mentioned. The obtained results are useful to understand the most relevant topics and communities and, specifically, how patients discuss, feel, and react to symptoms, changes in habits, and medication. That is, to learn how to enhance the dissemination of information and raise awareness among patients, which are the 2 main objectives of health-related stakeholders.

Methods

Twitter Communication Environment

Currently, the architecture of Twitter supports different user actions in response to a tweet (ie, replies, retweets, and favorites), each of them holding a specific meaning in terms of communication capabilities. Notably, replies represent the specific response to a sent tweet, retweets stand for the reposting of tweets (which is useful to quickly share and promote content), and favorites indicate that the content of a tweet is highly appreciated by the community and can be seen as a user tweet bookmark.

Figure 1 exemplifies the aforementioned communication modes. On the one hand, replies and retweets (ie, user relations) allow identifying users who support or discuss any sent message. For example, users B and C are interested in a tweet published by user A. This allows to identify how the information is spread and which are the most influential users. On the other hand, the number of retweets and favorites (ie, tweet interactions) helps to measure the relevance of the new content to the community. For example, user D likes the original tweet sent by user A and saves it for future reference.

General Workflow

Figure 2 depicts the workflow implemented in this study to retrieve, process, and analyze BD-related tweets, which consisted of 2 fundamental phases: (1) data collection and filtering and (2) corpus processing and analysis.

From February 1, 2018, to August 31, 2018, tweet data were retrieved via the Twitter application programming interface

XSL•FO
RenderX

(API). Tweet contents and associated information were processed: whenever possible, users were geolocated, their gender was determined, and they were identified as organization or patient; tweet contents were cleaned (hashtags and mentions to user accounts were removed) and prepared for further text processing; and, entity recognition and sentiment analysis were applied.

Data Collection and Filtering

Data collection accounted for tweets containing terms as *Inflammatory bowel disease*, *Irritable bowel disease*, *Irritable colon*, *Ulcerative colitis*, *Ileocolitis*, *Ileitis*, *Crohn*, *Granulomatous*, and *Jejunoileitis*. The Java library Twitter4J [22] was used to perform such collection. From all the retrieved tweets, only those written in English were considered to ensure the consistency of further examination. A total of 4.10% (1055/25,689) tweets written in other languages, such as French, Spanish, or Italian, were eliminated. The final dataset comprised 24,634 unique tweets written in English by 13,295 different users.

Corpus Processing and Analysis: User Characterization

All the tweets were automatically labeled by tweet creator. Data present in user profiles were collected and further validated (see details in the next subsections). Specifically, user characterization involved gender determination, differentiation of organizations and patients, and geolocation.

Unfortunately, it was not possible to determine the age of the users because of the low precision of current prediction models and because several studies point that most Twitter users fall into a small range of years [23].

Face and Gender Recognition

Gender identification entailed a 2-step strategy based on a gender-name dictionary [24] and a convolutional neural network model (this model was trained over more than 500,000 public face images extracted from IMDb and Wikipedia and was reported to have nearly a 90% of accuracy) [25]. First, the user name was checked against the dictionary. If there was a perfect match, that is, a unique gender associated with the name, the gender was resolved. Otherwise, and if there was a user profile picture, the deep learning model was applied. If there was no user profile picture, or the model could not recognize a single face in the image, the gender was set as unknown.

User Identification

Whenever possible, user accounts were categorized as belonging to organization, individual (ie, patient and medical expert), and unknown. The strategy to identify users was focused on the analysis of the user account (ie, name and description).

First, to identify users as organizations, several cascade steps were followed, namely, (1) the account had a country, country code, or a continent in the user name; (2) the account had a URL domain in the name (eg, .org); and (3) usage of regular expressions to check for nonpersonal keywords in the description (eg, official, news, info, or pharma).

If the user was not identified as an organization, the following steps were applied to check if the user was an individual: (1) the account had a recognized gender; (2) the account was recognized by Twitter as a contributor or a translator; (3) the description was written using first-person pronouns and their variant forms (ie, possessive and reflexive); (4) the description had emojis or emoticons; and (5) usage of regular expression to check for person abbreviations (eg, Ms or Mr).

Finally, to differentiate BD experts (eg, doctors, medical staff, or researchers) from patients, all accounts identified as individuals were processed with an additional recognition step to check for expert-related keywords (eg, Dr, Prof, MD, or PhD).

Whenever the strategy could not help determine the type, the user was labeled as an unknown type.

Geolocation

Twitter does not require users to specify the location. When users introduce such data, it is in the form of free text, which often raises consistency issues in further analysis (eg, a user can enter *NYC* and others may identify the same location as *New York City*). Another issue to take into account is the existence of cities in different countries that share the same name (eg, *Guadalajara* is a city both in Spain and Mexico).

Thus, the applied location identification method took into consideration information about the time zone, the Coordinated Universal Time (UTC) offset, and the location text. These data were used in combination with the GeoNames database [26], which contains over 10 million geographical names and is accessible through a free Web service. The data extracted from Twitter were searched against the GeoNames database. If the data were not accurate enough, that is, matching multiple database entries, the time zone and the UTC offset were used to help resolve the location. In those cases where different cities shared the same name and time zone but were located in different countries, the location was set as unknown.

Corpus Processing and Analysis: Tweet Characterization

To analyze the content of the generated corpus, it was essential to be able to properly recognize the relevant (topic related) terms mentioned in the tweets. For this purpose, several text preprocessing techniques were applied and then an in-house–developed named entity recognizer supported the annotation of terms pertaining to the selected BD-related semantic categories.

Data Cleaning

As a first step, the following preprocessing tasks were applied to the content of the tweets in the dataset:

- Removal of special characters that did not provide useful information (eg, &, *(*“, *“)*, *, +, <, or >)
- Identification of replies and mentions to other users (represented with @) and extraction of URLs
- Removal of the symbol # in hashtags, and split of hashtags in multiple words (if possible) with the goal of revealing relevant terms to the analysis (eg, *InflammatoryBowelDisease* to *Inflammatory Bowel*

XSL•FO
RenderX

Disease). All these operations were carried out using the Twitter text library [27]

- Deletion of repeated letters if there are 3, or more, consecutive and identical characters (eg, *haaaapppyy* to *haappyy*)
- Correction of spelling errors using the Hunspell dictionary [28], a collection of specific medical terms [29] obtained from the OpenMedSpel [30], and the MTH-Med-Spel-Check tool [31]. The correction was done automatically by selecting the suggested word with the highest similarity with the original (incorrect) term. The similarity was calculated using the normalized Levenshtein algorithm [32]
- Expansion of abbreviations and shorthand terms, which were not included in the Hunspell dictionary (eg, *SBBOS* to *small bowel bacterial overgrowth syndrome*). Although Twitter has increased the maximum length of tweets from 140 to 280 characters, the use of abbreviations is still very common. Therefore, a custom dictionary of abbreviations was constructed in house, comprising terms extracted from multiple locations [33-35].

Then, a new round of text preprocessing tasks prepared the tweets for named entity recognition, namely, tokenization (ie, breaking a stream of text up into words, phrases, or other meaningful elements), stop word removal (ie, removal of too frequent, not content-bearing tokens), part of speech tagging (ie, to assign a lexical category to each token), and lemmatization (ie, to obtain the lexeme form of the token). Beside single word tokens (unigrams), bigrams and trigrams, that is, contiguous of 2 or 3 sequences of tokens, were also considered in entity recognition. All the aforementioned tasks were implemented using the Stanford CoreNLP pipeline [36].

Named Entity Recognition

The semantic lexicon applied in named entity recognition was mostly retrieved from the repository of biomedical ontologies BioPortal [37] as follows: the Human Disease Ontology (DOID) [38], which provides descriptions of human disease terms, phenotype characteristics, and related medical vocabulary; the Ontology For Nutritional Studies (FoodOn) [39], which covers human food raw ingredients, food products, and product types and develops semantics to food production, culinary, nutritional, and chemical ingredients and processes; the Symptom Ontology (SYMP) [40], which covers disease symptoms, with symptoms encompassing perceived changes in function, sensations, or appearance reported by a patient indicative of a disease; and the branch *Intervention or Procedure* of the National Cancer Institute Thesaurus (NCIT) [41], which describes treatments or actions taken to prevent or treat disease or improve health in other ways. The DrugBank ontology supported the recognition of chemical, pharmacological, and pharmaceutical terminology, that is, approved small molecule drugs, approved biotech (protein and peptide) drugs, nutraceuticals, and experimental drugs [42].

As a whole, the lexicon supporting the entity recognition encompassed a total of 217,468 term entries. For the sake of simplicity, the results of the semantic annotations are presented and discussed in terms of the *meta* categories, that is, *Drug* encloses all the classes encompassed by DrugBank, *Food and Diet* refers to the food ingredients and food products classified by FoodOn, *Symptom* relates to the symptoms as presented by SYMP, *Treatment* refers to the treatments classified by NCIT, and *Disease* groups together disease terms, phenotype characteristics, and related medical vocabulary, as described by DOID.

The named entity recognition pipeline was implemented in house and entailed dictionary lookup, as well as pattern- and rule-based recognition. To be able to match the lexicon with tweet contents, the lexicon required some processing, namely, convert all terms to lowercase, remove extra whitespaces, remove small and long terms (ie, less than 2 characters and terms longer than the maximum tweet length), replace special characters by a whitespace, and remove terms associated with more than one category.

An inverted recognition technique was used in actual entity recognition [43]. This technique uses the words in the text as patterns to be matched against the lexicon. This was a valid approximation for this study because the number of words in the tweets were much smaller than the number of terms in the lexicon, that is, a fewer number of patterns to match. Recognition preference was given to the longest possible n-grams. In addition, the recognizer accepted perfect matches as well as lexical variations of the terms (ie, lemmatized entries and abbreviations).

Sentiment Analysis

The sentiment of the tweets was analyzed using the Valence Aware Dictionary and sEntiment Reasoner (VADER) API for Python [44]. VADER is a lexicon- and rule-based sentiment analysis tool that is specifically attuned to sentiments expressed in social media. The predicted sentiment (ie, compound score) is computed by summing the valence scores of each word in the lexicon, adjusted according to emotion-related rules, and then normalized to have values between −1 (most extreme negative emotion) and +1 (most extreme positive emotion).

Network Representation and Analysis

Graph analysis was performed to measure the relevance of individual terms as well as term-term pairs. Specifically, this analysis was applied to user interactions, via mentions and retweets (ie, directed mentions and retweets), and co-occurrence of semantically meaningful terms (ie, whenever 2 terms were found in the same tweet, these 2 terms were considered to share a link).

Networks were generally described in terms of the number of nodes and edges, as well as metrics of degree, characteristic path length, clustering coefficient, and the average number of neighbors [45]. In more detail, graph connectedness was measured by degree centrality, betweenness centrality, and closeness centrality [46-48]. Briefly, degree centrality measured the total amount of direct links with the other nodes (ie, higher degree implies the node is more central), betweenness centrality measured the *mediation* role of the nodes (ie, if other nodes have to go through the node to ensure communication, then the node is likely important and has a high betweenness centrality), and closeness centrality measured the convenience and ease of

XSL•FO
RenderX

connections between each node and the rest of the nodes (ie, if the average shortest path of the node is small, then the node has a high closeness centrality) [49,50].

The clustering coefficient was used to measure the degree to which nodes tend to cluster together. Evidence suggests that social network nodes tend to create tight groups, which are characterized by a relatively high density of links; this likelihood tends to be greater than the average probability of a link randomly established between 2 nodes [51,52].

Results

Overview

The Results section was structured following the logical order of the proposed questions. Figure 3 illustrates how particular research questions were answered by a certain analysis and identifies the main insights provided by each of these analyses. Most notably, results were structured such that the basic characterization of the BD-related tweets in terms of user activity and topics of interest were presented first and, then, the interrelation of topics in the conversations was detailed along with the visibility of external sources of information.

The ability to identify the most active users (ie, users that post more tweets) and topic-specific communities (eg, gluten-centered community) is pivotal to gain a better understanding about how to adapt or finetune the communication (ie, reach a broader audience or focus on a specific community) as well as discover influencers (ie, *vessels* of information distribution, within or across the communities). The analysis of user activity (different time windows may apply to different communities) is interesting as a means to plan the best time to post a new tweet (ie, when to expect major user attention).

The semantic annotation of tweet contents goes a step forward, providing knowledge about the topics being discussed (ie, individually and in combination). Likewise, the identification of tweets linking to external information sources is relevant to bring forward the visibility these sources are receiving through Twitter.

A health-related stakeholder is a typical example of someone that can benefit from the insights provided by the overall study. Say that the aim is to plan an awareness campaign about BD and food habits. Likely, this stakeholder wants to study the target audience in terms of topics that are attracting more attention and tweeting habits. The campaign may be planned to receive short-term attention (eg, announcing the launch of a novel drug or a new food supplement) or promote awareness throughout a longer period of time (eg, promote healthier food habits).

Bowel Disease Communities

User relations, that is, mentions and retweets, were represented in a network to study communication interplay, namely, to identify target communities and influential users (ie, individuals and/or institutions with high audiences). Figure 4 depicts this network such that the nodes denote the users (ie, account name), the node size is based on the node in-degree (ie, the users that received most communication are represented by bigger nodes), and the edges account for the number of mentions and retweets. The 4 different background areas in the figure denote the communities, mostly characterized by account descriptions: gluten-, nutritional-, BD-, and food-related relations. In this study, only communities with more than 5 connected users were considered.

In this network, the out-degree of the nodes (ie, users initiating communication with other users) is much lower than the in-degree of the nodes (ie, users receiving significant tweets). The user accounts @CrohnsColitisUK, @CrohnsColitisFn, and @HealioGastro, which belong to disease-specific organizations and have primarily informative/educational goals, were among the accounts receiving more communication (ie, higher in-degree). This conveys the rationale that individuals typically prefer to ask for health information to trusted organizations [6,53,54]. Conversely, accounts of BD patients (eg, @colitisandme) and experts in bowel conditions or nutrition (eg, @IBDMD and @charlie_lees) were among those showing highest out-degree.

Public lists of BD-related influencers [55] supported the identification of organization accounts, such as @CrohnsColitisUK, @HealioGastro, and @ACCUCatalunya, and personal accounts, such as @colitisandme, @IBDMD, and @EdwardLoftus2, that are typically reached for medical advice.

Figure 5 shows an example of a tweet exchange: the user @Crohnoid, a Crohn disease patient, asks the user @ibddoctor, an expert in BD, about 2 possible diagnostic techniques, that is, magnetic resonance imaging and computed tomography; @ibddoctor explained the advantages and disadvantages of these techniques, and another user @SandraZelinsky also intervened, pointing out limited access.

Demographic Distribution and Hot Zones

Knowledge about how the community is demographically distributed is important to carry out public information campaigns and study the impact of government and institutional actions in different demographic areas. In this analysis, 59.98% of the users (7975/13,295) were geographically distributed all around the world.

Figure 6 shows the geographical distribution of the BD communities. The size of the nodes represents the number of users located in each country (ie, the bigger the circle, the higher is the number of users) and colors identify the continents (ie, blue for America, purple for Europe, brown for Africa, green for Asia, and red for Australia).

In general, most of the users were located in the United States, the United Kingdom, Canada, and Australia, which is consistent with current knowledge about the prevalence of these conditions [56]. The highest reported prevalence is in Europe (with the highest prevalence of ulcerative colitis in Norway and of Crohn disease in Germany) and North America (with the highest prevalence of ulcerative colitis in the United States and Crohn disease in Canada). The prevalence of IBD exceeded 0.3% in North America, Oceania, and many countries in Europe.

The number of users from India, Bangladesh, and the Philippines was also noticeable and may be explained by the rising incidence

XSL•FO
RenderX

of BD conditions in newly industrialized countries in Africa, Asia, and South America.

Biomedical and Scientific Topics

Understanding what the user community is talking about was important to identify the topics that received more attention and to be able to better align new information/communication strategies with the interests of the community. Table 1 presents the top 25 most mentioned terms along with their corresponding semantic category (ie, *Food and Diet*, *Disease*, *Treatment*, *Symptom*, and *Drug*) and sorted by the number of including tweets. Explicit mentions to IBS and IBD (eg, *Inflammatory Bowel Disease* or *Ulcerative Colitis*) and noncontent–bearing, generalist terms (eg, *Disease* or *Food*), were at the top of term mentions but were not listed in the table.

The #Tweets column indicates the number of tweets in which a term is mentioned, the #Favorites column indicates the number of times that a tweet containing a term was selected as a favorite, the #Retweets column indicates the number of times that a tweet containing a term was retweeted, and the #Hashtags column indicates the number of times that a term was used in hashtags, including term variations (eg, *anxiety disorder* like *#anxiety*).

The volume of retweets is useful to understand how the information flows among users, whereas the number of favorites can be understood as a metric to measure the usefulness of the tweets containing the term [57,58]. In addition, the number of times a term appears in hashtags is interesting to understand how the user wishes the tweet to be indexed (to be easily found by others with similar interests as well as to bring attention to certain topics). This can be understood as a metric to measure the impact of the term in the community [59].

As illustrated in Table 1 (and in the topological analysis of the co-occurrence network of terms in Multimedia Appendix 1), the most mentioned terms related to *Disease* (35.64%, 11,688/32,794) and *Food and diet* (25.43%, 8342/32,794) categories, followed by terms from *Symptom* (17.71%, 5811/32,794), *Treatment* (14.83%, 4864/32,794), and *Drug* (6.37%, 2089/32,794) categories, respectively.

Diarrhea (10.33%, 1208/11,688 of the mentions to disease terms) and *constipation* (8.94%, 1046/11,688 of the mentions to disease terms) were the most mentioned disease-related terms. This was somewhat expected considering that many IBD and IBS patients have diarrhea as a side effect of the disease and many others claim to have problems of constipation [60]. Interestingly, constipation is one of the terms most often included in hashtags, which indicates that the topic is meaningful and timely to the community. Other diseases typically associated with BD conditions were also discussed, notably *cancer* (3.37%, 395/11,688 of the mentions to disease terms), *anxiety disorder* (3.04%, 356/11,688 of the mentions to disease terms), *depression* (2.49%, 292/11,688 of the mentions to disease terms), *arthritis* (1.71%, 200/11,688 of the mentions to disease terms), and *asthma* (1.46%, 171/11,688 of the mentions to disease terms). Chronic inflammation is known to be a major risk factor for the development of gastrointestinal malignancies and is often associated with inflammatory arthritis. Notably, as the population of patients with IBD grows older, with longer periods of chronic inflammation and longer exposure to immunosuppression, there is an increased risk of developing cancer [61] and arthritis [62]. Moreover, studies show that the rates of anxiety and depression tend to be higher among patients with Crohn disease or ulcerative colitis compared with those with other diseases as well as the general population [63,64]. Finally, interest in discussing asthma is justified by the association between this disease and early- and late-onset ulcerative colitis, particularly because of shared environmental risk factors [65]. Posts discussing BD and other diseases often show high retweet rates, and it is noticeable that depression is currently at the center of attention (ie, those tweets have the greatest number of favorites).

Symptoms typically associated with the different stages of the disease such as *bloating* (4.40%, 256/5811 of the mentions to symptom terms), *flatulence* (3.20%, 186/5811 of the mentions to symptom terms), and *abdominal pain* (2.97%, 173/5811 of the mentions to symptom terms) were also discussed with considerable frequency and the containing tweets were among the most retweeted, that is, people within this community find it relevant to broadcast information related with the symptomatology of BD conditions (eg, symptom-disease evidence and symptom relief therapies) [60].

The BD community is also sharing information about diet-based treatments and shows particular interest in *gluten-free dietary interventions* [66] and *probiotics* [67] (both with a 7.32%, 611/8342 and 544/8342 of the mentions to food and diet terms). Although the characteristics of gluten sensitivity in IBD remain unclear, gluten is known to generate peptides that can alter intestinal permeability and affect the immune system [68]. Recent studies investigated the effects of a wheat gluten–containing diet on the evolution of sodium dextran sulfate–induced colitis [69] and evaluated the usefulness of a low fermentable oligosaccharides, disaccharides, monosaccharides, and polyols diet on patients with IBS, nonactive IBD, and celiac disease compared with a gluten-free diet [70]. Likewise, some prebiotics, such as germinated barley foodstuff, psyllium, or oligofructose-enriched inulin, might provide some benefit in patients with active ulcerative colitis or ulcerative colitis in remission [71]. Other studies suggest that the VSL#3 probiotic may be effective in inducing remission in active ulcerative colitis [72]. The high number of favorites, retweets, and hashtags reflect the interest of the BD community in knowing more about these dietary interventions and discuss the pros and cons of different diets and foods.

XSL•FO
RenderX

Table 1. Top 25 mentioned terms sorted by the number of tweets.

Term	Category	Number of tweets	Number of favorites	Number of retweets	Number of hashtags
Diarrhea	Disease	1009	487	271	45
Constipation	Disease	971	359	213	152
Pain	Symptom	701	682	267	112
Medical cannabis	Drug	630	1134	614	100
Gluten	Food and Diet	519	1952	191	62
Probiotic	Food and Diet	498	936	344	142
Celiac disease	Disease	376	542	221	16
Cancer	Disease	367	691	424	50
Anxiety disorder	Disease	338	724	269	52
Depression	Disease	283	1107	501	51
Bloating	Symptom	264	158	80	52
Fibromyalgia	Disease	213	362	161	74
Dietary supplement	Food and Diet	196	602	146	9
Arthritis	Disease	190	872	375	32
Myocardial infarction	Disease	188	122	77	17
Allergic hypersensitivity disease	Disease	179	471	241	22
Flatulence	Symptom	175	482	35	6
Abdominal pain	Symptom	172	801	129	6
Multiple sclerosis	Disease	172	480	243	15
Asthma	Disease	170	486	259	12
Obesity	Disease	152	370	186	24
Vitamin D	Food and Diet	148	96	49	26
Heart disease	Disease	148	289	142	33
Autistic disorder	Disease	144	435	210	29
Hypnotherapy	Treatment	131	210	131	29

In general, drugs and nondiet-based treatments are less mentioned than the other categories. However, *medical cannabis* (36.38%, 760/2089 of the mentions to drug terms) and *hypnotherapy* (3.57%, 174/4864 of the mentions to treatment terms) raise some attention as an alternative, nonconventional treatments. Although these results are still inconclusive, recent studies show improvements in some BD-related symptoms. Medical cannabis is being tested for the treatment of gastrointestinal disorders such as abdominal pain and diarrhea. Experimental tests show that single ingredients from cannabis, such as tetrahydrocannabinol and cannabidiol, are responsible for these effects [73,74]. Conversely, the use of hypnotherapy is being validated in the treatment of gastrointestinal symptoms such as reducing fasting distal colonic motility or reducing systemic and rectal mucosal inflammatory responses [75]. Although, the results of using these alternative therapies remain unclear. As it stands, users show greater interest for cannabis-related posts, as denoted by the high number of tweet favorites and retweets as well as the inclusion of hashtags. In addition, users find these tweets interesting as they choose to disseminate them among their followers and make the topic indexation easier.

Topics of Patient Communications

A closer look into the tweets posted by BD patients was relevant to better understand how patients feel about and deal with BD symptoms, associated diseases, changes in habits, and medication. This analysis considered only the tweets posted by user accounts identified as patients (58.63%, 7795/13,295 of the total of user accounts). Moreover, it was possible to identify 2800 females (35.92%, 2800/7795 of the user accounts identified as patients) and 3860 males (49.51%, 3860/7795 of the user accounts identified as patients). Typically, the 11,098 tweets (45.05%, 11,098/24,634 of the total of tweets) posted by patients expressed a negative sentiment (51.99%, 5770/11,098), but there were also positive tweets (31.99%, 3551/11,098) and some tweets with a neutral sentiment (13.99%, 1553/11,098).

The high amount of negative opinions is explained by the fact that patients are known to use social platforms to *vent out* their emotions, namely, their frustrations when it comes to diseases that do not have a cure, such as in the case of IBD [76,77]. Table 2 describes the tweets of patients in terms of the recognized semantics terms and the tweet sentiment.

XSL•FO
RenderX

Table 2. Distribution of the sentiment of tweets by semantic category.

Category	Negative tweets, n (%)	Neutral tweets, n (%)	Positive tweets, n (%)	Total number of tweets, N
Disease	5691 (54.00)	1686 (16.00)	3161 (30.00)	10,538
Symptom	1152 (67.02)	103 (5.99)	464 (26.99)	1719
Food and Diet	1303 (47.01)	222 (8.01)	1247 (44.98)	2772
Drug	237 (36.02)	118 (17.93)	303 (46.05)	658
Treatment	710 (43.99)	339 (21.00)	565 (35.01)	1614

An important part of the tweets expressing positive emotions was related to gluten-free diets and probiotic supplements. In this line, there were more positive tweets posted by females than males (57.98%, 2274/3922 of the female tweets against 51.99%, 2390/4597 of the male tweets). However, in general, the tweets containing mentions to foods and diets (Figure 7) showed a similar distribution in terms of negative and positive sentiments. The main reason was that the effects of certain food components (eg, fiber, iron, and magnesium) were perceived differently, depending on the state of the disease and other personal conditions of the patient.

Disease and symptom (Figure 7) were the semantic categories with the highest number of mentions in negative tweets. These tweets discussed the *bad side* of BD conditions, that is, symptoms such as pain, fatigue, and migraines along with the co-occurrence of other conditions, namely, depression, diarrhea, and fibromyalgia [64,78]. In contrast, several tweets containing mentions of drugs (Figure 7) showed positive emotions, namely, the tweets highlighting the benefits of medical cannabis for the treatment of different chronic diseases.

As a means to look into these tweets from another perspective, Figure 8 depicts a subgraph of the semantic co-occurrence network reconstructed from the patient tweets (see details on the topological analysis of the complete network in Multimedia Appendix 2). This subgraph shows the relations between diseases and drugs. The size of the nodes is based on the node degree (ie, bigger nodes represent the terms mentioned in more tweets), whereas the edge size expresses the strength of co-occurrence (ie, thicker edges represent a higher number of term-term occurrences). Red nodes represent drugs and yellow nodes represent diseases, whereas the edge color stands for the tweet sentiment (ie, red indicates a majority of negative tweets, green indicates a majority of negative tweets, and black indicates neutral sentiment).

Although medical cannabis (and its components such as cannabidiol) was the most discussed drug, it was also possible to track down patient discussions about commercial drugs such as l-glutamine, lactulose, loperamide, and *Plantago* seed. For example, patients reported the positive effect of l-glutamine on diarrhea and muscular atrophy, but they also expressed their concern with glutamate leading to the occurrence of anxiety disorders. Glutamine is used to protect the mucous membrane of the esophagus and intestines and can boost immune cell activity in the gut [79,80]. Conversely, glutamine helps to create gamma-aminobutyric acid, a neurotransmitter that can stable the mind but can also produce glutamate, an excitatory neurotransmitter that can overstimulate the brain [81].

Patients discussed the use of laxatives in treating constipation, but not all of the mentioned laxatives were recommended. Notably, lactulose is a sugar that cannot be digested in the gut and thus, tends to cause or aggravate IBS symptoms, such as gas, bloating, discomfort, and cramping [82]. In turn, loperamide was suggested as an effective astringent to treat diarrhea and constipation. Indeed, a previous study reported a significant improvement in stool frequency and consistency [83].

Finally, patients were interested in the beneficial healing properties of medicinal plants such as *Plantago* (in various forms, such as roasted seeds, decoction, or syrup), namely, anti-inflammatory, laxative, and astringent properties [84].

Temporal Analysis of User Activity

Table 3 describes the number of tweets and the volume of tweet interactions (ie, retweets and favorites) from February 1, 2018, to August 31, 2018. It was interesting to identify the periods of time when users are more likely to tweet and, in particular, how specific events, such as the IBD day, a scientific conference, or an informational campaign, could affect such activity.

In particular, the celebration of the World IBD Day (on May 19, 2018), which is a worldwide event to raise awareness about BD conditions and to urge governments and health care professionals to take action and show support to the sufferers, motivated an increase in tweet interactions, that is, retweets and favorites (ie, 15,820 and 18,158 tweet interactions in May and June 2018, respectively). The average number of retweets during these months increased by 174% (3314/1902) compared with the average number of retweets during the rest of the year. Regarding tweet favorites, the increment was still noticeable (329%, 13,585/4130) compared with the average number of favorites during the rest of the year.

XSL•FO
RenderX

Table 3. Monthly activity of posting and interaction during the analyzed period.

Month	Number of tweets, n	Number of tweet interactions, n
February	2930	4285
March	4239	5822
April	3992	6693
May	4202	16,506
June	3422	18,298
July	2366	7955
August	3952	4468

External Sources of Information

The hyperlinks shared via tweet provided useful information about current health research and development initiatives (public and private), health promotion actions, and other events that are promoted via Twitter as means to reach out to and engage more people. Tables 4 and 5 summarize the results obtained by grouping the URLs into 4 categories: (1) informative or awareness campaigns about BD; (2) scientific sites, articles, and conferences, usually covering new treatments; (3) health Web pages related to a specific disease; and (4) commercial sites. The #Tweets column indicates the number of tweets in which the URL is mentioned, the #Retweets column indicates the number of times that a tweet containing the URL was retweeted, and the #Favorites column indicates the number of times that a tweet containing the URL was selected as a favorite.

Only a small number of external links included in tweets posted by organizations got highly retweeted and labeled as favorite. Notably, the Guts4life Web page (a portal about IBD) got the highest number of retweets and favorites. This sort of analysis can be of aid in identifying the external sources that the community finds most useful/interesting, especially considering that most of them are related to pages describing BD symptomatology and potential treatments.

Regarding the external links included in the tweets shared by patients, their information flow was in the BD community at the same level compared with that of links posted by organizations (an average of 49 retweets against an average of 49 retweets, respectively). In turn, the links shared by patients had a lower average number of favorites compared with those shared by organizations (ie, an average of 44 favorites against an average of 92 favorites, respectively).

Looking into the linked contents, most of the resources were related to posted articles that belonged to highly prestigious journals, namely, *Nature* and *British Medical Journal*, and reported recent research in BD topics (with no particular focus). The presence of commercial links to pages selling *stoner*- and other drug-related products that do not require a medical prescription was also noteworthy.

Table 4. Top 10 external sources of information mentioned in the bowel disease tweets posted by organizations. The URLs are sorted by the corresponding sum of the number of tweets, retweets, and favorites.

Source of information	Category	Number of tweets, n	Number of retweets, n	Number of favorites, n
Guts4life [85]	Health Web page	7	164	1918
Are Your Digestion Troubles Irritable Bowel Syndrome? [86]	Informative or campaign	5	128	231
About Crohn's Disease [87]	Informative or campaign	1	69	63
Can You Treat Irritable Bowel Syndrome with Cannabis? [88]	Informative or campaign	4	34	84
How to Manage Irritable Bowel Syndrome with Your Brain [89]	Informative or campaign	3	30	77
Inflammatory Bowel Disease (IBD) [90]	Health Web page	10	45	51
Why I Get Excited When You Say You Know Someone With IBD [91]	Informative or campaign	3	27	65
New Treatment Options for Inflammatory Bowel Diseases [92]	Scientific	1	30	57
Supporting Someone With IBD: A Guide For Friends and Family [93]	Health Web page	2	31	48
Chronic Inflammation [94]	Informative or campaign	5	48	26

XSL•FO
RenderX

Table 5. Top 10 external sources of information mentioned in the bowel disease tweets posted by patients. The URLs are sorted by the sum of the corresponding number of tweets, retweets, and favorites.

Source of information	Category	Number of tweets, n	Number of retweets, n	Number of favorites, n
Ginger for Nausea, Menstrual Cramps and Irritable Bowel Syndrome [95]	Informative or campaign	1	62	122
Symptoms of Ulcerative Colitis [96]	Commercial	1	178	1
Medicinal Marijuana as a Treatment for IBD Inflammatory Bowel Disease [97]	Commercial	8	45	94
A Starbucks barista called 911 [98]	Informative or campaign	1	24	99
Advances in Inflammatory Bowel Disease Pathogenesis: Linking Host Genetics and the Microbiome [99]	Scientific	1	47	70
Fungal Microbiota Dysbiosis in IBD [100]	Scientific	1	47	70
Murine Colitis Reveals A Disease-Associated Bacteriophage Community [101]	Scientific	1	47	70
I Am LOVING These Probiotics! [102]	Informative or campaign	1	83	27
Effects of Prebiotics vs a Diet Low in FODMAPs in Patients With Functional Gut Disorders [103]	Scientific	2	28	56
Acute GI Bleeding [104]	Health Web page	2	31	50

Discussion

Principal Findings

The objective of this paper was to characterize and study the BD community on Twitter. To do so, a dataset of tweets related to BD was collected, processed, and analyzed. The dataset covered a consecutive period of 8 months, from February 1, 2018, to August 31, 2018. As a whole, this analysis provided new insights into 6 main questions: Which are the main communities and most influential users?; Where are the main content providers from?; What are the key biomedical and scientific topics under discussion? How are topics interrelated in patient communications?; How do external events influence user activity?; What kind of external sources of information are being promoted?

Health organizations and BD experts (eg, @CrohnsColitisUK and @IBDMD) were the users that received more tweets, typically looking for trusted information about the conditions. Patients shared experiences among themselves or asked for medical advice. Moreover, the most active users were located in the United States and the United Kingdom, which are among the demographic regions with highest BD prevalence.

Most of the tweets talked about BD symptoms, related diseases, foods, and diets. Specifically, diarrhea, constipation, and pain were the symptoms that raised more concern (in general as well as among patients), whereas gluten and probiotics were among the most discussed dietary interventions (including a high number of favorites and retweets). In this line, females showed higher positive emotions about these dietary interventions than males. Medical cannabis was the most commented drug, notably in the tweets that raised the highest number of favorites, and patients actively discussed the beneficial effects of cannabis (and its components) in mitigating common BD symptoms. Regarding more commercial drugs, patients expressed positive emotions with the usage of l-glutamine on diarrhea and muscular atrophy but also reported negative sentiments because of the production of glutamate and its influence on anxiety disorders. Another notable group of drugs in the discussion were the laxatives, for example, the usage of lactulose was associated with negative emotions because it tends to cause or aggravate IBS symptoms, whereas loperamide was noted to be an effective astringent to treat diarrhea and constipation.

Users were more active during and after the World IBD Day, which shows that these types of initiatives are raising public awareness about these diseases as well as indicates that social networks are part of the routine communication of the BD community. The external resources being shared in tweets by organizations aim to draw people's attention to awareness/informational sites, whereas those shared by patients typically point to more scientific contents (eg, scientific articles on BD) and alternative treatments (such as cannabis).

Limitations

The most immediate limitation arises from the fact that the capture of raw data was carried out using the free Twitter API, that is, the identification of users that select one tweet as a favorite as well as the number of available tweets is restricted, with no assurance of a random or representative sample [105]. For this reason, it was not possible to perform a more exhaustive analysis of the social BD communities. Thus, a full data retrieval through automated dashboard vendors, or using a paid service of the Twitter API, may provide further insights.

It should also be noted that this study was based on the assumption that the data entered by Twitter users are true. It is not possible to detect if certain data (eg, the user profile picture or the user location) are reliable. This limitation impacts mainly the analysis of the demographic distribution and the conclusions inferred for a particular gender. That being said, the obtained

XSL•FO
RenderX

results were in accordance with the common knowledge of these communities.

Language is another aspect of analysis to take into consideration. This study was only focused on tweets written in English. If the applied techniques are extended to support a greater variety of languages, such as Chinese or Spanish, it may provide complementary findings.

Finally, this study was focused on Twitter. However, the set of social networks might be expanded to analyze a richer dataset from a wide variety of sources. Considering other studies in the literature [21,106,107], Facebook and Instagram would be also sources of interest, although public data access is greatly limited.

Conclusions and Further Research

In this study, tweets related to BD were analyzed to characterize the user community and the exchanged contents. According to the obtained results, it was possible to detect communities and to describe the most discussed topics among these communities. The large and increasing volume of tweets demonstrates that Twitter is becoming a space for online conversation about BD, namely, associated symptoms and alternative treatments. In addition, the location of users indicates that conversations are happening at a global scale and, motivated by this, health-related stakeholders are using the platform to reach out to a larger audience on a daily basis.

In terms of future research, it would be interesting to perform user classification, that is, being able to identify experts, researchers, and companies, as well as patients. Thus, it would be possible to apply different sentiment analysis and TM approaches to the tweets to explore user-specific motivations, questions, and concerns. For example, it would be interesting to discover the opinion of patients about different treatments and specific symptoms.

Acknowledgments

SING group thanks CITI (*Centro de Investigación, Transferencia e Innovación*) from the University of Vigo for hosting its information and technology infrastructure. This study was partially supported by the Consellería de Educación, Universidades e Formación Profesional (Xunta de Galicia) under the scope of the strategic funding of ED431C2018/55-GRC Competitive Reference Group, the Portuguese Foundation for Science and Technology (FCT) under the scope of the strategic funding of UID/BIO/04469/2013 unit and COMPETE 2020 (POCI-01-0145-FEDER-006684). The authors also acknowledge the PhD grants of MPP and GP-R, funded by the Xunta de Galicia.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Topological analysis of the co-occurrence network of terms for all types of users.
[PDF File (Adobe PDF File), 369KB - jmir_v21i8e12610_app1.pdf]

Multimedia Appendix 2
Topological analysis of the co-occurrence network of terms for patient users.
[PDF File (Adobe PDF File), 263KB - jmir_v21i8e12610_app2.pdf]

References

1. Miller D, Costa E, Haynes N, McDonald T, Nicolescu R, Sinanan J, et al. How the World Changed Social Media. London: UCL Press; 2016.
2. Mainka A, Hartmann S, Stock WG, Peters I. Government and Social Media: A Case Study of 31 Informational World Cities. In: Proceedings of the 47th Hawaii International Conference on System Sciences. 2014 Presented at: HICSS'14; January 6-9, 2014; Waikoloa, Hawaii p. 1715-1724. [doi: 10.1109/HICSS.2014.219]
3. Smith A, Anderson M. Pew Research Center. 2018. Social Media Use in 2018 URL: https://www.pewinternet.org/2018/03/01/social-media-use-in-2018/
4. European Commission. 2017. Eurostat Regional Yearbook: 2017 Edition URL: https://ec.europa.eu/eurostat/statistics-explained/index.php/Eurostat_regional_yearbook
5. Weller K, Bruns A, Burgess J, Mahrt M, Puschmann C. Twitter II: towards a news medium for event-following. In: Jones S, editor. Twitter and Society. New York: Peter Lang Publishing; 2014.
6. Moorhead SA, Hazlett DE, Harrison L, Carroll JK, Irwin A, Hoving C. A new dimension of health care: systematic review of the uses, benefits, and limitations of social media for health communication. J Med Internet Res 2013 Apr 23;15(4):e85 [FREE Full text] [doi: 10.2196/jmir.1933] [Medline: 23615206]
7. Walji M, Sagaram S, Meric-Bernstam F, Johnson CW, Bernstam EV. Searching for cancer-related information online: unintended retrieval of complementary and alternative medicine information. Int J Med Inform 2005 Aug;74(7-8):685-693. [doi: 10.1016/j.ijmedinf.2005.01.001] [Medline: 16027034]

XSL•FO
RenderX

8. Valdez RS, Brennan PF. Exploring patients' health information communication practices with social network members as a foundation for consumer health IT design. Int J Med Inform 2015 May;84(5):363-374. [doi: 10.1016/j.ijmedinf.2015.01.014] [Medline: 25704761]
9. Hwang KO, Ottenbacher AJ, Green AP, Cannon-Diehl MR, Richardson O, Bernstam EV, et al. Social support in an internet weight loss community. Int J Med Inform 2010 Jan;79(1):5-13 [FREE Full text] [doi: 10.1016/j.ijmedinf.2009.10.003] [Medline: 19945338]
10. Bedell SE, Agrawal A, Petersen LE. A systematic critique of diabetes on the world wide web for patients and their physicians. Int J Med Inform 2004 Sep;73(9-10):687-694. [doi: 10.1016/j.ijmedinf.2004.04.011] [Medline: 15325325]
11. Reich J, Guo L, Groshek J, Weinberg J, Chen W, Martin C, et al. Social media use and preferences in patients with inflammatory bowel disease. Inflamm Bowel Dis 2019 Feb 21;25(3):587-591. [doi: 10.1093/ibd/izy280] [Medline: 30203036]
12. M'Koma AE. Inflammatory bowel disease: an expanding global health problem. Clin Med Insights Gastroenterol 2013;6:33-47 [FREE Full text] [doi: 10.4137/CGast.S12731] [Medline: 24833941]
13. Molodecky NA, Soon IS, Rabi DM, Ghali WA, Ferris M, Chernoff G, et al. Increasing incidence and prevalence of the inflammatory bowel diseases with time, based on systematic review. Gastroenterology 2012 Jan;142(1):46-54.e42; quiz e30. [doi: 10.1053/j.gastro.2011.10.001] [Medline: 22001864]
14. -. Announcement: world IBD day - May 19, 2017. MMWR Morb Mortal Wkly Rep 2017 May 19;66(19):516 [FREE Full text] [doi: 10.15585/mmwr.mm6619a9] [Medline: 28520712]
15. Curtis JR, Chen L, Higginbotham P, Nowell WB, Gal-Levy R, Willig J, et al. Social media for arthritis-related comparative effectiveness and safety research and the impact of direct-to-consumer advertising. Arthritis Res Ther 2017 Dec 7;19(1):48 [FREE Full text] [doi: 10.1186/s13075-017-1251-y] [Medline: 28270190]
16. Roland D, Spurr J, Cabrera D. Preliminary evidence for the emergence of a health care online community of practice: using a netnographic framework for Twitter hashtag analytics. J Med Internet Res 2017 Dec 14;19(7):e252 [FREE Full text] [doi: 10.2196/jmir.7072] [Medline: 28710054]
17. Tsuya A, Sugawara Y, Tanaka A, Narimatsu H. Do cancer patients Tweet? Examining the Twitter use of cancer patients in Japan. J Med Internet Res 2014 May 27;16(5):e137 [FREE Full text] [doi: 10.2196/jmir.3298] [Medline: 24867458]
18. Bian J, Zhao Y, Salloum RG, Guo Y, Wang M, Prosperi M, et al. Using social media data to understand the impact of promotional information on laypeople's discussions: a case study of lynch syndrome. J Med Internet Res 2017 Dec 13;19(12):e414 [FREE Full text] [doi: 10.2196/jmir.9266] [Medline: 29237586]
19. Liu Y, Mei Q, Hanauer DA, Zheng K, Lee JM. Use of social media in the diabetes community: an exploratory analysis of diabetes-related tweets. JMIR Diabetes 2016 Nov 7;1(2):e4 [FREE Full text] [doi: 10.2196/diabetes.6256] [Medline: 30291053]
20. Guo L, Reich J, Groshek J, Farraye FA. Social media use in patients with inflammatory bowel disease. Inflamm Bowel Dis 2016 May;22(5):1231-1238. [doi: 10.1097/MIB.0000000000000713] [Medline: 26894839]
21. Keller MS, Mosadeghi S, Cohen ER, Kwan J, Spiegel BM. Reproductive health and medication concerns for patients with inflammatory bowel disease: thematic and quantitative analysis using social listening. J Med Internet Res 2018 Dec 11;20(6):e206 [FREE Full text] [doi: 10.2196/jmir.9870] [Medline: 29891471]
22. Yamamoto Y. Twitter4J. 2018. Twitter4J - A Java Library for the Twitter API URL: http://twitter4j.org/en/index.html [accessed 2018-07-10]
23. Smith A, Anderson M. Pew Research Center. 2018. The Demographics of Social Media Use in 2018 Internet URL: http://www.pewinternet.org/2018/03/01/social-media-use-in-2018/ [accessed 2018-07-10]
24. Raffo J. IDEAS/RePEc: Economics and Finance Research. 2016. Worldwide Gender-Name Dictionary URL: https://ideas.repec.org/c/wip/eccode/10.html
25. Rothe R, Timofte R, Van Gool L. Deep expectation of real and apparent age from a single image without facial landmarks. Int J Comput Vis 2016 Aug 10;126(2-4):144-157 [FREE Full text] [doi: 10.1007/s11263-016-0940-3]
26. GeoNames. URL: http://www.geonames.org/ [accessed 2018-07-10]
27. GitHub. 2017. Twitter-Text Library URL: https://github.com/twitter/twitter-text [accessed 2018-07-10]
28. Hunspell. URL: https://hunspell.github.io/ [accessed 2018-07-10]
29. GitHub. 2016. Hunspell-en-med-glut: Hunspell Dictionary of English Medical Terms URL: https://github.com/glutanimate/hunspell-en-med-glut [accessed 2018-07-10]
30. e-MedTools. 2017. Medical Spell Checker for Firefox and Thunderbird URL: https://e-medtools.com/openmedspel.html [accessed 2018-07-10]
31. HugeDomains. 2017. Medical Spell Checker for Microsoft Word URL: http://mtherald.com/free-medical-spell-checker-for-microsoft-word-custom-dictionary/ [accessed 2018-07-10]
32. Yujian L, Bo L. A normalized Levenshtein distance metric. IEEE Trans Pattern Anal Mach Intell 2007 Jun;29(6):1091-1095. [doi: 10.1109/TPAMI.2007.1078] [Medline: 17431306]
33. Beal V. Webopedia: Online Tech Dictionary for IT Professionals. 2004. Huge List of Texting and Chat Abbreviations URL: https://www.webopedia.com/quick_ref/textmessageabbreviations.asp [accessed 2018-07-10]
34. Rader W. The Online Slang Dictionary. 2018. Slang ('Urban') Thesaurus: Slang Words for Acronyms URL: http://onlineslangdictionary.com/thesaurus/words+meaning+acronyms+%28list+of%29.html [accessed 2018-07-10]

XSL•FO
RenderX

35. Beal V. Webopedia. 2010. Twitter Dictionary: A Guide to Understanding Twitter Lingo URL: https://www.webopedia.com/quick_ref/Twitter_Dictionary_Guide.asp [accessed 2018-07-10]
36. Manning CD, Surdeanu M, Bauer J, Finkel J, Bethard SJ, Mcclosky D. The Stanford CoreNLP Natural Language Processing Toolkit. In: Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations.: Association for Computational Linguistics; 2014 Presented at: ACL'14; June 22-27, 2014; Baltimore, Maryland p. 55-60 URL: https://stanfordnlp.github.io/CoreNLP/ [doi: 10.3115/v1/P14-5010]
37. Musen MA, Noy NF, Shah NH, Whetzel PL, Chute CG, Story MA, NCBO Team. The national center for biomedical ontology. J Am Med Inform Assoc 2012;19(2):190-195 [FREE Full text] [doi: 10.1136/amiajnl-2011-000523] [Medline: 22081220]
38. Kibbe WA, Arze C, Felix V, Mitraka E, Bolton E, Fu G, et al. Disease ontology 2015 update: an expanded and updated database of human diseases for linking biomedical knowledge through disease data. Nucleic Acids Res 2015 Jan;43(Database Issue):D1071-D1078 [FREE Full text] [doi: 10.1093/nar/gku1011] [Medline: 25348409]
39. Dooley DM, Griffiths EJ, Gosal GS, Buttigieg PL, Hoehndorf R, Lange MC, et al. FoodOn: a harmonized food ontology to increase global food traceability, quality control and data integration. NPJ Sci Food 2018;2:23 [FREE Full text] [doi: 10.1038/s41538-018-0032-6] [Medline: 31304272]
40. Schriml LM. NCBO BioPortal. 2019. Symptom Ontology URL: https://bioportal.bioontology.org/ontologies/SYMP [accessed 2019-04-01]
41. Haendel M. NCBO BioPortal. 2019. National Cancer Institute Thesaurus URL: https://bioportal.bioontology.org/ontologies/NCIT/?p=summary [accessed 2019-04-01]
42. Wishart DS, Feunang YD, Guo AC, Lo EJ, Marcu A, Grant JR, et al. DrugBank 5.0: a major update to the DrugBank database for 2018. Nucleic Acids Res 2018 Jan 4;46(D1):D1074-D1082 [FREE Full text] [doi: 10.1093/nar/gkx1037] [Medline: 29126136]
43. Couto FM, Lamurias A. MER: a shell script and annotation server for minimal named entity recognition and linking. J Cheminform 2018 Dec 5;10(1):58 [FREE Full text] [doi: 10.1186/s13321-018-0312-9] [Medline: 30519990]
44. Hutto CJ, Gilbert E. VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text. In: Proceedings of the Eighth International AAAI Conference on Weblogs and Social Media. 2014 Presented at: AAAI'14; June 1–4, 2014; Ann Arbor, Michigan, USA.
45. Assenov Y, Ramírez F, Schelhorn SE, Lengauer T, Albrecht M. Computing topological parameters of biological networks. Bioinformatics 2008 Jan 15;24(2):282-284. [doi: 10.1093/bioinformatics/btm554] [Medline: 18006545]
46. Freeman LC, Roeder D, Mulholland RR. Centrality in social networks: ii. experimental results. Soc Networks 1979 Jan;2(2):119-141. [doi: 10.1016/0378-8733(79)90002-9]
47. Freeman LC. Centrality in social networks conceptual clarification. Soc Networks 1978 Jan;1(3):215-239. [doi: 10.1016/0378-8733(78)90021-7]
48. Freeman LC. A set of measures of centrality based on betweenness. Sociometry 1977 Mar;40(1):35. [doi: 10.2307/3033543]
49. Nieminen J. On the centrality in a graph. Scand J Psychol 1974 Sep;15(1):332-336. [doi: 10.1111/j.1467-9450.1974.tb00598.x]
50. Bolland JM. Sorting out centrality: an analysis of the performance of four centrality models in real and simulated networks. Soc Networks 1988 Sep;10(3):233-253. [doi: 10.1016/0378-8733(88)90014-7]
51. Watts DJ, Strogatz SH. Collective dynamics of 'small-world' networks. Nature 1998 Jun 4;393(6684):440-442. [doi: 10.1038/30918] [Medline: 9623998]
52. Holland PW, Leinhardt S. Transitivity in structural models of small groups. Comp Group Stud 2016 Aug 18;2(2):107-124. [doi: 10.1177/104649647100200201]
53. Antheunis ML, Tates K, Nieboer TE. Patients' and health professionals' use of social media in health care: motives, barriers and expectations. Patient Educ Couns 2013 Sep;92(3):426-431. [doi: 10.1016/j.pec.2013.06.020] [Medline: 23899831]
54. Neiger BL, Thackeray R, Burton SH, Thackeray CR, Reese JH. Use of twitter among local health departments: an analysis of information sharing, engagement, and action. J Med Internet Res 2013 Aug 19;15(8):e177 [FREE Full text] [doi: 10.2196/jmir.2775] [Medline: 23958635]
55. Gut Microbiota for Health. 2015. Gut Microbiota and Inflammatory Bowel Disease URL: https://tinyurl.com/y5cszbtz
56. Ng SC, Shi HY, Hamidi N, Underwood FE, Tang W, Benchimol EI, et al. Worldwide incidence and prevalence of inflammatory bowel disease in the 21st century: a systematic review of population-based studies. Lancet 2018 Dec 23;390(10114):2769-2778. [doi: 10.1016/S0140-6736(17)32448-0] [Medline: 29050646]
57. Meier F, Elsweiler D, Wilson ML. School of Computer Science - The University of Nottingham. 2017. More than Liking and Bookmarking? Towards Understanding Twitter Favouriting Behaviour URL: http://www.cs.nott.ac.uk/~pszmw/pubs/icwsm2014-favouriting.pdf
58. Gorrell G, Bontcheva K. Classifying Twitter favorites: like, bookmark, or thanks? J Assoc Inf SciTechnol 2014 Dec 22;67(1):17-25. [doi: 10.1002/asi.23352]
59. Bruns A, Burgess J. The Use of Twitter Hashtags in the Formation of Ad Hoc Publics. In: Proceedings of the 6th European Consortium for Political Research. 2011 Presented at: ECPR'11; August 25-27, 2011; Iceland, Reykjavik URL: https://eprints.qut.edu.au/46515/

XSL•FO
RenderX

60. Lee AD, Spiegel BM, Hays RD, Melmed GY, Bolus R, Khanna D, et al. Gastrointestinal symptom severity in irritable bowel syndrome, inflammatory bowel disease and the general population. Neurogastroenterol Motil 2017 Dec;29(5):- [FREE Full text] [doi: 10.1111/nmo.13003] [Medline: 27981684]
61. Axelrad JE, Lichtiger S, Yajnik V. Inflammatory bowel disease and cancer: the role of inflammation, immunosuppression, and cancer treatment. World J Gastroenterol 2016 May 28;22(20):4794-4801 [FREE Full text] [doi: 10.3748/wjg.v22.i20.4794] [Medline: 27239106]
62. Arvikar SL, Fisher MC. Inflammatory bowel disease associated arthropathy. Curr Rev Musculoskelet Med 2011 Sep;4(3):123-131 [FREE Full text] [doi: 10.1007/s12178-011-9085-8] [Medline: 21710141]
63. Yongwen NJ, Chauhan U, Armstrong D, Marshall J, Tse F, Moayyedi P, et al. A comparison of the prevalence of anxiety and depression between uncomplicated and complex IBD patient groups. Gastroenterol Nurs 2018;41(5):427-435. [doi: 10.1097/SGA.0000000000000338] [Medline: 30272605]
64. Byrne G, Rosenfeld G, Leung Y, Qian H, Raudzus J, Nunez C, et al. Prevalence of anxiety and depression in patients with inflammatory bowel disease. Can J Gastroenterol Hepatol 2017;2017:6496727 [FREE Full text] [doi: 10.1155/2017/6496727] [Medline: 29181373]
65. Kuenzig ME, Barnabe C, Seow CH, Eksteen B, Negron ME, Rezaie A, et al. Asthma is associated with subsequent development of inflammatory bowel disease: a population-based case-control study. Clin Gastroenterol Hepatol 2017 Sep;15(9):1405-12.e3 [FREE Full text] [doi: 10.1016/j.cgh.2017.02.042] [Medline: 28344063]
66. Britto S, Kellermayer R. Carbohydrate monotony as protection and treatment for inflammatory bowel disease. J Crohns Colitis 2019 Jan 30:-. [doi: 10.1093/ecco-jcc/jjz011] [Medline: 30715243]
67. Ganji-Arjenaki M, Rafieian-Kopaei M. Probiotics are a good choice in remission of inflammatory bowel diseases: a meta analysis and systematic review. J Cell Physiol 2018 Mar;233(3):2091-2103. [doi: 10.1002/jcp.25911] [Medline: 28294322]
68. Limketkai BN, Sepulveda R, Hing T, Shah ND, Choe M, Limsui D, et al. Prevalence and factors associated with gluten sensitivity in inflammatory bowel disease. Scand J Gastroenterol 2018 Feb;53(2):147-151. [doi: 10.1080/00365521.2017.1409364] [Medline: 29216767]
69. Menta PL, Andrade ME, Leocádio PC, Fraga JR, Dias MT, Cara DC, et al. Wheat gluten intake increases the severity of experimental colitis and bacterial translocation by weakening of the proteins of the junctional complex. Br J Nutr 2019 Feb;121(4):361-373. [doi: 10.1017/S0007114518003422] [Medline: 30554574]
70. Testa A, Imperatore N, Rispo A, Rea M, Tortora R, Nardone OM, et al. Beyond irritable bowel syndrome: the efficacy of the low fodmap diet for improving symptoms in inflammatory bowel diseases and Celiac disease. Dig Dis 2018;36(4):271-280. [doi: 10.1159/000489487] [Medline: 29763907]
71. Orel R, Trop TK. Intestinal microbiota, probiotics and prebiotics in inflammatory bowel disease. World J Gastroenterol 2014 Sep 7;20(33):11505-11524 [FREE Full text] [doi: 10.3748/wjg.v20.i33.11505] [Medline: 25206258]
72. Derwa Y, Gracie DJ, Hamlin PJ, Ford AC. Systematic review with meta-analysis: the efficacy of probiotics in inflammatory bowel disease. Aliment Pharmacol Ther 2017 Dec;46(4):389-400 [FREE Full text] [doi: 10.1111/apt.14203] [Medline: 28653751]
73. Hasenoehrl C, Storr M, Schicho R. Cannabinoids for treating inflammatory bowel diseases: where are we and where do we go? Expert Rev Gastroenterol Hepatol 2017 Apr;11(4):329-337 [FREE Full text] [doi: 10.1080/17474124.2017.1292851] [Medline: 28276820]
74. Goyal H, Singla U, Gupta U, May E. Role of cannabis in digestive disorders. Eur J Gastroenterol Hepatol 2017 Feb;29(2):135-143. [doi: 10.1097/MEG.0000000000000779] [Medline: 27792038]
75. Peters SL, Muir JG, Gibson PR. Review article: gut-directed hypnotherapy in the management of irritable bowel syndrome and inflammatory bowel disease. Aliment Pharmacol Ther 2015 Jun;41(11):1104-1115 [FREE Full text] [doi: 10.1111/apt.13202] [Medline: 25858661]
76. Wynn R, Oyeyemi SO, Johnsen JA, Gabarron E. Tweets are not always supportive of patients with mental disorders. Int J Integr Care 2017 Jul 11;17(3):149. [doi: 10.5334/ijic.3261]
77. Shaw Jr G, Karami A. Computational Content Analysis of Negative Tweets for Obesity, Diet, Diabetes, and Exercise. In: Proceedings of the Association for Information Science and Technology Banner. 2017 Presented at: ASIST'17; October 27-November 1, 2017; Washington DC, USA p. 357-365. [doi: 10.1002/pra2.2017.14505401039]
78. Chen JH, Chen HJ, Kao CH, Tseng CH, Tsai CH. Is fibromyalgia risk higher among male and young inflammatory bowel disease patients? Evidence from a Taiwan cohort of one million. Pain Physician 2018 Dec;21(3):E257-E264 [FREE Full text] [Medline: 29871388]
79. Kim MH, Kim H. The roles of glutamine in the intestine and its implication in intestinal diseases. Int J Mol Sci 2017 May 12;18(5):pii: E1051 [FREE Full text] [doi: 10.3390/ijms18051051] [Medline: 28498331]
80. Salehian B, Mahabadi V, Bilas J, Taylor WE, Ma K. The effect of glutamine on prevention of glucocorticoid-induced skeletal muscle atrophy is associated with myostatin suppression. Metabolism 2006 Sep;55(9):1239-1247. [doi: 10.1016/j.metabol.2006.05.009] [Medline: 16919545]
81. Struzyńska L, Sulkowski G. Relationships between glutamine, glutamate, and GABA in nerve endings under Pb-toxicity conditions. J Inorg Biochem 2004 Jun;98(6):951-958. [doi: 10.1016/j.jinorgbio.2004.02.010] [Medline: 15149801]

XSL•FO
RenderX

82. Hunter J. Irritable Bowel Solutions: The Essential Guide To Irritable Bowel Syndrome, Its Causes And Treatments. First Edition. United Kingdom: Random House UK; 2009.
83. Hovdenak N. Loperamide treatment of the irritable bowel syndrome. Scand J Gastroenterol Suppl 1987;130:81-84. [doi: 10.3109/00365528709091004] [Medline: 3306904]
84. Najafian Y, Hamedi SS, Farshchi MK, Feyzabadi Z. Plantago major in traditional Persian medicine and modern phytotherapy: a narrative review. Electron Physician 2018 Feb;10(2):6390-6399 [FREE Full text] [doi: 10.19082/6390] [Medline: 29629064]
85. Guts4life - The Home of IBD Information and Support. 2014. URL: http://www.guts4life.com/
86. Health Essentials from Cleveland Clinic. 2017. Are Your Digestion Troubles Irritable Bowel Syndrome? URL: https://health.clevelandclinic.org/digestion-troubles-irritable-bowel-syndrome/?utm_campaign=cc [accessed 2019-04-05]
87. NHS Inform. 2019. Crohn's Disease URL: https://www.nhsinform.scot/illnesses-and-conditions/stomach-liver-and-gastrointestinal-tract/crohns-disease [accessed 2019-05-04]
88. Lindsey N. High Times. 2018. Can You Treat Irritable Bowel Syndrome With Cannabis? URL: https://hightimes.com/health/can-treat-irritable-bowel-syndrome-cannabis/ [accessed 2019-05-04]
89. Health Essentials from Cleveland Clinic. 2017. How to Manage Irritable Bowel Syndrome with Your Brain URL: https://health.clevelandclinic.org/manage-irritable-bowel-syndrome-ibs-brain/?utm_campaign=cc [accessed 2019-05-04]
90. Medical Symptoms Guide. 2019. Inflammatory Bowel Disease (IBD) URL: https://www.medicalsymptomsguide.com/inflammatory-bowel-disease-ibd.html [accessed 2019-05-04]
91. Garcia T. The Mighty. Making Health About People. 2018. Why I Get Excited When You Say You Know Someone With IBD URL: https://themighty.com/2018/07/inflammatory-bowel-disease-knowledge-and-awareness/ [accessed 2019-05-04]
92. Verstockt B, Ferrante M, Vermeire S, van Assche G. New treatment options for inflammatory bowel diseases. J Gastroenterol 2018 May;53(5):585-590 [FREE Full text] [doi: 10.1007/s00535-018-1449-z] [Medline: 29556726]
93. Crohn's & Colitis UK. Supporting Someone With IBD: A Guide For Friends and Family URL: https://www.crohnsandcolitis.org.uk/about-inflammatory-bowel-disease/publications/supporting-someone [accessed 2019-05-04]
94. National Cancer Institute. 2015. Chronic Inflammation URL: https://www.cancer.gov/about-cancer/causes-prevention/risk/chronic-inflammation [accessed 2019-05-04]
95. Greger M. Care2. 2018. Ginger for Nausea, Menstrual Cramps and Irritable Bowel Syndrome URL: https://www.care2.com/greenliving/ginger-for-nausea-menstrual-cramps-and-irritable-bowel-syndrome.html [accessed 2019-05-04]
96. Symptoms of Ulcerative Colitis. 2018. URL: http://symptomsofulcerativecolitissupport.com/ [accessed 2019-05-04]
97. Mr. Stinky's Green Garden. 2018. Medicinal Marijuana as a Treatment for IBD Inflammatory Bowel Disease URL: http://www.mrstinkysgreengarden.com/2018/04/medicinal-marijuana-as-treatment-for.html [accessed 2019-05-04]
98. Twitter. 2018. A Starbucks Barista Called 911 URL: https://twitter.com/theappeal/status/985342724197896194 [accessed 2019-05-04]
99. Knights D, Lassen KG, Xavier RJ. Advances in inflammatory bowel disease pathogenesis: linking host genetics and the microbiome. Gut 2013 Oct;62(10):1505-1510 [FREE Full text] [doi: 10.1136/gutjnl-2012-303954] [Medline: 24037875]
100. Sokol H, Leducq V, Aschard H, Pham HP, Jegou S, Landman C, et al. Fungal microbiota dysbiosis in IBD. Gut 2017 Dec;66(6):1039-1048 [FREE Full text] [doi: 10.1136/gutjnl-2015-310746] [Medline: 26843508]
101. Duerkop BA, Kleiner M, Paez-Espino D, Zhu W, Bushnell B, Hassell B, et al. Murine colitis reveals a disease-associated bacteriophage community. Nat Microbiol 2018 Dec;3(9):1023-1031 [FREE Full text] [doi: 10.1038/s41564-018-0210-y] [Medline: 30038310]
102. Instagram. 2018. I am Loving these Probiotics! URL: https://www.instagram.com/p/BlTBIWShu3R/ [accessed 2019-04-11]
103. Huaman JW, Mego M, Manichanh C, Cañellas N, Cañueto D, Segurola H, et al. Effects of prebiotics vs a diet low in FODMAPs in patients with functional gut disorders. Gastroenterology 2018 Dec;155(4):1004-1007. [doi: 10.1053/j.gastro.2018.06.045] [Medline: 29964041]
104. NEJM Resident 360. 2018. Gastroenterology: Acute GI Bleeding URL: https://resident360.nejm.org/pages/home?resource_collection_id=gastroenterology&subtopic=acute-gi-bleeding [accessed 2019-04-11]
105. Kim AE, Hansen HM, Murphy J, Richards AK, Duke J, Allen JA. Methodological considerations in analyzing Twitter data. J Natl Cancer Inst Monogr 2013 Dec;2013(47):140-146. [doi: 10.1093/jncimonographs/lgt026] [Medline: 24395983]
106. Cho H, Silver N, Na K, Adams D, Luong KT, Song C. Visual cancer communication on social media: an examination of content and effects of #melanomasucks. J Med Internet Res 2018 Dec 5;20(9):e10501 [FREE Full text] [doi: 10.2196/10501] [Medline: 30185403]
107. 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 Dec 22;20(6):e226 [FREE Full text] [doi: 10.2196/jmir.9355] [Medline: 29934290]

Abbreviations

API: application programming interface
BD: bowel disease

XSL•FO
RenderX

DOID: Human Disease Ontology
IBD: inflammatory bowel disease
NCIT: National Cancer Institute Thesaurus
TM: text mining
UTC: Coordinated Universal Time
VADER: Valence Aware Dictionary and sEntiment Reasoner

Edited by G Eysenbach; submitted 30.10.18; peer-reviewed by J Groshek, Z He; comments to author 13.12.18; revised version received 23.01.19; accepted 26.04.19; published 15.08.19.

Please cite as:
Pérez-Pérez M, Pérez-Rodríguez G, Fdez-Riverola F, Lourenço A
Using Twitter to Understand the Human Bowel Disease Community: Exploratory Analysis of Key Topics
J Med Internet Res 2019;21(8):e12610
URL: http://www.jmir.org/2019/8/e12610/
doi:10.2196/12610
PMID:31411142

©Martín Pérez-Pérez, Gael Pérez-Rodríguez, Florentino Fdez-Riverola, Anália Lourenço. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 15.08.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Stroke Survivors on Twitter: Sentiment and Topic Analysis From a Gender Perspective

Alejandro Garcia-Rudolph[1,2,3], PhD; Sara Laxe[1,2,3], PhD; Joan Saurí[1,2,3], PhD; Montserrat Bernabeu Guitart[1,2,3], MD

[1]Institut Guttmann Hospital de Neurorehabilitacio, Badalona, Spain

[2]Universitat Autònoma de Barcelona, Bellaterra (Cerdanyola del Vallès), Spain

[3]Fundació Institut d'Investigació en Ciències de la Salut Germans Trias i Pujol, Badalona, Spain

Corresponding Author:
Alejandro Garcia-Rudolph, PhD
Institut Guttmann Hospital de Neurorehabilitacio
Camí de Can Ruti, s / n
Badalona, 08916
Spain
Phone: 34 934977700
Email: alejandropablogarcia@gmail.com

Abstract

Background: Stroke is the worldwide leading cause of long-term disabilities. Women experience more activity limitations, worse health-related quality of life, and more poststroke depression than men. Twitter is increasingly used by individuals to broadcast their day-to-day happenings, providing unobtrusive access to samples of spontaneously expressed opinions on all types of topics and emotions.

Objective: This study aimed to consider the raw frequencies of words in the collection of tweets posted by a sample of stroke survivors and to compare the posts by gender of the survivor for 8 basic emotions (anger, fear, anticipation, surprise, joy, sadness, trust and disgust); determine the proportion of each emotion in the collection of tweets and statistically compare each of them by gender of the survivor; extract the main topics (represented as sets of words) that occur in the collection of tweets, relative to each gender; and assign happiness scores to tweets and topics (using a well-established tool) and compare them by gender of the survivor.

Methods: We performed sentiment analysis based on a state-of-the-art lexicon (National Research Council) with *syuzhet* R package. The emotion scores for men and women were first subjected to an F-test and then to a Wilcoxon rank sum test. We extended the emotional analysis, assigning happiness scores with the hedonometer (a tool specifically designed considering Twitter inputs). We calculated daily happiness average scores for all tweets. We created a term map for an exploratory clustering analysis using VosViewer software. We performed structural topic modelling with *stm* R package, allowing us to identify main topics by gender. We assigned happiness scores to all the words defining the main identified topics and compared them by gender.

Results: We analyzed 800,424 tweets posted from August 1, 2007 to December 1, 2018, by 479 stroke survivors: Women (n=244) posted 396,898 tweets, and men (n=235) posted 403,526 tweets. The stroke survivor condition and gender as well as membership in at least 3 stroke-specific Twitter lists of active users were manually verified for all 479 participants. Their total number of tweets since 2007 was 5,257,433; therefore, we analyzed the most recent 15.2% of all their tweets. Positive emotions (anticipation, trust, and joy) were significantly higher (P<.001) in women, while negative emotions (disgust, fear, and sadness) were significantly higher (P<.001) in men in the analysis of raw frequencies and proportion of emotions. Happiness mean scores throughout the considered period show higher levels of happiness in women. We calculated the top 20 topics (with percentages and CIs) more likely addressed by gender and found that women's topics show higher levels of happiness scores.

Conclusions: We applied two different approaches—the Plutchik model and hedonometer tool—to a sample of stroke survivors' tweets. We conclude that women express positive emotions and happiness much more than men.

(J Med Internet Res 2019;21(8):e14077) doi:10.2196/14077

KEYWORDS

stroke; emotions; Twitter; infodemiology; infoveillance; sentiment analysis; topic models; gender

XSL•FO
RenderX

Introduction

General Background

Tweets can contain information about the mood of their authors. Even when users are not specifically posting about their personal emotive status, the message can reflect their mood. As such, tweets are regarded as microscopic instantiations of emotions. Twitter has been extensively analyzed for health-related conditions. Nevertheless, to the best of our knowledge, no study has been conducted in chronic stroke, with a focus on the emotional aspects and topics addressed by stroke survivors.

Stroke in Young Adults

Stroke is the third leading cause of long-term disability and one of the leading causes of depression worldwide [1]. Evidence suggests that stroke incidence in young adults is increasing in high-income countries [2]. It has been recently reported that ischemic stroke is no longer a disease affecting just elderly people, and an estimated 3.6 million young people (age<55 years) are affected each year [3]. The burden of stroke in young people may be increasing further, since multiple recent studies have reported increasing incidence of ischemic strokes, particularly at younger ages, while the incidence at older ages has been declining during the same period [4].

Globally, almost half of the entire stroke burden is on young individuals, as they have a greater likelihood to survive strokes, with long life spans ahead, and because strokes occur at younger ages in low- and middle-income countries [5]. Moreover, the overall population burden of cerebrovascular disease in young people may be underestimated, since clinically silent infarcts and white-matter changes are prevalent even in young stroke patients [6].

About one-fourth of ischemic strokes occur in working-aged individuals in high-income countries, with the incidence increasing worldwide in this age group from the 1980s to present [3].

Gender Differences in Stroke Outcomes

After experiencing a stroke, women experience more activity limitations, worse health-related quality of life (HRQoL), and more poststroke depression than men, as recently reported in an updated systematic review of sex differences [7].

Recent research published in January 2019 in the European Journal of Neurology reported that women are twice as likely to suffer from severe depression following a stroke than men. Ayis and colleagues [8] followed the progression of symptoms over 5 years after stroke onset in 2313 people (1275 men and 1038 women) from the South London Stroke Register and found that 20% of women suffered from severe depression compared to 10% of men [8].

The higher prevalence of depression among women may reflect a higher prevalence in the general population, where depression was identified as the leading cause of disease burden in women worldwide [9].

#Stroke

The expansion of social media has changed the way in which patients, physicians, and other health care stakeholders interact [10]. Twitter has led to the development of disease-specific communities that can categorize and aggregate their interactions using “hashtags.” These Twitter communities serve as readily accessible, no-cost platforms that provide significant educational and professional benefits.

Within stroke medicine, social media, specifically, Twitter has been recently highlighted for its potential to benefit patients, stroke organizations, and medical education [11].

The stroke-related Twitter network has been recently studied [12], through 621,653 tweets containing the #Stroke hashtag from March 20, 2012, to January 31, 2018, in relation to tweet content, activity metrics, engagement, and user characteristics. The most commonly discussed topics were prevention, diabetes, atrial fibrillation, aphasia, dementia, thrombectomy devices, thrombolysis, and tobacco. Specifically, the content of discussions included recognition of the signs of a stroke, associated risk factors (eg, atrial fibrillation, heart disease, and diabetes), and findings of peer-reviewed journals regarding stroke treatment. Tweets were mainly composed by advocacy/support organizations (21.5%), physicians (8.4%), individuals not known to be directly working in the health care industry (14.0%), other health care professionals (5.5%), organizations related to research/academia (2.3%), and academics (2.2%), while stroke patients contributed to 6.7% of tweets (n=41,822). There was a similar proportion of total tweets with the #Stroke hashtag generated by physicians (8.4%) and patients (6.7%) during the study period and apparent minimal network communication between physicians and patients, as reported in the study conclusions [12].

Emotional Distress in the Adjustment Process for Stroke Survivors

Brennan emphasizes the importance of assumptions in adapting to the world around us. According to Brennan’s model, we each have a cognitive map or representation of the world, resulting from our social and cultural context and the accumulation of our life experience. This highly complex “assumptive world” is biologically adaptive in that it allows us to anticipate and plan for the future [13].

In the case of a typical stroke patient, their assumptive world will almost always be challenged or disconfirmed by the experience of stroke and its immediate repercussions [14]. As Brennan states, “adjusting core assumptions involves huge amounts of cognitive processing and emotional distress, this often leads to acute emotional difficulties, such as feelings of confusion, loss, sadness and anger.”

Moreover, the experience of stroke and disability may also confirm previously held negative beliefs for some individuals (eg, “I am worthless” or “Others see me as weak”) and may lead to emotional distress in this manner [14].

XSL•FO
RenderX

Spontaneous, Emotional Language, and Everyday Topic Discussions on Twitter

Over the last few years, Twitter has become a notable data source in sociolinguistics, as it captures opinions and sentiments on a wide range of topics. Although Twitter users are a self-selected group, it has been argued that analyses of Twitter data produce results congruent with those obtained using standard research methods and data sources [15].

Considering the frequent use of emotional language in tweets that relate to everyday experiences [16], for a large proportion of the population, Twitter provides unobtrusive access to time-sensitive and ecologically valid samples of spontaneously expressed emotions [17].

Sentiment analysis in the health care setting is not a new phenomenon, for example, in previous research, greater positive sentiment within discharge summaries was associated with a significantly decreased risk of readmission [18].

This Study

In the following subsections, we describe the specific characteristics and objectives of our study.

Twitter Lists

Previous studies have shown that topical experts are often the primary drivers of interesting discussions on Twitter [19]. In contrast to random sampling for gathering Twitter data, alternative sampling methods have been put forward; one of them proposed to retrieve content only from topical experts, that is, Twitter users whose followers consider them to be knowledgeable on some topic, to reduce the number of unwanted tweets in the sampled data while still gathering useful tweets related to a specific topic. The key challenge, however, lies in identifying a good set of experts [20].

Twitter users can organize the accounts that they follow into Twitter user lists. These lists are used in a variety of ways. In some cases, they may correspond to personal lists of a given user's friends and families, but frequently, lists are employed to group together Twitter accounts based on a common topic or theme. In this way, every Twitter user can effectively become a community curator. Therefore, previous research has proposed that we consider a Twitter user a "topical expert" if the user belongs to several lists on a particular topic [20].

In our study, we propose to take advantage of user lists in the field of stroke. To the best of our knowledge, lists have not been used in studies related to chronic health conditions.

Plutchik's Human Emotions

Currently, there is no single accepted psychological theory of basic human emotions; nevertheless, there is an agreement that a simple positive-negative dichotomy is not enough to capture the full range of emotions [21].

In this work, we use the Plutchik [22] approach, which postulates the following eight basic human emotions: joy, sadness, anger, fear, trust, disgust, anticipation, and surprise. There have been extensive applications of this approach, for example, the National Research Council (NRC) Word-Emotion Association Lexicon, which contains 10,170 lexical items that are coded for Plutchik's basic human emotions [23], and has been applied in several sentiment analysis studies [24].

Plutchik's categories also have the advantage of providing a balanced list of positive (trust, joy, anger, and anticipation) and negative (disgust, sadness, fear, and surprise) emotions, which, to the best of our knowledge, have not been applied in chronic conditions, in general, or stroke, in particular.

Hedonometer

After performing emotional analysis based on Putchik's model, we propose another point of view, by assigning happiness scores to tweets with the hedonometer tool. The hedonometer [25] was developed from Twitter, Google Books, music lyrics, and the New York Times for measuring expressed happiness—positive and negative sentiment—in large-scale text corpora. Since its development, the hedonometer has been applied to studies on predictive markers of depression on Instagram [26] or the climate change sentiment on Twitter [27]. The hedonometer calculates a happiness score based on the happiness of the individual words used in the text. A total of 10,222 of the most frequently used English words in four disparate corpora were given happiness ratings using Amazon's Mechanical Turk online marketplace.

Adding Covariate Information With Structural Topic Models

Although Latent Dirichlet Allocation (LDA) is, perhaps, the most common form of topic modeling, a number of associated techniques now exist, including dynamic topic models, correlated topic models, and hierarchical topic models. One of the most increasingly popular techniques to emerge in recent years, however, is structural topic modeling (STM). STM provides a flexible way to incorporate "metadata" associated with the text, such as when the text was written, where (eg, which country) it was written, who wrote it, and characteristics of the author, into the analysis using document-level covariates. In turn, it allows analysis of relationships between metadata and topics in the text corpus.

Study Objectives

As Brennan states [13], the adjusting process involves huge amounts of emotional distress. This often leads to acute emotional difficulties such as feelings of confusion, loss, sadness, and anger. Considering that women experience more activity limitations, worse health-related quality of life, and more poststroke depression, in this study, we propose to take advantage of unobtrusive access to samples of spontaneously expressed emotions and opinions provided by Twitter and to analyze them from a gender perspective using two different, well-established approaches (Plutchik model and the hedonometer tool), with the following specific aims:

- To compare tweets by gender of stroke survivor for the 8 basic emotions (anger, fear, anticipation, surprise, joy, sadness, trust, and disgust) while considering the raw frequencies of words in the collection of tweets posted by a stroke survivors' sample.
- To determine the proportion of each emotion in the collection of tweets and statistically compare each of them

XSL•FO
RenderX

by gender. This measurement thus allows us to track the proportion of each emotion for each individual tweet and is less affected by single outliers.
- To extract the main topics (represented as sets of words) that occur in the collection of tweets, related to each gender.
- To assign happiness scores to tweets and topics (using the hedonometer) and compare them by gender.

Methods

Data Collection

We considered the network analysis from previous research [12] (see #Stroke in the Introduction) as the starting point. Node size is related to user influence, which is directly correlated to the amount a user is mentioned. The top identified nodes and their corresponding number of followers are as follows: @TheStrokeAssoc (102 million), @signagnststroke (68 million), @StrokeHope (93.8 million), @PeterCoghlan1 (7.2 million), @strokefdn (11.5 million), @StrokeAssocNW (5.5 million), @StrokeAHA_ASA (10.6 million), and @HeartandStroke (45.3 million), @HeartandStroke (45.3 million).

Twitter data collection was performed using the *rtweet* R package [28] via Twitter's REST (representational state transfer) and stream application program interfaces (APIs). We initially applied the *lists_users* function to obtain all lists that the top nodes subscribe to, including their own. Subsequently, we used the *lists_members* function to obtain Twitter list members (users on a given list). To retain a list member, we imposed the condition that it should appear in at least 3 different lists.

For each identified user, we retrieved tweets with the get_timelines() function (it retrieves the most recent 3200 tweets for each Twitter user, without any time restriction). We collected 1,300,845 tweets from a thousand users and further classified them in tweets from particulars (woman or man) and from organizations (institutions and associations); the last step was to collect only tweets from particulars where the gender could be clearly determined, as explained in the next section.

Participant Selection Process

We modeled our data collection methods on prior studies that have used the Twitter platform for generating a convenience sample of users with publicly available accounts, who self-identify as stroke survivors in their profile or tweets.

We then confirmed the self-reported stroke diagnosis by having one researcher generate this initial list of Twitter users and a second researcher check the details for each Twitter user on the list to ensure correct identification of stroke survivors users.

We then employed a stepwise process for coding each Twitter user's gender as male, female, or unknown/insufficient data. Two researchers independently used these codes, beginning with each Twitter user's username, followed by profile name, profile description, profile photo, and tweets. Both researchers then reviewed their final gender codes for each Twitter user to ensure consistency and resolve disagreements.

Data Cleaning

The final sample was prepared for analysis by using the *quanteda* R package. This included the process of basic normalization (eg, remove punctuation and lowercase all text), stop word removal (eg, the words "a" and "the"), normalization of Twitter user mentions (eg, "@janedoe" is converted to "@user"), lemmatization (eg, "dog," "dogs," and "dog's" are all converted to "dog"), and nonprintable character removal (eg, emojis). All analyses relied on public, anonymized data; adhered to the terms and conditions, terms of use, and privacy policies of Twitter; and were performed under Institutional Review Board approval from the authors' institution.

We do not report any specific tweets that could be used to identify the original Twitter user who posted the content online, as this is an important concern that has been discussed extensively in recent literature on the ethics of using Twitter data for research [29].

Sentiment Analysis

We calculated the overall frequencies of emotion words for each Plutchik category for each user (and therefore gender) by using the *syuzhet* R package [30]. The NRC Word-Emotion Association Lexicon is available via open access and has been implemented in the get_nrc_sentiment() function of the *syuzhet* R package. Finally, the data were subjected to statistical analyses: For each tweet, given an emotion X, an emotion proportion score was calculated as:

$$proportion_X = frequency\ of\ words\ with\ emotion\ X\ in\ a\ tweet\ /\ (frequency\ of\ negative\ words\ in\ a\ tweet + frequency\ of\ positive\ words\ in\ a\ tweet)\ (equation\ 1)$$

The emotion proportion scores for men and women were then subjected to a Wilcoxon rank sum test in R, since the *F*-test had revealed that the two distributions did not meet the criterion of variance homogeneity [31].

Structural Topic Models

Considering the final sample of tweets from the data cleaning phase presented above as the starting point, we proceeded as follows:

1. Convert cleaned tweets to tm corpus and create a term document matrix (TDM) using the *tm* R Package [32].
2. Calculate the term frequency inverse document frequency (TF-IDF) for all the words in TDM.
3. Exclude all the words with TF-IDF≤0.1 to remove all the words that are less frequent.
4. Calculate the optimal number of topics (K) in the corpus using the log-likelihood method for the calculated TDM using Gibbs sampling and exploring different metrics: "Griffiths2004," "CaoJuan2009," "Arun2010," and "Deveaud2014" using the *FindTopicsNumber* function from the *ldatuning* R package [33].
5. Apply the spectral method using the *stm* package to discover topics.
6. Topic validation (semantic coherence and exclusivity).
7. Visualization and interpretation of results from the calculated model.

XSL•FO
RenderX

A unique feature of STM, implemented by the *stm* R package [34], is that it can model how the document level covariates affect the topical prevalence parameter μ with a generalized linear model. As mentioned in the Sentiment Analysis section above, our covariate is the gender factor with two levels ("Woman" and "Man").

Besides the inclusion of the gender covariate, the *stm* R package supports the explicit estimation of correlations among topics. This feature provides further information on the corpus structure. Correlations are estimated by replacing the Dirichlet distribution in the standard LDA framework with a logistic normal distribution as in the Correlated Topic Model [35].

This allows us to identify when two topics are likely to cooccur within tweets (here, we focus on both positive and negative correlations, which are also useful to identify gender differences).

Hedonometer

We applied the hedonometer tool to all tweets and to the main identified topics as follows: For each word in each tweet, we obtained a happiness score, calculated the mean happiness score for each day, and plotted it by date grouping by gender; STM allows us to identify the main topics and label the topics as "More likely Women" and "More likely Men." As each topic is defined by a set of words, we obtained the happiness score of each word using the hedonometer, and therefore, we are able to compare topics according to their happiness score. This also allows us to select, for example, the top 25 words with the highest levels of happiness and identify if such words belong to female or male topics.

Results

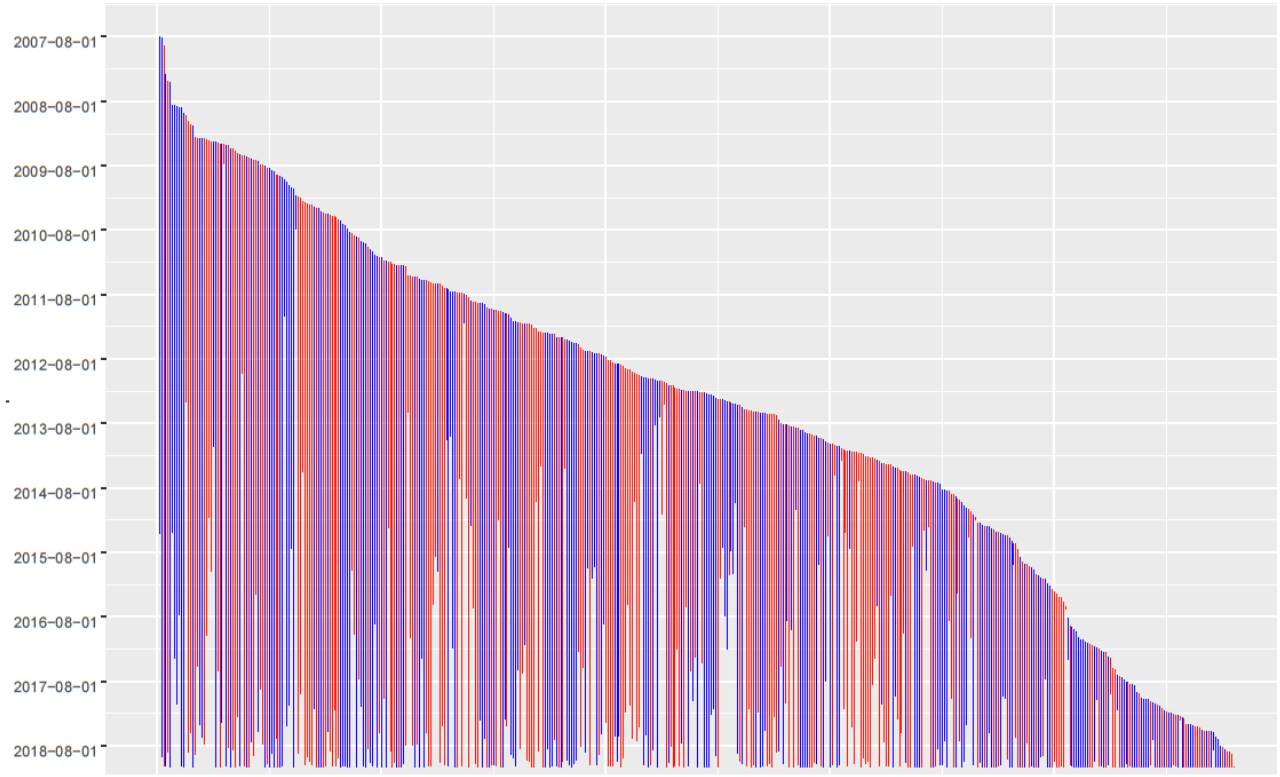
Sample Description

After the selection process, a final sample of 479 Twitter users who posted 800,424 tweets between August 1, 2007, and December 1, 2018, were selected. Women (n=244) posted a total of 3,788,069 tweets; from them, we collected 396,898 tweets (the most recent ones, up to December 2018), and the mean number of tweets posted by our selected sample was 1620. In addition, 54% of the selected sample posted more than 1000 tweets and 71% posted more than 500 tweets. The total number of followers of the selected sample was 182,807.

Men (n=235) posted a total of 1,469,364 tweets, from which we collected 403,526 tweets (the most recent ones, up to December 2018), and the mean number of tweets posted by our selected sample was 1717. In addition, 59% of the selected sample posted more than 1000 tweets and 73% posted more than 500 tweets. The total number of followers of the selected sample was 255,053.

Figure 1 shows the date of the first and last posted tweets for each selected participant included in our sample (women in red, men in blue; same code colors throughout the analysis). Each vertical line in the plot represents a participant whose first tweet was posted at the top of the vertical line and last tweet was posted at the bottom of it. We ordered participants in the plot from left to right, where the earliest date of the first tweet is shown leftmost for each participant. For example, the leftmost participant is a man whose first tweet was posted in 2007 and last tweet was posted in 2014.

Figure 1. Topics and gender covariate obtained with spectral structural topic modeling.



XSL•FO
RenderX

In Multimedia Appendix 1, we present the number of tweets posted by year; a larger number of tweets was posted in 2018 (about 300,000) and the other 500,000 posts were distributed with growing tendency since 2007, as presented in previous research (described in the #Stroke section).

For each of the 479 participants, we reviewed their profiles to verify their geographic locations, obtained by means of the *rtweet* library. We were able to identify the geographic locations of 378 of the 479 users (78.91%).

In Multimedia Appendix 1, we present the total number of users (N=378) by country, showing that most of the users are from four countries: 95% are from Australia, Canada, the United Kingdom, or the United States.

United States had the most users (206/378; 55%). The United Kingdom had 113 users (29.89%). As such, both countries together accounted for more than 85% of the participants.

In Multimedia Appendix 1, we present wordclouds of the top 500 words in all participants' profile description. Most words are repeated in both wordclouds, but some distinctive characteristics can be observed (women clearly refer to Music, Live, and Time, while men do not).

Sentiment Analysis

The NRC Word-Emotion Association Lexicon, which contains 10,170 lexical items that are coded for Plutchik's basic human emotions [23] and implemented in the *syuzhet* R package, associates an emotion (or more than one emotion) to each of the 10,170 lexical items. Given a word and emotion X, the NRC Word-Emotion Association Lexicon associates a score (range: 0 to 1) with it. A score of 1 indicates that the word conveys the highest amount of emotion X. A score of 0 indicates that the word conveys the lowest amount of emotion X.

We then identified (via the get_nrc_sentiment() function) the number of words that, according to the NRC, express positive or negative sentiment as well as one (or more than one) of Plutchik's eight basic emotions.

Table 1 summarizes the raw number of words (and their percentages) obtained with the get_nrc_sentiment() function of the *syuzhet* R package.

Among both men and women, the most frequent emotions were trust, anticipation, and joy (top 3), as shown in Figure 2.

Women used considerably more words from all positive categories (except anger), and men used more words in all negative categories (except surprise), as shown in Table 1.

When considering negative or positive words, women used 12% of negative words, while men used 13.6% of negative words. In contrast, women used 21.8% of positive words, while men used 20.5% of positive words. Positive and negative labels for words are also obtained from the NRC lexicon using the get_nrc_sentiment() function of the *syuzhet* R package.

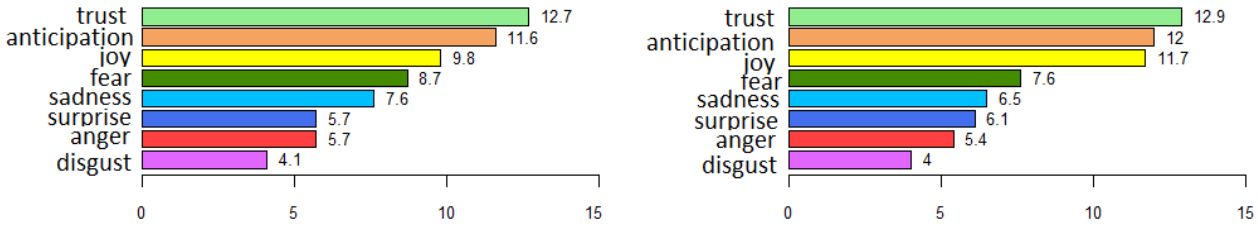
Table 1. Raw frequencies of words identified for each emotion.

Emotion	Men, n (%)	Women, n (%)
Anger	76,650 (5.7)	74,858 (5.4)
Anticipation	155,608 (11.6)	166,150 (12.0)
Disgust	55,512 (4.1)	54,785 (4.0)
Fear	117,221 (8.7)	104,826 (7.6)
Joy	131,243 (9.8)	161,933 (11.7)
Sadness	101,475 (7.6)	89,868 (6.5)
Surprise	77,109 (5.7)	83,663 (6.1)
Trust	170,0176 (12.7)	178,718 (12.9)
Negative	182,288 (13.6)	166,000 (12.0)
Positive	276,124 (20.5)	300,751 (21.8)

[a]Not applicable because .

[b]N/A: not applicable.

Figure 2. Ranking of emotions (in percentage of the total words. Men: left; women: right). Each bar represents the percentage of total words presented in Table 1.



XSL•FO
RenderX

We then calculated the emotion proportion score for each emotion X, as shown in equation 1 in the Methods section.

Table 2 reports statistical comparisons; for example, for the global positive emotion, women (median=100%, mean=65.57%) used considerably more positive words than men (median=66.67%, mean=60.73%). Since the *F*-test indicated that the two distributions have a significantly different variance ($F_{237040,242190}$=1.0468, *P*<.001), they were subjected to a Wilcoxon rank sum test. This test showed that the difference between men and women is highly statistically significant (W=2.6817e+10, *P*<.001). Similar results are shown in Table 2 for global negative emotion: Men used considerably more negative words than women; in addition, each individual positive emotion (joy, anticipation, and trust, except surprise) was favorable to women and each individual negative emotion (fear, sadness, and disgust) was preferred by men.

Global negative-positive proportion comparisons are presented in Figure 3. Women used considerably fewer negative words and more positive words than men (shown at the top and bottom of Figure 3, respectively)

Plutchik's eight emotions are subdivided into four complementary pairs, namely, joy–sadness, anticipation–surprise, trust–disgust, and anger–fear [23].

Table 2. Statistical comparison of words identified for each emotion.

Emotion, participants	Median	Mean	*F (df)*	*P* value	W	*P* value
Joy			0.9030 (237040,242190)	<.001	2.63e+10	<.001
Men	0	0.2972	—	—	—	—
Women	0	0.3611	—	—	—	—
Negative			1.0468 (237040,242190)	<.001	3.05e+10	<.001
Men	0.3333	0.3927	—	—	—	—
Women	0	0.3443	—	—	—	—
Fear			1.1183 (237040,242190)	<.001	3.01e+10	<.001
Men	0	0.2481	—	—	—	—
Women	0	0.2121	—	—	—	—
Positive			1.0468 (237040,242190)	<.001	2.68e+10	<.001
Men	0.6667	0.6073	—	—	—	—
Women	1.0000	0.6557	—	—	—	—
Sadness			1.0261 (237040,242190)	<.001	2.45e+10	<.001
Men	0	0.2204	—	—	—	—
Women	0	0.1868	—	—	—	—
Anger			1.0399 (237040,242190)	<.001	2.90e+10	<.001
Men	0	0.1573	—	—	—	—
Women	0	0.1495	—	—	—	—
Anticipation			0.9837 (237040,242190)	<.001	2.79e+10	<.001
Men	0	0.3299	—	—	—	—
Women	0	0.3488	—	—	—	—
Surprise			1 (237040,242190)	>.99	N/A[b]	N/A
Men	0	0.1672	—	—	—	—
Women	0	0.1779	—	—	—	—
Trust			1.0134 (237040,242190)	<.001	2.81e+10	<.001
Men	0.1667	0.3628	—	—	—	—
Women	0.2500	0.3755	—	—	—	—
Disgust			1.0559 (237040,242190)	<.001	2.88e+10	<.001
Men	0	0.1153	—	—	—	—
Women	0	0.1109	—	—	—	—

XSL•FO
RenderX

Figure 3. Topics and gender covariate obtained with spectral structural topic modeling.

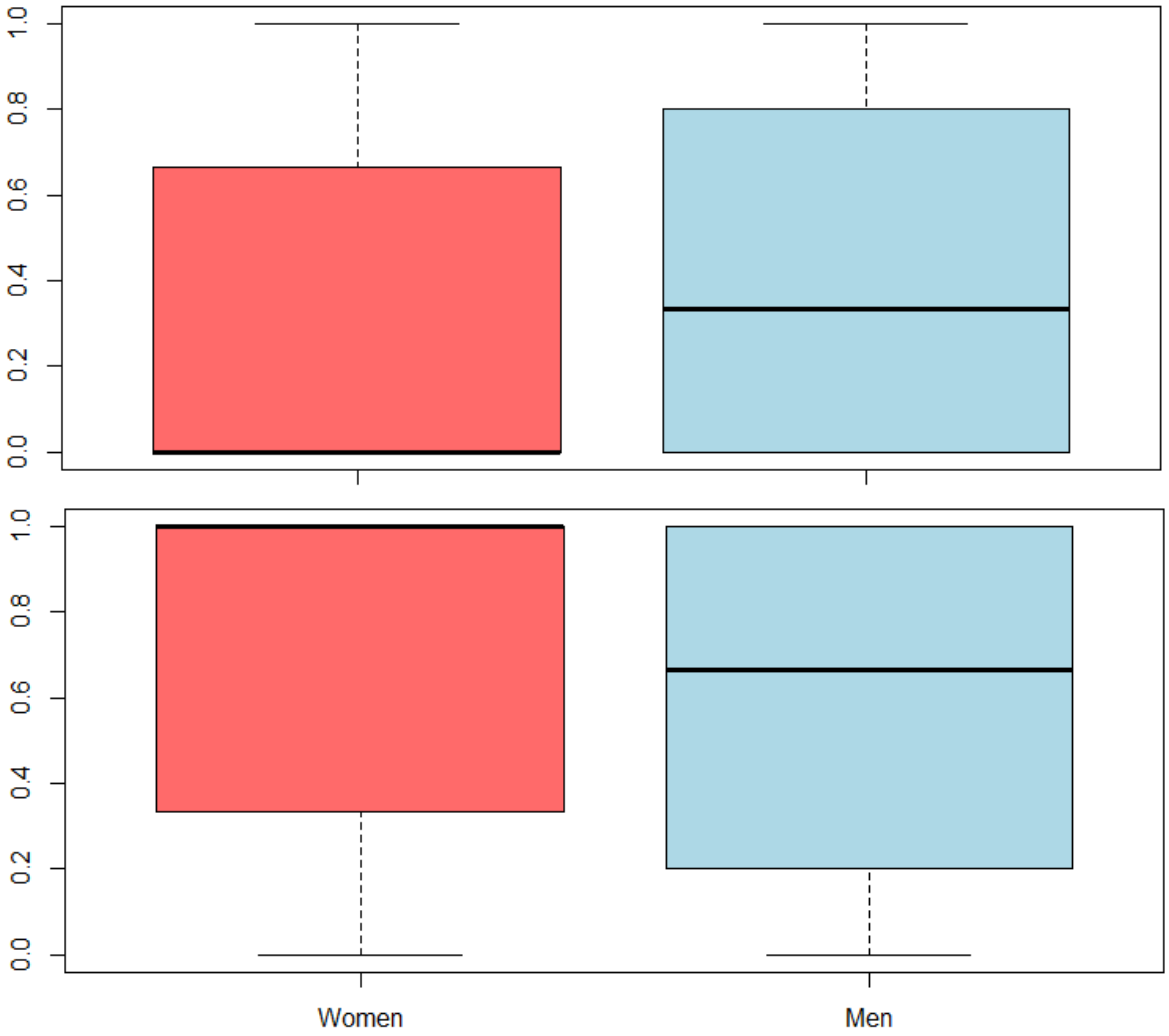
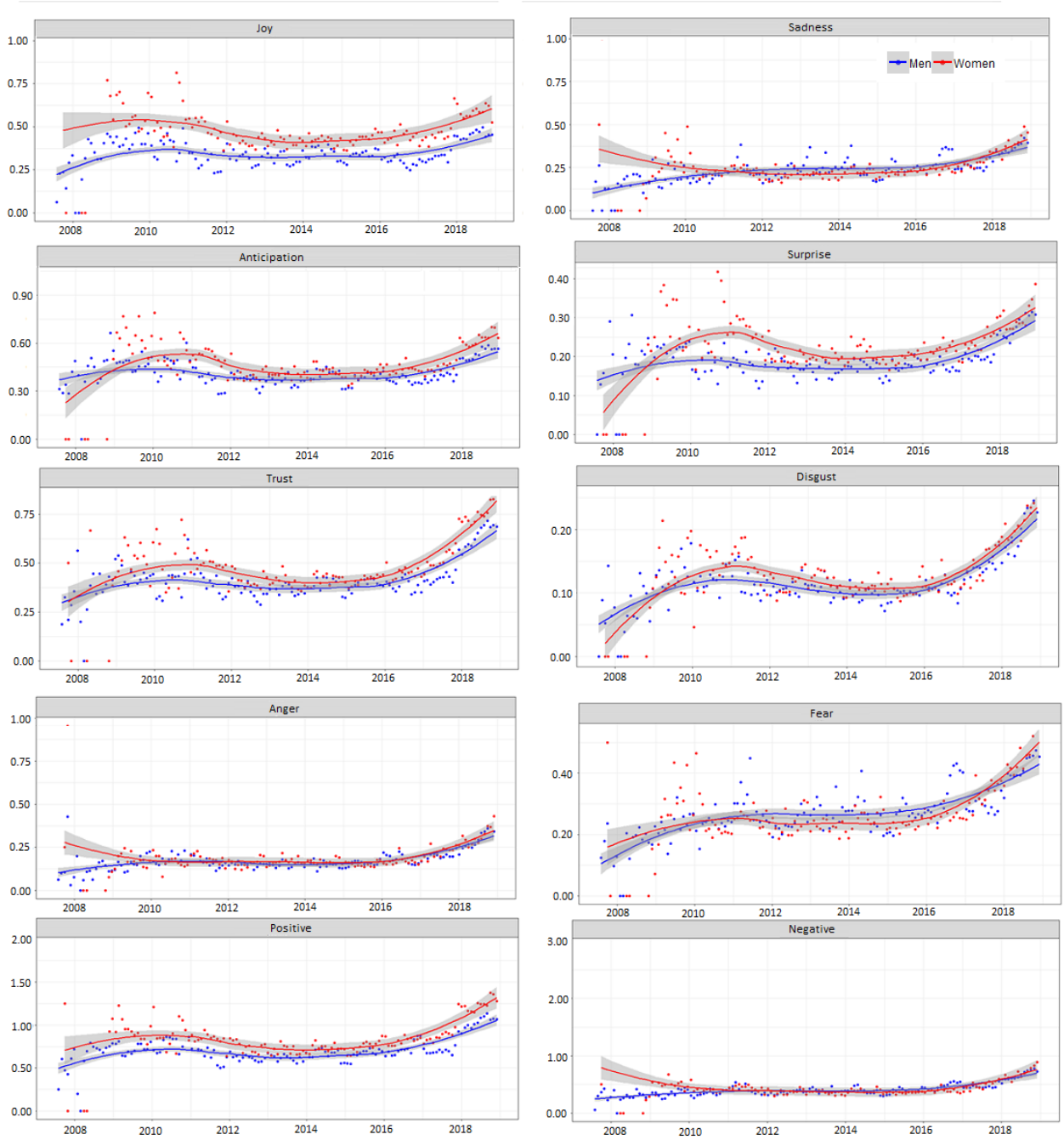


Figure 4 plots such scores for each emotion summarized monthly along all the time periods in the study of emotion words for each pair of emotions, obtained with *syuzhet* R package and plotted with the *ggplot2* R package. It clearly shows higher scores for women in positive emotions along time and lower scores for men in almost every emotions throughout the considered period.

As shown in Figure 4, joy and global positive words clearly present higher values for women throughout the considered period.

XSL•FO
RenderX

Figure 4. Monthly mean scores for NRC emotions in the 2008-2018 period.



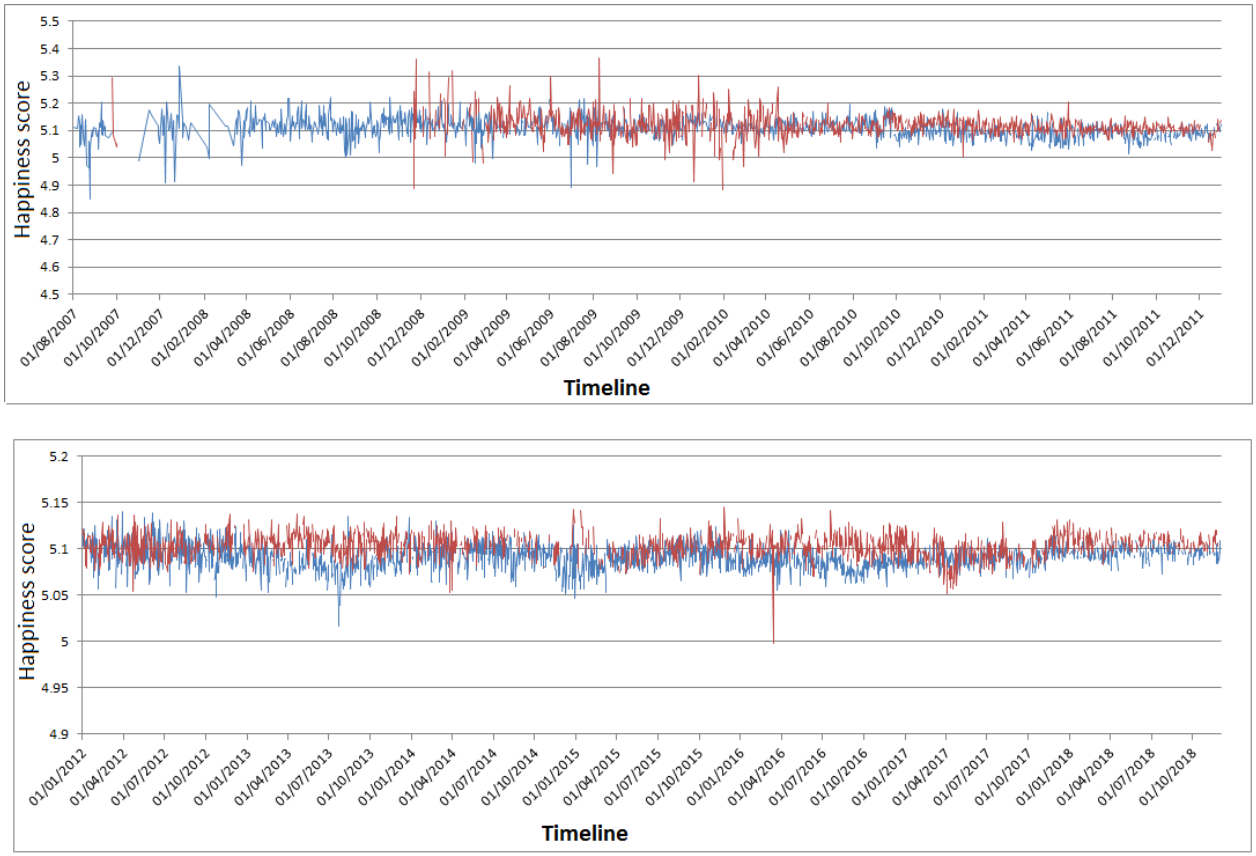
Assigning Happiness Scores With the Hedonometer

We then calculated the happiness score using the hedonometer for each word in each tweet, summarized the mean happiness score for each day during the whole period under study, and plotted it by date, grouping by gender.

As shown in Figure 5, happiness ratings obtained by hedonometer summarized on a daily basis for each user are also higher for women than for men, almost throughout the considered time period, with remarkable differences in favor of women, for example, in the 2013-2014 period, 2016, and 2018.

XSL•FO
RenderX

Figure 5. Topics and gender covariate obtained with spectral structural topic modeling.



Structured Topic Modelling

Before the application of STM, we performed an exploratory cluster analysis using VosViewer [36]. As defined by VosViewer, a term map is a two-dimensional representation, in which strongly related terms are located close to each other and less strongly related terms are located further away from each other. Each point in a term map has a color that depends on the density of items at that point. It is argued that the VOS mapping technique yields more satisfactory term maps than popular multidimensional scaling–based approaches to bibliometric mapping. Maps constructed using these multidimensional scaling–based approaches are shown to suffer from certain artifacts. Maps constructed using the VOS mapping technique do not have this problem, as reported by Waltman et al [36]. Details are presented in Multimedia Appendix 1 (VosViewer Cluster Analysis).

We tested different parameter configurations to increase intercluster distances and reduce intracluster distances. VosViewer allowed us to identify seven clusters for men and five clusters for women (Multimedia Appendix 1). In the obtained clusters for the most relevant 250 words for men and women, we highlighted words that are common to clusters obtained by men and women. Unfortunately, this is the case for most of the words; therefore, it did not allow us to visually identify gender differences.

Nevertheless, in Multimedia Appendix 1, we present the clusters for the words that are not common to both men and women, and we applied the hedonometer to each of them; the happiness scores are shown in brackets for each word.

We summarized the happiness scores and obtained mean happiness scores of 5.31 (SD 1.31) for all the words present only in men clusters and 6.25 (SD 1.07) for those present only in women clusters. We then considered each of the largest clusters separately and obtained the following for men: mean happiness score in cluster 1=5.52 (SD 0.99) and mean happiness score in cluster 2=5.06 (SD 1.25). For women, the mean happiness score was 6.30 (SD 1.12) in cluster 1, 5.70 (SD 1.13) in cluster 2, and 6.68 (SD 0.81) in cluster 3. Again, happiness scores of women were higher than those of men when considering the scores at the cluster level.

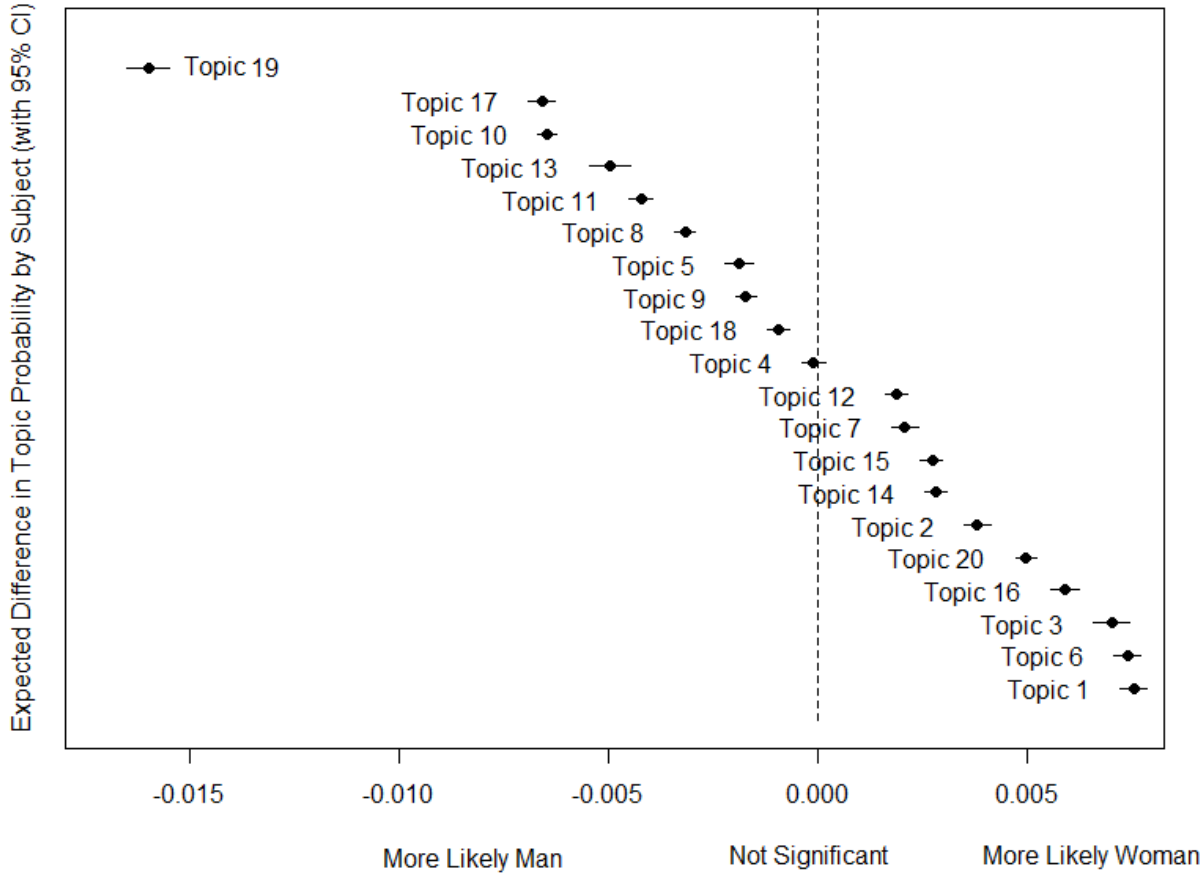
Before STM, we also performed LDA analysis for seven topics; the number of topics was determined as shown in Multimedia Appendix 1 by using different metrics and the FindTopicsNumber function from the *ldatuning* R package.

The obtained topics are presented in Multimedia Appendix 1, but as with cluster analysis, we could not identify topics clearly related to men or women.

Therefore, we applied STM to associate covariates (Gender) to the identified topics and plot results as presented in Figure 6. As with most topic models, the objective function maximized by STM is multimodal. Therefore, the way we choose the starting values for the variational EM algorithm can affect our final solution.

XSL•FO
RenderX

Figure 6. Topics and gender covariate obtained with spectral structural topic modeling.



We applied LDA initialization (the default option), which uses several passes of collapsed Gibbs sampling to initialize the algorithm.

Table 3 shows the top three topics selected for men and women, with 3 different types of word profiles: highest probability, frequency-exclusivity (FREX), and score values. Detailed descriptions of all identified topics are presented in Multimedia Appendix 1.

FREX measures exclusivity of the words to the topic in a way that balances it with word frequency. The score metric divides the log frequency of the word in the topic by the log frequency of the word in other topics, and highest probability considers words within each topic with the highest probability, inferred directly from topic-word distribution.

In Multimedia Appendix 1, we present the evaluation of the obtained topics. Semantic coherence is a criterion developed by Mimno and colleagues [37]; it is maximized when the most probable words in a given topic frequently cooccur together, and it has been shown that the semantic coherence metric correlates well with human judgment of topic quality [37].

Topics 19, 17, and 10 show all semantic coherence values to the right side of the plot, with topic 19 to the rightmost possible position and close to topic 10. Topics 1, 6, and 3 are in the central positions, while upper right side positions are six optimal selected topics showing acceptable values.

We then assigned happiness scores to topics with the hedonometer tool (we proceeded as was done with VosViewer clusters described in Multimedia Appendix 1). As shown in Figure 6, STM allows us to label the topics as “More likely Women” and “More likely Men.”

As presented in Table 3, each topic is defined as a list of 20 words. To assign happiness scores, we selected the words with the highest probability (first row of each topic in Table 3). Therefore, we applied the hedonometer tool to assign a happiness score to each word with the highest probability of each topic.

In the Multimedia Appendix 1, we present the complete list of all words defining each topic. For each word, we present the happiness score and its corresponding topic (and therefore the associated gender to the topic). We selected a subset of these words (Tables 4 and 5). We show the 25 words with the highest happiness scores and those with the lowest happiness scores along with the corresponding gender.

Table 4 shows that 19 of the 25 words with the highest happiness scores correspond to women’s topics and only 6 correspond to men’s topics. Table 5 shows the 25 words with the lowest happiness scores, and only 7 corresponded to women, while 18 corresponded to men.

XSL•FO
RenderX

Table 3. Top three identified topics and percentages for women (topics 1, 3,6) and men (topics 10,17,19).

Topic (%)	Highest probability	FREX[a]	Score
1 (5.23)	year, happy, tomorrow, open, birthday, take, come, busy, christmas, baby, sleep, friday, sunday, monday, list, bed, smile, market, treat, guess	merri, birthday, appl, ang, eve, awak, con, clay, decemb, angel, happi, est, syracus, ako, relax, closet, lang, store, carousel	happi, year, birthday, tomorrow, open, christma, sleep, friday, come, busi, babi, sunday, take, bed, store, monday, holiday, list, date, market
3 (7.21)	good, great, video, hope, night, morn, lol, tonight, long, done, head, weekend, fun, readi, celebrate, citi, movie, luck, earli, forget	playlist, chicken, grill, peter, egg, chees, movi, delici, potato, cooki, bbq, soup, recip, video, kitti, cup, chili, luck, pan, belli	good, video, hope, night, morn, great, lol, weekend, playlist, movi, tonight, don, luck, fun, long, sweet, forget, gonna, saturday, dinner
6 (5.73)	time, life, thing, world, god, famili, twitter, hear, power, hate, pass, speak, human, step, posit, bless, super, continu, messag, creat	god, lord, pray, faith, amen, bless, prayer, psalm, soul, negat, holi, heal, thank, charl, nchousingbuild, merci, evil, accomplish, yea, compass	time, god, thing, life, famili, twitter, world, lord, bless, hear, power, step, super, hate, pray, prayer, congrat, pop, faith, posit
10 (3.82)	end, heart, walk, news, stop, run, hand, pay, mile, rate, worth, success, dead, offer, singl, reach, staff, fail, snow, hero	mile, rate, bpm, attitud, anthem, bioness, hawk, failur, shoulder, flaw, casual, complic, tattoo, zombi, hero, pinterest, hand, virus, vancouv	heart, walk, end, news, mile, stop, run, rate, bpm, hand, pay, dead, success, attitud, hero, snow, worth, bbc, offer, reach
17 (5.25)	back, play, game, team, job, place, boy, man, point, won, lost, black, park, perfect, act, lose, john, footbal, film, player	yard, hole, playoff, player, joe, nfl, eagl, cowboy, bronx, kiss, dalla, theater, doodl, lewi, cunt, throw, golden, barn, korea, brave	game, team, play, back, boy, job, footbal, player, park, perfect, black, act, place, north, beat, test, film, lose, tour, kick
19 (6.25)	stroke, support, find, survivor, learn, lot, brain, care, health, patient, help, aware, money, raise, children, research, import, risk, experience, hospital	aware, raise, foundat, risk, research, patient, region, medic, lot, recoveri, donat, increas, factor, studi, resourc, rehab, treatment, cancer, rehabilit, learn	stroke, survivor, learn, lot, find, support, brain, patient, awar, care, rais, health, research, risk, foundat, injuri, studi, region, recoveri, disease

[a]FREX: frequency-exclusivity.

XSL•FO
RenderX

Table 4. Top 25 words with highest happiness scores, topics, and gender of participants.

Word	Participant	Score	Topic
Love	Women	8.42	T16
Happy	Women	8.3	T1
Win	Women	8.12	T15
Smile	Women	8.1	T1
Won	Men	8.1	T17
Music	Women	8.02	T2
Weekend	Women	8.0	T3
Celebrate	Women	7.98	T3
Christmas	Women	7.96	T1
Fun	Women	7.96	T3
Free	Men	7.96	T11
Great	Women	7.88	T3
Success	Men	7.86	T10
Award	Women	7.86	T15
Positive	Women	7.8	T6
Hero	Men	7.8	T10
Sun	Men	7.8	T11
Birthday	Women	7.78	T1
Winner	Women	7.78	T15
Beauty	Men	7.76	T5
Family	Women	7.72	T6
Gift	Women	7.72	T15
Brilliant	Women	7.68	T2
Super	Women	7.68	T6
Amazing	Women	7.66	T16

XSL•FO
RenderX

Table 5. Top 25 words with the lowest happiness scores, topics, and gender of participants.

Word	Participant	Score	Topic
Death	Men	1.54	T18
Kill	Women	1.56	T12
Die	Men	1.74	T18
Fail	Men	1.96	T10
Dead	Men	2.0	T10
Pain	Men	2.1	T4
Hell	Men	2.22	T9
Poor	Men	2.32	T9
Hate	Women	2.34	T6
Sad	Women	2.38	T12
Attack	Men	2.42	T8
Shot	Women	2.5	T2
Shit	Men	2.5	T18
Aphasia	Men	2.58	T11
Stroke	Men	2.58	T19
Lie	Men	2.6	T13
Bad	Women	2.64	T16
Fight	Women	2.7	T16
Lost	Men	2.76	T17
Lose	Men	2.76	T17
Disabled	Men	2.82	T18
Problem	Men	2.98	T4
Wrong	Men	3.14	T18
Forget	Women	3.22	T3
Cut	Men	3.42	T9

We then calculated the boxplots of the happiness scores for each topic (Figure 7), ordered from “More Likely Men” to “More Likely Women”; the means and regression line are shown in red circles and a red line, respectively (P<.001).

Figure 7 shows higher happiness scores from Topic 4 to the right (ie, women’s topics) with the exception of Topic 16, which contains several words with low happiness scores (eg, “bad” or “lone”; Multimedia Appendix 1). The regression line shows a positive slope in the direction of women’s topics (P<.001).

We then compared happiness scores by pairs from the leftmost and rightmost topics in Figure 6 to the center (Topic 19-Topic 1, Topic 17-Topic 6, Topic 10-Topic 3, etc). We found significant differences in favor of women in 4 of the 10 pairs of topics (and none in favor of men) when comparing the happiness scores by pairs of topics (Figure 8; men blue, women red). The complete list of comparisons is presented in Multimedia Appendix 1.

STM also permits correlations between topics. Positive correlations indicate that both topics are likely to be discussed in a tweet. In Figure 9, we plot both positive and negative correlations for all identified topics.

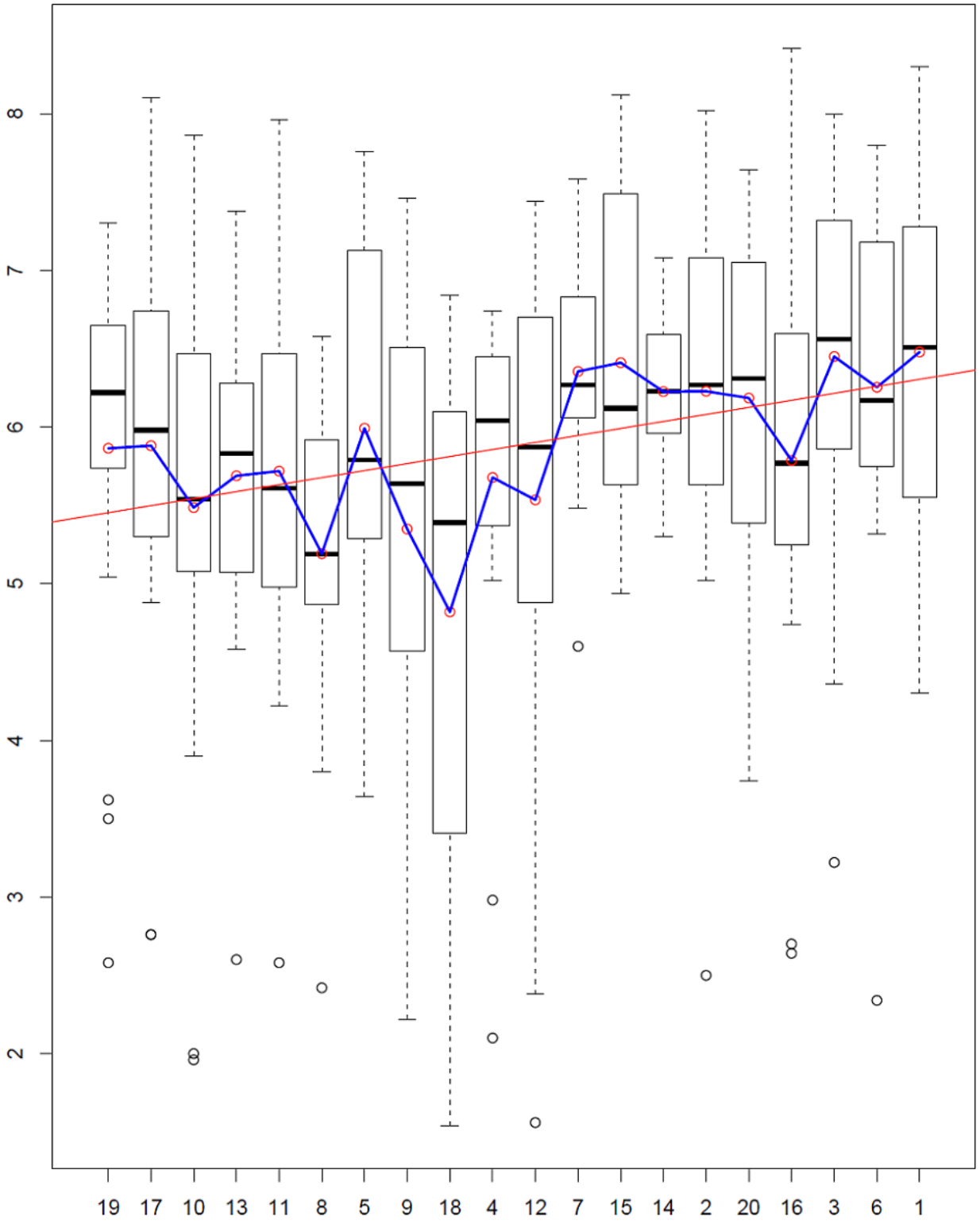
Topic 1 shows the highest positive correlation with Topic 3. This can be further confirmed in Table 3, as both topics address actual positive everyday life situations like celebrations (birthday, Christmas, holiday, merry), and Topic 1 was strongly negatively correlated with Topic 19, which refers to research, studies, risks, factors, hospital, disease, stroke, and care.

Topic 3, therefore, is also strongly negatively correlated with Topic 19 and Topic 10.

Topic 10 refers to running, beats per minute, heart rate, attitude, stop, walk, and reach, while Topic 3 refers to fun, celebrate, movie, Saturday, dinner, barbeque, chicken, grill, egg, cheese, delicious, potato, and cooking. Topic 6 addresses religion—god, lord, pray, faith, amen, bless, prayer, psalm, soul—while Topic 17 addresses sports—playoff, nfl, game, yard, football, player—showing clear differences in topics of interest addressed by men and women.

XSL•FO
RenderX

Figure 7. Topics and gender covariate obtained with spectral structural topic modeling.



XSL•FO
RenderX

Figure 8. Topics and gender covariate obtained with spectral structural topic modeling.

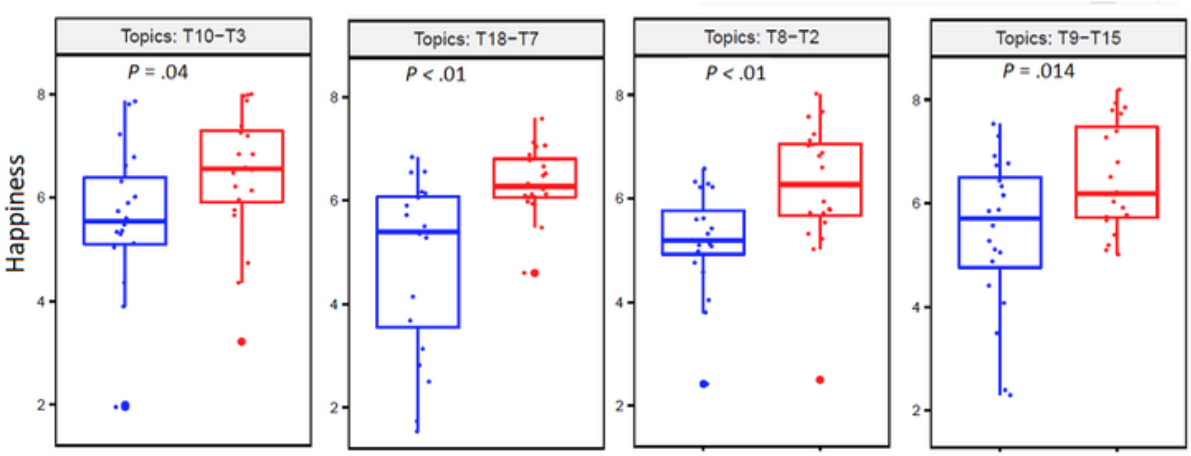
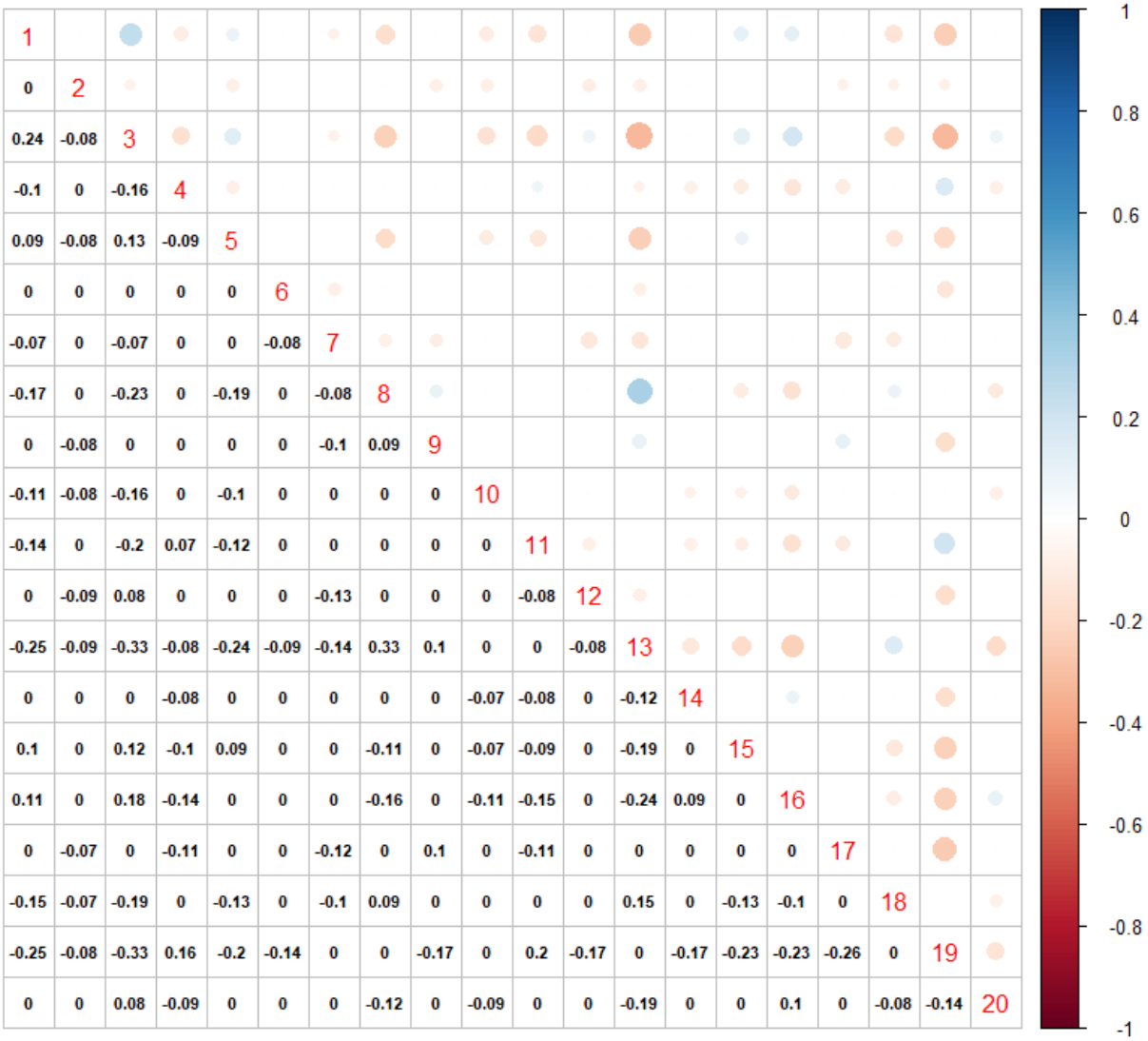


Figure 9. Topics and gender covariate obtained with spectral structural topic modeling.



XSL•FO
RenderX

Discussion

Principal Findings

In this work, we proposed the application of Twitter lists to a chronic health condition in a GNU framework (R-3.5.1). We applied a set of publicly available R libraries for collecting and organizing Twitter data via Twitter's REST and stream API (*rtweet*), sentiment analysis (*syuzhet*), text mining (*tm, quanteda*), and structural topic model (*stm*). We also applied the hedonometer tool to assign happiness scores to topics by gender.

According to our findings, men use significantly more words while expressing negative emotions in their tweets than women, while women use significantly more words when expressing positive emotions.

The results also show that the top three most frequent emotions evoked by both men and women are trust, anticipation, and joy. Besides, the statistical analysis of the basic emotions detects significant preferences for each gender: While words from the emotional fields of trust, anticipation, and joy appear significantly more often in women's tweets, men's tweets significantly exhibit a preference for evoking disgust, anger, fear, and sadness.

We also applied another tool that was specifically designed for considering Twitter inputs—the hedonometer. Happiness ratings obtained by the hedonometer, summarized on a daily basis for each user, are also higher for women than for men almost all along the considered time period.

Finally, we applied structural topic modelling (to the best of our knowledge, for the first time to a chronic health condition) to identify main topics addressed by gender and determined positive and negative correlations between topics by gender.

Topics in this context are defined as sets of words; therefore, we assigned happiness scores to the words with highest probabilities in the identified topics and found that the topics women talk about show higher happiness scores than the topics addressed by men.

A common stereotype in both Western and Eastern cultures suggests that women are more emotional than men, particularly when responding to negative emotions [38]. As remarked in the Introduction section, after stroke, women experience more activity limitations, worse health-related quality of life, and more poststroke depression than men [7] and are twice as likely to suffer from severe depression following a stroke than men. We identify several explanations for our findings, listed below.

First, according to Ayis et al [8], women draw larger components of their sense of self and self-worth from interpersonal relationships and networks, and they are more sensitive to adversities of these. Therefore, female stroke survivors may experience (to a larger extent in comparison to men) the interpersonal and intrapersonal benefits of sharing positive events and emotions on social network sites (SNSs). The intrapersonal benefits of sharing positive events and emotions on SNSs consist of re-experiencing and prolonging these positive events; the interpersonal benefits comprise positive social interaction and positive feedback from other SNS users (according to the results of an ethnographic diary study on Facebook use from Sas et al [39]).

Second, prior research indicates that the positivity of self-presentation on SNSs has an influence on both the quantity and quality of reactions from other SNS contacts. For example, Utz [40] found that SNS users were least likely to receive reactions from their online friends when they expressed sadness in their postings. Similarly, Forest and Wood [41] demonstrated that more positive status updates on Facebook received more positive and favorable feedback from friends than negative status updates.

A third explanation to our findings can be related to the existence of "the positivity bias in SNS communication," which states that "while the SNS environment generally enables authentic self-presentation, it favors positive forms of authenticity over the presentation of negative aspects of the true self" [42].

Therefore, according to Reinecke et al, due to the positivity bias in SNS communication, individuals with higher levels of psychological well-being have a higher chance of experiencing authenticity through the use of SNSs than SNS users with low psychological well-being.

The fourth explanation is related to a recent Facebook analysis involving 15,000 users [43]. The authors concluded that "language used more by self-identified females was interpersonally warmer, more compassionate, polite, and—contrary to previous findings—slightly more assertive in their language use, whereas language used more by self-identified males was colder, more hostile, and impersonal." In fact, the following text from their publication, can also be applied to our own findings:

> *The most strongly female-linked topics included words describing positive emotions (e.g., "excited", "happy", "<3", "love",), social relationships (e.g., "friends", "family", "sister"), and intensive adverbs (e.g., "sooo", "sooooo", "ridiculously"). Strongly male-linked topics included words related to politics (e.g., "government", "tax", "political"), sports and competition (e.g.,"football", "season", "win", "battle"), and specific interests or activities, such as shooting guns, playing musical instruments, or playing video games.*

Therefore, according to this fourth explanation, our findings in another SNS such as Twitter are similar to those involving users not necessarily identified as stroke survivors on Facebook.

Limitations

The collected sample was not intended to be representative or a comprehensive set of all tweets posted by stroke survivors during the period under study. Although the collected data also included tweets directed at other users (ie, conversational tweets), the results cannot be considered to reflect all topics of conversation appearing in Twitter for stroke survivors.

Data collection relied on Twitter's streaming API, which prevents collection of tweets from private Twitter accounts. As

XSL•FO
RenderX

a result, findings may not represent individuals with private accounts.

Furthermore, recent analysis [44] shows that 62% of all Twitter users are less than 49 years old; our participants are skewed toward such an age range, and most of them from the United States.

Nevertheless, as discussed in the Introduction section, multiple recent studies have reported a sustained increasing incidence of stroke at younger ages and the included participants were randomly selected after checking their membership to several Twitter stroke–related lists and manually double checked in relation to gender and stroke survivor condition.

We analyzed women (n=244) who posted a total of 3,788,069 tweets. From them, we included 396,898 tweets in our analysis (the most recent ones, up to December 2018); therefore, we analyzed 10.5% of all posted tweets by women participating in this study.

We analyzed men (n=235) who posted a total of 1,469,364 tweets. From them, we included 403,526 tweets in our analysis (the most recent ones, up to December 2018). Therefore, we analyzed 27.4% of all posted tweets by men participating in this study.

The total number of tweets posted by women from whom we extracted our sample is clearly larger than tweets posted by men. This seems to be coincidental with general Twitter use statistics: Women are usually more active, and each month, 40 million more women than men visit Twitter [45].

Other relevant factors to be mentioned as limitations to our study are related to geographic location, spatial trajectory, or the time of the day a tweet has been posted. As remarked by Padilla et al [46] and Gore et al [47], such factors may affect tweets' sentiments. We observed that 85% of our participants profiles are from the United Kingdom and United States, but spatiotemporal aspects are not controlled in our study.

Finally, the individual psychological differences that stroke survivors may experience must also be mentioned. Certain individuals might have personality traits that make them more predisposed to positive or negative sentiments. The degree to which sentiment reflects variance in psychological traits versus the situational context in which those traits were expressed is unclear. Possible users affected by severe depression may not be active on Twitter; this could be a source for another significant bias in the data sample.

Comparison with Prior Work

One of the scarce previous research about tweet topics or sentiment analysis on chronic health conditions was recently conducted by Brunner and colleagues [48]. Tweets tagged with traumatic brain injury (TBI)-related hashtags were harvested over a one-month period in 2016 and analyzed qualitatively and quantitatively. A total of 29,199 tweets included tweets sent by 893 users, 219 of whom had a brain injury. Twitter was used to discuss health issues, raise awareness of TBI, talk about life after TBI, talk about sport and concussion, and communicate inspirational messages.

In relation to depression, Lachmar and colleagues [49] captured 3225 original tweets for the hashtag #MyDepressionLooksLike that circulated in May 2016. Cleaning resulted in a total of 1978 tweets. Using qualitative content analysis revealed seven themes: dysfunctional thoughts, lifestyle challenges, social struggles, hiding behind a mask, apathy and sadness, suicidal thoughts and behaviors, and seeking relief. Contrary to Lachmar and colleagues' [49] analysis or the #Stroke analysis (the one presented in the Introduction section), our analysis is not linked to a specific hashtag.

It is important to remark the need for further research from a gender perspective, as promoted by initiatives such as the Women's Brain Project [50].

Conclusions

This study explored emotional expressivity for eight specific types of emotion and identified 20 main topics of interest through Twitter posts in stroke survivors from a gender perspective. Numerous studies have shown that, compared with men, women usually experience more frequent and stronger negative emotions. Nevertheless, our results show that men present more frequent and stronger negative emotions in their tweets, when considering both globally positive-negative or individual tweets and analyzing them using two different well-established approaches: the Plutchik model and the hedonometer tool.

Acknowledgments

This research was partially funded by EU H2020 PRECISE4Q - Personalized Medicine by Predictive Modeling in Stroke for better Quality of Life (Grant Agreement 777107 – Research and Innovation Action).

Authors' Contributions

AG-R and SL conceived the study. AG-R, JS, and SL collected, selected, and cleaned the data. AG-R and JS analyzed the data. AG-R drafted the initial manuscript. SL, JS, and MBG revised the manuscript critically for important intellectual content and approved the final manuscript. AG-R, SL, JS, and MBG received funding for the study.

Conflicts of Interest

None declared.

Multimedia Appendix 1

XSL•FO
RenderX

Demographics, wordclouds, VosViewer cluster analysis, latent Dirichlet allocation topics, correlation analysis, STM topics, hedonometer scores, and the Plutchik psychoevolutionary model.
[PDF File (Adobe PDF File), 2 MB - jmir_v21i8e14077_app1.pdf]

References

1. Johnson W, Onuma O, Owolabi MS. Bulletin of the World Health Organization. 2016. Stroke: a global response is needed URL: https://www.who.int/bulletin/volumes/94/9/16-181636.pdf [accessed 2019-07-22]
2. Béjot Y, Daubail B, Jacquin A, Durier J, Osseby G, Rouaud O, et al. Trends in the incidence of ischaemic stroke in young adults between 1985 and 2011: the Dijon Stroke Registry. J Neurol Neurosurg Psychiatry 2014 May;85(5):509-513. [doi: 10.1136/jnnp-2013-306203] [Medline: 24249786]
3. Putaala J. Ischemic stroke in the young: Current perspectives on incidence, risk factors, and cardiovascular prognosis. Eur Stroke J 2016 Mar 08;1(1):28-40 [FREE Full text] [doi: 10.1177/2396987316629860] [Medline: 31008265]
4. Vangen-Lønne AM, Wilsgaard T, Johnsen SH, Carlsson M, Mathiesen EB. Time trends in incidence and case fatality of ischemic stroke: the tromsø study 1977-2010. Stroke 2015 May 07;46(5):1173-1179. [doi: 10.1161/STROKEAHA.114.008387] [Medline: 25851772]
5. Krishnamurthi RV, Moran AE, Feigin VL, Barker-Collo S, Norrving B, Mensah GA, et al. Stroke Prevalence, Mortality and Disability-Adjusted Life Years in Adults Aged 20-64 Years in 1990-2013: Data from the Global Burden of Disease 2013 Study. Neuroepidemiology 2015 Oct 28;45(3):190-202. [doi: 10.1159/000441098]
6. Putaala J, Kurkinen M, Tarvos V, Salonen O, Kaste M, Tatlisumak T. Silent brain infarcts and leukoaraiosis in young adults with first-ever ischemic stroke. Neurology 2009 May 26;72(21):1823-1829. [doi: 10.1212/WNL.0b013e3181a711df]
7. Gall S, Phan H, Madsen TE, Reeves M, Rist P, Jimenez M, et al. Focused Update of Sex Differences in Patient Reported Outcome Measures After Stroke. Stroke 2018 Mar;49(3):531-535. [doi: 10.1161/STROKEAHA.117.018417]
8. Ayis SA, Rudd A, Ayerbe L, Wolfe CDA. Sex differences in trajectories of depression symptoms and associations with 10-year mortality in stroke patients: The South London Stroke Register. London: The South London Stroke Register; 2019.
9. Mathers C. The global burden of disease update. In: The Global Burden of Disease: 2004 Update. Geneve: World Health Organization; 2008:2004.
10. Radu G, Solomon M, Gheorghe CM, Hostiuc M, Bulescu IA, Purcarea VL. The adaptation of health care marketing to the digital era. J Med Life 2017;10(1):44-46 [FREE Full text] [Medline: 28255375]
11. Cabrera-Maqueda JM, Minhas JS. New Horizons for Stroke Medicine: Understanding the Value of Social Media. Stroke 2018 Feb;49(2):25-27. [doi: 10.1161/STROKEAHA.117.020068]
12. Bundy JJ, Chick JFB, Hage AN, Srinivasa RN, Chaudhary N, Srinivasa RN, et al. #Stroke. J NeuroIntervent Surg 2018 Apr 20;10(12):e33-e33. [doi: 10.1136/neurintsurg-2018-013877]
13. Brennan J. Adjustment to cancer - coping or personal transition? Psychooncology 2001;10(1):1-18. [Medline: 11180573]
14. Taylor GH, Todman J, Broomfield NM. Post-stroke emotional adjustment: A modified social cognitive transition model. Neuropsychological Rehabilitation 2011 Dec;21(6):808-824. [doi: 10.1080/09602011.2011.598403]
15. Sylwester K, Purver M. Twitter Language Use Reflects Psychological Differences between Democrats and Republicans. PLoS One 2015 Sep;10(9):e0137422 [FREE Full text] [doi: 10.1371/journal.pone.0137422] [Medline: 26375581]
16. Kramer A. The spread of emotion via Facebook. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. USA: ACM; 2012 Presented at: Presented atIGCHI Conference on Human Factors in Computing Systems; 2012; Austin, Texas p. 05-10.
17. Kern ML, Park G, Eichstaedt JC, Schwartz HA, Sap M, Smith LK, et al. Gaining Insights From Social Media Language: Methodologies and Challenges. Psychol Methods 2016 Aug 8:507-525. [doi: 10.1037/met0000091] [Medline: 27505683]
18. Gohil S, Vuik S, Darzi A. Sentiment Analysis of Health Care Tweets: Review of the Methods Used. JMIR Public Health Surveill 2018 Apr 23;4(2):e43 [FREE Full text] [doi: 10.2196/publichealth.5789] [Medline: 29685871]
19. Ardon S. Spatio-Temporal Analysis of Topic Popularity in Twitter. arXiv 2011 Apr 05:2904-2908.
20. Ghosh S, Zafar M, Bhattacharya P, Sharma N, Ganguly N. On sampling the wisdom of crowds: random vs. expert sampling of the twitter stream. 2013. URL: http://ama.liglab.fr/wikiDeAMA/DOCS/randomvsexpert.pdf [accessed 2019-08-06]
21. Hoffmann T. “Too many Americans are trapped in fear, violence and poverty”: a psychology-informed sentiment analysis of campaign speeches from the 2016 US Presidential Election. Linguistics Vanguard 2018;4(1):1-9. [doi: 10.1515/lingvan-2017-0008]
22. Plutchik R. The Nature of Emotions: Human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice. American Scientist 2001:344-350.
23. Mohammad S, Turney P. Emotions evoked by common words and phrases: Using Mechanical Turk to create an emotion lexicon. 2010 Presented at: NAACL-HLT Workshop on computational approaches to analysis and generation of emotion in text; 2010; California, US.
24. Schweinberger M. A sociolinguistic analysis of emotives in Irish English. 2016 Presented at: Annual Meeting of the Society for Text & Discourse; 2016; Kassel.

XSL•FO
RenderX

25. 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 Dec;6(12):e26752 [FREE Full text] [doi: 10.1371/journal.pone.0026752] [Medline: 22163266]
26. Reece A, Danforth C. Instagram photos reveal predictive markers of depression. EPJ Data Sci 2017 Aug 8;6(1):15. [doi: 10.1140/epjds/s13688-017-0110-z]
27. Cody E, Reagan A, Mitchell L, Dodds P, Danforth CM. Climate Change Sentiment on Twitter: An Unsolicited Public Opinion Poll. PLoS One 2015;10(8):e0136092 [FREE Full text] [doi: 10.1371/journal.pone.0136092] [Medline: 26291877]
28. Kearney M. Package 'rtweet'. URL: https://cran.r-project.org/web/packages/rtweet/rtweet.pdf [accessed 2019-03-19]
29. Webb H, Jirotka M, Stahl B, Housley W, Edwards A, Williams M, et al. The ethical challenges of publishing Twitter data for research dissemination. 2017 Presented at: ACM on Web Science Conference; 2017; Troy, NY p. 25-28.
30. Jockers M. Introduction to the Syuzhet package. URL: https://cran.r-project.org/web/packages/syuzhet/vignettes/syuzhet-vignette.html [accessed 2019-08-06]
31. Gries S. Statistics for linguistics with R: A practical introduction. NY: Mouton de Gruyter; 2009.
32. Feinerer I. Introduction to the tm Package: Text Mining in R. URL: https://cran.r-project.org/web/packages/tm/vignettes/tm.pdf [accessed 2019-03-19]
33. Murzintcev N. ldatuning: Tuning of the Latent Dirichlet Allocation Models Parameters. URL: https://cran.r-project.org/web/packages/ldatuning/index.html [accessed 2019-03-19]
34. Roberts M. Journal of Statistical Software. stm: R Package for Structural Topic Models URL: https://cran.r-project.org/web/packages/stm/vignettes/stmVignette.pdf [accessed 2019-03-19]
35. Blei DM, Lafferty JD. A correlated topic model of Science. Ann Appl Stat 2007 Jun;1(1):17-35. [doi: 10.1214/07-AOAS114]
36. Waltman L, van Eck NJ, Noyons EC. A unified approach to mapping and clustering of bibliometric networks. Journal of Informetrics 2010 Oct;4(4):629-635. [doi: 10.1016/j.joi.2010.07.002]
37. Mimno D, Wallach H, Talley E, Leenders M, McCallum A. Optimizing semantic coherence in topic models. In: In Proceedings of the Conference on Empirical Methods in Natural Language Processing, EMNLP ?11, pp. 262?272. Association for Computational Linguistics, Stroudsburg, PA, USA. 2011 Presented at: Conference on Empirical Methods in Natural Language Processing; 2011; Stroudsburg, PA, US.
38. Deng Y, Chang L, Yang M, Huo M, Zhou R. Gender Differences in Emotional Response: Inconsistency between Experience and Expressivity. PLoS ONE 2016 Jun 30;11(6):e0158666. [doi: 10.1371/journal.pone.0158666]
39. Sas C, Dix A, Hart J, Su R. Dramaturgical capitalization of positive emotions: The answer for Facebook success? 2009 Presented at: 23rd British HCI group annual conference on people and computers: Celebrating people and technology; 2009; Swinton, UK.
40. Utz S. Social network site use among Dutch students: Effects of time and platform. In: Networked Sociability and Individualism: Technology for Personal and Professional Relationships. Hershey, PA: IGI Global; 2011:103-125.
41. Forest AL, Wood JV. When Social Networking Is Not Working. Psychol Sci 2012 Feb 07;23(3):295-302. [doi: 10.1177/0956797611429709]
42. Reinecke L, Trepte S. Authenticity and well-being on social network sites: A two-wave longitudinal study on the effects of online authenticity and the positivity bias in SNS communication. Computers in Human Behavior 2014 Jan;30:95-102. [doi: 10.1016/j.chb.2013.07.030]
43. Park G, Yaden DB, Schwartz HA, Kern ML, Eichstaedt JC, Kosinski M, et al. Women are Warmer but No Less Assertive than Men: Gender and Language on Facebook. PLoS One 2016;11(5):e0155885 [FREE Full text] [doi: 10.1371/journal.pone.0155885] [Medline: 27223607]
44. Omnicore. Twitter by the Numbers: Stats, Demographics & Fun Facts URL: https://www.omnicoreagency.com/twitter-statistics/ [accessed 2019-03-18]
45. Who is more active on Social Media?. URL: https://www.digitalinformationworld.com/2014/10/who-is-more-active-on-social-media-men-or-women-infographic.html [accessed 2019-03-19]
46. Padilla J, Kavak H, Lynch C, Gore R, Diallo SY. Temporal and spatiotemporal investigation of tourist attraction visit sentiment on Twitter. PLoS One 2018;13(6):e0198857 [FREE Full text] [doi: 10.1371/journal.pone.0198857] [Medline: 29902270]
47. Gore R, Diallo S. You Are What You Tweet: Connecting the Geographic Variation in America?s Obesity Rate to Twitter Content. PLoS ONE 2015;You Are What You Tweet:1-16 [FREE Full text]
48. Brunner M, Hemsley B, Dann S, Togher L, Palmer S. Hashtag #TBI: A content and network data analysis of tweets about Traumatic Brain Injury. Brain Injury 2017 Dec 08;32(1):49-63. [doi: 10.1080/02699052.2017.1403047]
49. Lachmar EM, Wittenborn AK, Bogen KW, McCauley HL. #MyDepressionLooksLike: Examining Public Discourse About Depression on Twitter. JMIR Ment Health 2017 Oct 18;4(4):e43 [FREE Full text] [doi: 10.2196/mental.8141] [Medline: 29046270]
50. Womens Brain Project. URL: http://www.womensbrainproject.com/ [accessed 2019-06-10]

XSL•FO
RenderX

Abbreviations

API: application program interface
EM: expectation maximization
FREX: frequency-exclusivity
HRQoL: health-related quality of life
LDA: latent Dirichlet allocation
NRC: National Research Council
REST: representational state transfer
SNS: social network site
STM: structural topic modeling
TDM: term document matrix
TF-IDF: term frequency – inverse document frequency

Edited by G Eysenbach; submitted 23.03.19; peer-reviewed by R Gore, L Subirats, S Kiritchenko; comments to author 18.04.19; revised version received 11.06.19; accepted 16.06.19; published 26.08.19.

Please cite as:
Garcia-Rudolph A, Laxe S, Saurí J, Bernabeu Guitart M
Stroke Survivors on Twitter: Sentiment and Topic Analysis From a Gender Perspective
J Med Internet Res 2019;21(8):e14077
URL: http://www.jmir.org/2019/8/e14077/
doi:10.2196/14077
PMID:31452514

©Alejandro Garcia-Rudolph, Sara Laxe, Joan Saurí, Montserrat Bernabeu Guitart. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 26.08.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Patient Commitment to Health (PACT-Health) in the Heart Failure Population: A Focus Group Study of an Active Communication Framework for Patient-Centered Health Behavior Change

Daniella Meeker[1], PhD; Jordan Goldberg[2], BA; Katherine K Kim[3], PhD, MPH, MBA; Desi Peneva[4], MS; Hugo De Oliveira Campos[5]; Ross Maclean[4], MD; Van Selby[6], MD; Jason N Doctor[7], PhD

[1]Department of Preventive Medicine, University of Southern California, Los Angeles, CA, United States

[2]stickK.com, New York City, NY, United States

[3]Betty Irene Moore School of Nursing, University of California Davis, Sacramento, CA, United States

[4]Precision Health Economics, Los Angeles, CA, United States

[5]Kaiser Permanente, Oakland, CA, United States

[6]Cardiology Division, Department of Medicine, University of California San Francisco, San Francisco, CA, United States

[7]Schaeffer Center for Health Policy and Economics, University of Southern California, Los Angeles, CA, United States

Corresponding Author:
Jason N Doctor, PhD
Schaeffer Center for Health Policy and Economics
University of Southern California
635 Downey Way
Los Angeles, CA, 90089
United States
Phone: 1 213 821 8142
Email: jdoctor@usc.edu

Abstract

Background: Over 6 million Americans have heart failure, and 1 in 8 deaths included heart failure as a contributing cause in 2016. Lifestyle changes and adherence to diet and exercise regimens are important in limiting disease progression. Health coaching and public commitment are two interactive communication strategies that may improve self-management of heart failure.

Objective: This study aimed to conduct patient focus groups to gain insight into how best to implement health coaching and public commitment strategies within the heart failure population.

Methods: Focus groups were conducted in two locations. We studied 2 patients in Oakland, California, and 5 patients in Los Angeles, California. Patients were referred by local cardiologists and had to have a diagnosis of chronic heart failure. We used a semistructured interview tool to explore several patient-centered themes including medication adherence, exercise habits, dietary habits, goals, accountability, and rewards. We coded focus group data using the a priori coding criteria for these domains.

Results: Medication adherence barriers included regimen complexity, forgetfulness, and difficulty coping with side effects. Participants reported that they receive little instruction from care providers on appropriate exercise and dietary habits. They also reported personal and social obstacles to achieving these objectives. Participants were in favor of structured goal setting, use of online social networks, and financial rewards as a means of promoting health lifestyles. Peers were viewed as better motivating agents than family members.

Conclusions: An active communication framework involving dissemination of diet- and exercise-related health information, structured goal setting, peer accountability, and financial rewards appears promising in the management of heart failure.

(J Med Internet Res 2019;21(8):e12483) doi:10.2196/12483

KEYWORDS

heart failure; behavioral economics; motivational interviewing

XSL•FO
RenderX

Introduction

Background

Cardiovascular disease is the leading cause of death in the United States and a source of rapidly increasing annual health expenditures [1]. Over 6 million Americans have heart failure, and 1 in 8 deaths included heart failure as a contributing cause in 2016 [2,3]. In addition, heart failure costs the nation an estimated US $32 billion each year and is projected to increase substantially over the next 20 years [1]. The Affordable Care Act targeted the economic burden of cardiovascular mortality by imposing financial penalties for cardiovascular-related hospital readmissions and offering financial incentives for numerous cardiovascular quality measures [4]. Given the significant burden of this disease and unmet need for improvement, a stronger focus on encouraging healthy behaviors among persons with and at risk for heart failure may be one of the best ways to improve patient outcomes.

Symptom burden and mortality risk are associated with behavioral factors among persons with or at risk for heart failure. Past studies have shown that moderate to high levels of physical activity are associated with a reduced risk of heart failure [5-7], whereas poor exercise capacity is associated with increased mortality risk in this population [8]. In addition, the American College of Cardiology and American Heart Association recommend low-sodium diets as a primary prevention for cardiovascular disease [9]. Some estimates suggest that as many as 1 in 20 heart failure admissions are precipitated by failure to follow dietary guidelines [10]. As physical activity and dietary habits are modifiable health-related behaviors that are directly associated with the risk of heart failure, the PAtient CommiTment to Health (PACT-Health) in the heart failure population study has high potential to improve health outcomes among persons with or at risk for heart failure by encouraging healthy behaviors through health coaching and public commitments.

Although evidence around managing heart failure has been established, guidelines are complex and often require behavioral change outside of the clinician visit. An important gap in evidence remains as to whether better communication of patients' preferences and goals can improve adherence to plans developed with clinicians. Healthy lifestyle changes may improve clinical outcomes but require patients to engage in multiple health behaviors. Patient education as part of disease management programs delivered by multidisciplinary health care teams has been the primary communication and dissemination medium used to address disease self-management in this population. However, recent systematic reviews do not provide clear evidence for the effectiveness of this approach [11,12]. In general, the data indicate that self-care interventions can improve symptoms or prognosis [5] but often fail because of lack of persistence in their application [13].

With the premise that care planning requires thoughtful identification and communication of patient goals and mechanisms to motivate adherence to goals, we study the receptivity of patients to two research areas in behavior change: health coaching and public commitments. Health coaching is focused on evidenced-based approaches to setting goals and overcoming barriers that might prevent the achievement of goals, as well as in motivating lifestyle changes [14-16]. Health coaching is based on motivational interviewing, a procedure developed in clinical psychology to help people work through ambivalence about making changes in their life [17]. The simple idea is that patients do not enter into a dialogue ready to tackle changing their behavior. Giving direct advice often leads to defensiveness, resistance, and may exacerbate the behavior of concern. Health coaching asks the patients to articulate their own reasons for concern about their behavior and asks them for their own arguments for making a change. Strategy is matched to the patient's readiness to change in brief educational modules. Information generation through health coaching is a patient-centered activity that requires the patient to generate all reasons for changes and concern over their behavior patterns and actions. By clearly developing these intentions, patients are able to more effectively represent their own interests and concerns when given the opportunity to discuss plans with clinicians.

Commitments follow readiness to change in health coaching. Since 2010, our group and others have been conducting studies showing that uptake of health-related behaviors is improved with the use of principles of psychology and economics to *nudge* persons toward making healthy choices [18-20]. We focus here on public commitment nudges—statements of dedication to a course of action that are visible to others [21-23]. People are more likely to follow through with their expressed intentions when that intention was framed as a commitment [24-26]. Public commitments have been shown to increase recycling [27-29], heighten participation in hotel towel reuse programs [30], boost monetary contributions to organizations serving the disabled [31], enhance the likelihood of voting in an upcoming election [32], reduce the prescription of inappropriate antibiotics [19], and lead dieters to lose more weight [33]. Public commitment is more effective than education as a tool to prompt greater personal motivation to perform a behavior [24,32,34]. Notably, public commitments are not forced by health providers nor formally negotiated agreements with another party but rather voluntary statements that strengthen dedication to an action and which later are identified with one's values [35]. Commitment interventions are highly pragmatic. They minimize disruption of usual routines because they can be elicited in many different ways and take less than a minute to complete. Following through with a commitment makes a person's actions consistent with their stated goals. Commitments then establish contingencies that enhance motivation to engage in self-management.

Objective

This study aimed to conduct focus groups to gain insight into how best to implement health coaching and public commitment strategies within the heart failure population.

Methods

Before commencement of focus group activities, we submitted a request to the Western Institutional Review Board for an exemption determination; the request was approved.

XSL•FO
RenderX

Recruitment

Patients with chronic heart failure were recruited from a convenience sample through two cardiologists and two leaders of patient advocacy groups in Oakland, California, and Los Angeles, California. Participants were sourced during patient visits, through email and telephonic outreach, and via various social media posts. The primary obstacle to securing patient participation was physical difficulty (ie, health) or logistical difficulty (ie, can no longer drive or do not own a car) in leaving the home for an extended period of time. Patients were told that the discussion would last approximately 2 hours and would be recorded for transcription and follow-up analysis. Patients were given a US $100 cash incentive for participation in addition to a US $75 cash reimbursement for any travel expenses incurred.

Interview Format

We used a semistructured interview tool (Multimedia Appendix 1) to explore several patient-centered themes: medication adherence, exercise habits, dietary habits, goals, accountability, and rewards. Following the interview portion, mockups of the Web-based user experience were presented to participants for their feedback. The mockups included both sample commitments (Multimedia Appendix 2) and sample rewards (Multimedia Appendix 3).

Coding

A total of 2 reviewers coded the focus group data using codes developed *a priori* and evaluated responses from 2 reviews of each transcript. The Oakland transcript was 38 pages in length; the Los Angeles transcript was 44 pages in length. There were a total of 32 codes used (see Multimedia Appendix 4 for codebook). Published guidelines were followed to plan and report the study (Multimedia Appendix 5) [36].

The interviewers used several strategies to mitigate potential investigator bias. First, to address reflexivity, the interviewers discussed with other researchers on the team their assumptions, potential biases, and concerns before conducting the focus groups [37]. Second, the interviewers debriefed findings utilizing the transcripts with 2 peer researchers who did not attend the focus groups. Finally, quotes and findings were reviewed with 1 focus group participant to assess the accuracy of researchers' interpretations of the findings.

Results

Sample Characteristics

We conducted two in-person focus groups. Table 1 displays the sample characteristics for participants from both groups.

The first focus group was held in Oakland, California, on March 16, 2015. There were 2 participants, both men with heart failure, ages 51 and 64 years. One of the men was Hispanic and both have lived for more than 5 years with heart failure. The intimate setting enabled us to essentially conduct two one-on-one interviews simultaneously, with both participants answering each question posed by the moderator. The second focus group was held in Los Angeles, California, on March 17, 2015. There were 5 participants of mixed sex, age, and race. The larger, open dynamic fostered more dialogue among participants, yielding more of a *conversational* tone than interview.

Table 1. Demographics of focus group participants.

Age (years)	Sex	Years with heart failure	Race, ethnicity	Location
51	Male	9	White, Hispanic	Oakland, California
64	Male	>5	White, non-Hispanic	Oakland, California
59	Female	<1	Other	Los Angeles, California
78	Female	4	White, Hispanic	Los Angeles, California
66	Female	17.5	White, non-Hispanic	Los Angeles, California
36	Male	1	White, non-Hispanic	Los Angeles, California
45	Female	9	White, non-Hispanic	Los Angeles, California

Qualitative Findings

Participants readily shared their perspectives on living with heart failure. We began with a discussion of family history of heart disease and the moment of discovery of heart failure for each patient. A total of 1 participant was born with a heart condition and had corrective surgery at the age of 4 years and largely lived a normal childhood before experiencing heart failure episodes again at the ages of 26 and 34 years. Several reported having a pacemaker or a defibrillator. Several also shared a provocative event that began their lives living with heart failure. A total of 2 participants stated in no uncertain terms that their spouses left them as a result of the difficulties and stress of becoming a caregiver. Both participants have since found new significant others:

> *Nine years ago on my way to work before I even got to my car, I suffered a – sweating real bad, and couldn't understand that. And by the time I got to my car, I passed out. When I woke up, I was in Kaiser Hospital. And that's when Kaiser Hospital found out that I had CHF. I was angry. I was really angry.*

> *I couldn't get up the jet way [to the plane], and ended up in the local hospital. And, you know, one thing led to another, and the next thing you know, they slice me open when I got back to Austin, which has, by the way, fantastic doctors, and they put a device in me.*

> *I was driving one day and I just got all sweaty and my heart was beating really fast and I went into a V-tach and went to the doctors and they hit me with the paddles and all that. Was there for three days and*

XSL•FO
RenderX

one of the doctors said, "We're gonna install an ICD in you." And I was like, "I don't know." It took some convincing but I got one.

Medication Adherence

The conversation continued with a discussion of medication compliance. Most felt that compliance was not a serious concern—participants understand the severity of their condition and the necessity for adherence. However, participants cited complexity of the regimen, forgetfulness, or difficulty adjusting to side effects as a reason for noncompliance at times:

Well, I've kind of gotten a schedule and I don't necessarily have approval from my doctors because it is complicated and some of them say to take it at night, some say to take it with food, some say to take it without food.

It's almost like the devil...which one do you want? You take it, you get so sick, and you don't take it, you get sick because it's in your system.

Current Exercise Habits

There was a spectrum of physical capacity to exercise, although most felt significant limitations in their abilities as a result of their condition. Exercise appeared largely limited to walking, biking, yoga, and short, light jogging. Los Angeles participants received very limited advice from their physicians on exercise regimens to follow. A total of 1 participant watched video tutorials for different exercise routines provided by her physician. The Oakland participants, on the contrary, both stated that they received clear advice from their physicians on exercise routines.

However, many felt uneasy about exercise because of fears and safety concerns that were neither endorsed nor discredited by their physicians. Many experienced tightness in their chest, rising heart rates, shortness of breath, and other forms of discomfort. A firm grasp of boundaries did not appear well understood. Outdoor temperatures (high heat) was cited as a big concern and an obstacle to exercise. In fact, one participant relocated from Texas to Northern California primarily for this reason.

Several participants mentioned that they make a habit of telling a loved one of their intentions to exercise before beginning, to make sure someone is aware in case there is a medical emergency that requires immediate attention as a result of exercising. Others indicated that even basic activities such as getting the mail were strenuous. Depression that stems from their condition was also cited as a hurdle to exercise:

...I find that walking, I can do it fine for a short time but then I get short of breath and I feel that tightness, so I just slow down.

My exercise is just daily life. There's no regimen. And I always feel guilty about it but it doesn't seem quite enough to push me into getting it done.

...if you don't have that guidance, it's really hard because the doctors sent me out of the hospital and said, "Okay, now you've got to exercise. You've got to eat right." I was scared to death to exercise.

I was on all this medication and every time I exercised, I felt horrible. I thought I was going to die, or have a heart attack, or something.

I'm careful to keep it under – well, basically when I'm at the gym, I'm working at about just the heart rate of a hundred, and I'm shooting for that maximum, and leaving it there.

...I always let my wife know that I'm leaving, and she's timing me to make sure. Because she knows that if it's a certain time, she'll come out looking for me. It's just when I do walk, I get to the part that I have to stop a certain time to get some breath, and go back and walk; a certain time, catch a breath, or sit down because that's basically my routine, and then I come back the same way.

I mean there is only a certain amount of things that the doctor can tell you because they don't read inside your body. They can only go by the reports and how you feel and stuff, but it all depends on how my body takes it.

Current Dietary Habits

Similar to exercise habits, many participants received little guidance on diet from their physicians. One mentioned the Mediterranean diet as a specific recommendation from her doctor, but other participants received little specifics other than managing liquid intake and reducing sodium. Many participants claimed to still feel well informed on what to eat and what not to eat. Almost universally, participants cited reduction and strict limits on sodium intake to be a keystone practice in managing their health.

Difficulty sticking to a diet regimen was noted by several as a concern. Reasons cited included convenience of fast food, limited healthy options when dining out with friends and family, flavorful temptations of less healthy foods, and proximity to children's or grandchildren's unhealthy food. One participant mentioned boredom with the monotony of healthier foods. Many reiterated the importance of greater attention to healthy eating and a need to eat more fruits and vegetables specifically.

Another distraction to sticking with a diet plan is family. Participants cited family meals and gatherings as opportunities to eat unhealthy food; many found that family members are not as understanding or supportive as they would like them to be about their dietary needs:

I've been following the Mediterranean diet, a lot of grains, fruits, vegetables, and fish, and in general, just everything in moderation, so not a lot of one thing...But my diet's not perfect. I drink too much coffee, and I eat chocolate sometimes, and I have young kids, too, and they want macaroni, and I sometimes eat their macaroni.

For me, I find that I'm not hungry, so I have to force myself to eat. I do eat vegetables. I'm not a meat eater.

I drink a lot of water. Your 64 ounces, that's by 7:00 in the morning for me. I drink a lot of water.

XSL•FO
RenderX

[My Fitness Pal] lets me keep track of my liquids and my sodium. It's awesome for tracking your sodium.

Well, bread is just so easy because you can put almond butter on it or when I cook fish, I cook extra so I can make a fish salad 'cause you can't do tuna, canned tuna is too high in salt. But bread just makes portable meals and quick, easy, don't have to fuss meals.

I eat a lot more salad now than I ever have. I also watch what I eat. Even when we go to the store and look at the box to decide – I was always wondering why my father was doing it, now I see why because you have to know how much salt you can take in. We're the type of patients that we can't have no salt. The doctor tells you, "No salt."

Basically I eat a lot of fish, a lot of chicken and turkey. I very seldom have a piece of meat.

You wouldn't believe how many times my nurse practitioner has been frustrated with me because the way I ate or the way I took my medications. Now, we understand each other.

I get contradictory recommendations because I also have type II diabetes. There are times when the heart doctor tells me this, and the diabetes doctor tells me that, and they don't coincide.

Goals

Several participants noted the importance of goal setting in achieving new exercise, weight, or dietary milestones, but few have taken concrete actions to set goals. Most rely on resolve and willpower. One participant remarked that she is a "big list maker," whereas two others noted the use of mobile apps to monitor and track fitness activity, eating habits, and glucose level.

A third participant mentioned that she has purchased a Fitbit, which she intends to use, but has not yet taken the time to set it up and begin using it. The remaining participants seemed largely unfamiliar with connected, wearable devices, but all expressed enthusiasm at the notion of what they can provide once it was explained to them. One participant expressed reluctance to use devices because it would serve as a reminder that he has a handicap and perhaps set off a depressive episode.

One participant noted his desire to get from "70 percent" to "80 or 90 percent," while recognizing that 100 percent would be unlikely. He did not elaborate upon what it means to be 70 percent versus 90 percent. Another participant stated his intentions of doing more biking by increasing his mileage, not his pace, and has set concrete goals (in his mind) on how many miles he would like to do on a regular basis.

Several participants discussed goals more broadly. They shared their desire to spend active days with family and specifically grandchildren. Many also repeatedly stated their desire to be able to travel more regularly, or at all, and without the same degree of daily burden that their current condition imposes. One participant would like to get his band back together to play more regularly once he is able to confidently travel for longer periods of time:

I'm at 233 [pounds], and I want to go as far as making it to 180-200 [pounds]. That's my goal right now.

I nudge – if I go four blocks, I try to get at least five in before I come in. So I move myself a little bit more and more to my goal.

I talk to myself a lot. I don't seem to be very successful. Maybe I need to be talking to somebody else.

I write a lot of things down, but today I forgot my notebook, so I'm kind of bummed. That's what I do. I usually just write everything down.

Motivation and Accountability

When discussing motivation and accountability, there was a vibrant conversation that covered many domains. A few common themes bore out. Participants appeared motivated to manage their health to be around for family, specifically children and grandchildren. Participants also noted on several occasions that at the end of the day, "it was on them" and that they should do it for themselves:

So here's my family, I need to be there for them. I need to keep myself strong enough for them.

My grandkids...mean the world to me. I mean they know I can't run around chasing them or anything like that, but I see them, and I see me. I only have one daughter, and she wonders if I took my medicines, or have I done this or done that. So she, in so many words, concerned also. So is my family…my mom, my dad, you know, my immediate family. It keeps me motivated.

My granddaughter turns ten next month and I really would like to see her at least graduate from high school but I'd like to see her graduate from college. She's a pretty awesome little girl. But I don't think I'm setting a lot of goals other than I want a quality of life.

The most commonly cited obstacle to feeling motivated to manage their health were bouts of depression. There was not a common theme with respect to triggers. Many participants had difficulty identifying triggers or consistency in their types of triggers. A total of 1 participant noted that at times if he noticed someone treating him too well, likely as a result of his condition, it would trigger a depressive episode. On the opposite end of the spectrum, the same participant noted that if someone is not treating him very well at all, it could trigger the same effect. Others said being reminded of their limitations in various forms can act as a trigger.

A total of 2 participants noted the ability that technology can play in motivation:

I just dig the numbers, the numbers that show up on the things that I'm logging, you know. That's tangible enough for me to keep me going.

XSL•FO
RenderX

When the conversation progressed from motivation to accountability, there was a distinct shift in attitudes toward friends and family. Although many noted the importance they play in their desire to be healthy, many found family members to be a hindrance. Often family members are a source of temptation:

> *Because there are times when my family would love to force feed me barbecue, or any number of other things that can do me harm just because it's a familiar pattern.*

Although some noted that family members can be a basis for temptation, an underlying sentiment among many participants was simply that family members cannot be relied upon to hold them accountable in managing their daily habits. This did not appear to be a significant source of consternation. Rather, family members simply do not understand what sufferers are going through because it is a personal journey. Those who do not live with their conditions on a daily basis by definition cannot have an appreciation for it:

> *You know, like my brothers, they couldn't understand.*

Similar attitudes toward friends were pervasive but not to the same degree as family. Many participants shared the fact that they discuss and find comfort from confiding in contemporaries and would like to do so more. Spending time with friends and family can sometimes be seen as an escape.

At the moment, however, there is a similar issue of a lack of understanding with the condition. In addition, a couple of participants expressed their desire to not burden friends and family with their condition, or make the time they spend together depressing:

> *Yeah, and as much as my church friends, they try to be supportive, "Oh, you look so good and aren't you glad that you're all better now?" Sometimes I would say, "No. This is chronic and progressive."*

Though many participants felt friends did not understand their situation, they also noted that their friends often suffer from other ailments (more so than family members, particularly younger family members); as a result, there is some common ground for discussion and support:

> *Fortunately, I have many friends I've known for many, many years. We talk about anything, not just about the heart. They have their issues and we just support each other in that way.*

With respect to tangible support and accountability, several participants discussed at length the importance and influence that their peer support groups provided. Specifically, a few participants were active members in online support groups on Facebook and other forums, and others were active in support groups that were local and had regular, in-person meetings. These participants were particularly vocal in their advocacy of the value of these networks and openly encouraged other participants who were not in such groups to join:

> *We're not going to be normal, you know, 100 percent, but we need to hear from people who understand us.*

> *The Facebook group...there is a lot of people that join right after they get sliced open and have a device put in... they're all over the world. There are a lot of Canadians and English people.*

> *Family loves you...But the group helps you.*

> *Definitely the support group, 'cause they're the peers who really get the struggle and they're living it and they can relate. 'Cause friends who haven't had congestive heart failure don't understand.*

One participant successfully formed a personal bond with her cardiac nurse practitioner and found her to be a powerful form of accountability and support. Another participant found that goals set by his doctor for his next visit was a strong form of accountability:

> *Well, I would have to say Peggy, who's my cardiac nurse practitioner...she's my biggest supporter when I have any kind of a health issue, or something I'm afraid of.*

Another participant, who lives a somewhat remote life on a farm, eschewed the notion of sharing progress with anyone but herself. This is the same participant who uses My Fitness Pal, which she said keeps her focused and accountable:

> *No, I don't have a caregiver and it's not up to anybody else to hold me accountable. No, this is me and my life and I get to make the decisions.*

The youngest participant, a male aged 36 years who was born with a heart condition, seemed content with not sharing his feelings with anyone, in part because of his gender and in part because of his age. He mentioned feeling isolated because his understanding is that few others his age have this condition. Two other participants sought to debunk the myth:

> *I guess, I live with my girlfriend, so I guess maybe her. I don't really talk. I go to the doctor's, I do my thing. I might come home and then I'll talk about anything, and I don't want to talk about my feelings, and I don't talk to my friends about it, but I've also been like that for 36 years about anything in life that is troubling. I'm also a guy. We tend to do that all the time...It's not a young person's game, I guess, and so I don't have a lot in common with people that are older. I guess that's why I just kind of keep it to myself.*

Rewards and Incentives

The conversation concluded with a discussion of what types of tangible rewards and incentives would motivate behavior change. There was a lively discussion in both groups, and there were many common ideas. When discussing rewards of significant value, travel as a reward was the most widely cited idea—from discounted airfare and other transportation, hotels, to activities at the destination.

An important insight is that the appeal of the reward was not merely the reward itself (ie, a fun trip) but the fact that traveling signifies a freedom and a loosening of restrictions that heart failure places on their lives. Rewards that speak to this would be quite powerful:

XSL•FO
RenderX

> *I would like to be able to go to places that I used to go.*
>
> *I've been talking to the guys in the band about expanding our range, and buying a bus.*
>
> *I would go to Vegas. I haven't been to Vegas.*
>
> *I'd go to Arizona and visit some friends.*

When discussing lower cost rewards, an outing with friends or family to a restaurant was a popular theme:

> *I think I would go to a restaurant with my significant other, and enjoy the beach, and a nice meal – just a getaway, something simple.*
>
> *Go out and eat and have a margarita.*
>
> *Have a hamburger and a glass of wine.*
>
> *I'd definitely have a beer.*
>
> *Yeah, I think that would be the first thing is to go out to a nice dinner, not necessarily McDonald's or In-N-Out, but to a nice place.*

One male and one female participant suggested clothes shopping. This played into a similar notion that it was not just the reward itself (ie, new clothes) that would be motivating but the significance of what purchasing new clothes might represent (ie, weight loss and a renewed sense of self and freedom):

> *I enjoy doing that, and I don't mind the shopping either because I can go out or buy a new outfit in a minute. But the thing is, I'm not ready. I'm not ready. And when I'm ready, that means the weight loss; I can be able to do that.*

Two male participants suggested tickets to a sports game, specifically the Oakland Raiders in the Oakland focus group and the L.A. Dodgers in the Los Angeles focus group. Other participants suggested Japanese fountain pens that run approximately US $20.00, T-shirts, books, house plants, or simply anything fun that was not a simple cash reward:

> *I would spend it on something fun. My mom, she always gave us money for birthdays and Christmas and it was always, "You're not to pay bills with this. You are to spend it on you." And that always felt good.*

Several participants asserted to their desire to spend the incentive on loved ones, in particular, on grandchildren:

> *I'd like to get my kid – her birthday is coming up, you know, so I thought I'd get her something nice for a change, maybe a nice purse or something. And then my grandkids, another toy, you know, they're crazy about toys. So that's basically it. Just to see smiles on their faces.*

At this point in the focus group discussion, the mockups of the user experience were presented to the participants. Afterward, additional suggestions for goals and rewards were sought above and beyond what were shown on screen. Additional rewards suggestions included specific examples of activities for grandchildren, including museum tickets, movie tickets, aquarium passes (Long Beach and Monterey), Disneyland tickets, or an ice cream outing.

Other specific recommendations for rewards included Amazon gift cards, dinner cruises for 2, flowers, credits for cardiac rehab, and a high-end walker. One participant offered a conceptual suggestion of access to exclusive events, products, or services, one that money could not ordinarily purchase. Examples of this may include an invitation to a private dinner party, a local film screening, or literary reading:

> *I think the only thing that I would think of is it's always nice to have access to something that you wouldn't normally have access to, so something that you can get on your own? I don't know what that would be, but some sort of exclusive behind-the-scenes kind of program or something.*

Commitments

After the mockups were shown to the participants, many thought the list of commitments was fairly comprehensive and additional suggestions were tangential to the standard health and wellness goals offered on the platform mockups.

The participants suggested additional goals not shown included commitments to complete household tasks or hobbies such as laundry or putting together photograph albums. A few participants noted the ability to build a network of support through the site and how helpful that may prove to be:

> *Yeah. I would want it to be a network so that I can connect with others who have my same goal. The more you connect with others who have similar experiences, the more helpful it can be.*
>
> *What I like about it is I find that I function better with accountability 'cause if I say something or do something, it's important for me to complete it 'cause I feel better about it...if I'm accountable to someone and telling the truth, 'cause you can't lie about your weight.*

Survey

At the conclusion of the focus group, participants were asked to complete a survey rating their interest in various rewards that would serve as an incentive to behavior change. Rewards fell into one of the following categories: health and fitness, luxury services, sports and leisure, home and family, or other items. Responses were recorded on a scale of 1 to 5, with a 1 denoting *I don't want this at all*, a 3 denoting indifference, and a 5 denoting *I really want this*. All participants completed the survey.

Table 2 displays the medians and interquartile ranges for each reward, which are listed in order of most preferred to least preferred. Items with a median score of 5 include blood pressure or heart rate monitor, theater or symphony or opera tickets, and Fitbit, pedometer, or other activity tracker. Those with a median score of 2 or less included tension ropes and a round of golf at a local course.

XSL•FO
RenderX

Table 2. Preference for incentives.

Possible reward	Participant response (scale 1-5)[a]							Response, mean (SD)
	P1	P2	P3	P4	P5	P6	P7	
Health and fitness								
Blood pressure or heart rate monitor	4	5	5	4	5	5	5	4.71 (0.49)
Fitbit, pedometer, or other activity tracker	4	5	5	4	4	5	5	4.57 (0.53)
Digital scale with body fat calculator	3	5	4	3	3	5	5	4.00 (1.00)
Healthy dinner vouchers (4)	3	5	4	3	3	5	5	4.00 (1.00)
Yoga mat	3	5	3	3	3	3	5	3.57 (0.98)
Nike Plus (workout or run tracking device)	2	4	3	3	2	5	5	3.43 (1.27)
Personal training session at gym	5	3	3	2	4	2	5	3.43 (1.27)
Mountain or racing bike	4	5	4	1	1	5	4	3.43 (1.72)
1-year subscription to *EatingWell* magazine	3	5	3	4	4	2	2	3.29 (1.11)
1-month gym membership	3	3	3	5	3	3	2	3.14 (0.90)
Track suit	2	5	3	3	3	2	4	3.14 (1.07)
Flex ball	2	3	3	3	3	2	2	2.57 (0.53)
Tension ropes	2	3	3	4	2	2	2	2.57 (0.79)
Home and family								
House-cleaning service for 4 weeks	4	5	5	5	4	1	3	3.86 (1.46)
Personal chef for dinner party for your family and friends	2	5	5	1	4	3	5	3.57 (1.62)
Fruit or gift basket sent to your home	3	5	4	2	3	1	1	2.71 (1.50)
Luxury services								
Theater or symphony or opera tickets (2)	5	4	5	4	5	2	5	4.29 (1.11)
Wine tasting tour (2)	4	5	5	2	2	4	5	3.86 (1.35)
1-hour massage	4	5	5	1	4	1	5	3.57 (1.81)
Dinner cruise (2)	3	5	3	4	1	4	4	3.43 (1.27)
Other items								
iPod Touch (16g)	4	5	4	4	4	5	5	4.43 (0.53)
Gift card (eg, Target, Best Buy, Whole Foods, Nordstrom, and Sports Authority)	4	5	4	4	3	5	5	4.29 (0.76)
Kindle or other electronic reader	4	5	4	5	3	5	1	3.86 (1.46)
Recognition of your success in a spotlight section	3	5	4	1	3	4	5	3.57 (1.40)
Compact digital camera	2	5	4	1	3	2	3	2.86 (1.35)
Sports and leisure								
Movie tickets (2)	4	5	2	4	3	4	5	3.86 (1.07)
Tickets to local sports event (eg, Angels, Giants, Lakers, Clippers, Kings) (2)	4	5	2	3	1	5	4	3.43 (1.51)
Weekend spa passes (2)	5	4	5	1	2	1	5	3.29 (1.89)
Disneyland park entries (4)	3	5	5	1	1	1	5	3.00 (2.00)
Weekend organized bike trip (2)	4	4	2	1	1	3	4	2.71 (1.38)
Round of golf at local course (2)	3	2	1	1	1	1	1	1.43 (0.79)

[a]Responses were recorded on a scale of 1 to 5, with a 1 denoting *I don't want this at all*, a 3 denoting indifference, and a 5 denoting *I really want this*.

XSL•FO
RenderX

Discussion

Principal Findings

The focus groups yielded many insights relevant to the development of a patient-centered intervention for the management of heart failure. A few stood out as having particular importance. With regard to medication adherence, regimen complexity and logistics posed a problem for patients as did forgetfulness. These issues may present barriers to individuals' readiness to consider behavior change, a prerequisite to making an actual commitment to change. For instance, a systematic review of past studies concluded that health coaching is one type of intervention that has produced positive effects on patients' physiological, behavioral, and psychological conditions related to chronic illness in general [38] and heart failure specifically [39]. Commitment contracts are congruent with the health coaching process and can help reinforce action once an individual is ready to change. Commitments reinforce mindfulness through report reminders and creating layers of accountability, which will help with forgetfulness. The act of writing down a commitment (on paper or digitally), and the requisite details, will compel patients to put together a plan of action for compliance, which should help manage the complexity and logistical difficulty. Many participants cited the need to push through the first few months of a new regimen to either reduce or get acquainted with the side effects. Commitments that focus specifically on managing the side effects associated with the first few months of a new drug regimen can help improve adherence.

With regard to exercise, participants felt that heart failure imposes significant limitations on their ability to exercise. They also indicated that they receive little specific instruction from care providers. They were unsure of their own boundaries for safe exercise. Likewise, depression is often an obstacle to working out.

Commitment contracts, which address specific types of exercises most heart failure patients can perform, may be useful. Educating physicians on the need to improve communication about exercise is important. Developing a campaign that targets commitments around behaviors that address stress, anxiety, and depression management is also important.

As with exercise, participants reported receiving little instruction from providers with respect to diet. However, participants still felt well informed about what to eat and what not to eat. Limits on sodium intake and handling liquid intake are critical in managing their condition. Dining out with friends and family and temptations of unhealthy food nearby are primary drivers for noncompliance with a diet plan. Family members were not perceived as always understanding of the dietary restrictions that patients need to follow. Having users make a commitment to create a diet plan and share it with close friends and family so that they can better hold them accountable and not be a source of temptation may be an important strategy to improve diet among persons with heart failure.

Interestingly, few participants regularly practiced goal setting. Furthermore, there was little consistency in the goal setting process among the few that said they do practice goal setting. The use of technology, such as wearable devices, was of significant interest to the focus group participants. The development of commitments around device goals may be helpful. For instance, if a user has a commitment to walk 50,000 steps per week, a wearable device can validate this.

Family was considered a poor resource for structuring accountable commitments by the group. Patients often found that friends and family do not understand their conditions sufficiently to hold them accountable. Peer networks (both online and offline), particularly people who suffer from the same ailments, are very powerful centers of influence and accountability. They are seen as partners in their journey toward managing their disease. Use of a peer referee may facilitate greater and more intimate interaction with both online and real-world networks. As a lack of understanding and responsiveness from loved ones is apparent in the responses, creating a framework for better educating loved ones about heart failure in an effort to boost their ability to hold their loved ones accountable may also be needed. Similarly, no participant mentioned favoring personal accountability, which may suggest that a patient's motivation to manage their health is at odds with the helplessness they feel in managing their condition. Encouraging autonomy and personal accountability may facilitate better health management.

Finally, participants were very amenable to leveraging incentives for motivation. Expectations that rewards will encourage better health are fair and justified. Careful development of incentive choice sets seems important. Particularly, there will be a need to develop material rewards that speak to the diverse tastes of individuals with heart failure.

Study Limitations

There are several limitations to this study. Focus group research is qualitative in nature, often providing insufficient insight into the general pattern of behavior to develop of a strong set of intuition, which can be used to refine concepts or assist in the development of new research ideas. As in the case of any qualitative research, the findings presented herein may have been influenced by the context of the focus group. Finally, our findings are limited given the small sample size. We were unable to recruit a greater number of participants because many of the referred patients were too ill or did not have the physical ability to travel to the focus group locations owing to their condition. We hope future work in the chronic heart failure population can address this limitation.

Conclusions

In summary, we evaluated patients' overall outlook concerning their heart condition. We asked about the steps that participants have taken (eg, medication adherence, exercise, and diet) and the specific goals they have set to manage their condition. We studied what and who motivates their behavior. Overall, we found that medication regimens were complex and involved adverse side effects. We also found that providers did not provide exercise and diet recommendations. There was an interest in goal setting but a lack of practice in doing so. Participants favored peer accountability for goals rather than

XSL•FO
RenderX

family member accountability. Finally, financial incentives were viewed favorably as a means to motivate behavior change.

Acknowledgments

Funding support for this study was provided by Novartis Pharmaceuticals. The sponsor was involved in the conception of the study and in the decision to submit the manuscript for publication. The funding agreement ensured the authors' independence in writing the paper.

Authors' Contributions

JND, DM, JG, KKK, DP, HDOC, and RS were responsible for the conception and development of research design. JG and DP drafted the focus group interview content. JG conducted the focus groups, coded the focus group transcripts, and conducted the qualitative analysis. DP managed the recruitment and implementation logistics. All authors contributed to the writing of the manuscript at all stages.

Conflicts of Interest

DP and RM are employees of and JND and DM are consultants for Precision Health Economics, which received funding from Novartis Pharmaceuticals to conduct this work.

Multimedia Appendix 1

Interview tool.

[DOCX File, 24KB - jmir_v21i8e12483_app1.docx]

Multimedia Appendix 2

Sample commitments.

[DOCX File, 23KB - jmir_v21i8e12483_app2.docx]

Multimedia Appendix 3

Sample rewards.

[DOCX File, 22KB - jmir_v21i8e12483_app3.docx]

Multimedia Appendix 4

Codebook.

[DOCX File, 24KB - jmir_v21i8e12483_app4.docx]

Multimedia Appendix 5

Consolidated Criteria for Reporting Qualitative Studies (COREQ): 32-item Checklist.

[DOCX File, 28KB - jmir_v21i8e12483_app5.docx]

References

1. Heidenreich PA, Trogdon JG, Khavjou OA, Butler J, Dracup K, Ezekowitz MD, American Heart Association Advocacy Coordinating Committee, Stroke Council, Council on Cardiovascular Radiology Intervention, Council on Clinical Cardiology, Council on Epidemiology Prevention, Council on Arteriosclerosis, Thrombosis Vascular Biology, Council on Cardiopulmonary, Critical Care, Perioperative Resuscitation, Council on Cardiovascular Nursing, Council on the Kidney in Cardiovascular Disease, Council on Cardiovascular Surgery Anesthesia,Interdisciplinary Council on Quality of Care Outcomes Research. Forecasting the future of cardiovascular disease in the United States: a policy statement from the American Heart Association. Circulation 2011 Mar 01;123(8):933-944. [doi: 10.1161/CIR.0b013e31820a55f5] [Medline: 21262990]
2. Benjamin EJ, Muntner P, Alonso A, Bittencourt MS, Callaway CW, Carson AP, American Heart Association Council on EpidemiologyPrevention Statistics CommitteeStroke Statistics Subcommittee. Heart disease and stroke statistics-2019 update: a report from the American Heart Association. Circulation 2019 Dec 05;139(10):e56-528. [doi: 10.1161/CIR.0000000000000659] [Medline: 30700139]

XSL•FO
RenderX

3. Benjamin EJ, Virani SS, Callaway CW, Chamberlain AM, Chang AR, Cheng S, American Heart Association Council on EpidemiologyPrevention Statistics Committee Stroke Statistics Subcommittee. Heart disease and stroke statistics-2018 update: a report from the American Heart Association. Circulation 2018 Dec 20;137(12):e67-e492. [doi: 10.1161/CIR.0000000000000558] [Medline: 29386200]
4. Government Publishing Office [US]. 2010 Mar 15. Patient Protection and Affordable Care Act Pub. L. No. 111-148, §2702, 124 Stat 119 URL: http://www.gpo.gov/fdsys/pkg/PLAW-111publ148/pdf/PLAW-111publ148.pdf [accessed 2019-03-15]
5. Rahman I, Bellavia A, Wolk A. Relationship between physical activity and heart failure risk in women. Circ Heart Fail 2014 Nov;7(6):877-881. [doi: 10.1161/CIRCHEARTFAILURE.114.001467] [Medline: 25236882]
6. Myers J, McAuley P, Lavie CJ, Despres JP, Arena R, Kokkinos P. Physical activity and cardiorespiratory fitness as major markers of cardiovascular risk: their independent and interwoven importance to health status. Prog Cardiovasc Dis 2015;57(4):306-314. [doi: 10.1016/j.pcad.2014.09.011] [Medline: 25269064]
7. Andersen K, Mariosa D, Adami HO, Held C, Ingelsson E, Lagerros YT, et al. Dose-response relationship of total and leisure time physical activity to risk of heart failure: a prospective cohort study. Circ Heart Fail 2014 Sep;7(5):701-708. [doi: 10.1161/CIRCHEARTFAILURE.113.001010] [Medline: 25185250]
8. Stewart RA, Szalewska D, She L, Lee KL, Drazner MH, Lubiszewska B, et al. Exercise capacity and mortality in patients with ischemic left ventricular dysfunction randomized to coronary artery bypass graft surgery or medical therapy: an analysis from the STICH trial (Surgical Treatment for Ischemic Heart Failure). JACC Heart Fail 2014 Aug;2(4):335-343 [FREE Full text] [doi: 10.1016/j.jchf.2014.02.009] [Medline: 25023813]
9. Arnett DK, Blumenthal RS, Albert MA, Buroker AB, Goldberger ZD, Hahn EJ, et al. 2019 ACC/AHA guideline on the primary prevention of cardiovascular disease: executive summary: a report of the American College of Cardiology/American Heart Association task force on clinical practice guidelines. J Am Coll Cardiol 2019 Mar 17 Epub ahead of print. [doi: 10.1016/j.jacc.2019.03.009] [Medline: 30894319]
10. Fonarow GC, Abraham WT, Albert NM, Stough WG, Gheorghiade M, Greenberg BH, OPTIMIZE-HF InvestigatorsHospitals. Factors identified as precipitating hospital admissions for heart failure and clinical outcomes: findings from OPTIMIZE-HF. Arch Intern Med 2008 Apr 28;168(8):847-854. [doi: 10.1001/archinte.168.8.847] [Medline: 18443260]
11. Jaarsma T, van Veldhuisen DJ. When, how and where should we "coach" patients with heart failure: the COACH results in perspective. Eur J Heart Fail 2008 Apr;10(4):331-333 [FREE Full text] [doi: 10.1016/j.ejheart.2008.02.017] [Medline: 18353718]
12. Savard LA, Thompson DR, Clark AM. A meta-review of evidence on heart failure disease management programs: the challenges of describing and synthesizing evidence on complex interventions. Trials 2011 Aug 16;12:194 [FREE Full text] [doi: 10.1186/1745-6215-12-194] [Medline: 21846340]
13. Lainscak M, Blue L, Clark AL, Dahlström U, Dickstein K, Ekman I, et al. Self-care management of heart failure: practical recommendations from the Patient Care Committee of the Heart Failure Association of the European Society of Cardiology. Eur J Heart Fail 2011 Feb;13(2):115-126 [FREE Full text] [doi: 10.1093/eurjhf/hfq219] [Medline: 21148593]
14. Ngo V, Hammer H, Bodenheimer T. Health coaching in the teamlet model: a case study. J Gen Intern Med 2010 Dec;25(12):1375-1378 [FREE Full text] [doi: 10.1007/s11606-010-1508-5] [Medline: 20862560]
15. Bodenheimer T, Laing BY. The teamlet model of primary care. Ann Fam Med 2007;5(5):457-461 [FREE Full text] [doi: 10.1370/afm.731] [Medline: 17893389]
16. Alm M, Soroudi N, Wylie-Rosett J, Isasi CR, Suchday S, Rieder J, et al. A qualitative assessment of barriers and facilitators to achieving behavior goals among obese inner-city adolescents in a weight management program. Diabetes Educ 2008;34(2):277-284. [doi: 10.1177/0145721708314182] [Medline: 18375777]
17. Miller WR. Motivational interviewing with problem drinkers. Behav Psychother 2009 Jun 16;11(2):147-172 [FREE Full text] [doi: 10.1017/S0141347300006583]
18. Bourdeaux CP, Davies KJ, Thomas MJ, Bewley JS, Gould TH. Using 'nudge' principles for order set design: a before and after evaluation of an electronic prescribing template in critical care. BMJ Qual Saf 2014 May;23(5):382-388. [doi: 10.1136/bmjqs-2013-002395] [Medline: 24282310]
19. Meeker D, Knight TK, Friedberg MW, Linder JA, Goldstein NJ, Fox CR, et al. Nudging guideline-concordant antibiotic prescribing: a randomized clinical trial. JAMA Intern Med 2014 Mar;174(3):425-431 [FREE Full text] [doi: 10.1001/jamainternmed.2013.14191] [Medline: 24474434]
20. Persell SD, Friedberg MW, Meeker D, Linder JA, Fox CR, Goldstein NJ, et al. Use of behavioral economics and social psychology to improve treatment of acute respiratory infections (BEARI): rationale and design of a cluster randomized controlled trial [1RC4AG039115-01]--study protocol and baseline practice and provider characteristics. BMC Infect Dis 2013 Jun 27;13:290 [FREE Full text] [doi: 10.1186/1471-2334-13-290] [Medline: 23806017]
21. Heider F. The Psychology of Interpersonal Relations. New York: Psychology Press; 1958.
22. Festinger L, Carlsmith J. Cognitive consequences of forced compliance. J Abnorm Soc Psychol 1959 Mar;58(2):203-210. [doi: 10.1037/h0041593]
23. Newcomb TM. An approach to the study of communicative acts. Psychol Rev 1953 Nov;60(6):393-404. [Medline: 13112341]
24. Cialdini RB, Cacioppo JT, Bassett R, Miller JA. Low-ball procedure for producing compliance: commitment then cost. J Pers Soc Psychol 1978;36(5):463-476. [doi: 10.1037/0022-3514.36.5.463]

XSL•FO
RenderX

25. Deutsch M, Gerard HB. A study of normative and informational social influences upon individual judgment. J Abnorm Soc Psycho 1955 Nov;51(3):629-636. [doi: 10.1037/h0046408]
26. Cialdini RB. Influence: Science And Practice (5th Edition). In: Allyn and Bacon. Boston: Allyn And Bacon; 2010.
27. Burn SM, Oskamp S. Increasing community recycling with persuasive communication and public commitment. J Appl Soc Psychol 1986;16(1):29-41 [FREE Full text]
28. De Leon IG, Fuqua RW. The Effects of Public Commitment and Group Feedback on Curbside Recycling. Environment and Behavior 2016 Jul 26;27(2):233-250. [doi: 10.1177/0013916595272007]
29. Krendl KA, Olson B, Burke R. Preparing for the environmental decade: a field experiment on recycling behavior. J Appl Commun Res 1992 Feb 01;20(1):19-36. [doi: 10.1080/00909889209365317]
30. Baca-Motes K, Brown A, Gneezy A, Keenan EA, Nelson L. Commitment and behavior change: evidence from the field. J Consum Res 2013 Feb;39(5):1070-1084. [doi: 10.1086/667226]
31. Schwarzwald J, Bizman A, Raz M. The foot-in-the-door paradigm: effects of second request size on donation probability and donor generosity. Pers Soc Psychol Bull 2016 Jul 02;9(3):443-450. [doi: 10.1177/0146167283093015]
32. Greenwald AG, Carnot CG, Beach R, Young B. Increasing voting behavior by asking people if they expect to vote. J Appl Psychol 1987 May;72(2):315-318. [doi: 10.1037/0021-9010.72.2.315]
33. Nyer PU, Dellande S. Public commitment as a motivator for weight loss. Psychol Mark 2010 Jan;27(1):1-12. [doi: 10.1002/mar.20316]
34. Allison ST, Messick DM. The feature-positive effect, attitude strength, and degree of perceived consensus. Pers Soc Psychol Bull 1988 Jun;14(2):231-241. [doi: 10.1177/0146167288142002] [Medline: 30045470]
35. Payne R, Anderson E, Arnold R, Duensing L, Gilson A, Green C, et al. A rose by any other name: pain contracts/agreements. Am J Bioeth 2010 Nov 29;10(11):5-12. [doi: 10.1080/15265161.2010.519425] [Medline: 21104545]
36. 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. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
37. Finlay L. Negotiating the swamp: the opportunity and challenge of reflexivity in research practice. Qual Res 2016 Aug 17;2(2):209-230. [doi: 10.1177/146879410200200205]
38. Kivelä K, Elo S, Kyngäs H, Kääriäinen M. The effects of health coaching on adult patients with chronic diseases: a systematic review. Patient Educ Couns 2014 Nov;97(2):147-157. [doi: 10.1016/j.pec.2014.07.026] [Medline: 25127667]
39. Brodie D, Inoue A. Motivational interviewing to promote physical activity for people with chronic heart failure. J Adv Nurs 2005 Jun;50(5):518-527. [doi: 10.1111/j.1365-2648.2005.03422.x] [Medline: 15882368]

Edited by G Eysenbach; submitted 12.10.18; peer-reviewed by R Wu, D Wang; comments to author 03.02.19; revised version received 30.04.19; accepted 14.05.19; published 06.08.19.

Please cite as:
Meeker D, Goldberg J, Kim KK, Peneva D, Campos HDO, Maclean R, Selby V, Doctor JN
Patient Commitment to Health (PACT-Health) in the Heart Failure Population: A Focus Group Study of an Active Communication Framework for Patient-Centered Health Behavior Change
J Med Internet Res 2019;21(8):e12483
URL: http://www.jmir.org/2019/8/e12483/
doi:10.2196/12483
PMID:31389339

©Daniella Meeker, Jordan Goldberg, Katherine K Kim, Desi Peneva, Hugo De Oliveira Campos, Ross Maclean, Van Selby, Jason N Doctor. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 06.08.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

From Information Seekers to Innovators: Qualitative Analysis Describing Experiences of the Second Generation of E-Patients

Therese Scott Duncan[1*], MSc; Sara Riggare[1*], MSc; Sabine Koch[1*], PhD; Lena Sharp[2,3*], PhD, RN; Maria Hägglund[1,4*], PhD

[1]Health Informatics Centre, Department of Learning, Informatics, Management and Ethics, Karolinska Institutet, Stockholm, Sweden

[2]Division of Innovative Care Research, Department of Learning, Informatics, Management and Ethics, Karolinska Institutet, Stockholm, Sweden

[3]Regional Cancer Centre Stockholm - Gotland, Stockholm County Council, Stockholm, Sweden

[4]Department of Womens and Childrens Health, Uppsala University, Uppsala, Sweden

[*]all authors contributed equally

Corresponding Author:
Therese Scott Duncan, MSc
Health Informatics Centre
Department of Learning, Informatics, Management and Ethics
Karolinska Institutet
Tomtebodavägen 18 A
Stockholm, 17177
Sweden
Phone: 46 73 512 40 63
Fax: 46 8 311101
Email: therese.scott.duncan@ki.se

Abstract

Background: Current health care systems are rarely designed to meet the needs of people living with chronic conditions. However, some patients and informal caregivers are not waiting for the health care system to redesign itself. These individuals are sometimes referred to as e-patients. The first generation of e-patients used the internet for finding information and for communicating with peers. Compared with the first generation, the second generation of e-patients collects their own health data and appears to be more innovative.

Objective: The aim of this study was to describe the second generation of e-patients through exploration of their active engagement in their self-care and health care.

Methods: Semistructured interviews were conducted with 10 patients with chronic conditions and 5 informal caregivers. They were all recruited through a Web-based advertisement. Data were analyzed according to the framework analysis approach, using the 3 concepts of the self-determination theory—autonomy, relatedness, and competence—at the outset.

Results: Study participants were actively engaged in influencing their self-care and the health care system to improve their own health, as well as the health of others. This occurred at different levels, such as using their own experience when giving presentations and lectures to health care professionals and medical students, working as professional peers in clinical settings, performing self-tracking, contributing with innovations, and being active on social media. When interaction with health care providers was perceived as being insufficient, the participants sought support through their peers, which showed strong relatedness. Competence increased through the use of technology and learning experiences with peers. Their autonomy was important but was sometimes described as involuntary and to give up was not an option for them.

Conclusions: Like the first generation of e-patients, the participants frequently searched for Web-based information. However, the second generation of e-patients also produce their own health data, which they learn from and share. They also engage in the innovation of digital tools to meet health-related needs. Utilizing technological developments comes naturally to the second generation of e-patients, even if the health care system is not prepared to support them under these new circumstances.

(J Med Internet Res 2019;21(8):e13022) doi:10.2196/13022

KEYWORDS

consumer health informatics; eHealth; qualitative research; self-care; motivation

XSL•FO
RenderX

Introduction

Background

Many patients, especially those with chronic or long-term conditions, and their informal caregivers experience a need to be actively involved in care provision to co-ordinate contacts with health care professionals and navigate the health care system [1]. Unfortunately, health care systems today are rarely designed to meet the current needs. However, some patients and informal caregivers are not waiting for the health care system to redesign itself. They take matters into their own hands and create innovative approaches and solutions to manage their care and interactions with health care providers, often with the use of electronic health (eHealth) solutions [2]. These individuals are sometimes referred to as e-patients, and in this study conducted in Sweden, we explored how their activities have evolved.

First Generation of E-Patients—Information Seekers

The first generation of e-patients was described by Ferguson and Frydman [3] as citizens who use the internet for searching for health information or use electronic communication tools to solve a personal health-related need. The concept of e-patients includes both patients and informal caregivers who receive *better health information and services, and different (but not always better) relationships with their doctors* [3]. The term *informal caregiver* is used in this study to represent a family member or other person, for example, a close friend, who supports and cares for a patient without being formally employed or reimbursed to do so. The *e* in e-patients primarily stands for *electronic;* however other, more descriptive attributes are mentioned, describing these digitally literate patients and informal caregivers as equipped, empowered, enabled, and engaged in their self-care and in health care [4]. We will use this concept of e-patients in this study, which includes both patients and informal caregivers who have these descriptive attributes. E-patients are still an underutilized resource. However, embracing e-patients' ideas and engagement to a higher degree could potentially improve collaboration with health care providers [3,5] as well as quality of care.

Moving From the First to the Second Generation of E-Patients

Different concepts, sometimes with overlapping definitions, are used in the literature when trying to better understand why patients and informal caregivers actively engage in their self-care and in health care, and how eHealth plays a part. The *highly informed patient* —such as the first generation of e-patients—is described as using the internet to search for information regarding a specific problem and to seek support from peers in online communities [6]. *Expert patients* are described in the early 21st century as patients with chronic conditions who are confident, informed, and knowledgeable and have the skills to take a central role in their self-care and management of care [7]. With a substantial understanding about their condition and context, expert patients have skills in self-care and in working as partners to health care professionals [7,8]. The *digitally engaged patient* is a concept describing that patient engagement encourages the use of digital media technologies for self-care [9]. The concept of a *lead patient* has its roots in the term *lead users* used in design sciences to describe users who face a need before the general market and create their own solutions for this need [10]. The term lead patients is hence used to describe patients or informal caregivers who benefit considerably from finding solutions to their health- or health care–related needs. Lead patients have also been described to contribute to development and important innovations in their self-care and within the health care system [8]. Both digitally engaged and lead patients are examples of the second generation of e-patients. Patients and informal caregivers are also increasingly engaging in the improvement of health care systems. In Sweden today, there are examples of patients employed by health care providers [11]. Their lived experiences and expert knowledge regarding health and care are thereby captured and used to provide a more patient-centered, integrated model of care. Employments of this kind mostly occur within mental health care, for example, the *peer support workers* concept in the United Kingdom [12]. However, our research focuses on a Swedish context. With almost 50% of the Swedish population having chronic conditions [13] and 94% of the population using the internet (with a large variation regarding frequency) [14], there is a great potential for the digitalization of self-care and health care.

Second Generation of E-Patients—Innovators

Eysenbach and Diepgen described a potential for citizens to manage their own self-care and collaboration with health care with the use of different eHealth solutions [15]. These solutions are suggested to contribute to increased information and control, resulting in more empowered patients. Research suggests that when designing new technological solutions, it is important to have an understanding of the needs of the end user [16,17]. These needs are often connected to the users' context and personal capabilities and behaviors [18]. Patients' care experiences may depend on their needs and can be highly individual [19]. Research has also demonstrated that patients are often capable of long-term self-care [20,21] as well as being important contributors to health care development and care delivery [20,22-24]. We see a growing movement of patients and informal caregivers using or creating technological innovations adapted to their specific needs, such as the diabetes patient Dana Lewis who created an artificial pancreas with an open-source approach [25]. They could be described as lead patients or e-patients who have taken their engagement to a new level. However, there is a lack of knowledge about how this second generation of e-patients goes beyond searching for and sharing health information on the Web.

We therefore aimed to describe the second generation of e-patients through exploration of their active engagement in self-care and health care.

Methods

A qualitative approach with semistructured interviews was used to get a deeper understanding of the second generation of e-patients to describe them.

XSL•FO
RenderX

Recruitment and Sampling

We purposely recruited participants who could be considered as being part of the second generation of e-patients. The recruitment was performed through convenience and snowball sampling [26]. A Web-based advertisement was published (a website reaching 9054 unique visitors), and through a newsletter (reaching 1500 subscribers). Patients and informal caregivers in Sweden were targeted. Active patients or informal caregivers could either volunteer or be suggested by someone else. A short description of the characteristics of an e-patient was provided (Multimedia Appendix 1). We received 67 suggestions in total, whereof 12 were suggestions from someone else (other patients). A purposeful selection was performed to cover different chronic conditions, different ways of being actively engaged, different locations in Sweden, being a patient or informal caregiver, gender, and age. Overall, 7 out of 15 selected participants were suggested by someone else.

Data Collection

An interview guide was developed by 3 of the authors (TSD, SR, and MH) based on the literature regarding e-patients. The interview guide consisted of 4 themes: *Background*, *Your health journey*, *Health behavior*, and *Your role in self-care and health care*. The questions were open ended, with the aim of letting the participants talk about what matters most to them. This enabled the researchers to capture narratives that were not foreseen [27]. The interview guide was pilot tested with 5 patients and informal caregivers. The purpose of this pilot testing was to improve the interview guide, and the questions in the guide were slightly altered based on feedback from the pilot interviews. The data from the pilot interviews were not included in the analysis or in the results of the study. Data from the 15 interviews were collected from October to December 2017 through interviews conducted face-to-face (n=6) or over telephone (n=9), depending on the participants' preferences. The first author and 4 coworkers from a project called "Lead patients", all with experience of qualitative research interviewing, conducted the interviews. The semistructured interviews took an average of 40 min each. All the interviews were recorded, and the recorded files were transcribed verbatim.

Data Analysis

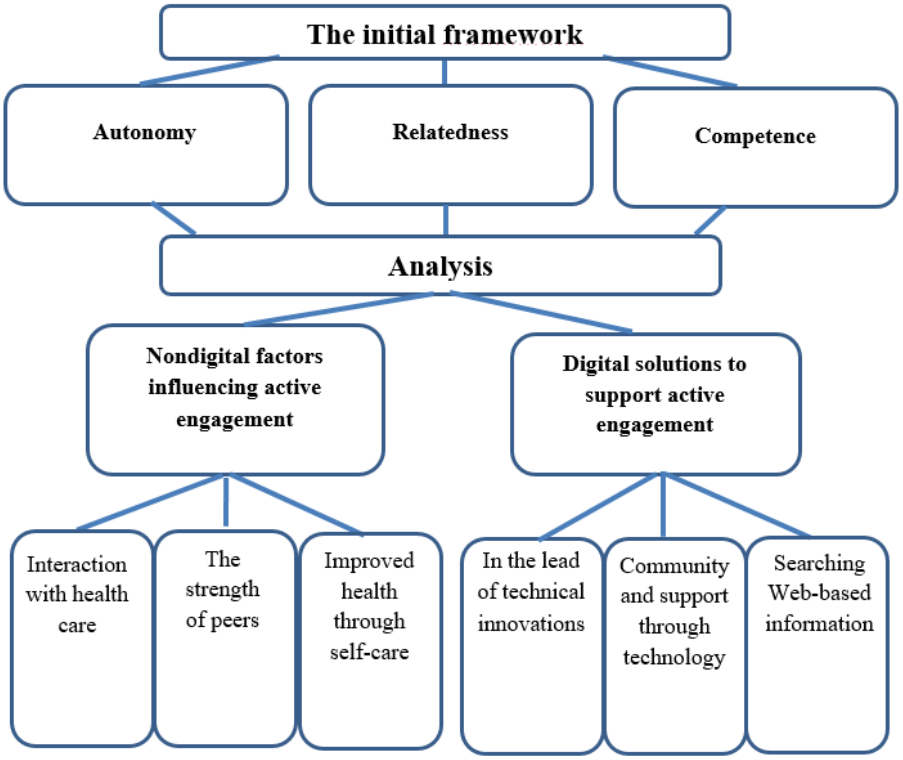
The collected data were analyzed according to the framework analysis (FA) approach, with a focus on identifying the occurrence of different concepts regarding e-patients' active engagement in health and care [28,29]. The analysis process began by producing an initial framework from what was already known to the authors: the second generation of e-patients is highly motivated to be actively engaged in their self-care and in health care. Therefore, we used the self-determination theory (SDT) as a motivational theory and initial framework for our study.

Initial Framework

The 3 basic psychological needs from SDT—autonomy, relatedness, and competence—provided a start to describe e-patients' motivation [30,31] (Figure 1). Using SDT as an initial framework enabled us to explore how participants engage to meet these psychological needs [32]. Autonomy can be described as the capacity and independence to make decisions. Relatedness implies feeling a meaningful connection with others. Competence comprises feeling efficient and able to master difficult situations. The theory argues that these basic needs provide the energy and direction for individuals to act to satisfy their psychological needs, a long-lasting motivation that facilitates persons to value the activity itself and not only an outcome of the activity [30-33].

XSL•FO
RenderX

Figure 1. Themes and categories building on the onset of the 3 basic psychological needs from the self-determination theory.



Organizing Data

We followed the 5 steps according to the framework analysis model (familiarization, identifying a thematic framework, coding and indexing, charting, and mapping and interpretation) [28,29] using SDT as an initial framework. As a first step, we familiarized ourselves with the transcribed data, and each interview was given a label. To categorize the content from the interviews, the data were organized into themes, categories, and subcategories (Multimedia Appendix 2). It was an iterative process as we did not want to force the data to fit into the initial framework. The themes were updated multiple times during the analysis process. The meaning, relevance, and importance of the themes were considered and compared with the aim of the study. Sections of data were identified and indexed to match the different themes. All data were indexed, while keeping the links to the specific interview by labels. To ensure that all data were included, color coding was applied to the raw data in the transcribed material. By mapping and interpreting the coding, it was possible to define concepts and find relationships. The beliefs of the participants regarding their reasons for being e-patients and what support they gained from different eHealth solutions were recited [28,29]. To ensure trustworthiness in the analyzed data, at least 2 of the authors were involved at each stage of the analysis. Translation into English was done after analyzing the data, and the translation was verified by all the authors.

Ethical Considerations

The participating e-patients were provided with written information concerning the purpose of the study and how the researchers would manage their data before signing the informed consent. It was emphasized that participation was voluntary and that it was possible to withdraw at any time without explanations. The requirement of confidentiality was fulfilled. On the basis of a decision from the ethical board in Stockholm, legislation regarding ethical review was not applicable to this study (decision 2015/1572-31/4).

Results

Demography and Description of Themes

Semistructured interviews (n=15) were conducted with patients with chronic conditions (n=10) and informal caregivers (n=5). The participants had different diagnoses (Table 1), gender (10 females, 5 males), and age (range 34-77 years; average age per category was as follows: female patients, 50 years; male patients, 61 years; female informal caregivers, 52 years; and male informal caregivers, 64 years).

Subcategories and categories were formed from the data, as described in the Methods section, and 2 overall themes emerged (Figure 1; Multimedia Appendix 2). The theme *nondigital factors influencing active engagement* has 3 categories describing the second generation of e-patients. Their active engagement was influenced by the interaction with health care providers, the strength of peers, and the focus on their own well-being through self-care. The theme *digital solutions to support active engagement* represents the participants' relationship to eHealth solutions. This theme includes innovations by the second generation of e-patients and how support and learning aspects can be achieved from online communities as well as when finding information on the internet.

XSL•FO
RenderX

Within the themes, we found the 3 basic psychological needs from the SDT: autonomy, relatedness, and competence. Going from the initial framework with motivated second generation of e-patients, we found more components to why, and how, they actively engage in their own or others' health and care.

Table 1. Demography of participants. Participant identification consists of M: male; F: female; P: patient; and I: informal caregiver.

Participant identification	Type of informal caregiver	Diagnosis
FP1	—[a]	Connective tissue disease
FP2	—	Fatigue syndrome + fibromyalgia
FP3	—	Systemic sclerosis
FP4	—	Mental illness
FP5	—	Irritable bowel syndrome + motility disorder
FP6	—	Parkinson disease
FP7	—	Rheumatic disease
MP1	—	Myocardial infarction
MP2	—	Kidney failure + kidney cancer
MP3	—	Multiple sclerosis
FI1	Mother	Hypersensitivity
FI2	Wife	Thymus neoplasms
FI3	Mother	Down syndrome + heart failure
MI1	Husband	Mental illness
MI2	Husband	Pulmonary fibroses + liver cancer

[a]Missing data, as they are not informal caregivers.

Nondigital Factors Influencing Active Engagement

Looking into how the participants actively engaged in their self-care and in health care, we found that they strive for an adaptable health care system, wished to collaborate with health care providers (regardless of whether it was primary or specialized care), and wanted to receive feedback regarding their self-care. These 3 aspects were often described as problems needed to be solved. Other important aspects in the patients' and informal caregivers' lives were also mentioned: to take a greater responsibility regarding health, to be engaged with peers, to tell their story, and to increase their knowledge. These aspects were often considered as solutions for the problems they experienced. Some participants also described how they strived to be diagnosed correctly and to be believed and taken seriously when in contact with primary health care and the society. All 3 basic psychological needs of SDT (autonomy, relatedness, and competence) were brought up as motivational factors for their active engagement. However, the most consistent factor was to care for others—relatedness. The participants expressed how they wanted to share their experiences with both their peers and the health care professionals. In addition to the 3 SDT needs, all informal caregivers and some patients described how they felt that their motivation to be active was not even a choice.

Interaction With Health Care

The participants experienced that practice and knowledge regarding different diseases and conditions was rather low at different health care providers, especially within primary care. The participants expressed that they understood how difficult it may be for health care professionals, regardless of primary or specialized care, to have specific knowledge regarding all diseases. However, these patients meant that they were often confronted with disbelief regarding their specific situation and wished to be approached with more respect regarding their symptoms. Hence, the patients expressed how they needed a more adaptive system that could meet their different requests in a better way. Some of the participants had observed unsafe situations when incorrectly diagnosed and treated both in primary and specialized care, resulting in long-winding situations with misunderstandings. In addition, they sometimes described the health care system as not being able to meet different expectations of engagement from the patients' side:

> *The interaction with healthcare is a lot about calling someone on the phone. However, many of the people I've met in psychiatry have affective and social difficulties, and for them that is very difficult.* [MI1]

Most of the participants, during some limited period, experienced lack of support and sustainability in their relationship with primary or specialized health care:

> *I got a lot of help [from healthcare], but it wasn't the kind of help that I needed.* [FP4]

This together with previous experiences of incorrect diagnosis and treatment, motivated a few of the participants to regularly read their medical records online, to ensure correct diagnosis and treatment.

> *I always log into my health record and check the content. I have also started to record all my conversations in healthcare, to confirm what we actually discussed...* [FP5]

XSL•FO
RenderX

Being actively engaged was described as a means for the patients to solve health-related problems or misunderstandings at different health care providers. However, it seemed important not to dwell on their own situation or resort to self-pity and apathy. Several participants mentioned how they searched for research regarding their conditions, to increase their knowledge as well as to provide their physician with more information. Participants wanted to share information with their health care professionals as they considered them not having the time to search for new or existing research:

> *I have read a lot of research, and when I had a physician that was interested in research, we could use that new research to decide on my medication together.* [FP3]

Some of the participants stressed the importance of making their own judgements and not only listening to the health care professionals. These participants wished that health care professionals could be more constructive by encouraging patients to propose their own ideas and solutions and to give support in this process. To influence the direction of future research, some patients wanted to contribute to new ways of generating research ideas. Therefore, they had chosen to be research partners. However, their agenda was often not met:

> *The idea was really that I would give input on how the result can be communicated in a way so that patients understand... My own agenda was to see how much influence I can have. Where is the line for how critical I can be as a patient and layman, albeit knowledgeable?* [MP1]

One mentioned aspect was also to learn from each other, both from health care professionals and from peers, as well as to learn from different situations. Several participants experienced that when their gained knowledge was acknowledged and valued and they felt listened to, they increased their relatedness and continued to be active. Most of the participants expressed that they gained a lot of their knowledge from going through problematic and demanding situations. They would find meaning from these difficulties and challenges as they provided opportunities for learning:

> *My strategy is to learn from what happens... and as long as I learn from these difficult situations I go through, I just can't lose.* [FI3]

Some of the participants described how they are frequently lecturing or teaching in different contexts, with the goal of improving the health care system. With a perspective that comes from a different side of the health care system, they expressed how they gave inspiration and were highlighting new perspectives of existing problems:

> *My role is to be a catalyst for inspiration and change... I call it an "Accident Investigation Authority."* [MI2]

Some patients described that they were working as professional patient peers in specialized clinical contexts. They explained how they assisted other patients in clinical settings, as it could be easier to tell professional peer patients about their needs. One of the patients explained that her role as a patient peer was to encourage others to use different digital solutions for their self-care and to be prepared for their clinical encounter:

> *When new patients arrive to the clinic, I can help them with their questions, since I have different knowledge and perception than healthcare professionals... That means a lot for the patients.* [FP7]

Some participants also mentioned the importance of being a patient representative and collaborating with health care providers in quality improvement work:

> *...it's a part of how I lead my life, we have to help each other out. That is what keeps me going.* [FP3]

Strength of Peers

Several participants explained how they found inspiration and strength in the relationship with peers through different online communities. This included support on how to navigate the health care system and coordinate their care within different health care contexts or to get suggestions and support for their self-care. The participants were sometimes looking for this support when they experienced insufficient collaboration or interaction with health care professionals or providers.

One patient reasoned about the importance of peer support before an encounter with health care professionals. If several peers had experienced the same problem, it was easier to discuss it with health care professionals, knowing they were not alone regarding their problems legitimized the issue:

> *...all these patients I talk to in social media... their experiences, reflections and stories are really important. They make me feel less lonely.* [FP6]

Just as important as finding strength in peers, the participants also found strength in helping others. Several participants found meaning in telling their story to the rest of the world, and by doing so, they indirectly helped other peers. By being peers themselves, the participants stated that they can make a difference with their story.

Improved Health Through Self-Care

Several patients expressed that self-care was often part of a life-changing process. They believed that self-care would help and even be necessary in reaching their goals of increasing or maintaining their health:

> *Healthcare professionals had given up on me before I even entered the room... After that I met a very competent doctor and I asked: What can I do myself? That is the message I have communicated for the last 15 years...* [MP2]

Both patients and informal caregivers often had concerns for their families, and sometimes being active was described as a result of guilt:

> *It was a defining moment... I realised that there is actually something I could do to improve my chances of a good life. And I felt that I owe my kids that, at least to give it a try.* [MP3]

XSL•FO
RenderX

Through extensive experience in self-care, several participants had learned how to recognize signs that needed attention. These signs could be the key for a correct diagnosis or dose of medication or for assessing when acute help is needed. Some participants also described how they learned about their condition by solving a puzzle for their specific situation:

> *They [health care providers] had answers for each separate issue, but I could see that all these separate issues are somehow connected to each other.* [FP3]

The informal caregivers described taking on heavy responsibility to ensure good quality of care and self-care for their next of kin. They expressed how they felt that they did not have a choice and that they needed to be in control, otherwise they believed everything would collapse:

> *I don't believe they would make it without me... It is always me taking care of health care and school, while my husband is doing things I don't have time for at home. I'm the one with the whole responsibility.* [FI1]

To track health, well-being, and medications, some patients described how they used self-tracking to accomplish self-care. Furthermore, some participants indicated that self-care focused on keeping track of their limits and adjusting their daily activities to their current ability. These activities included work, taking care of family, food, physical activity, and other aspects of daily living. If it exceeded their limits, there could be setbacks or relapses. The participants also increased their expert knowledge through research within the field and acted accordingly in relation to lifestyle, for example, diet and physical activity. Several participants had the impression that it is important to be in control of their lives, both in relation to their disease and to life in general:

> *It's more about recovering and learning how to deal with your life so that it doesn't consume you.* [FP4]

All e-patients explained how they chose not to give up and to use their coping strategies to the best of their ability. To do so, some of them separated their chronic condition from their personality and others went into a role facing the most difficult situations:

> *For me those are two different things: how I am as a person, my personality, and how my body works and the limitations my disease brings. They are not one and the same.* [FP2]

However, there were also descriptions of different periods in the participants' lives when they chose not to be in control of their disease.

Digital Solutions to Support Active Engagement

Several participants expressed that they need to be able to use different kinds of technical solutions as these solutions played an important role in their self-care, in communication with health care professionals, and when reaching their peers. The participants all used different eHealth solutions or had ideas for new solutions to achieve all of the above.

In the Lead of Technical Innovations

All e-patients explained how they saw the potential for future innovations—technical or nontechnical—that would help them and others. Among them, 4 participants expressed ideas of new improvements or had already developed new technical innovations. Several e-patients reported that they were digitally helping peers with their self-care and care coordination or that they were supporting health care providers in their improvement work. In addition, 2 participants had developed their own digital solutions:

> *I have digitalized a questionnaire for primary care to be able to decide where to send a referral for rare diseases to specialized care, in order for the patients to get the correct diagnosis faster.* [FP1]

> *I programmed a web page that I run for my peers... I do it to facilitate for people to share experiences to help each other. I actually have such a close relationship with my doctor that he answers questions from the community on the web page.* [MP3]

Two informal caregivers had ideas for innovations that could better satisfy their needs. These were models for improved engagement and information exchange between patients, informal caregivers, and health care professionals that displayed and took into account the patients' daily experience over time:

> *...I have a great need for an easy method to keep track of side effects. If I had the strength and the energy to do it, I would have created an app for it...* [FI2]

> *There are many things to remember since the last encounter, and that is completely hopeless for many people. It's impossible... I have an idea of using activity trackers for people with mental health issues, to register important aspects of the disease automatically, as an objective measurement...* [MI1]

Community and Support Through Technology

The participants often mentioned social media as a means to find more or less formal communities of peers. In these communities, some participants wrote to inspire or drive for improvements or communicated knowledge gained during their time as patients and informal caregivers. The participants described their use of social media, writing blogs, producing debate articles, writing newspapers and books, and creating a webpage for their peers as support for their activity. Often, their narratives were related to a health care system or societal concern or a problem regarding their health. It could also be aimed at educating peers in self-care:

> *I have educated many peers through the Internet, so they can answer the most common questions regarding our condition. However, communities still want me to answer more complicated issues... I have become a source of information.* [FI1]

Other participants explained how they used their knowledge and perceptions to educate others through presentations at different meetings and conferences. However, this is described as physically challenging for some, and they explained how

XSL•FO
RenderX

they used video conferences instead to communicate their knowledge:

> *I'm bedridden six to nine months a year, so technology is crucial for me in order to be active... I use video calls a lot.* [FP1]

Different types of solutions were used to facilitate self-care, such as devices that track health-related changes, for instance, blood pressure, oxygen uptake, heart rate variability; use of technical support for disabilities; and accessing their electronic health records on the Web.

Searching Web-Based Information

Most patients and informal caregivers searched for information on the internet. They were pursuing and finding relevant information all over the world thanks to the internet, which required the participants to learn different languages and understand different contexts. The information could be research or life experiences from peers in different communities or social media. However, the participants explained how they avoided blindly following any information they found, instead they searched for patterns regarding specific health issues. This selective approach was reported by some to be a result of previous negative experiences of being misdiagnosed:

> *I went home from the health care visit and started googling Fibromyalgia. However, I felt that I didn't belong there—it was not my diagnosis.* [FP3]

The participants explained how they found information on the Web that increased their understanding of how the body works. Despite all the support from Web-based communities and information they found when searching the internet, the e-patients reported that they also needed feedback and collaboration with health care professionals regarding this information:

> *The answer can never be not to google. It has to be: let's talk about this—how can we relate to this?* [FI2]

Discussion

Principal Findings

This qualitative study contributes to a better understanding of the second generation of e-patients and what they do in their active engagement. The participants engage with others in mutual concerns—family members, peers, and sometimes health care providers. These e-patients describe how they move from problems to finding solutions for their interaction with health care providers and self-care. When interaction with health care providers is perceived as being insufficient, they seek support from their peers. This shows that their relatedness seems to be strong, thanks to the strength of peers, as their collaboration with health care professionals is rather weak. Competence increases through the use of technology and from learning experiences with peers. Their autonomy is important; however, it is sometimes described as involuntary and that it is not really an option to give up. This is because of social contexts and what is expected of them within these contexts, which can be seen as powerful interventions for becoming motivated [32], as the participants sometimes experience that they do not have a choice. There are always different aspects of the social context—the health care system, eHealth solutions, and peers—that will enhance or undermine the individuals' possibilities to be engaged and active. The relationship between contextual factors and the individuals' psychological needs will thus affect their well-being and development. One of the factors influencing the participants' active engagement is striving for an adaptable health care system to support them in their self-care and to be listened to. The use of different eHealth solutions are described as seeking information on the Web, use of apps for managing disabilities or to be involved with health care providers, self-tracking, writing blogs, Web-based access to medical records, social media, or programming their own solutions.

The second generation of e-patients is actively engaged in influencing both their own self-care and the health care system to improve the health of themselves and others. This is done at different levels, such as using their own knowledge and experience and lecturing for health care providers, working as professional peers in clinical settings, and performing self-tracking to give their health care providers a possibility to understand their situation by interpreting the patients' own data. The first generation of e-patients searched for health information, or resources, on the internet or used online communication with peers [3]. Shaw et al described the use of eHealth such as *health in our hands*, *interacting for health*, and *data enabling health* [21], which lead us to the second generation of e-patients. They still search for health information and communicate health information in different contexts. However, the second generation also generates information themselves, such as described in *health in our hands* [21]. They are enabled to use different digital devices for performing self-tracking, share their knowledge on the Web, have ideas about innovations for solving their needs, and some of the patients have used their programming skills to develop innovations both for their peers and for health care providers. Even though health care is a complex system, and not always suitable for individual solutions, the second generation of e-patients has the knowledge of the system through their lived experience—generally being part of the system for a long period of time. The second generation of e-patients is not only looking for solutions through the internet, this generation is also equipped with providing solutions by sharing information from their own data and creating new interventions for their direct needs. We can compare the evolution of e-patients with that of the World Wide Web, where the Web 2.0 represents the move from static Web pages to dynamic pages with user-generated content and a participatory culture [34]. Similarly, the early e-patient movement has evolved from searching for content and support online to generating knowledge and solutions that are shared. They are engaged, equipped, enabled, and empowered early adopters and innovators that meet their own needs as well as the needs of others. We can see how the participants make purposive choices and act upon these choices in the sense of how the World Bank describes empowerment [35].

The use of digital solutions is one aspect of the social context that might enhance or undermine the possibility to act to satisfy the psychological needs. The results of this study do not provide

XSL•FO
RenderX

us with information of whether the use of digital solutions encourages intrinsic motivation or if already perceived motivation lead to an increased use of digital solutions. There are examples from the literature where digital solutions do not increase motivation. Choi et al explored how well a smoking cessation app met the psychological needs and found that it did not encourage autonomous motivation [36]. Still, we argue that the use of digital solutions for self-care should be supported and encouraged by health care providers [37,38] to give a stronger sense of autonomy. This could include guiding patients and informal caregivers within the fragmented health information on the Web [39]. However, building a care relationship can be more or less difficult as actively engaged patients can be challenging for health care professionals [40]. We found that the motivation for the participants varied over time and often depended on context and could therefore not be considered a persistent state. Differences in biology, personality, and individual differences also influence consistency [30]. Threats or opposition could challenge any kind of motivation [41]. Therefore, it is important to have encouragement from both health care professionals and peers [37,38].

Strengths

The findings of this study provide a broader view of the second generation of e-patients and their active engagement. Most of the 15 participants could be seen as representatives for other digital pioneers within the population of the second generation of e-patients. The concept of e-patients includes informal caregivers as participants. By including informal caregivers, we also obtained views from a group of people often not included in research, and it gave us a broader view by identifying the needs of the severely ill patients that the informal caregivers care for. To learn about the second generation of e-patients and their active engagement enabled us to see how they are slowly transforming the health care system through innovations and being early adopters. Here, we found an interesting interplay between 3 types of actors: the health care system, peers, and the e-patient him or herself.

Limitations

The result could be affected by the participants' different chronic conditions as different diseases may give different incentives to actively engage. However, we consider the variety of health issues represented in the study as a strength, and we were able to capture similarities across diagnoses. Regardless of diagnoses, several participants had experienced difficulties in being correctly diagnosed. This could be seen as a need specific for patients with rare diseases; however, our data indicate that it also happened to patients with more common diseases. This was a reason for engaging regardless of condition, such as for patients with mental illness, fibromyalgia, rheumatic disease, cancer, or rare diseases. We did not divide the result between patients and informal caregivers and compared them, even though their needs ought to be different [32]. This is because of the individual differences within the group of informal caregivers as they all had different prerequisites: being a mother to an adult or to a child or being a husband to someone still alive or not. This could however be relevant with another study design, to see the differences within the group of e-patients. There was a rather small group of informal caregivers volunteering (n=5); however, they were all included in the study. Participants were mostly represented by female patients (n=7), which was the dominant group in the recruitment process (n=56). An obvious limitation with a study population of 15 participants is that it is rather challenging to generalize the result to a larger population, even though saturation was reached [19]. Further research is therefore necessary to provide more knowledge about the second generation of e-patients, to further explore individual differences within the concept, and to look into different aspects of possible impact on health care settings. It is also important to see how the health care system, the society, eHealth developers, and research can meet the needs of the second generation of e-patients. This will lead to better solutions, both organizational and digital or technical, for all different kinds of patients and their informal caregivers.

Conclusions

This study contributes to a better understanding of the second generation of e-patients and how they actively engage in their self-care and in health care. This enabled us to see how they are seeking collaboration with health care professionals and peers and how they interact through innovations. The participants are actively engaged in influencing health care to improve the health care system and their own and others' health. This indicates that their relatedness seems to be strong. Their competence increases through the use of technology and self-care and from learning experiences with peers and sometimes with health care providers. Their autonomy is sometimes described as involuntary and not possible to give up. It is described as important to have a sufficient collaboration with health care professionals and peers and to have support for their active engagement to continue being motivated and to be an e-patient. Similar to the first generation of e-patients, the participants frequently search for Web-based information. However, the second generation of e-patients also produces their own health data, which they learn from and share, as well as comes up with ideas for digital innovations to meet health-related needs. Taking advantage of the technological development comes naturally for the second generation of e-patients, even if the health care system is not prepared to meet them on these new terms.

Acknowledgments

The authors wish to thank Hans Lindqvist (Quality Registry Center Stockholm), Cristin Lind (Quality Registry Center Stockholm), Jon Engström (Stockholm University), and Anna Thies (Karolinska Hospital) from the project *Lead patients* for assistance with the interviews and analyses. This work was financially supported by Vinnova, the Swedish Governmental Agency for Innovation Systems, through their support of the project *Lead patients* (grant number 2017-01221), which authors TSD, SR, LS, and MH are part of.

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1
The recruitment text.
[PDF File (Adobe PDF File), 95KB - jmir_v21i8e13022_app1.pdf]

Multimedia Appendix 2
The categorization process of the themes.
[PDF File (Adobe PDF File), 184KB - jmir_v21i8e13022_app2.pdf]

References

1. Holman H, Lorig KR. Patient self-management: a key to effectiveness and efficiency in care of chronic disease. Public Health Rep 2004;119(3):239-243 [FREE Full text] [doi: 10.1016/j.phr.2004.04.002] [Medline: 15158102]
2. Graffigna G, Barello S, Triberti S, Wiederhold BK, Bosio AC, Riva G. Enabling ehealth as a pathway for patient engagement: a toolkit for medical practice. Stud Health Technol Inform 2014;199:13-21. [doi: 10.3233/978-1-61499-401-5-13] [Medline: 24875682]
3. Ferguson T, Frydman G. The first generation of e-patients. Br Med J 2004 May 15;328(7449):1148-1149 [FREE Full text] [doi: 10.1136/bmj.328.7449.1148] [Medline: 15142894]
4. Hoch D, Ferguson T. What I've learned from e-patients. PLoS Med 2005 Aug;2(8):e206 [FREE Full text] [doi: 10.1371/journal.pmed.0020206] [Medline: 16060721]
5. Riggare S. E-patients hold key to the future of healthcare. Br Med J 2018 Feb 26;360:k846. [doi: 10.1136/bmj.k846] [Medline: 29483151]
6. Roper L, Jorm C. Please leave your phone outside: policymakers and medical app providers are encouraging patient participation, but doctors in the consult room are not. BMJ Innov 2017 Mar 24;3(2):65-70. [doi: 10.1136/bmjinnov-2016-000134]
7. Shaw J, Baker M. 'Expert patient'--dream or nightmare? Br Med J 2004 Mar 27;328(7442):723-724 [FREE Full text] [doi: 10.1136/bmj.328.7442.723] [Medline: 15044270]
8. Engström J. DiVA Portal. 2014. Patient Involvement and Service Innovation in Healthcare URL: http://www.diva-portal.org/smash/get/diva2:717875/FULLTEXT01.pdf [accessed 2018-07-27]
9. Lupton D. The digitally engaged patient: self-monitoring and self-care in the digital health era. Soc Theory Health 2013 Jun 19;11(3):256-270. [doi: 10.1057/sth.2013.10]
10. von Hippel C. A next generation assets-based public health intervention development model: the public as innovators. Front Public Health 2018;6:248 [FREE Full text] [doi: 10.3389/fpubh.2018.00248] [Medline: 30234092]
11. Krey J. Dagens Medicin. 2018. Norrbotten Vill Anställa en Patient URL: http://www.dagensmedicin.se/artiklar/2018/10/31/norrbotten-vill-anstalla-en-patient/ [accessed 2019-04-23]
12. Repper J. Recovery Context Inventory. 2013. Peer Support Workers: Theory and Practice URL: https://recoverycontextinventory.com/images/resources/ImROC_peer_support_workers_theory_practice.pdf [accessed 2019-04-23]
13. Vårdanalys. 2018. Stöd På Vägen: En Uppföljning Av Satsningen På Att Förbättra Vården För Personer Med Kroniska Sjukdomar URL: https://www.vardanalys.se/wp-content/uploads/2018/06/2018-3-St%C3%B6d-p%C3%A5-v%C3%A4gen_webb.pdf [accessed 2019-04-17]
14. Svenskarna Och Internet. 2018. Super-Connected and Rarely Connected – Online Divide in 2018 URL: https://2018.svenskarnaochinternet.se/the-swedes-and-the-internet-2018-summary/ [accessed 2019-04-17]
15. Eysenbach G, Diepgen TL. The role of e-health and consumer health informatics for evidence-based patient choice in the 21st century. Clin Dermatol 2001;19(1):11-17. [doi: 10.1016/S0738-081X(00)00202-9] [Medline: 11369478]
16. Koch S, Vimarlund V. Critical advances in bridging personal health informatics and clinical informatics. Yearb Med Inform 2012;7:48-55. [doi: 10.1055/s-0038-1639430] [Medline: 22890341]
17. Eysenbach G. Consumer health informatics. Br Med J 2000 Jun 24;320(7251):1713-1716 [FREE Full text] [doi: 10.1136/bmj.320.7251.1713] [Medline: 10864552]
18. Valdez RS, Holden RJ, Novak LL, Veinot TC. Transforming consumer health informatics through a patient work framework: connecting patients to context. J Am Med Inform Assoc 2015 Jan;22(1):2-10 [FREE Full text] [doi: 10.1136/amiajnl-2014-002826] [Medline: 25125685]
19. Hägglund M, Bolin P, Koch S. Living with lung cancer--patients' experiences as input to ehealth service design. Stud Health Technol Inform 2015;216:391-395. [doi: 10.3233/978-1-61499-564-7-391] [Medline: 26262078]
20. Clancy CM. Patient engagement in health care. Health Serv Res 2011 Apr;46(2):389-393 [FREE Full text] [doi: 10.1111/j.1475-6773.2011.01254.x] [Medline: 21371026]

XSL•FO
RenderX

21. Shaw T, McGregor D, Brunner M, Keep M, Janssen A, Barnet S. What is ehealth (6)? Development of a conceptual model for ehealth: qualitative study with key informants. J Med Internet Res 2017 Oct 24;19(10):e324 [FREE Full text] [doi: 10.2196/jmir.8106] [Medline: 29066429]
22. Crawford MJ, Rutter D, Manley C, Weaver T, Bhui K, Fulop N, et al. Systematic review of involving patients in the planning and development of health care. Br Med J 2002 Nov 30;325(7375):1263 [FREE Full text] [doi: 10.1136/bmj.325.7375.1263] [Medline: 12458240]
23. Davis K, Schoenbaum CS, Audet AM. A 2020 vision of patient-centered primary care. J Gen Intern Med 2005 Oct;20(10):953-957 [FREE Full text] [doi: 10.1111/j.1525-1497.2005.0178.x] [Medline: 16191145]
24. Barello S, Graffigna G. Patient engagement in healthcare: pathways for effective medical decision making. Neuropsychol Trends 2015 Apr;17(1):53-65. [doi: 10.7358/neur-2015-017-bare]
25. Litchman ML, Lewis D, Kelly LA, Gee PM. Twitter analysis of #OpenAPS DIY artificial pancreas technology use suggests improved A1C and quality of life. J Diabetes Sci Technol 2019 Mar;13(2):164-170. [doi: 10.1177/1932296818795705] [Medline: 30198751]
26. Marshall C, Rossman GB. Designing Qualitative Research. Thousand Oaks, CA: Sage Publications; 2014.
27. Kvale S. Doing Interviews. Volume 2. Thousand Oaks, CA: Sage Publications; 2008.
28. Ritchie J, Spencer L. Qualitative data analysis for applied policy research. In: Bryman A, Burgess RG, editors. Analyzing Qualitative Data. London, UK: Routledge; 1994:173-194.
29. Ward DJ, Furber C, Tierney S, Swallow V. Using framework analysis in nursing research: a worked example. J Adv Nurs 2013 Nov;69(11):2423-2431. [doi: 10.1111/jan.12127] [Medline: 23517523]
30. Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. Am Psychol 2000 Jan;55(1):68-78. [doi: 10.1037/0022-3514.34] [Medline: 11392867]
31. Deci EL, Ryan RM. Intrinsic Motivation and Self-Determination in Human Behavior. New York: Plenum Press; 1985.
32. Deci EL, Ryan RM. Levels of analysis, regnant causes of behavior and well-being: the role of psychological needs. Psychol Inq 2011 Feb 28;22(1):17-22. [doi: 10.1080/1047840X.2011.545978]
33. Reeve J. Understanding Motivation and Emotion. Seventh Edition. Hoboken, New Jersey: John Wiley & Sons Inc; 2018.
34. Fox S, Madden M. Pew Research Center. 2006. Riding the Waves of 'Web 2.0' URL: http://www.pewinternet.org/2006/10/05/riding-the-waves-of-web-2-0/ [accessed 2018-11-19]
35. Research and Publications - World Bank Group. 2007. Empowerment in Practice: Analysis and Implementation URL: http://siteresources.worldbank.org/WBI/Resources/EmpowermentLearningModulebody.pdf [accessed 2019-05-13]
36. Choi J, Noh GY, Park DJ. Smoking cessation apps for smartphones: content analysis with the self-determination theory. J Med Internet Res 2014 Feb 12;16(2):e44 [FREE Full text] [doi: 10.2196/jmir.3061] [Medline: 24521881]
37. Lie SS, Karlsen B, Oord ER, Graue M, Oftedal B. Dropout from an ehealth intervention for adults with type 2 diabetes: a qualitative study. J Med Internet Res 2017 May 30;19(5):e187 [FREE Full text] [doi: 10.2196/jmir.7479] [Medline: 28559223]
38. O'Cathain A, Drabble SJ, Foster A, Horspool K, Edwards L, Thomas C, et al. Being human: a qualitative interview study exploring why a telehealth intervention for management of chronic conditions had a modest effect. J Med Internet Res 2016 Jun 30;18(6):e163 [FREE Full text] [doi: 10.2196/jmir.5879] [Medline: 27363434]
39. Lee K, Hoti K, Hughes JD, Emmerton L. Dr Google and the consumer: a qualitative study exploring the navigational needs and online health information-seeking behaviors of consumers with chronic health conditions. J Med Internet Res 2014 Dec 2;16(12):e262 [FREE Full text] [doi: 10.2196/jmir.3706] [Medline: 25470306]
40. Grünloh C, Myreteg G, Cajander A, Rexhepi H. 'Why do they need to check me?' Patient participation through ehealth and the doctor-patient relationship: qualitative study. J Med Internet Res 2018 Jan 15;20(1):e11 [FREE Full text] [doi: 10.2196/jmir.8444] [Medline: 29335237]
41. Sheldon KM, Filak V. Manipulating autonomy, competence, and relatedness support in a game-learning context: new evidence that all three needs matter. Br J Soc Psychol 2008 Jun;47(Pt 2):267-283. [doi: 10.1348/014466607X238797] [Medline: 17761025]

Abbreviations

eHealth: electronic health
SDT: self-determination theory

XSL•FO
RenderX

Edited by G Eysenbach; submitted 19.02.19; peer-reviewed by G Myreteg, K Groth, L Roper, A Waruru, M Ghajarzadeh; comments to author 04.04.19; revised version received 23.05.19; accepted 19.07.19; published 15.08.19.

Please cite as:
Scott Duncan T, Riggare S, Koch S, Sharp L, Hägglund M
From Information Seekers to Innovators: Qualitative Analysis Describing Experiences of the Second Generation of E-Patients
J Med Internet Res 2019;21(8):e13022
URL: http://www.jmir.org/2019/8/e13022/
doi:10.2196/13022
PMID:31418421

©Therese Scott Duncan, Sara Riggare, Sabine Koch, Lena Sharp, Maria Hägglund. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 15.08.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Evaluation of a Mobile Telesimulation Unit to Train Rural and Remote Practitioners on High-Acuity Low-Occurrence Procedures: Pilot Randomized Controlled Trial

Jennifer Jewer[1], BCom, BEd, MASc, PhD; Michael H Parsons[2], BSc, MD, CCFP(EM), FCFP; Cody Dunne[2], BSc, MD; Andrew Smith[2], MEng, MD, CCFP(EM); Adam Dubrowski[3], PhD

[1]Faculty of Business, Memorial University, St John's, NL, Canada

[2]Faculty of Medicine, Memorial University, St John's, NL, Canada

[3]Faculty of Health Sciences, Ontario Tech University, Oshawa, ON, Canada

Corresponding Author:
Jennifer Jewer, BCom, BEd, MASc, PhD
Faculty of Business
Memorial University
300 Prince Phillip Drive
St John's, NL, A1B 3X5
Canada
Phone: 1 709 864 3094
Fax: 1 709 864 7680
Email: jenniferj@mun.ca

Abstract

Background: The provision of acute medical care in rural and remote areas presents unique challenges for practitioners. Therefore, a tailored approach to training providers would prove beneficial. Although simulation-based medical education (SBME) has been shown to be effective, access to such training can be difficult and costly in rural and remote areas.

Objective: The aim of this study was to evaluate the educational efficacy of simulation-based training of an acute care procedure delivered remotely, using a portable, self-contained unit outfitted with off-the-shelf and low-cost telecommunications equipment (mobile telesimulation unit, MTU), versus the traditional face-to-face approach. A conceptual framework based on a combination of Kirkpatrick's Learning Evaluation Model and Miller's Clinical Assessment Framework was used.

Methods: A written procedural skills test was used to assess Miller's learning level— *knows* —at 3 points in time: preinstruction, immediately postinstruction, and 1 week later. To assess procedural performance (*shows how*), participants were video recorded performing chest tube insertion before and after hands-on supervised training. A modified Objective Structured Assessment of Technical Skills (OSATS) checklist and a Global Rating Scale (GRS) of operative performance were used by a blinded rater to assess participants' performance. Kirkpatrick's *reaction* was measured through subject completion of a survey on satisfaction with the learning experiences and an evaluation of training.

Results: A total of 69 medical students participated in the study. Students were randomly assigned to 1 of the following 3 groups: comparison (25/69, 36%), intervention (23/69, 33%), or control (21/69, 31%). For *knows*, as expected, no significant differences were found between the groups on written knowledge (posttest, P=.13). For *shows how*, no significant differences were found between the comparison and intervention groups on the procedural skills learning outcomes immediately after the training (OSATS checklist and GRS, P=1.00). However, significant differences were found for the control versus comparison groups (OSATS checklist, P<.001; GRS, P=.02) and the control versus intervention groups (OSATS checklist, P<.001; GRS, P=.01) on the pre- and postprocedural performance. For *reaction*, there were no statistically significant differences between the intervention and comparison groups on the satisfaction with learning items (P=.65 and P=.79) or the evaluation of the training (P=.79, P=.45, and P=.31).

Conclusions: Our results demonstrate that simulation-based training delivered remotely, applying our MTU concept, can be an effective way to teach procedural skills. Participants trained remotely in the MTU had comparable learning outcomes (*shows how*) to those trained face-to-face. Both groups received statistically significant higher procedural performance scores than those in the control group. Participants in both instruction groups were equally satisfied with their learning and training (*reaction*). We

XSL•FO
RenderX

believe that mobile telesimulation could be an effective way of providing expert mentorship and overcoming a number of barriers to delivering SBME in rural and remote locations.

(J Med Internet Res 2019;21(8):e14587) doi:10.2196/14587

KEYWORDS

medical education; distributed medical education; simulation training; emergency medicine; rural health; remote-facilitation; assessment; chest tubes

Introduction

Challenges Accessing Simulation-Based Medical Education

The provision of acute care in rural and remote areas presents unique challenges. Skills related to high-acuity low-occurrence procedures and clinical encounters are particularly susceptible to degradation over time and are inadequately served through on-the-job experience alone [1]. Therefore, a systematic approach to training personnel for these procedures is required. In recent years, an increasing proportion of this training has made use of simulation-based modalities. Simulation-based medical education (SBME) has been shown to be an effective training approach because it can provide opportunities to practice infrequently encountered procedures [2-5] without compromising patient safety [6]. However, SBME often takes place in urban centers, and it can be difficult for rural and remote acute care practitioners to access these centers because of geographic, cost, and time constraints [7,8].

SBME delivered through technologies such as telesimulation and mobile simulation has been shown to be an effective means of training medical practitioners and has helped to address some of the above constraints [4,7-17]. However, use of these technologies is accompanied by their own challenges. Telesimulation involves delivering SBME over the internet, but effective delivery of telesimulation training can be limited if the trainees are unable to access simulation equipment or an efficient training setup. Mobile simulation can address constraints by delivering an immersive simulation environment in a purposefully designed unit. However, mobile simulation often involves bringing an expert to rural and remote sites to facilitate the session. This can often prove to be quite expensive and prohibitive because of time constraints.

Through an iterative design process, our multidisciplinary group has developed an MTU that explores many of the challenges to the delivery of SBME to rural and remote acute care practitioners. The intention is the deployment of the MTU at a rural or remote location that could house the skills training session through communication with an off-site, skilled mentor. Such a deployment would provide trainees with the appropriate simulation equipment, a standardized training environment, and access to an experienced mentor to guide the training. To our knowledge, this is one of the few such units, which combines telecommunication and mobile simulation to deliver such training.

A rigorous, theory-based, iterative approach was followed to develop the MTU and evaluate the acceptability and feasibility of delivering training remotely using the unit. Details on the development of the MTU and training materials have been published elsewhere [18-23].

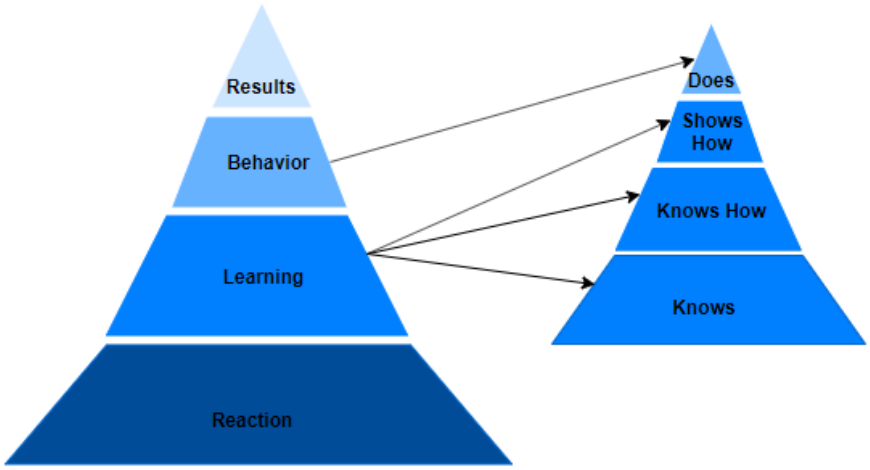
The objective of this study was to compare the educational efficacy of face-to-face versus remote delivery of educational content with respect to learner's perceptions and objective assessment of procedural performance.

Framework for Learning Assessment

This study uses a conceptual framework based on a combination of Kirkpatrick's Learning Evaluation Model [24] and Miller's Clinical Assessment Framework [25] to guide the assessment of the MTU. This model (Figure 1; adapted from Dubrowski et al [26]) is based on the work of Moore et al [27] who developed a framework "of an ideal approach to planning and assessing continuing medical education that is focused on achieving desired outcomes" (pg 3). The new model incorporates Kirkpatrick's 4 levels, which represent a sequence of ways to evaluate a program, with Miller's assessment tools for each level of competence.

XSL•FO
RenderX

Figure 1. Framework for Learning Assessment, based on Kirkpatrick (left) and Miller (right). Adapted from Dubrowski et al [26].



The base of Kirkpatrick's model relates to subject *reaction*, measuring how participants react to or perceive program content. There is no direct correlation of this feature to a level on Miller's framework. The second level of the Kirkpatrick model, *learning*, corresponds to the bottom 3 levels of Miller's framework (*knows*, *knows how*, and *shows how*), whereas the third level of Kirkpatrick's framework, *behavior*, is closely related to the top of Miller's framework, *does*. Finally, the top level of Kirkpatrick's model, *results*, does not relate to Miller's framework. This study examines Kirkpatrick's *reaction* and *learning*, consisting of *knows* and *shows how*. We do not examine *knows how* because of anticipated challenges of subject retention and expected loss to follow-up during the study. Rather, we decided to measure the higher level *shows how* because we could evaluate the participants' performance of the procedure during the study. We do not examine Kirkpatrick's *behavior* and, consequently, do not examine Miller's *does*. We also do not examine Miller's *results*, as these are assessments of practice in a clinical setting, and this study is limited to an experimental setting. This paper discusses the findings in relation to Kirkpatrick's *reaction* and *learning* (consisting of Miller's *knows* and *shows how*) levels.

Methods

Research Setting

This study was conducted at Memorial University of Newfoundland. Training of rural and remote acute care practitioners is of particular interest in the province, as 40% of the population lives in rural areas, and the province has a relatively small population (525,000) distributed across a large geographic area (405,000 km^2). Acute care is delivered at a variety of health centers and hospitals across the province. These sites are staffed by physicians, nurses, and nurse practitioners with varying levels of experience. Access to SBME opportunities is often limited. Health Research Ethics Board of Memorial University of Newfoundland approved this study.

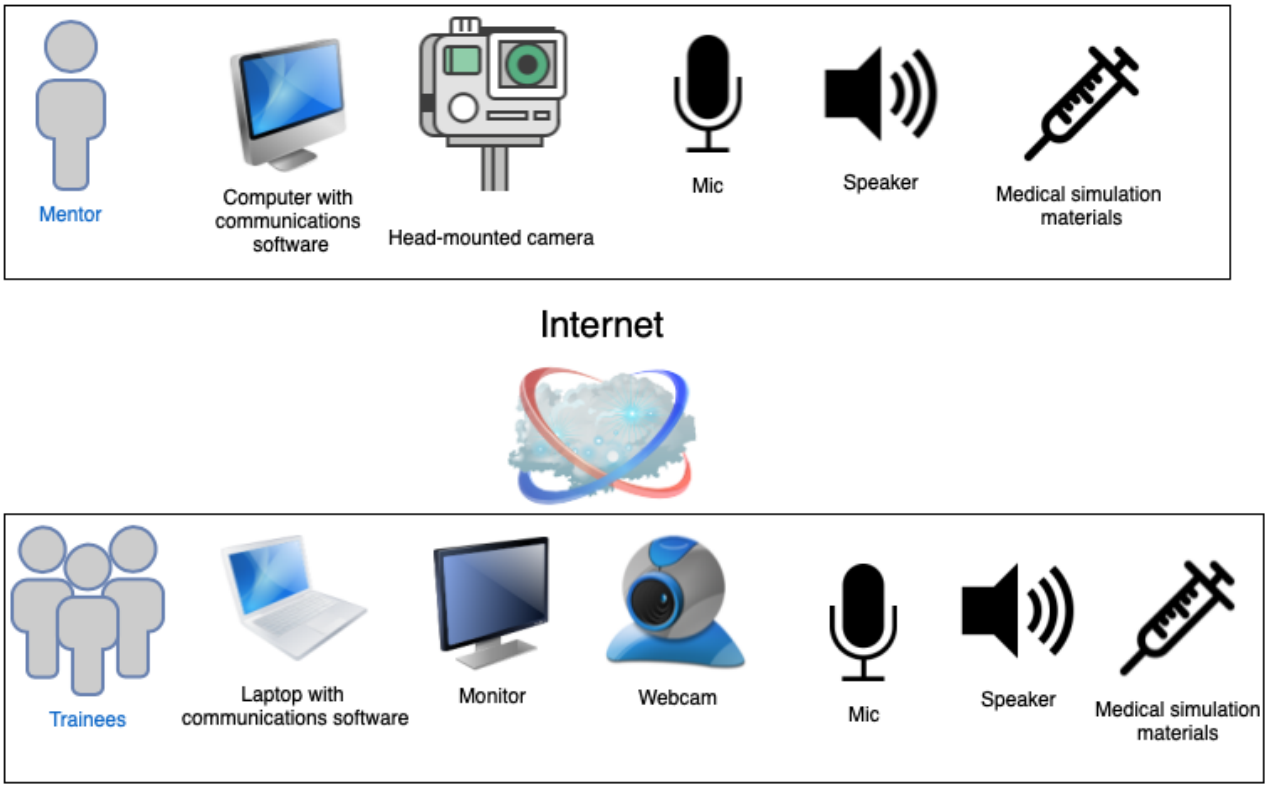
The MTU consists of an inflatable rapid deployment tent (Figure 2), which is outfitted with portable technology necessary to allow for 2-way communication between the trainees and the mentor: laptop with communications software, monitor, camera, speaker, and microphone and a portable wireless internet hub. The mentor uses comparable software, a camera, speaker, and microphone to communicate with the trainees. Off-the-shelf and low-cost equipment was used to keep the design of the MTU accessible and practical. Both the trainees and the mentor would have similar simulation supplies and setup to enable efficient demonstration and instruction (Figures 3 and 4). Studies by Jewer et al provide more information on the MTU [18,21].

The eventual goal was to deliver simulation-based training remotely through the use of a self-contained vehicle outfitted with simulation equipment necessary for delivery of a number of scenarios. However, for the purpose of our test-of-concept approach, a portable and rapid deployment tent was used.

XSL•FO
RenderX

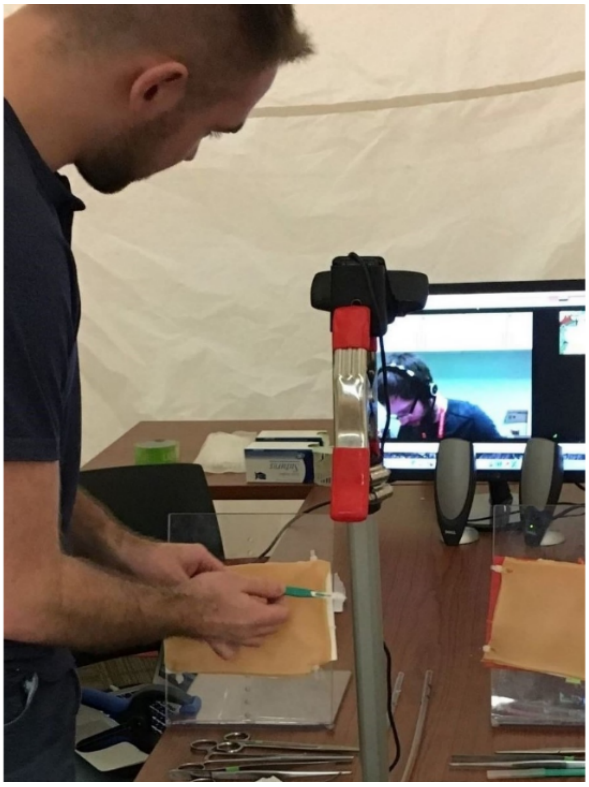
Figure 2. The mobile telesimulation unit rapid deployment tent.

Figure 3. Overview of the setup for the mentor and the trainees in the mobile telesimulation unit.



XSL•FO
RenderX

Figure 4. The interior of the mobile telesimulation unit demonstrating setup for procedural training.



Study Design

A randomized controlled trial design was followed. A total of 3 sessions were held to compare the learning outcomes of participants who received training remotely in the MTU versus those who received the same training face-to-face. To minimize variables affecting study outcomes, face-to-face training sessions also took place in the MTU space. A control group (ie, received no training) was included to show that the intervention group (ie, remote) was not inferior to the comparison group (ie, face-to-face), and that both instructional approaches are actually effective [28].

The sessions focused on teaching an important high-acuity low-occurrence procedure, chest tube insertion, using a low-fidelity setup: 3D-printed ribs, secured to a plexiglass stand, covered with low-cost simulated skin, and subcutaneous tissue (Figure 4). The chest tube insertion was selected as a representative procedure because it is an essential skill in acute care settings requiring precision [29], and it is a multistep procedure amenable to objective scoring. The training sessions were 20-min long and consisted of simulation-based training, with deliberate hands-on practice and mentor feedback.

Figure 5 depicts the flow of the study procedure. A week before the procedural session, participants were emailed presession information consisting of a Web-based New England Journal of Medicine video, demonstrating proper performance of the procedure and important details about chest tube insertion including indications, contraindications, complications, and necessary equipment [30]. This was to help ensure that participants started with a similar base level of knowledge.

Participants were randomly assigned to 1 of 3 groups: intervention, comparison, and control. Testing procedures were conducted before the training (pretest), after the training (posttest), and 1 week later (retention test). During the pretest, participants completed a questionnaire on demographic information, the number of times they performed or witnessed a chest tube insertion before this session, their previous experience with SBME, and their previous experience with telemedicine. Next, participants completed a written procedural skills knowledge test on a number of chest tube procedure-specific questions. The demographic questionnaire and the procedural skills knowledge test were written components used to assess whether there were differences in the baseline knowledge about the chest tube procedure within or between the groups at the start of the study. The procedural skills knowledge test was also used to measure learning after the session. This corresponds to the *knows* level of learning. These materials were reviewed by an experienced emergency medicine physician to determine if differences existed.

To measure *shows how*, during the pretest, participants were video recorded performing a chest tube insertion on a low-fidelity simulated model (Figure 6). A modified Objective Structured Assessment of Technical Skills (OSATS) checklist and a Global Rating Scale (GRS) of operative performance were used to assess procedural performance [31].

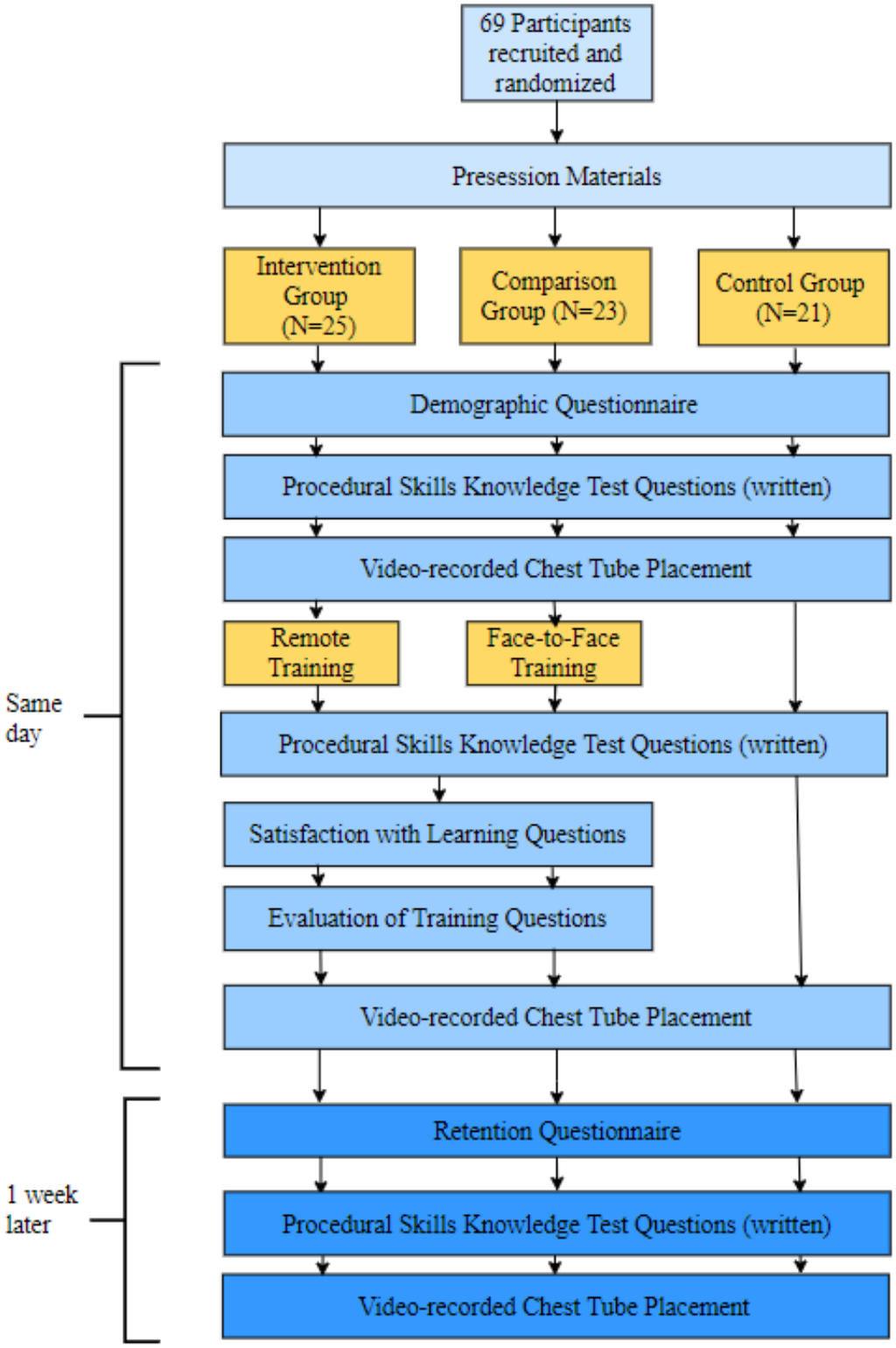
After the training session, during the posttest, participants in the intervention and comparison groups were asked to evaluate their satisfaction with learning and their evaluation of the training. This corresponds to Kirkpatrick's *reaction* level of the learning framework. Participants also completed the written procedural skills knowledge test again (ie, *knows*). All participants were then once again video recorded performing a chest tube insertion (ie, *shows how*).

XSL•FO
RenderX

Furthermore, 1 week after the training session (retention test), the participants completed a questionnaire on their experiences with the procedure in the past week. They also completed the written procedural skills knowledge test again (ie, *knows*), and they were video recorded for the third time performing a chest tube insertion (ie, *shows how*).

An emergency medicine physician with 11 years of clinical emergency room experience used the modified OSATS checklist and GRS to assess the participants' performance on the video recordings. The reviewing physician was blinded to participants' identity and was unaware of the phase of the study (pretest, posttest, or retention test). Overall, 12% of the videos were randomly selected for review by a second experienced emergency medicine physician. The modified OSATS checklist and GRS scores were used as the primary indicators of learning outcomes (ie, *shows how*).

Figure 5. Study design.



XSL•FO
RenderX

Figure 6. Setup used in the video recording of the chest tube procedure (A) and example of a completed chest tube insertion (B).

Participants

Medical students during their first and second year of training (approximately 80 students per cohort) were invited to participate in the study. Participation was voluntary and was limited by the number of slots available at a scheduled data collection time (Multimedia Appendix 1). These medical students were novices in the chest tube procedure, and using such subjects with similar background knowledge and skills enabled us to more clearly measure learning. Participants provided informed consent before enrollment in the study.

Measures

Learning—Knows

To measure the *knows* dimension of learning, participants were asked a set of chest tube procedure–specific questions (Textbox 1).

Textbox 1. Procedural skills knowledge test questions (possible score: 15).

- Name 3 indications for chest tube placement.
- Name 3 contraindications to chest tube placement.
- Name 4 potential complications of chest tube placement.
- Name 5 essential pieces of equipment for chest tube placement.

Learning—Knows How

Participants' performance of the chest tube procedure was evaluated using a modified OSATS checklist to measure the *knows how* dimension of learning. The OSATS checklist was originally developed and validated to assess the performance of multiple surgical procedures at different stations [31]. It has since been used to assess the performance of a single surgical procedure [32]. Research has demonstrated that the OSATS has high reliability and construct validity for measuring technical abilities outside of the operating room [31].

This study used a modified OSATS checklist and a GRS of operative performance. The checklist consists of 10 items that are scored as done correctly or not (Textbox 2). For the purposes of this study, 1 item of the scale was removed because it was not relevant to our training scenario (ie, item #9—a Pleur-evac setup was not available to participants). The GRS is composed of 9 items, each measuring a different aspect of operative performance. Each item was graded on a 5-point Likert scale from 1, poor performance, to 5, good performance (Figure 7). Again, for the purposes of this study, 2 items of the scale were removed because they were not appropriate for the training scenario (removed items included *use of assistants* and *knowledge of instruments* because there was no assistant in the study design and knowledge of instruments implied participants were *asking* for the right things or *saying* the right names, something which was not part of the study design). Thus, the maximum GRS score attainable is 35 points, and the minimum is 7 points.

XSL•FO
RenderX

Textbox 2. Checklist for chest tube insertion (not done, incorrect=0; done, correct=1).

- Injects local anesthetic
- Cuts skin with scalpel to subcutaneous tissue plane (no scything)
- Uses blunt dissection to enter chest cavity
- Enters pleural space above rib
- Checks position with digit before inserting chest tube
- Inserts chest tube safely using Kelly at the tip of the tube
- Inserts correct length of chest tube into chest
- Secures chest tube to chest wall with silk or nylon
- Connects tube and secures to drainage system with tape
- Applies airtight dressing

XSL•FO
RenderX

Figure 7. Global Rating Scale of operative performance.

Please circle the number corresponding to the candidate's performance **regardless of the candidate's level of training**.

Respect for tissue				
1 Frequently used unnecessary force on tissue or caused damage by inappropriate use of instruments	2	3 Careful handling of tissue but occasionally caused inadvertent damage	4	5 Consistently handled tissue appropriately with minimal damage to tissue
Time and motion				
1 Many unnecessary moves	2	3 Efficient time/motion but some unnecessary moves	4	5 Clear economy of movement and maximum efficiency
Instrument Handling				
1 Repeatedly makes tentative or awkward moves with instruments through inappropriate use	2	3 Competent use of instruments but occasionally appeared stiff or awkward	4	5 Fluid movements with instruments and no stiffness or awkwardness
Knowledge of Instruments				
1 Frequently asked for wrong instrument or used inappropriate instrument	2	3 Knew names of most instruments and used appropriate instrument	4	5 Obviously familiar with instruments and their names
Flow of Operation				
1 Frequently stopped operating and seemed unsure of next move	2	3 Demonstrated some forward planning with reasonable progression of procedure	4	5 Obviously planned course of operation with effortless flow from one move to the next
Use of Assistants				
1 Consistently placed assistants poorly or failed to use assistants	2	3 Appropriate use of assistants most of the time	4	5 Strategically used assistants to the best advantage at all times
Knowledge of Specific Procedure				
1 Deficient knowledge. Required specific instruction at most steps of operation	2	3 Knew all important steps of operation	4	5 Demonstrated familiarity with all steps of the operation
OVERALL PERFORMANCE				
1 Very poor	2	3 Competent	4	5 Clearly superior
QUALITY OF FINAL PRODUCT				
1 Very poor	2	3 Competent	4	5 Clearly superior

Reaction

To measure participants' reactions to the training, participants in the remote and comparison groups were asked to evaluate the training by indicating whether they thought the MTU could play an important role in rural and remote medical training, how satisfied they were with their overall experience in the MTU, and if they would recommend the MTU approach to their colleagues. Participants were also asked to indicate their satisfaction with the learning experiences. These measures were adapted from the National League of Nursing (NLN) Student Satisfaction and Self-Confidence in Learning scales [33]. These NLN scales have been widely used and have been found to have sufficient reliability and validity to be used in education research [33,34].

Data Analysis

Participants were assigned a unique identifier, and this was used to anonymize the data before analysis with respect to their training group. Data analysis was completed using SPSS version

XSL•FO
RenderX

25. Descriptive statistics were computed for the demographic variables.

Learning—Knows

Because our data did not enable us to use the parametric repeated measures analysis of variance to analyze the pretest, posttest, and retention written procedural skills tests, we created 2 new variables (pretest minus posttest score and posttest minus retention test score). The Kruskal-Wallis test (nonparametric equivalent) was then used to compare participants' performance on the procedural skill test between the groups.

Learning—Knows How

There was acceptable interrater reliability between the 2 raters who evaluated the performance of the chest tube procedure. An excellent intraclass correlation coefficient (ICC) of 0.909 was found for the GRS, and a good ICC of 0.757 was found for the checklist. Again, limited to nonparametric techniques, we created 2 new variables: 1 variable to calculate the difference between the pre- and postchecklist and GRS scores, and the second to calculate the difference between the postchecklist and retention checklist and GRS scores. A Kruskal-Wallis test was then used to compare pretest, posttest, and retention test scores for the 3 groups (ie, intervention, comparison, and control) on the modified OSATS checklist and GRS scores.

Reaction

The Mann-Whitney *U* test (nonparametric equivalent) was used to compare the intervention with the comparison groups on satisfaction with learning and their evaluation of the training.

For all tests, a *P* value less than .05 was considered statistically significant.

Results

In total, 69 medical students participated in the study across the 3 different sessions (Table 1). Participants were randomly assigned to their study group: intervention, comparison, or control groups.

Participants' Experience

The groups were very similar—mean age in low to mid-20s and relatively equally mixed between the first and second year of medical school. If there was any impact on results of students being in the first or second year of medical school, it would probably negatively influence the intervention group because a slightly higher percentage of participants in this group were in their first year. However, training on chest tube insertion is not part of the standard curriculum in the first 2 years of medical school, and most participants indicated that they had never performed or even witnessed a chest tube placement before; therefore, the presession materials and this training were the first exposures to the skill for most participants. The majority had participated in low-fidelity SBME using task trainers before, between 1 and 10 times, and the majority had never received training using telemedicine.

Table 1. Participants' experience.

Characteristics	Intervention group (n=25)	Comparison group (n=23)	Control group (n=21)
Age (years), mean	25	23	21
Level of medical training, n (%)			
1st year	16 (64)	6 (26)	9 (43)
2nd year	9 (36)	17 (74)	12 (57)
Performed a chest tube insertion before, n (%)			
Never	24 (96)	22 (96)	20 (95)
Yes	1 (4)	1 (4)	2 (5)
Witnessed a chest tube insertion before, n (%)			
Never	22 (88)	20 (87)	15 (71)
Yes	3 (12)	3 (13)	6 (29)
Participated in simulation-based medical education[a] , n (%)			
Never	2 (8)	5 (22)	4 (19)
1-10 times	21 (84)	18 (78)	15 (71)
>10 times	2 (8)	0 (0)	2 (9.5)
Past exposure to telemedicine, n (%)			
Never	25 (100)	18 (78)	19 (91)
At least quarterly	0 (0)	5 (22)	2 (10)

[a]Low-fidelity task trainers (eg, suturing pads, airway models, and chest tube placement).

XSL•FO
RenderX

Table 2. Questionnaire responses at the time of retention test (1 person from the comparison group and 2 from the control group did not complete the retention test).

Characteristics	Intervention group (n=25)		Comparison group (n=22)		Control group (n=19)	
	No, n (%)	Yes, n (%)	No, n (%)	Yes, n (%)	No, n (%)	Yes, n (%)
Performed a chest tube in the past week	23 (92)	2 (8)	22 (100)	0 (0)	19 (100)	0 (0)
Witnessed a chest tube in the past week	25 (100)	0 (0)	22 (100)	0 (0)	19 (100)	0 (0)
Received any training or done further reading on chest tube insertions in the past week	24 (96)	1 (4)	21 (96)	1 (4)	17 (90)	2 (11)

Similarly, the retention test survey, assessing exposure to chest tube insertions in the week since the training, showed no real differences between the groups. Most had not performed a chest tube since the training, witnessed a chest tube, or received any training or done any further reading on chest tube insertions (Table 2).

Learning—Knows

A Kruskal-Wallis test was used to compare the results of the procedural skills knowledge test. This was a brief written test completed after receiving the presession materials but before the training session. The mean test score (out of a possible score of 15) and SD were 11.52 (2.07) for the intervention group, 10.91 (2.02) for the comparison group, and 10.76 (2.56) for the control group. There was no significant difference between the groups before starting the session (χ^2_2=1.9; *P*=.39). This indicates that the participants in the 3 groups had similar levels of written knowledge about chest tube insertion before the training.

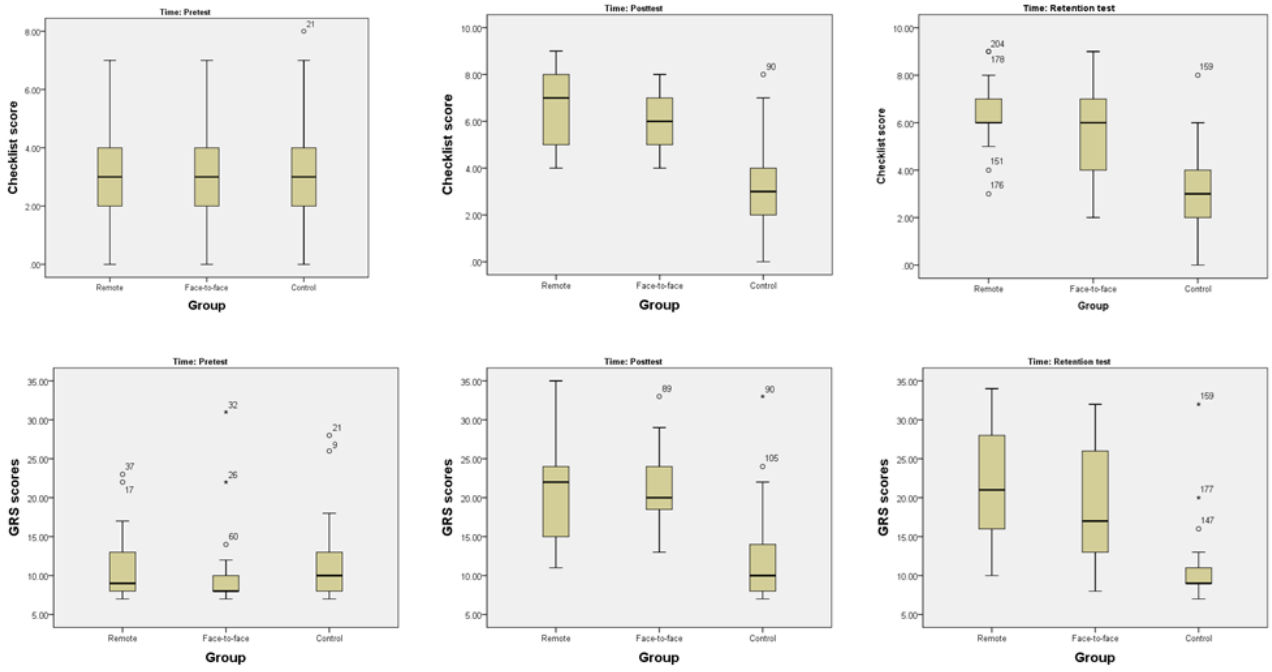
Subsequent Kruskal-Wallis tests revealed that there were no significant differences between groups from the pretest to the posttest (χ^2_2=4.1; *P*=.13) or from the posttest to the retention test (χ^2_2=1.6; *P*=.46; Multimedia Appendix 2).

Learning—Knows How

A total of 204 videos of procedural performance were included in the analysis, with 3 videos per participant (3 participants did not complete the retention test). Results of the modified OSATS checklist and GRS assessment for the 3 groups (pretraining, posttraining, and 1 week after the training) are shown in Multimedia Appendix 3. Box plots of the scores are shown in Figure 8.

A Kruskal-Wallis test revealed that there were statistically significant differences between the groups on the pre- and post-OSATS checklist and GRS scores (Multimedia Appendix 4). Pairwise comparisons were performed using Dunn [35] procedure with a Bonferroni correction for multiple comparisons. This post hoc analysis revealed statistically significant differences in median OSATS checklist and GRS scores differences between the control and comparison and the control and intervention groups, but not between the comparison and intervention groups. There was no difference between the posttest and retention scores.

Figure 8. Box plots of the modified Objective Structured Assessment of Technical Skills checklist and GRS scores. GRS: Global Rating Scale.



XSL•FO
RenderX

Reaction

Satisfaction with learning and evaluation of the training measures was used to examine participants' reaction to the training.

Satisfaction with Learning

The results of the satisfaction with learning questions (adapted from the NLN scales) that were asked in the posttest for the intervention and comparison groups are shown in Table 3. On average, participants rated the teaching methods as helpful and effective for the intervention and comparison groups, with scores 4.52 and 4.65, respectively, out of 5. Averaged responses also indicated that they enjoyed how the teacher taught the session for the intervention and comparison groups with scores 4.40 and 4.52, respectively, out of 5. A Mann-Whitney *U* test revealed that there were no statistically significant differences between the intervention and comparison groups on these items.

Table 3. Self-reported learning—scale of 1 (strongly disagree) to 5 (strongly agree).

Measurement item (satisfaction with learning)	Intervention group (n=25), mean (SD)	Comparison group (n=23), mean (SD)	Mann-Whitney *U* test		
			U	z	*P* value
The teaching methods used were helpful and effective.	4.52 (0.71)	4.65 (0.49)	306.5	0.47	.65
I enjoyed how the teacher taught the session.	4.40 (0.82)	4.52 (0.59)	299.0	0.27	.79

Participant Evaluation of Training

Participants in the intervention and comparison groups were asked to evaluate their experiences with the training session that took place physically in the MTU space. Participants indicated that the MTU could play an important role in rural medical training (4.32 and 4.48 out of 5 for the intervention and comparison groups, respectively), they were satisfied with their overall experience in the MTU (4.32 and 4.43 out of 5 for the intervention and comparison groups, respectively), and they would recommend the MTU to their colleagues for SBME (4.32 and 4.43 out of 5 for the intervention and comparison groups, respectively). A Mann-Whitney *U* test revealed that there were no statistically significant differences between the intervention and the comparison groups on any of these questions (Table 4).

Table 4. Participants' evaluation of training modality—scale 1 (strongly disagree) to 5 (strongly agree).

Evaluation of training modality	Intervention group (n=25), mean (SD)	Comparison group (n=23), mean (SD)	Mann-Whitney *U* test or *t* test		
			U	z	*P* value
Do you think the MTU[a] could play an important role in rural medical training?	4.32 (1.11)	4.48 (0.51)	276.0	−0.27	.79
How satisfied are you with your overall experience in the MTU?	4.32 (0.56)	4.43 (0.59)	319.5	0.38	.45
Would you recommend the MTU to your colleagues for simulation-based medical training?	4.32 (0.56)	4.43 (0.73)	331.0	1.02	.31

[a]MTU: mobile telesimulation unit.

Discussion

Principal Findings

Using a conceptual framework based on Kirkpatrick's and Miller's works [24,25], we examined learning based on *knows* and *shows how* levels and also studied the *reactions* of the participants to the training. We found this framework useful to help ensure a thorough evaluation of the training delivered using an MTU. The results from this study indicate comparable learning (*knows* and *shows how*) and *reactions* of participants who received the procedural skills training remotely with those who received the training face-to-face.

Consistent with the literature, we found that subject's knowledge level (*knows*) remained unchanged after the training. This is what was expected as there are 2 distinct key areas of knowledge with respect to competent procedural skills performance—one related to factual background information (*knows*) and the second being the ability to complete all necessary steps (*shows how*). Our study focused on *shows how*, as the ability to physically and capably complete a procedural skill relies on deliberate practice of that skill [36]. Nevertheless, it was important to measure the procedural skill knowledge (*knows*), as it enabled us to ensure there was a consistent knowledge level across all groups. This is particularly important, as procedural skills training sessions aim to enable participants' performance of the procedure (ie, *shows how*), which is a higher level than *knows*.

With respect to the *shows how* learning, our study supports previous findings related to telesimulation and mobile simulation [7,8,37,38]. We found that the learning outcomes for the participants who received training remotely through the MTU, as assessed using modified OSATS checklist and GRS, are comparable with those of the face-to-face simulation-based training group. Furthermore, participants who received training, either remotely or face-to-face, received statistically significantly better scores than those who did not receive instruction (ie, the control group). The average scores on the checklist more than

XSL•FO
RenderX

doubled from the pretests to posttests for the intervention (from 3 to 6.54) and comparison groups (from 2.96 to 6.22). However, the increase in the scores for the control group was negligible, increasing by only 0.33 points (from 2.91 to 3.24). This indicates that training resulted in similar acquisition of skills-based knowledge for both the remote training and face-to-face groups.

Retention tests indicated that there were no statistically significant differences in skills retention between all 3 of the groups. On average, differences between the scores on retention test–modified and posttest-modified OSATS checklist and GRS scores either stayed the same or decreased slightly for all groups. From this, we conclude that the manner of instructional delivery (either remote or face-to-face) does not impact retention.

In addition to the comparable learning outcomes, participants had similarly high levels of satisfaction with learning in the MTU. Rating the teaching methods as helpful and effective, participants indicated that, on average, they enjoyed instruction during the session. This is encouraging as satisfaction with the training, in the case of the MTU concept facilitated through a local healthcare facility, could influence commitment and readiness to transfer learning to the workplace at their own site [39,40].

Overall, participants evaluated their training experience with the MTU as positive. There were no statistically significant differences in evaluations between those who received training remotely versus those who received it face-to-face. Participants felt that the MTU could play an important role in rural medical training, they indicated that they were satisfied with their overall experience in the MTU, and they would recommend the MTU to their colleagues for SBME.

The primary limitation of this study is the relatively small sample size and the inclusion of research subjects from a single institution. However, several things help make the study more robust: (1) the inclusion of a control group; (2) the study design including pretest, posttest, and retention tests; and (3) the triangulation of the results of the modified OSATS checklist and GRS scores with 2 blinded raters demonstrating a favorable interrater reliability provides reassurance of the robustness of the study results [41]. The second limitation is the fact that the physician who was involved in the design of the MTU is the one who led the training sessions for all subjects. It would be interesting to examine the impact on training if a physician not directly involved in the study delivered the training.

There are a number of implications for future SBME and research. First, there is a shift from delivery of medical education in large urban academic centers toward distributed medical education. Technologies such as video conferencing and digital library collections have enabled this advancement and are tied to social, health, and economic benefits [42,43]. There is potential for the MTU concept to play a role in this area, and further research is needed to determine how best to incorporate this concept into practice. Here, it would be particularly important to consider 2 significant time challenges faced by rural practitioners; the maintenance of busy clinical practice, often with limited backup, in addition to the invaluable contributions made in teaching a variety of learners, often with limited resources. A collaborative approach, drawing on local expertise, along with distance-guided mentorship could facilitate valuable advances. The second and related contribution is the potential MTU-enabled cost-savings for trainees and mentors. Cost is often a major barrier to accessing SBME, but it is often not considered in SBME research [44]. Traditional delivery models will have a course, and associated expenses brought to a particular site, such as the related costs of travel, equipment, mentors, and time. The alternative being that the rural practitioner must travel to a central location to teach and is left to address the challenges of patient coverage, time off, and expenses relating to the training and travel. By making cost an important consideration in the development of the MTU, the intention is to make this novel approach more accessible. An economic impact evaluation relating to the use of the MTU in practice is recommended. Third, further studies should be conducted to validate the utility and effectiveness of the MTU concept for skills training that is important to the practice needs of the target audience. Through collaborative discussion and targeted needs assessments with rural practitioners, the specific clinical and educational needs would best be determined. This would enable the examination of Kirkpatrick's *behavior* (and Miller's *does*), as well as the *results* levels of the learning evaluation model. As a broader range of skills sessions are delivered remotely through mobile telesimulation, opportunities to study validity and reliability will be more readily available. Fourth, further exploration of skill and scenario characteristics that make them amenable to the remote-mentoring approach to training is necessary, including ability to observe key performance features and maneuvers. This study demonstrated equivalent learning outcomes on assessment of procedural skills for chest tube insertion. This should be further explored for other procedural skills. Fifth, training sessions for this study were conducted in areas with reliable, high-speed internet access; however, rural and remote areas may have limited internet connectivity, which will impede the delivery of remote training and may particularly affect how learners perceive and rate their remote mentoring experience. Future research will explore the use of purpose-built efficient communications systems designed for low bandwidth. Sixth, as proposed by others [3], future research should compare different forms of simulation. Using mobile telesimulation, this would involve comparison of training delivered remotely in an MTU using different levels of fidelity simulators. Finally, the unavailability of mentors comfortable with using simulation-based teaching delivered through telecommunication may present a barrier to expanding this novel approach to SBME [45]. Therefore, the use of the MTU for the remote assessment of skills should also be examined, especially in domains that are poorly covered by traditional written and oral examinations.

Conclusions

SBME is a well-established training approach, particularly for high-acuity, low-occurrence procedures and scenarios. Practitioners located in rural and remote locations particularly stand to benefit as they face a number of unique challenges with respect to simulation resources, including geographic, cost, and time constraints. This study describes an evaluation of educational efficacy comparing remote versus face-to-face mentoring for procedural skills training. To our knowledge, this

XSL•FO
RenderX

study is one of a few to develop and assess SBME combining the concepts of telesimulation and mobile simulation.

We used a conceptual framework based on the combination of Kirkpatrick's Learning Evaluation Model and Miller's Clinical Assessment Framework to guide the study. We found that training delivered remotely through the MTU is an effective way to conduct a skills session. Those who were remotely trained had comparable learning outcomes (*shows how*) to subjects who received face-to-face instruction. Participants were also satisfied (*reaction*) with their learning and training experiences. Such remote mentor–led SBME expands opportunities for health practitioners to more easily access the training and mentor-guided practice that they require. Future investigation is needed to examine the utility of the MTU approach in practice, with different skills and level of fidelity, and as a means to provide remote assessment of skills.

Acknowledgments

This project has been supported by an Ignite grant awarded by the Research and Development Corporation of Newfoundland and Labrador. The authors thank the following organizations at the Memorial University of Newfoundland: the Tuckamore Simulation Research Collaborative for research support and advice, the Clinical Learning and Simulation Center for equipment and operational support, and Memorial University of Newfoundland MED 3D for the provision of simulation models. The authors also thank the following people for their assistance during this research project: Dr Chrystal Horwood for clinical expertise in video review; Kristopher Hoover for technical assistance and involvement in early MTU prototype development; research assistant, Megan Pollard, Samantha Noseworthy, Sarah Boyd, and Krystal Bursey; Tate Skinner (technical support); Joanne Doyle (Discipline of Emergency Medicine senior secretary); and Memorial University's Emergency Medicine Interest Group.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Consolidated Standards of Reporting Trials flow diagram.

[PDF File (Adobe PDF File), 103KB - jmir_v21i8e14587_app1.pdf]

Multimedia Appendix 2

Differences between pre, post, and retention procedural skills knowledge tests (written).

[PDF File (Adobe PDF File), 113KB - jmir_v21i8e14587_app2.pdf]

Multimedia Appendix 3

Modified Objective Structured Assessment of Technical Skills checklist and Global Rating Scale assessment of chest tube performance, mean (standard deviation) reported.

[PDF File (Adobe PDF File), 80KB - jmir_v21i8e14587_app3.pdf]

Multimedia Appendix 4

Differences between pre, post, and retention-modified Objective Structured Assessment of Technical Skills checklist and Global Rating Scale test scores.

[PDF File (Adobe PDF File), 137KB - jmir_v21i8e14587_app4.pdf]

Multimedia Appendix 5

CONSORT‐EHEALTH checklist (V 1.6.1).

[PDF File (Adobe PDF File), 2MB - jmir_v21i8e14587_app5.pdf]

References

1. Williams JM, Ehrlich PF, Prescott JE. Emergency medical care in rural America. Ann Emerg Med 2001 Sep;38(3):323-327. [doi: 10.1067/mem.2001.115217] [Medline: 11524654]
2. Roy KM, Miller MP, Schmidt K, Sagy M. Pediatric residents experience a significant decline in their response capabilities to simulated life-threatening events as their training frequency in cardiopulmonary resuscitation decreases. Pediatr Crit Care Med 2011 May;12(3):e141-e144. [doi: 10.1097/PCC.0b013e3181f3a0d1] [Medline: 20921919]

XSL•FO
RenderX

3. Cook DA, Hatala R, Brydges R, Zendejas B, Szostek JH, Wang AT, et al. Technology-enhanced simulation for health professions education: a systematic review and meta-analysis. J Am Med Assoc 2011 Sep 7;306(9):978-988. [doi: 10.1001/jama.2011.1234] [Medline: 21900138]
4. Scott DJ, Dunnington GL. The new ACS/APDS skills curriculum: moving the learning curve out of the operating room. J Gastrointest Surg 2008 Feb;12(2):213-221. [doi: 10.1007/s11605-007-0357-y] [Medline: 17926105]
5. Issenberg SB, McGaghie WC, Petrusa ER, Lee GD, Scalese RJ. Features and uses of high-fidelity medical simulations that lead to effective learning: a BEME systematic review. Med Teach 2005 Jan;27(1):10-28. [doi: 10.1080/01421590500046924] [Medline: 16147767]
6. Ziv A, Wolpe PR, Small SD, Glick S. Simulation-based medical education: an ethical imperative. Acad Med 2003 Aug;78(8):783-788. [doi: 10.1097/01.SIH.0000242724.08501.63] [Medline: 12915366]
7. Rosen MA, Hunt EA, Pronovost PJ, Federowicz MA, Weaver SJ. In situ simulation in continuing education for the health care professions: a systematic review. J Contin Educ Health Prof 2012;32(4):243-254. [doi: 10.1002/chp.21152] [Medline: 23280527]
8. Ikeyama T, Shimizu N, Ohta K. Low-cost and ready-to-go remote-facilitated simulation-based learning. Simul Healthc 2012 Feb;7(1):35-39. [doi: 10.1097/SIH.0b013e31822eacae] [Medline: 22228281]
9. Bischof JJ, Panchal AR, Finnegan GI, Terndrup TE. Creation and validation of a novel mobile simulation laboratory for high fidelity, prehospital, difficult airway simulation. Prehosp Disaster Med 2016 Oct;31(5):465-470. [doi: 10.1017/S1049023X16000534] [Medline: 27530816]
10. Ullman E, Kennedy M, di Delupis FD, Pisanelli P, Burbui AG, Cussen M, et al. The Tuscan mobile simulation program: a description of a program for the delivery of in situ simulation training. Intern Emerg Med 2016 Sep;11(6):837-841. [doi: 10.1007/s11739-016-1401-2] [Medline: 26861702]
11. Xafis V, Babidge W, Field J, Altree M, Marlow N, Maddern G. The efficacy of laparoscopic skills training in a mobile simulation unit compared with a fixed site: a comparative study. Surg Endosc 2013 Jul;27(7):2606-2612. [doi: 10.1007/s00464-013-2798-6] [Medline: 23389073]
12. Weinstock PH, Kappus LJ, Garden A, Burns JP. Simulation at the point of care: reduced-cost, in situ training via a mobile cart. Pediatr Crit Care Med 2009 Mar;10(2):176-181. [doi: 10.1097/PCC.0b013e3181956c6f] [Medline: 19188878]
13. Ireland S, Gray T, Farrow N, Danne P, Flanagan B. Rural mobile simulation-based trauma team training-an innovative educational platform. Int Trauma Care 2006;16:6-12 [FREE Full text]
14. Ohta K, Kurosawa H, Shiima Y, Ikeyama T, Scott J, Hayes S, et al. The effectiveness of remote facilitation in simulation-based pediatric resuscitation training for medical students. Pediatr Emerg Care 2017 Aug;33(8):564-569. [doi: 10.1097/PEC.0000000000000752] [Medline: 27261952]
15. Mikrogianakis A, Kam A, Silver S, Bakanisi B, Henao O, Okrainec A, et al. Telesimulation: an innovative and effective tool for teaching novel intraosseous insertion techniques in developing countries. Acad Emerg Med 2011 Apr;18(4):420-427 [FREE Full text] [doi: 10.1111/j.1553-2712.2011.01038.x] [Medline: 21496146]
16. Schulman CI, Levi J, Sleeman D, Dunkin B, Irvin G, Levi D, et al. Are we training our residents to perform open gall bladder and common bile duct operations? J Surg Res 2007 Oct;142(2):246-249. [doi: 10.1016/j.jss.2007.03.073] [Medline: 17631907]
17. Strongwater AM. Transition to the eighty-hour resident work schedule. J Bone Joint Surg Am 2003 Jun;85(6):1170-1172. [doi: 10.2106/00004623-200306000-00048] [Medline: 12784026]
18. Jewer J, Dubrowski A, Dunne C, Hoover K, Smith A, Parsons M. Piloting a mobile tele-simulation unit to train rural and remote emergency health care providers. In: Wickramasinghe N, Bodendorf F, editors. Delivering Superior Health and Wellness Management with IoT and Analytics. New York: Springer; 2019.
19. Parsons M, Wadden K, Pollard M, Dubrowski A, Smith A. P098: development and evaluation of a mobile simulation lab with acute care telemedicine support. Can J Emerg Med 2016 Jun 2;18(S1):S111. [doi: 10.1017/cem.2016.274]
20. Parsons M, Smith A, Hoover K, Jewer J, Noseworthy S, Pollard M, et al. P100: iterative prototype development of a mobile tele-simulation unit for remote training: an update. Can J Emerg Med 2017 May 15;19(S1):S112. [doi: 10.1017/cem.2017.302]
21. Jewer J, Dubrowski A, Hoover K, Smith A, Parsons M. Development of a Mobile Tele-Simulation Unit Prototype for Training of Rural and Remote Emergency Health Care Providers. In: Proceedings of the 51st Hawaii International Conference in System Sciences. 2018 Presented at: HICSS'18; January 3-6, 2018; Hawaii, United States p. 2894-2903 URL: https://aisel.aisnet.org/hicss-51/hc/ict_for_health_equity/2/ [doi: 10.24251/HICSS.2018.367]
22. Parsons M, Smith A, Rogers P, Hoover K, Pollard M, Dubrowski A. Outcomes of Prototype Development Cycle for a Mobile Simulation Lab With Acute Care Telemedicine Support- Work in Progress. In: 17th International Meeting on Simulation in Healthcare. 2017 Presented at: IMSH'17; January 26-30, 2017; Orlando, FL.
23. Dunne C, Jewer J, Parsons M. P039: application of the Delphi method to refine key components in the iterative development of a mobile tele-simulation unit (MTU). Can J Emerg Med 2018 May 11;20(S1):S70-S71. [doi: 10.1017/cem.2018.237]
24. Kirkpatrick DL, Kirkpatrick JD. Evaluating Training Programs: The Four Levels. Second Edition. San Francisco: Berrett-Koehler Publishers; 1998.
25. Miller GE. The assessment of clinical skills/competence/performance. Acad Med 1990 Sep;65(9 Suppl):S63-S67. [doi: 10.1097/00001888-199009000-00045] [Medline: 2400509]

XSL•FO
RenderX

26. Dubrowski A, Morin M. Evaluating pain education programs: an integrated approach. Pain Res Manag 2011;16(6):407-410 [FREE Full text] [doi: 10.1155/2011/320617] [Medline: 22184548]
27. Moore Jr DE, Green JS, Gallis HA. Achieving desired results and improved outcomes: integrating planning and assessment throughout learning activities. J Contin Educ Health Prof 2009;29(1):1-15. [doi: 10.1002/chp.20001] [Medline: 19288562]
28. Greene CJ, Morland LA, Durkalski VL, Frueh BC. Noninferiority and equivalence designs: issues and implications for mental health research. J Trauma Stress 2008 Oct;21(5):433-439 [FREE Full text] [doi: 10.1002/jts.20367] [Medline: 18956449]
29. Friedrich M, Bergdolt C, Haubruck P, Bruckner T, Kowalewski K, Müller-Stich BP, et al. App-based serious gaming for training of chest tube insertion: study protocol for a randomized controlled trial. Trials 2017 Dec 6;18(1):56 [FREE Full text] [doi: 10.1186/s13063-017-1799-5] [Medline: 28166840]
30. Dev SP, Nascimiento Jr B, Simone C, Chien V. Videos in clinical medicine. Chest-tube insertion. N Engl J Med 2007 Oct 11;357(15):e15. [doi: 10.1056/NEJMvcm071974] [Medline: 17928590]
31. Reznick R, Regehr G, MacRae H, Martin J, McCulloch W. Testing technical skill via an innovative 'bench station' examination. Am J Surg 1997 Mar;173(3):226-230. [doi: 10.1016/s0002-9610(97)89597-9] [Medline: 9124632]
32. Haubruck P, Nickel F, Ober J, Walker T, Bergdolt C, Friedrich M, et al. Evaluation of app-based serious gaming as a training method in teaching chest tube insertion to medical students: randomized controlled trial. J Med Internet Res 2018 Dec 21;20(5):e195 [FREE Full text] [doi: 10.2196/jmir.9956] [Medline: 29784634]
33. National League for Nursing. 2005. Descriptions of Available Instruments URL: http://www.nln.org/professional-development-programs/research/tools-and-instruments/descriptions-of-available-instruments
34. Franklin AE, Burns P, Lee CS. Psychometric testing on the NLN student satisfaction and self-confidence in learning, simulation design scale, and educational practices questionnaire using a sample of pre-licensure novice nurses. Nurse Educ Today 2014 Oct;34(10):1298-1304. [doi: 10.1016/j.nedt.2014.06.011] [Medline: 25066650]
35. Dunn OJ. Multiple comparisons using rank sums. Technometrics 1964 Aug;6(3):241-252. [doi: 10.2307/1266041]
36. Ericsson KA. Deliberate practice and acquisition of expert performance: a general overview. Acad Emerg Med 2008 Nov;15(11):988-994 [FREE Full text] [doi: 10.1111/j.1553-2712.2008.00227.x] [Medline: 18778378]
37. Treloar D, Hawayek J, Montgomery JR, Russell W, Medical Readiness Trainer Team. On-site and distance education of emergency medicine personnel with a human patient simulator. Mil Med 2001 Nov;166(11):1003-1006. [doi: 10.1093/milmed/166.11.1003] [Medline: 11725312]
38. Okrainec A, Vassiliou M, Kapoor A, Pitzul K, Henao O, Kaneva P, et al. Feasibility of remote administration of the fundamentals of laparoscopic surgery (FLS) skills test. Surg Endosc 2013 Nov;27(11):4033-4037. [doi: 10.1007/s00464-013-3048-7] [Medline: 24018759]
39. Mansour JB, Naji A, Leclerc A. The relationship between training satisfaction and the readiness to transfer learning: the mediating role of normative commitment. Sustainability 2017 May 16;9(5):834. [doi: 10.3390/su9050834]
40. Lim DH, Morris ML. Influence of trainee characteristics, instructional satisfaction, and organizational climate on perceived learning and training transfer. Hum Resour Dev Q 2006;17(1):85-115. [doi: 10.1002/hrdq.1162]
41. Lineberry M, Walwanis M, Reni J. Comparative research on training simulators in emergency medicine: a methodological review. Simul Healthc 2013 Aug;8(4):253-261. [doi: 10.1097/SIH.0b013e31828715b1] [Medline: 23508094]
42. Ellaway R, Bates J. Distributed medical education in Canada. Can Med Educ J 2018 Mar;9(1):e1-e5 [FREE Full text] [Medline: 30140329]
43. Lemky K, Gagne P, Konkin J, Stobbe K, Fearon G, Blom S, et al. A review of methods to assess the economic impact of distributed medical education (DME) in Canada. Can Med Educ J 2018 Mar;9(1):e87-e99 [FREE Full text] [Medline: 30140340]
44. Zendejas B, Wang AT, Brydges R, Hamstra SJ, Cook DA. Cost: the missing outcome in simulation-based medical education research: a systematic review. Surgery 2013 Feb;153(2):160-176. [doi: 10.1016/j.surg.2012.06.025] [Medline: 22884087]
45. Hayden EM, Navedo DD, Gordon JA. Web-conferenced simulation sessions: a satisfaction survey of clinical simulation encounters via remote supervision. Telemed J E Health 2012 Sep;18(7):525-529. [doi: 10.1089/tmj.2011.0217] [Medline: 22827475]

Abbreviations

GRS: Global Rating Scale
ICC: intraclass correlation coefficient
MTU: mobile telesimulation unit
NLN: National League of Nursing
OSATS: Objective Structured Assessment of Technical Skills
SBME: simulation-based medical education

XSL•FO
RenderX

Edited by G Eysenbach; submitted 06.05.19; peer-reviewed by M Reade, C Knopp; comments to author 20.06.19; revised version received 04.07.19; accepted 05.07.19; published 06.08.19.

Please cite as:
Jewer J, Parsons MH, Dunne C, Smith A, Dubrowski A
Evaluation of a Mobile Telesimulation Unit to Train Rural and Remote Practitioners on High-Acuity Low-Occurrence Procedures: Pilot Randomized Controlled Trial
J Med Internet Res 2019;21(8):e14587
URL: http://www.jmir.org/2019/8/e14587/
doi:10.2196/14587
PMID:

©Jennifer Jewer, Michael H Parsons, Cody Dunne, Andrew Smith, Adam Dubrowski. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 06.08.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Potential for Integrating Mental Health Specialist Video Consultations in Office-Based Routine Primary Care: Cross-Sectional Qualitative Study Among Family Physicians

Mariell Hoffmann[1], BA, MA; Mechthild Hartmann[1], MA; Michel Wensing[2], PhD; Hans-Christoph Friederich[1], MD; Markus W Haun[1], MD, BSc, MSc

[1]Department of General Internal Medicine and Psychosomatics, Heidelberg University, Heidelberg, Germany

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

Corresponding Author:
Markus W Haun, MD, BSc, MSc
Department of General Internal Medicine and Psychosomatics
Heidelberg University
Im Neuenheimer Feld 410
Heidelberg,
Germany
Phone: 49 622156 ext 8774
Email: markus.haun@med.uni-heidelberg.de

Abstract

Background: Although real-time mental health specialist video consultations have been proposed as an effective care model for treating patients with mental health conditions in primary care, little is known about their integration into routine practice from the perspective of family physicians.

Objective: This study aimed to determine the degree to which family physicians advocate that mental health specialist video consultations can be integrated into routine primary care, where most patients with mental health conditions receive treatment.

Methods: In a cross-sectional qualitative study, we conducted 4 semistructured focus groups and 3 telephonic interviews in a sample of 19 family physicians from urban and rural districts. We conducted a qualitative content analysis applying the Tailored Implementation in Chronic Diseases framework in a combined bottom-up (data-driven) and top-down strategy for deriving key domains.

Results: Family physicians indicated that mental health specialist video consultations are a promising and practical way to address the most pressing challenges in current practice, that is, to increase the accessibility and co-ordination of specialized care. Individual health professional factors were the most frequently discussed topics. Specifically, family physicians valued the anticipated clinical outcomes for patients and the anticipated resources set for the primary care practice as major facilitators (16/19, 84%). However, family physicians raised a concern regarding a lack of facial expressions and physical interaction (19/19, 100%), especially in emergency situations. Therefore, most family physicians considered a viable emergency plan for mental health specialist video consultations that clearly delineates the responsibilities and tasks of both family physicians and mental health specialists to be essential (11/19, 58%). Social, political, and legal factors, as well as guideline factors, were hardly discussed as prerequisites for individual family physicians to integrate mental health specialist video consultations into routine care. To facilitate the implementation of future mental health specialist video consultation models, we compiled a checklist of recommendations that covers (1) buy-in from practices (eg, emphasizing logistical and psychological relief for the practice), (2) the engagement of patients (eg, establishing a trusted patient-provider relationship), (3) the setup and conduct of consultations (eg, reliable emergency plans), and (4) the fostering of collaboration between family physicians and mental health specialists (eg, kick-off meetings to build trust).

Conclusions: By leveraging the primary care practice as a familiar environment for patients, mental health specialist video consultations provide timely specialist support and potentially lead to benefits for patients and more efficient processes of care. Integration should account for the determinants of practice as described by the family physicians.

Trial Registration: German Clinical Trials Register DRKS00012487; https://www.drks.de/drks_web/navigate.do?navigationId=trial.HTML&TRIAL_ID=DRKS00012487

XSL•FO
RenderX

(J Med Internet Res 2019;21(8):e13382) doi:10.2196/13382

KEYWORDS

video consultations; videoconferencing; integrated behavioral health; primary care; health services research; mental health

Introduction

Background

For decades, primary care has been the de facto mental health care outpatient system in most countries. In recent years, the management of mental health conditions in primary care has further increased because of the higher prevalence and regional shortages of mental health care providers [1]. Collaborative mental health care integrating mental health specialists in primary care practices has been proposed as a possible solution. However, its coordination with primary care has been found to be a persistent challenge [2,3]. To address this challenge, clinicians and policy makers have recently proposed real-time mental health specialist video consultations as a service delivery model [4-6]. Telepsychiatry technologies such as mental health specialist video consultations are integrated, patient-centered care models that can be classified into low-, moderate-, and high-intensity models. These levels depend on the complexity of the intervention as well as the amount of resources, such as the extent to which the primary care provider or specialist is involved [7]. In this study, we examined the potential for integrating mental health specialist video consultations as a moderate-intensity model into office-based routine primary care. Within our model, family physicians referred patients presenting in primary care with depression and/or anxiety to a mental health specialist. The mental health specialist will then schedule and conduct the mental health specialist video consultation with the patients in the primary care practice. The mental health specialist intervention is time limited and aimed at diagnostics, care planning, and crisis management or brief psychotherapy. To date, there have been some pioneering studies indicating that video consultations work well for health care settings [8-12]. Specifically, in randomized controlled trials that were conducted within the unique context of the Veterans Health Administration system and federally qualified health centers, video consultations were effective for depression [8,9]. Furthermore, commercial telepsychiatry providers have established nonclinical physician extenders as virtual care navigators for referring patients from primary care to a telepsychiatrist [10]. There is also preliminary evidence that video consultation–based collaborative care decreases costs and improves access to mental health care, especially in rural areas, by reducing travel time [11,12].

Implementation of Mental Health Specialist Video Consultations Into Primary Care

Although these findings are promising, we do not know whether and to what extent practicing family physicians themselves are convinced that mental health specialist video consultations can be integrated into routine practice. Moreover, although Powell et al recently reported that primary care patients are quite willing to engage in video consultations, the prerequisites for the amount of time needed for their adoption from the family physicians' perspective are unknown [13,14]. However, for the development of successful implementation strategies for routine care, it is crucial to allow family physicians, who are the central stakeholders in routine care, to evaluate the anticipated benefit of the proposed models. Therefore, the purpose of this qualitative focus group study was to assess the potential for the integration of mental health specialist video consultations into office-based routine primary care based on the analysis of family physicians' perceptions of current practice.

Methods

Study Design and Conceptual Framework

We conducted a cross-sectional qualitative study with semistructured focus groups as a naturalistic data collection method to assess the collective sense making of family physicians and to explore experiences and perceptions through the interaction of the participants [15]. We opted for focus groups, given that they are more appropriate than interviews when people do not have a personal stake in the topic [16] and that focus group participants usually have less constrained discussions than those in individual interviews [17]. We conducted telephonic interviews with family physicians who were not able to participate in a focus group for organizational reasons. By integrating focus groups and interviews, we were able to (1) avoid the withdrawal of family physicians who were particularly interested in study participation and (2) capture complementary perspectives on mental health specialist video consultations [18] and therefore gather broader information. To capture both current real-world practice and the potential integration of video consultations, we followed a pragmatic approach based on the critical realist position. This allowed us to derive recommendations for future interventional studies on mental health specialist video consultations [19]. This study was approved by the ethics committee of the medical faculty at Heidelberg University (S-197/2017) and preregistered with the German Clinical Trials Register (DRKS00012487). We followed the COnsolidated criteria for REporting Qualitative research guidelines for reporting qualitative study results [20].

Participants and Recruitment

We applied purposeful sampling to account for a diverse range of participants and information [21]. Specifically, in an initial survey study (results not presented here) on the integration of mental health specialist video consultations, we invited all 788 family physicians registered with the Association of Statutory Health Insurance Physicians in 1 urban and 4 rural districts (from a total of 35 districts in Baden-Wuerttemberg, 1 of 16 German federal states) to participate in a focus group. Apart from registration, there were no other eligibility criteria for family physicians. Overall, 107 family physicians (13.6%, 107/788) responded to the survey, 41 of whom showed interest in participating in a focus group. We conducted phone calls with 35 family physicians to schedule a focus group appointment with a maximum of 6 participants each. One family physician

XSL•FO
RenderX

took part in a focus group without a formal invitation but rather following another participating family physician's suggestion. Overall, 16 family physicians refused to participate, most frequently because of holiday leave (n=4) and lack of interest (n=4). In addition, 7 family physicians were not contacted because of earlier-than-expected data saturation. The latter was achieved when no additional insights emerged from data, and the content began to repeat. Overall, we conducted 4 focus groups (range: 2-6 participants, 90-120 min) involving 16 family physicians at Heidelberg University Hospital alongside individual telephonic interviews (40-55 min) with 3 family physicians. We offered a nonadvertised individual monetary compensation of €50.

Data Collection

We developed a semistructured question guide to prompt group discussions and interviews (Multimedia Appendix 1). The questions focused on how family physicians perceived current health care for patients with mental disorders, the potential for integrating mental health specialist video consultations into office-based routine primary care, and the determinants of the implementation of mental health specialist video consultations. We piloted the guide on 1 family physician and 1 senior health services researcher. It was also reviewed after the first focus group. After obtaining written informed consent from all participants, the first author (Doctor of Philosophy student, sociologist, and expertise in qualitative research) and the last author (Doctor of Medicine, internal medicine specialist, senior researcher, and content expert for mental health services) moderated the focus groups. To stimulate the discussion, the moderators presented a 7-min video clip illustrating the mental health specialist video consultation model. The interviews were conducted by the first author who verbally described the model. All focus groups and interviews were audio-recorded and uploaded to a secure server of Heidelberg University Hospital, which was only accessible to the research team. We analyzed the participants' sociodemographic and primary care practice data that were collected in the initial mail survey.

Data Analysis

First, before anonymizing the data, a professional transcription service conducted verbatim audio transcriptions of the recordings. Second, 2 authors (MH and MWH) independently conducted a qualitative content analysis of the first focus group in MAXQDA 12 (VERBI Software). To develop the code system, we followed a combined bottom-up (data-driven) and top-down strategy [22]. For the latter, we applied the 7 domains of the Tailored Implementation in Chronic Diseases (TICD) framework as an analytical lens to identify the determinants of practice [23]. The TICD accounts for multiple levels (micro- and macrolevels) and stakeholder perspectives (eg, patients and health care providers). Third, both researchers compared their code systems and resolved disagreements in a final version. Fourth, we applied the code system to the remaining transcripts. To ensure representation of all key aspects, we modified the codes when new aspects emerged. We summarize the key domains, including definitions and supporting quotes in Multimedia Appendix 2.

Results

Sample

Table 1 shows the sociodemographic characteristics of the 19 participating family physicians.

XSL•FO
RenderX

Table 1. Sample description (N=19).

Variable	Value
Female sex, n (%)	9 (48)
Age (years), mean (SD)	57.5 (7.3)
Years in office-based practice, mean (SD)	18.3 (9.7)
Type of practice, n (%)	
Solo practice	12 (63)
Shared practice	6 (32)
Group practice	1 (5)
Areas of recruitment, n (%)	
Cities (densely populated areas)	2 (11)
Towns and suburbs (intermediate density areas)	14 (74)
Rural areas (thinly populated areas)	3 (16)
Additional qualification in psychotherapy, n (%)[a]	4 (21)
Average number of patients per quarter, n (%)	
<500	1 (5)
501-1000	6 (32)
1001-1500	6 (32)
>1500	6 (32)
Patients with mental health conditions per week, n (%)	
1-5	1 (5)
5-10	3 (16)
10-15	7 (37)
>10-15	8 (42)

[a]Multiple responses possible.

Depiction of Current Practice

For family physicians, the coordination of care for patients presenting with mental health disorders was a persistent challenge in current practice. Waiting times and an insufficient number of available mental health specialists (14/19), as well as a lack of professional exchanges between family physicians and mental health specialists (12/19), were major difficulties:

> *It is very difficult because we have such poor access to psychotherapists. [...] They [the patients] must wait for a quarter of a year to half a year to get an appointment.* [Interview 3]

These barriers consistently produced a precarious psychological dilemma. On the one hand, family physicians did not have enough resources (eg, because of time restrictions) to provide mental health care themselves for all patients in need (12/19):

> *Psychotherapy cannot be integrated into everyday primary care practice. However, the first step to getting a feeling or becoming pretty sure that this patient would need to be referred to someone, to a psychotherapist...this takes more than the usual ten- or fifteen-minutes during the consultation...Um...to open Pandora's box, that takes more time. It...um...the patient must feel that he can open up himself. That is quite a time challenge because during flu season, I must admit, that I just don't listen, I just don't go there. That is not possible then.* [Focus group 1]

Family physicians could not refer patients to mental health specialists because of unavailability. On the other hand, observing that patients remained undertreated was at odds with the professional ethics of the family physicians, who generally commit themselves to providing comprehensive care to their patients (6/19, 32%):

> *Interviewer: How do you handle the less urgent cases?*
>
> *Bm1: Um, those are, um, mostly unserved, or I improvise and um, I am not so happy with that.* [Focus group 3]

Determinants for Implementing Mental Health Specialist Video Consultations

Family physicians perceived mental health specialist video consultations as a potentially effective approach to address the gaps in current practice. Specifically, family physicians felt that mental health specialist video consultations would increase low-threshold access and improve coordination to provide mental health care in primary care. In our analysis, we identified

XSL•FO
RenderX

44 subdomains and grouped them into the 7 domains of the TICD framework: individual health professional factors; patient factors; professional interactions; incentives and resources; capacity for organizational change; social, political and legal factors; and guideline factors. In the following sections, we present our results along these 7 domains arranged in descending order by code frequency. We report code frequency in brackets behind each code to provide information about data saturation and complement our findings. Two domains, namely, social, political, and legal factors and guideline factors, were hardly discussed and are not reported.

Individual Health Professional Factors

Family physicians' perceptions of the integration of mental health specialist video consultations into primary care were strongly influenced by both the anticipated clinical outcomes and the anticipated resources made available for primary care practice. Specifically, family physicians considered mental health specialist video consultations to potentially facilitate seamless and low-threshold access to specialized mental health care (7/19, 37%):

> *However, I do see an advantage in having a low-threshold option, to refer a patient to a consultation with a professional, without having to enlist him somewhere, without having him to drive anywhere and so on.* [Focus group 2]

By enabling timely referrals, family physicians argued, mental health specialist video consultations would lead to relatively rapid clinical improvement in patients, which emerged as the foremost purpose of family physicians' professional identity:

> *I could offer something to my patients, not for the purpose of advertisement, but I could actually propose something that makes patients feel better. And that's why I attend to the patient in the first place.* [Focus group 4]

Moreover, several family physicians expected that, to some extent, mental health specialist video consultations would make resources available for the entire primary care team and, in the long term, lead to more efficient daily workflows (9/19, 47%). They underscored that the major benefit would entail psychological relief for them as they would now be able to trust that the adequate treatment has been initiated:

> *They [mental health specialist video consultations] may produce some psychological relief by letting me know that I did something good for the patient [...] Of course, the feeling that I did something meaningful for the patient would be relieving.* [Focus group 4]

Furthermore, considering patient outcomes, some family physicians underscored the difference between virtual and face-to-face consultations (6/19, 32%). Some family physicians stated that with certain patient groups, there may be a risk of not being able to establish a sufficiently stable therapeutic relationship through mental health specialist video consultations. Consequently, family physicians argued that the establishment and monitoring of a solid patient-provider relationship should be the top priority when implementing mental health specialist video consultations. Otherwise, mental health specialist video consultations might alienate patients from care providers. One particularly concerned family physician stated the following:

> *Well, as I said, I think they [mental health specialist video consultations] may work with some people, but in principle, it is different from sitting across from someone. There you get information you won't get through a screen.* [Focus group 2]

Specifically, family physicians worried that nonverbal communication may be constricted in mental health specialist video consultations and that this anticipated limitation would become particularly apparent in emergency situations:

> *So, is the therapist actually able to assess someone through the camera in an empathetic manner? What happens if one raises a particular issue and he [the patient] leaves the room? [...] If he screams and pulls the cables out* [Focus group 1]

In every focus group and interview, family physicians noted the importance of facial expressions, gestures, and even physical interaction in such situations and advocated for technical solutions optimally supporting nonverbal communication. In addition, most family physicians considered a viable contingency or emergency plan for mental health specialist video consultations, clearly delineating responsibilities and tasks for both family physicians and mental health specialists as essential (11/19, 58%). For instance, they mentioned collaborative debriefing between the family physician and the mental health specialist as a consultant or expert after a mental health video consultation as well as hotline support for patients as potential solutions:

> *Yes, I would welcome, um, an interdisciplinary debriefing or a time window for resolving open questions.* [Focus group 1]

> *At least he [the patient] has to have some sort of hotline number [to call]. If he still had problems, and if his primary care physician and office were closed, he would still have a contact person to turn to. Something like this perhaps.* [Focus group 4]

Patient Factors

Family physicians discussed several patient-related themes, namely, target groups for mental health specialist video consultations, potential preferences, and barriers for technology-based interventions such as mental health specialist video consultations that patients might anticipate.

First, family physicians regarded certain environmental conditions as prerequisites for fostering patient involvement (12/19, 63%). For instance, family physicians wanted to have a designated room available to ensure confidential consultations. Second, family physicians considered mental health specialist video consultations to be generally suitable for patients but were ambivalent concerning certain groups. A few family physicians expected the elderly to be more hesitant in accepting professional psychosocial support in general (5/19, 26%). Family physicians suspected that elderly patients are less familiar with computer technology and therefore more likely to refuse video consultations:

XSL•FO
RenderX

> *Many older people state that they do not want to see a psychotherapist; these people consequently won't go for a video consultation either.* [Interview 1]

Nevertheless, some family physicians regarded elderly patients as potentially open-minded toward and interested in mental health specialist video consultations (6/19, 32%). To this end, 1 family physician gave an illustrative example:

> *I am always surprised to see my 90-year-old grandmas who use Skype without any difficulties. They have just acquired the necessary skills. From my perspective, the barrier to adopt the technology is not large at all.* [Focus group 1]

To facilitate patient engagement, family physicians felt responsible for introducing the patient to the mental health specialist as part of the first session and for being available to follow up with the patient beyond the mental health specialist video consultation. From the family physicians' perspectives, patients (1) who had physical-mental health comorbidities or medically unexplained symptoms (eg, unspecific gastrointestinal complaints; 4/19, 21%), (2) who hesitated to seek mental health care, (eg, because of stigma; 5/19), and (3) who were immobile, particularly in rural areas (4/19, 21%), would most likely benefit from mental health specialist video consultations. Some family physicians considered it necessary to explicitly encourage patients to try video consultations and highlighted that each patient would have to continuously consult with the same mental health specialist (5/19, 26%).

Professional Interactions

Family physicians eagerly discussed collaborative aspects and responsibilities related to mental health specialist video consultations. First, they highly appreciated the possibility of collaborating with the mental health specialists (10/19, 53%). Specifically, family physicians regarded brief case discussions with mental health specialists via colleague-to-colleague video calls as an opportunity to validate or revisit their initial diagnostic assessment and allow for treatments to be tailored according to patients' needs and social environments:

> *I think that they [mental health specialist video consultations] support you in your work as a family physician. After the consultation, it can be discussed what has been done and whether there remains anything urgent to manage or clarify* [Focus group 2]

Second, concerning the distribution of tasks among the involved health care personnel, according to family physicians, medical assistants would have to be responsible for administrative and organizational tasks (eg, appointment allocation and follow-up calls; 10/19, 53%). The family physicians themselves would be responsible for referring patients to the mental health specialist video consultations, whereas family physicians described the role of mental health specialists mainly as consultants with secondary care expertise. To foster a good relationship with the mental health specialists, family physicians demanded that there would be an initial kick-off meeting to meet each other in person.

Incentives and Resources

First and most importantly, most participants stated that reimbursement for mental health specialist video consultations should at least cover their costs for the use of the room where mental health specialist video consultations would be conducted and for additional personnel resources (13/19, 68%). The latter refers to family physicians' statements that family physicians or their medical staff might carry out some tasks related to the intervention, such as initializing the mental health specialist video consultations via the Web platform. Second, family physicians considered the unknown amount of spatial resources (eg, a room in the family physicians' practice to provide a confidential environment for conducting mental health specialist video consultations), personnel, and time resources initially necessary for the integration of mental health specialist video consultations into their practice as a main barrier because of the already tightly organized day-to-day routine. For instance, family physicians wanted themselves or their medical assistants to be responsible for initializing the individual mental health specialist video consultation via the Web platform. Therefore, the setup and implementation of mental health specialist video consultations should account for existing structures and workflows in the given practice. For instance, family physicians stated that the mental health specialist video consultations should be conducted outside the usual consultation hours so that a designated room can be guaranteed. Fixed time slots were expected to facilitate the integration. Third, with respect to technology, family physicians named several basic requirements: stable network connectivity, high visual definition, minimized speech delay, and instant technical service support, alongside training sessions for both practice staff and mental health specialists.

Capacity for Organizational Change

Family physicians rarely addressed this aspect spontaneously. Obviously, the anticipated capacity for change in practices was linked to the intrinsic motivation of the individual family physician. However, family physicians emphasized that the prospect of workload relief resulting from the intervention might foster readiness for change within the medical profession (2/19, 10.5%). One family physician also suspected that the family physician's age might determine his or her intent to adopt technology-based interventions with digital natives assumed to be more open minded (1/19, 5.3%):

> *It has something to do with being curious. And I can imagine that younger colleagues may be even more curious.* [Focus group 2]

Distinctions

First, a comparison of data from cities (2/19, 10.5%), towns or suburbs (14/19, 74%), and rural areas (3/19, 16%) did not reveal any major distinctions. Second, a comparison between solo (12/19, 63%) and group/shared practices (7/19, 37%) indicated that participants from group or shared practices slightly less frequently discussed help for patients or relief for family physicians as expected benefits or outcomes of the intervention. In fact, 5 of 7 participants from group or shared practices valued help for patients or relief for family physicians as potential

XSL•FO
RenderX

outcomes, whereas 12 of 12 participants from solo practices expected 1 of the 2 potential outcomes.

Discussion

Principal Findings

In this study, we investigated the potential for integrating real-time mental health specialist video consultations in primary care among family physicians working in urban and rural practices. Family physicians perceive current practice as fragmentary and deficient. From their perspective, mental health specialist video consultations are a promising and practical way to address the most pressing gaps, that is, to increase the access to and coordination of specialized care. With respect to the implementability of mental health specialist video consultations in primary care, we were able to derive specific recommendations that cover (1) buy-in from practices (eg, emphasizing potential logistical and psychological relief for the practice), (2) the involvement of patients (eg, establishing and securing a trusted patient-provider relationship), (3) the setup and conduct of consultations (eg, solid emergency plans in place), and (4) the fostering of collaboration between family physicians and mental health specialists (eg, in person kick-off meetings to build trust). With respect to both future video consultation applications in routine care as well as feasibility and full-scale intervention trials, we have summarized these recommendations in a checklist that supports stakeholders in accounting for determinants of implementation (Multimedia Appendix 3).

Previous work on video consultations has been limited to efficacy trials and postimplementation studies on perceptions, acceptance, and satisfaction [8,9,13,24-26]. In 2 efficacy trials, Fortney et al found video consultations to be a promising mode of delivery of mental health care [8,9]. However, both trials were in unique contexts (ie, the Veterans Health Administration system and rural federally qualified health centers) and did not apply any participatory assessments to capture the perspective of the professionals and patients. Studies assessing the determinants of the integration of technology-based mental health care models into primary care prospectively have been missing. To the best of our knowledge, our study is the first to provide in-depth qualitative findings on the anticipated benefit of video consultations for patients with mental health conditions in office-based primary care. In the following paragraphs, we therefore discuss our results against the background of more general findings on applying video consultations to tackle medical problems. First, from the patients' perspective, video consultations are welcomed in settings as varied as primary care [13], emergency medicine, and radiology [24]. Concerns about the practicability of video consultations have rarely been explored and have therefore remained rather unspecific [13]. From the perspective of family physicians, we were able to characterize more specific challenges, such as handling emergency situations virtually. We were also able to address potentially sustainable solutions for these challenges with family physicians who were very familiar with routine care conditions. Second, from the health care personnel's perspective, staff and financial resources constitute the main barriers for the integration of Web-based interventions in general [27] and video consultations in particular [25,28]. Our work adds that, in addition to these organizational factors, family physicians focus on establishing a reliable therapeutic relationship in the mental health specialist video consultation setting. The latter and patient satisfaction are substantial prerequisites for the effectiveness and acceptance of telemental health models such as mental health specialist video consultations for patients as well as health care professionals [26]. Hence, family physicians, as persons of trust, could support particularly skeptical patients in trying out mental health specialist video consultations, and family physicians could perform *warm handoffs* to refer patients by means of a personal introduction [29]. At this point, the unique advantage of embedding mental health specialist video consultations directly into primary care practice, an environment with which patients often have been familiar for decades, becomes apparent in our study. Only 1 study focused on collaborative aspects from the perspective of health care personnel using video consultations in primary care [30]. The main result, namely, that staff mostly want to consult with specialists in cases of diagnostic uncertainty, is in accordance with our findings. However, we also found that collaboration should be based on an initial personal encounter between family physicians and specialists. Similar to previous investigations in patients and family physicians, the likelihood of positive outcomes for patients was linked to the patients' literacy in modern technologies [31,32]. However, family physicians in our study underscored readily available specialist support and the family practice as a familiar environment for the patient as crucial determinants of clinical benefits. Finally, participants in our study rarely addressed health system factors. This observation is in accordance with other research on barriers for implementation, namely, that health system factors seem to be outside the perception of health care providers [33].

Strengths and Limitations

Our study has some limitations. First, because of the qualitative nature of our study, family physicians' anticipated barriers and facilitators of mental health specialist video consultations may not be generalizable to a larger population of family physicians or other health care providers. Thus, our findings may be biased in favor of participants willing to implement video consultations. However, a qualitative approach is most suitable to arrive at an in-depth exploration of the family physicians' perceptions, and controversial aspects of video consultations were consistently brought up in our study. Second, we used the TICD framework as an analytical concept and applied its key domains to the data. Nevertheless, the original TICD framework subdomains seemed unsuitable for covering the data. Consequently, we decided to generate the subdomains with a bottom-up approach by processing the entire text material available. We think that in doing so, we have further fostered the validity of the final code system. To limit the selection and confirmation bias for the codes generated bottom-up, we reviewed the preliminary code system along the entire material and modified the subdomains when needed. Third, we provide preliminary recommendations for facilitating mental health specialist video consultations in routine primary care derived from the data as a first practical guidance for initial feasibility studies in telepsychiatry for

XSL•FO
RenderX

integrated care. However, these recommendations have not yet been evaluated in practice and are therefore not yet evidence based. Finally, none of the participants had any previous practical experience with mental health specialist video consultations. This finding may imply that the anticipated determinants of implementation differ from those that family physicians would mention after having conducted mental health specialist video consultations. Nevertheless, we supported participants in being able to fully visualize the implementation and practice of mental health specialist video consultations by introducing the care model in a video clip (focus groups) or verbal description (telephonic interviews).

Conclusions

In conclusion, our study points to the precarious situation for patients and family physicians addressing mental health conditions in current everyday practice. Our findings suggest that mental health specialist video consultations show great potential to address the perceived challenges. Given the potential benefits outlined in our qualitative results, we have now embarked on a feasibility study (trial registration number: DRKS00015812) testing a tailored mental health specialist video consultation model for integration into primary care [34]. Specifically, the results presented here have informed the feasibility study protocol and intervention. From the family physicians' perspective, mental health specialist video consultations hold promise for the future by potentially increasing access to and coordination of specialized care, encouraging cross-sectoral collaboration and providing benefits for patients and family physicians alike.

Acknowledgments

Financial support for this study was provided entirely by a grant from the German Federal Ministry of Education and Research (grant number 01GY16129). The funding agreement ensured the authors' independence in designing the study, interpreting the data, writing, and publishing the report. The funder was not actively involved in conduct of the study. The authors had full access to all the data in this study and take complete responsibility for the integrity of the data and the accuracy of the data analysis.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Semistructured guide for focus groups and telephone interview.
[DOCX File, 15KB - jmir_v21i8e13382_app1.docx]

Multimedia Appendix 2
Summary of domains and subdomains.
[DOCX File, 17KB - jmir_v21i8e13382_app2.docx]

Multimedia Appendix 3
Recommendations for facilitating mental health specialist video consultations in routine primary care.
[DOCX File, 19KB - jmir_v21i8e13382_app3.docx]

References

1. Crowley RA, Kirschner N, Health and Public Policy Committee of the American College of Physicians. The integration of care for mental health, substance abuse, and other behavioral health conditions into primary care: executive summary of an American college of physicians position paper. Ann Intern Med 2015 Aug 18;163(4):298-299. [doi: 10.7326/M15-0510] [Medline: 26121401]
2. Wang PS, Demler O, Kessler RC. Adequacy of treatment for serious mental illness in the United States. Am J Public Health 2002 Jan;92(1):92-98. [doi: 10.2105/ajph.92.1.92] [Medline: 11772769]
3. Katon WJ, Seelig M. Population-based care of depression: team care approaches to improving outcomes. J Occup Environ Med 2008 Apr;50(4):459-467. [doi: 10.1097/JOM.0b013e318168efb7] [Medline: 18404019]
4. Hilty DM, Rabinowitz T, McCarron RM, Katzelnick DJ, Chang T, Bauer AM, et al. An update on telepsychiatry and how it can leverage collaborative, stepped, and integrated services to primary care. Psychosomatics 2018;59(3):227-250. [doi: 10.1016/j.psym.2017.12.005] [Medline: 29544663]
5. Liddy C, Rowan MS, Afkham A, Maranger J, Keely E. Building access to specialist care through e-consultation. Open Med 2013;7(1):e1-e8 [FREE Full text] [doi: 10.1097/HCM.0b013e318196def8] [Medline: 23687533]
6. Yellowlees P, Shore JH. Telepsychiatry and Health Technologies: A Guide for Mental Health Professionals. First Edition. Arlington, VA: American Psychiatric Publishing; 2018.

XSL•FO
RenderX

7. Hilty DM, Sunderji N, Suo S, Chan S, McCarron RM. Telepsychiatry and other technologies for integrated care: evidence base, best practice models and competencies. Int Rev Psychiatry 2018 Dec;30(6):292-309. [doi: 10.1080/09540261.2019.1571483] [Medline: 30821540]
8. Fortney JC, Pyne JM, Edlund MJ, Williams DK, Robinson DE, Mittal D, et al. A randomized trial of telemedicine-based collaborative care for depression. J Gen Intern Med 2007 Aug;22(8):1086-1093 [FREE Full text] [doi: 10.1007/s11606-007-0201-9] [Medline: 17492326]
9. Fortney JC, Pyne JM, Mouden SB, Mittal D, Hudson TJ, Schroeder GW, et al. Practice-based versus telemedicine-based collaborative care for depression in rural federally qualified health centers: a pragmatic randomized comparative effectiveness trial. Am J Psychiatry 2013 Apr;170(4):414-425 [FREE Full text] [doi: 10.1176/appi.ajp.2012.12050696] [Medline: 23429924]
10. Johnston B, Yellowlees P. Telepsychiatry consultations in primary care coordinated by virtual care navigators. Psychiatr Serv 2016 Jan;67(1):141-142. [doi: 10.1176/appi.ps.660905] [Medline: 26725496]
11. Backhaus A, Agha Z, Maglione ML, Repp A, Ross B, Zuest D, et al. Videoconferencing psychotherapy: a systematic review. Psychol Serv 2012 May;9(2):111-131. [doi: 10.1037/a0027924] [Medline: 22662727]
12. Richardson LK, Frueh BC, Grubaugh AL, Egede L, Elhai JD. Current directions in videoconferencing tele-mental health research. Clin Psychol (New York) 2009 Sep 1;16(3):323-338 [FREE Full text] [doi: 10.1111/j.1468-2850.2009.01170.x] [Medline: 20161010]
13. Powell RE, Henstenburg JM, Cooper G, Hollander JE, Rising KL. Patient perceptions of telehealth primary care video visits. Ann Fam Med 2017 May;15(3):225-229 [FREE Full text] [doi: 10.1370/afm.2095] [Medline: 28483887]
14. Gardner MR, Jenkins SM, O'Neil DA, Wood DL, Spurrier BR, Pruthi S. Perceptions of video-based appointments from the patient's home: a patient survey. Telemed J E Health 2015 Apr;21(4):281-285 [FREE Full text] [doi: 10.1089/tmj.2014.0037] [Medline: 25166260]
15. Morgan DL, Krueger RA. When to use focus groups and why. In: Morgan DL, editor. Successful Focus Groups: Advancing the State of the Art. Thousand Oaks, CA: SAGE Publications Inc; 1993:3-19.
16. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
17. Barbour RS. Making sense of focus groups. Med Educ 2005 Jul;39(7):742-750. [doi: 10.1111/j.1365-2929.2005.02200.x] [Medline: 15960795]
18. Lambert SD, Loiselle CG. Combining individual interviews and focus groups to enhance data richness. J Adv Nurs 2008 Apr;62(2):228-237. [doi: 10.1111/j.1365-2648.2007.04559.x] [Medline: 18394035]
19. Madill A, Jordan A, Shirley C. Objectivity and reliability in qualitative analysis: realist, contextualist and radical constructionist epistemologies. Br J Psychol 2000 Feb;91(Pt 1):1-20. [doi: 10.1348/000712600161646] [Medline: 10717768]
20. 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. [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
21. Patton MQ. Qualitative Research & Evaluation Methods. Third Edition. Thousand Oaks, CA: Sage Publications; 2002.
22. Kuckartz U. Qualitative Text Analysis: A Guide to Methods, Practice and Using Software. Thousand Oaks, CA: Sage Publications; 2014.
23. Flottorp SA, Oxman AD, Krause J, Musila NR, Wensing M, Godycki-Cwirko M, et al. A checklist for identifying determinants of practice: a systematic review and synthesis of frameworks and taxonomies of factors that prevent or enable improvements in healthcare professional practice. Implement Sci 2013 Mar 23;8:35-45 [FREE Full text] [doi: 10.1186/1748-5908-8-35] [Medline: 23522377]
24. Powell RE, Stone D, Hollander JE. Patient and health system experience with implementation of an enterprise-wide telehealth scheduled video visit program: mixed-methods study. JMIR Med Inform 2018 Feb 13;6(1):e10 [FREE Full text] [doi: 10.2196/medinform.8479] [Medline: 29439947]
25. Greenhalgh T, Shaw S, Wherton J, Vijayaraghavan S, Morris J, Bhattacharya S, et al. Real-world implementation of video outpatient consultations at macro, meso, and micro levels: mixed-method study. J Med Internet Res 2018 Apr 17;20(4):e150 [FREE Full text] [doi: 10.2196/jmir.9897] [Medline: 29625956]
26. Jenkins-Guarnieri MA, Pruitt LD, Luxton DD, Johnson K. Patient perceptions of telemental health: systematic review of direct comparisons to in-person psychotherapeutic treatments. Telemed J E Health 2015 Aug;21(8):652-660. [doi: 10.1089/tmj.2014.0165] [Medline: 25885491]
27. Montero-Marín J, Prado-Abril J, Botella C, Mayoral-Cleries F, Baños R, Herrera-Mercadal P, et al. Expectations among patients and health professionals regarding web-based interventions for depression in primary care: a qualitative study. J Med Internet Res 2015 Mar 10;17(3):e67 [FREE Full text] [doi: 10.2196/jmir.3985] [Medline: 25757358]
28. Johansson AM, Lindberg I, Söderberg S. The views of health-care personnel about video consultation prior to implementation in primary health care in rural areas. Prim Health Care Res Dev 2014 Apr;15(2):170-179 [FREE Full text] [doi: 10.1017/S1463423613000030] [Medline: 23402617]

XSL•FO
RenderX

29. Pace CA, Gergen-Barnett K, Veidis A, D'Afflitti J, Worcester J, Fernandez P, et al. Warm handoffs and attendance at initial integrated behavioral health appointments. Ann Fam Med 2018 Jul;16(4):346-348 [FREE Full text] [doi: 10.1370/afm.2263] [Medline: 29987084]
30. Johansson AM, Lindberg I, Söderberg S. Healthcare personnel's experiences using video consultation in primary healthcare in rural areas. Prim Health Care Res Dev 2017 Jan;18(1):73-83. [doi: 10.1017/S1463423616000347] [Medline: 27640522]
31. Eichenberg C, Wolters C, Brähler E. The internet as a mental health advisor in Germany--results of a national survey. PLoS One 2013;8(11):e79206 [FREE Full text] [doi: 10.1371/journal.pone.0079206] [Medline: 24278121]
32. Randhawa RS, Chandan JS, Thomas T, Singh S. An exploration of the attitudes and views of general practitioners on the use of video consultations in a primary healthcare setting: a qualitative pilot study. Prim Health Care Res Dev 2019 Jan;20:e5 [FREE Full text] [doi: 10.1017/S1463423618000361] [Medline: 29909798]
33. Huntink E, Wensing M, Klomp MA, van Lieshout J. Perceived determinants of cardiovascular risk management in primary care: disconnections between patient behaviours, practice organisation and healthcare system. BMC Fam Pract 2015 Dec 15;16:179 [FREE Full text] [doi: 10.1186/s12875-015-0390-y] [Medline: 26666285]
34. Tönnies J, Hartmann M, Wensing M, Szecsenyi J, Icks A, Friederich HC, et al. Mental health specialist video consultations for patients with depression or anxiety disorders in primary care: Protocol for a randomised controlled feasibility trial (in press). BMJ Open 2019. [doi: 10.1136/bmjopen-2019-030003]

Abbreviations

TICD: Tailored Implementation in Chronic Diseases

Edited by G Eysenbach; submitted 12.01.19; peer-reviewed by A Bauer, F Mold; comments to author 29.04.19; revised version received 18.06.19; accepted 21.07.19; published 19.08.19.

Please cite as:
Hoffmann M, Hartmann M, Wensing M, Friederich HC, Haun MW
Potential for Integrating Mental Health Specialist Video Consultations in Office-Based Routine Primary Care: Cross-Sectional Qualitative Study Among Family Physicians
J Med Internet Res 2019;21(8):e13382
URL: http://www.jmir.org/2019/8/e13382/
doi:10.2196/13382
PMID:31429419

©Mariell Hoffmann, Mechthild Hartmann, Michel Wensing, Hans-Christoph Friederich, Markus W Haun. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 19.08.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

The Association Between Willingness of Frontline Care Providers' to Adaptively Use Telehealth Technology and Virtual Service Performance in Provider-to-Provider Communication: Quantitative Study

Hyeyoung Hah[1*], PhD; Deana Goldin[2*], PhD, DNP; Sejin Ha[3*], PhD

[1]Department of Information Systems and Business Analytics, Florida International University, Miami, FL, United States

[2]Nicole Wertheim College of Nursing & Health Sciences, Florida International University, Miami, FL, United States

[3]Retail, Hospitality, and Tourism Management, College of Education, Health, and Human Sciences, The University of Tennessee, Knoxville, Knoxville, TN, United States

[*]all authors contributed equally

Corresponding Author:
Hyeyoung Hah, PhD
Department of Information Systems and Business Analytics
Florida International University
11200 SW 8th Street
Miami, FL, 33199
United States
Phone: 1 3053484342
Email: hyeyoung.hah@gmail.com

Related Article:

This is a corrected version. See correction statement: https://www.jmir.org/2019/11/e17123

Abstract

Background: Telehealth technology can create a disruptive communication environment for frontline care providers who mediate virtual communication with specialists in electronic consultations. As providers are dealing with various technology features when communicating with specialists, their flexible attitude and behaviors to use various telehealth-related technology features can change the outcome of virtual care service.

Objective: The objective of this study is to examine frontline care providers' technology adaptation behaviors in the electronic consultation context. From the perspective of frontline care providers, we reapply and retest a theoretical model, reflecting a mechanism through which technology users' personal characteristics and technology adaptation behavior enhance virtual service performance, which is an important performance enabler in this online meeting context. In provider-to-provider communication, particularly, we explore the association among providers' information technology (IT)–related personal characteristics, adaptive telehealth technology use, and virtual service performance.

Methods: An online survey was administered to collect individual providers' personal traits, IT adaptation, and perception on virtual service performance. Partial least squares-structural equation modeling was used to estimate our predictive model of personal traits—IT adaptation, such as exploitative use (use the telehealth technology in a standard way), and exploratory use (use the telehealth technology as innovative way)—and virtual service performance.

Results: We collected 147 responses from graduate nursing students who were training to be nurse practitioners in their master's program, resulting in 121 valid responses from the cross-section online survey. Our theoretical model explained 60.0% of the variance in exploitative use of telehealth technology, 44% of the variance in exploratory use of telehealth technology, and 66% of the variance in virtual service performance. We found that exploitative IT use is an important driver to increase virtual service performance (β=0.762, *P*<.001), and personal characteristics such as habit are positively associated with both exploitative (β=0.293, *P*=.008) and exploratory use behaviors (β=0.414, *P*=.006), while computer self-efficacy is positively associated with exploitative use of telehealth technology (β=0.311, *P*=.047).

XSL•FO
RenderX

Conclusions: This study discusses the unique role of frontline care providers in a virtual care service context and highlights the importance of their telehealth adaptation behavior in provider-to-provider communication. We showed that providers perceive that telehealth technologies should function as intended, otherwise it may create frustration or avoidance of the telehealth technology. Moreover, providers' habitual use of various technologies in daily lives also motivates them to adaptively use telehealth technology for improving virtual care service. Understanding providers' technology habit and adaptation can inform health care policy and further provide a better view of the design of telehealth technology for online communication.

(J Med Internet Res 2019;21(8):e15087) doi:10.2196/15087

KEYWORDS

telehealth technology; adaptive technology use; frontline care providers; virtual care service; daily habit of technology use; PLS modeling; telehealth; mhealth; ehealth; digital health; adaptive technology; frontline care; virtual care

Introduction

Background

As telehealth technologies enable virtual and timely communication among care providers, frontline care providers particularly face challenges in enhancing service performance while using such technologies. In the primary care setting, care provider groups such as doctors, nurse practitioners, and nurses have been the first point of contact for people who seek health care services [1] within close proximity of patients in the location [2]. As the use of telehealth technology in electronic consultation (e-consultation) has expanded care providers' role to managing some specialty care work [3,4] beyond locational boundaries [5], it has become visible how they broker specialty visits between primary and specialty care by using telehealth technology [6]. Telehealth technology is thus supposed to enhance frontline care providers' virtual communication electronically. For example, care providers speak to patients via scheduled or on-demand/urgent visits; in addition, they communicate with specialists to ask questions and help patients avoid further face-to-face consultation with specialists [7]. Our main focus in this paper is on the latter case, often termed telespecialty consultation or e-consultation [8], which is the interaction between frontline care providers and specialists. In this environment, a patient typically does not see a specialist and relies solely on care providers' intervention to gain access to specialists, and thus, care providers' need to manage each patient's case in a timely manner while communicating with specialists. Thus, care providers are challenged to act as care moderators of the relationships between patients and specialists to manage expanded care responsibilities and improve service performance with the use of telehealth technology.

Such care providers' moderating role requires them to adequately select and use telehealth technologies for successful virtual care outcome. Telehealth technology does not refer to a single technology artifact, but to a number of electronic information and communication technologies (ICTs) to facilitate long-distance clinical care, patient and professional health-related education, and public health administration [9]. Accordingly, prior literature has noted that care providers have increasingly used multiple technologies to manage not only the new form of health care, but also virtual communication simultaneously [10-12]. However, such use of multiple technologies for virtual services has led to mixed results. Care providers perceived the use of a single telehealth technology to be beneficial to the timely management of referrals to specialists, but at the same time, they felt burdened by the additional workload that had shifted from specialists [13]. In such processes, care providers' increased use of other relevant technologies may create frustration and avoid adoption of new technology when certain technological features with which they are familiar do not perform as intended [14] or supplement the role of care providers [5]. More specifically, in the e-consultation context, care providers can feel constrained in sending messages to specialists if this familiar use of messaging technology is not integrated with other services or health systems [15]. Thus, it is fair to say that care providers' prior use of various technologies and features and their expectation of telehealth technology influence their telehealth use behavior.

However, little attention has been devoted to understanding the telehealth-driven provider-to-provider communication in which individual providers' technology use behavior as care moderators can influence virtual service performance. Several prior studies have mainly focused on the antecedents of virtual service performance [16], and yet, the mechanisms that influence care moderators' perception on the use of telehealth technology and service performance in the process of e-consultation are unknown. Given that telehealth technology shares similarities with other health technologies on various devices such as smartphones, tablets, and desktop/laptop computers [17], care providers' pre-existing experience and self-confidence in dealing with similar or new features from other technologies may not only affect their attitude about using telehealth technology [18], but also the way in which they use the telehealth technology for virtual communication [19]. In other words, a care provider may select and use a set of related telehealth technologies to manage online communication with specialists, which may allow them to enjoy familiar system features or may alter the intended capabilities of the telehealth technology artifacts (referred to as "adaptive use of IT" [20]).

Two aspects distinguish this study from prior studies. First, this study explicitly focuses on care providers' postadoption behavior using a set of telehealth technologies. Thus far, extensive research has been performed at the intersection of human computer interaction (HCI), health informatics, design science, and information technology (IT) adoption strands, with the main focus on the cognitive/psychological aspect of technology use [21,22] and interactions with technology artifacts [23,24]. For example, prior research noted mechanisms through which individual care providers' intrinsic and extrinsic motivation,

XSL•FO
RenderX

gratification, and use environment influence adoption intention [25]. Additionally, exploration of how humans interact with social, organizational, and contextual environments has developed theoretical foundations to capture users' technology adoption intention across nonhealth domains such as business, marketing, education, engineering, and agriculture (eg, [26]). This paper, however, examines an unexplored area of technology use—telehealth technologies and users' level of flexibility—to mix, match, and use them for the success of telehealth care communication. This adds value to telehealth management and relevant research, given that the current technology market for telehealth is led by care providers, tech firms, and payers, and telehealth technologies may not share common features or include all necessary features for various care regimen [27]. For example, remote patient monitoring, as one of the important aspects of telehealth service, needs multiple technologies such as videoconferencing software, peripheral devices, telemedicine carts (filed kits) for the patient site, and remote patient monitoring kits [28]. In addition, care providers use audio and video technologies for live patient care [10]. In such a telehealth care environment, care providers play crucial roles in not only mediating as the first virtual contact for patients but also cocreating care plans with specialists. Their performance can thus affect the outcome of telehealth care services, which, in turn, influence telehealth technology adoption and use of other stakeholders in rural or medically underserved areas [29]. As a consequence, comprehending users' (ie, care providers') experiences with related technologies is essential to understand their telehealth technology behavior.

Second, this paper calls for an explicit focus on virtual communication between care providers and specialists. As telehealth service consists of multilateral communications among care stakeholders, primary care providers should moderate the encounters between those parties in case of virtual specialty care needed and enhance timely and quality of telehealth care. At the time of innovation, primary care providers are known to adapt their daily routines of care management to technological innovation, which leads to decreased productivity [30]. However, little is known about care providers' ability to adapt telehealth technology for virtual communication and team-based care services [31]. From the perspective of frontline care providers who manage care processes in the location and connect to specialists remotely, the manner in which they evaluate the use of telehealth technologies and perform in this new care format can determine the success of telehealth care services in the long run. Figure 1 summarizes the unique role of frontline care providers in a telehealth care process and is the focus of this study. In this study, we limit our focus on variation in the provider-to-provider communication while making patient-side inputs fixed to better isolate care providers' technology adaptation behavior.

Taken together, our aims are to examine whether care providers' adaptive technology use behavior improves virtual service performance in e-consultation and to explore whether individual characteristics in relation to technology (personal innovativeness, computer self-efficacy, and habit) influence their adaptive use behavior. To this end, we examined two adaptive use behaviors: exploitative and exploratory use of telehealth technology. Exploitative technology use refers to the use of telehealth technology under the existing norms, while exploratory technology use involves the use of IT in a novel or unprecedented way. We hypothesized that these two adaptive IT behaviors enhance task performance on the part of care providers in the process of relaying patient information and managing specialists' diagnoses and that personal traits may affect such adaptive IT use behaviors.

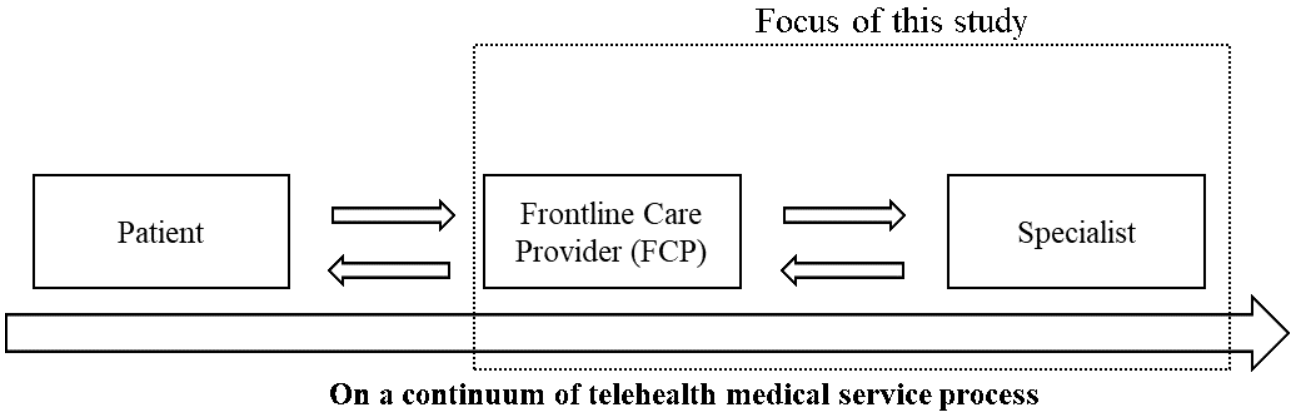
Theoretical Background

Adaptive Use of Telehealth Technology

In this paper, we define telehealth technology as electronic ICTs that support both care management and various modalities of virtual care meetings. To explore care providers' adaptive use of telehealth technology, we build on the Adaptive Structuration Theory (AST) at the level of individual users [20]. In Information Systems (IS) literature, AST explains constituents' adaptive responses to technological changes and decision outcomes in an organization. The theory describes the mechanisms through which constituents make sense of organizational, technology-driven changes by selecting, adapting, and altering existing "social structures," all of which lead either to group decision outcomes or create new structures within the organizational context [32-35]. Recently, Schmitz et al [20] extended these adaptation behaviors to the level of individual users within organizations by proposing individual-level social interactions with the focal technology and tasks. This theory states that social interaction processes can occur through two structuration episodes, including technology adaptation and task adaptation, where adaptation can be in two modes: exploitative and exploratory adaptation. Exploitative adaptation reflects the use of technology in line with existing norms and interpretations (expected use), while exploratory adaptation indicates technology use based on nonstandard interpretations (unexpected use). The dynamic effects of adaptive behaviors on individuals' performance have been demonstrated in various research settings [20] such as job performance and satisfaction in nonhealth domains [36].

XSL•FO
RenderX

Figure 1. Focus of this study.



In health care, care providers' adaptive use of telehealth technologies can play an important role in provider-to-provider communication. Weigel at al [37] defined individuals' adaptive use of health information technology (HIT) as "temporary or permanent modifications that a user makes to his or her behaviors or norms due to the limitations of the HIT." Thus, if technologies do not fully support e-consultation, this condition can elicit users' adaptation behaviors [15]. In theory, the adaptive use of technology takes place when users have experienced various technologies and are expected to use existing or new technologies adaptively. In this case, the possible responses are either that they use a system feature as it is or modify some features to produce better outcomes. In the telehealth context, the adaptive use of telehealth technology is also expected, as health stakeholders increasingly use multiple technologies to communicate virtually with one another for care management beyond office visits. For example, care providers use personal messaging apps (eg, Whatsapp) and social network sites [11] to moderate communications between patients and care providers in addition to the designated telehealth technologies [7,38]. Hence, care providers' adequate selection and use of other communication technologies may change the outcomes of telehealth services. In IS literature, such a moderating role of care providers is analogous to that of online meeting facilitators, whose behaviors not only affect the meeting outcomes, but also the other meeting attendees' behaviors [39]. More specifically, there are two important factors that affect online meeting outcomes: one is facilitators' personal characteristics such as their level of experience and facilitation [40], and the other is their technology-based skills [41-43]. In this paper, we focus on care providers' technology-based skills, ie, technology adaptation behaviors and technology-driven personal traits as key determinants to influence virtual service performance.

Virtual Service Performance

This study considers virtual care performance to be one of the virtual meeting outcomes that captures care providers' expected outcomes in response to adequate use of telehealth-related technologies in the e-consultation context. In general, the success of IT use or performance has been considered at multiple levels, such as at the individual, group, and organizational levels [44,45], and individual performance has been widely studied as a key dependent variable to measure individual users' postadoptive IT use behavior [46-48]. Applied to our research context, care performance from primary care services has been measured by patients' satisfaction [49], care providers' evaluation of care coordination [50], or organizational performance [51]. In fact, telehealth technologies need to be able to support various communication modalities via video-conferencing, texting, and a combination of both during care provision [52]. Hence, it is critical that care providers apply effective virtual communication skills when managing electronic patient data [53] and use various technological features, all of which depend upon their proper choice and actual use of the technologies. Thus, we suggest that care providers' virtual service performance be determined by the adaptive use of telehealth technologies, and their performance can be improved with respect to effectiveness of care, care management, quality of care tasks, decreased error rates in communication, and sharing information [54].

Personal Innovativeness in Information Technology, Computer Self-Efficacy, and Habit

Care providers' responses to these challenges can be reinforced or redirected by their personal characteristics. Prior studies on information technology use have documented the importance of personal characteristics that help users experiment with and control new technology based on their beliefs and experience. When IT users communicate with others online, in general, and when they act as meeting facilitators in a virtual meeting, in particular, the user who presides over the meeting becomes more important, because individual characteristics such as technology-based skills, capabilities, and level of experience [55] affect the success of online meetings and meeting members' use of technologies [11,38]. Given that care providers need to facilitate virtual communication with specialists, these users' characteristics and beliefs about the use of telehealth and relevant technologies simultaneously help to predict the online communication outcome, which is virtual service performance in this study. Thus, care providers' willingness and capability to use multiple telehealth-related technologies (including familiar and new features) and whether care providers possess characteristics that allow them to introduce more innovative ways of technology use can influence virtual service performance [56].

XSL•FO
RenderX

The literature on postadoptive technology, behavior has acknowledged that users' personal characteristics are important antecedents that explain their postadoption behavior [57,58]. In this model, we identified personal innovativeness in IT use, computer self-efficacy, and habit as determinants to explain care providers' adaptive telehealth technology use behaviors. First, personal innovativeness in IT use is defined as "the willingness of an individual to try out any new information technology" [57]. This concept has been widely used in IT use studies to capture individuals' intention to adopt technology, both generally [59] and in health care domains [60-62]. In the adaptive use context, Chow et al [63] found that personal innovativeness positively influenced adaptive IT use behavior. The higher innovativeness a user has, the more likely that he or she is to try new features and mix and match system features that are relevant to tasks (eg, by replacing some existing features with new ones, combining features, or inventing new ways to use certain features for tasks for which they were not intended). Second, computer self-efficacy—referring to a user's belief about his or her capability to control telehealth technologies—is also likely to influence IT users' motivation and outcomes [64,65]. Users' self-judgments about technology efficacy influence their beliefs about a focal technology's ease of use [65]. In postadoption IT use, individual users' beliefs about their ability to use new technology are associated with the technology's deep structural use [66,67]. Lastly, habit concerns the notion that the "habitualization of action occurs more or less automatically via a subconscious response to a work situation" [68]. Thus, people may be willing to adopt a new workplace technology when they understand other technologies in their lives. As automatic reactions to certain tasks that are attributable to prior learning, habits have been identified as predictors of technology adoption or moderators that interact with other factors in postadoption IT use [69-71]. Moreover, habits have been associated with continued use of IT [72,73] and its adaptive use [74]. Schmitz [20] used experience of technology as a personality trait, whereas we used habit instead. This is because habit captures an automatic reaction to certain tasks due to prior learning from technology, while experience reflects users' exposure to a focal technology in the passage of time [68]. As we focus on users' prior learning from the use of various technologies, habit is more applicable to our telehealth care context. Taken together, individual characteristics act as key antecedents that predict care providers' adaptive use behaviors.

Research Model and Hypothesis Development

Exploitative Use of Telehealth Technologies and Virtual Service Performance

According to Schmitz [20], exploitative use of focal technology occurs "when a user modifies technology features to facilitate usage of the technology consistent with how s/he perceives is intended or standard for the technology." Thus, exploitative IT use reflects the routine use of IT under existing norms and expectations [36]. Exploitative use of technology occurs in various settings. For example, users employ IT in repetitive tasks to improve efficiency [75], complete tasks [76], and maximize task performance [77]. In a provider-to-provider context, care providers use personal messaging apps (eg, Whatsapp) to expedite communication with specialists after submitting e-consultation requests on telehealth technology platforms [11]. In this case, combined use of the telehealth technology and personal messaging applications that accomplish repetitive tasks can facilitate instantaneous communication with specialists, all of which influence virtual care performance. Thus, we hypothesize the following:

> *Hypothesis 1: Exploitative use of telehealth positively affects virtual care performance.*

Exploratory Use of Telehealth Technologies and Virtual Service Performance

Exploratory technology use takes place "when a user develops new technology features to facilitate usage of the technology that s/he perceives is unusual or non-standard for the technology" [20]. Nontraditional IT use allows users to identify certain new capabilities of IT, such as exploring new skills and experimentation [76], and to make nonstandard interpretations of the focal phenomenon, leading to divergent consequences [20]. Accordingly, exploratory use of telehealth technology for e-consultation indicates an innovative use of telehealth technology that fosters deviation from existing tasks and the search for alternatives [77]. For example, to achieve timely communication with specialists, care providers need to be capable of managing images and reports and interacting with their electronic health records, which are typically accessed through their mobile phones [78]. Finding ways to quickly process multimedia images or connect hospital systems via interface applications installed on care providers' devices might be examples of the exploratory use of telehealth technologies. Therefore, we made the following hypothesis:

> *Hypothesis 2: Exploratory use of telehealth positively affects virtual care performance.*

Individual Characteristics as Antecedents of the Adaptive Use of Telehealth Technologies

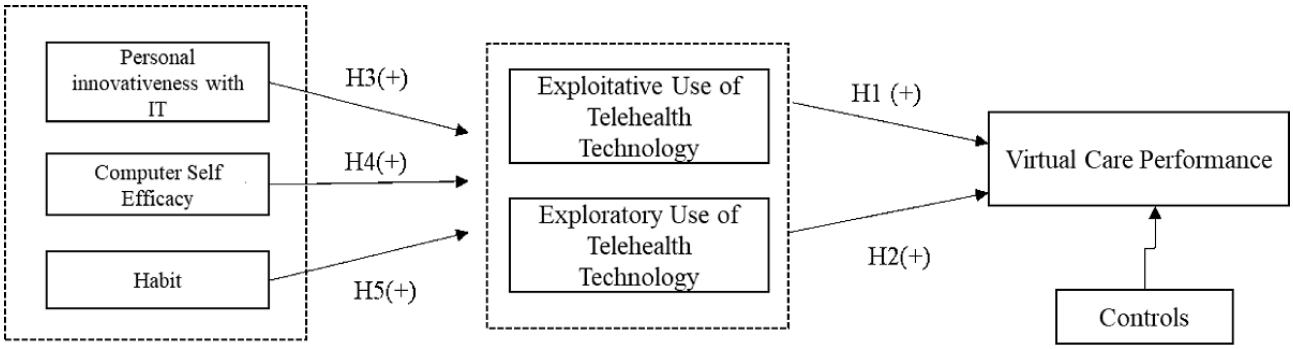
This paper proposes that personal traits are formulated via cumulative exposure to various technologies across multiple life domains (such as the workplace and at home) and hypothesizes that such traits can expand care providers' capabilities to use both existing and new features in telehealth-related technologies. Individual care providers' existing beliefs about self-innovativeness, self-judgment about telehealth technology use, and accumulated habits from using daily technologies across multiple life domains can influence the way in which they adaptively use new features of telehealth technologies in both expected and reconfigured ways, particularly for communication with providers (Figure 2). Therefore, we made the following hypotheses:

> *Hypothesis 3: Care providers' personal innovativeness with IT positively affects their adaptive (H3a: exploitative; H3b: exploratory) use of telehealth technologies.*
>
> *Hypothesis 4: Care providers' computer self-efficacy positively affects their adaptive (H4a: exploitative; H4b: exploratory) use of telehealth technologies.*

XSL•FO
RenderX

Hypothesis 5: Care providers' habits with regard to technology positively affect their adaptive (H5a: exploitative; H5b: exploratory) use of telehealth technologies.

Figure 2. Research model. H: hypothesis; IT: information technology.



Methods

Recruitment

We recruited graduate nursing students at a university in the Southeastern United States who were studying to become family nurse practitioners and had experienced various types of technologies across different contexts. In telehealth care contexts, in particular, the demand of nurse practitioners is growing because they can collaborate with other specialists to comanage patient care cases for rural or disadvantaged population [42]. With incorporation of such trends and our research focus on nurse practitioners' role as frontline care providers in e-consultations with specialists, we had to use purposive sampling for those who could provide unique information about telehealth care services that could not be obtained elsewhere [79]. The participants had been trained in the use of telehealth technologies in their education program and their communication with other care members in the simulation rooms, in which the nursing faculty had recorded and evaluated their electronic communication. In summary, our student sample could be legitimately used to capture their adaptive use of telehealth technologies. This nonprobability sampling met our goals to identify a representative sample of nurse practitioners who would be at the front line of telehealth care services as care providers (see [78] for the taxonomies of purposive sampling). To test our hypotheses, a Web-based survey was administered in the spring and fall semesters in 2018. In the online survey, each participant read a case scenario and responded to the questionnaire in terms of their perception on prior experience, adaptive use of telehealth technology, and care service performance (Multimedia Appendix 1).

Statistical Analysis

Instrument Development

The measurement items were adopted from well-established IS literature and adapted to the telehealth context (Multimedia Appendix 2). To measure the dependent variable, we adapted Goodhue's [47] five-item scale of task performance. This measure has been used to explain IT postadoption performance in various contexts, including health care [58]. Two independent variables—exploitative and exploratory use—were measured using items from a previous study [20], in which each item was assessed with four indicators. With regard to the input variables (antecedents to postadoptive IT use), personal innovativeness in IT was measured using three items adapted from a prior study [57]. Computer self-efficacy (a user's perception of his or her competency in using computers) was measured using four items developed by Compeau and Higgins [64]. Habit reflects care providers' automatic reactions to a new technology based on prior experience and was captured using four items [68]. Finally, we collected demographic variables (age, income, education, gender, occupation status, and experience in using health apps) and mobile technology usage experience and treated them as control variables in the data analysis. All research variables were reflective constructs and measured on a seven-point Likert scale that ranged from 1 ("strongly disagree") to 7 ("strongly agree"). To assess face validity, the researchers contextualized the survey scale to make it pertinent to the telehealth care situation. A pretest was performed to ensure content validity. The researchers and two nurse practitioners evaluated and refined each survey item. A total of 37 initial responses were used to revise and finalize the questionnaire for the final survey.

Analysis

This study used composite-based partial least squares-structural equation modeling (PLS-SEM) in SmartPLS (Version 3.0. Boenningstedt, Germany: SmartPLS GmbH) to assess the measurement and structural models [80]. PLS-SEM is a causal model that has been frequently used in IS literature [81,82] and is used to maximize the variance that the dependent latent constructs explain [83]. We applied a bias-corrected and accelerated bootstrapping procedure with replacement using 5000 subsamples. Our hypotheses were tested using a one-tailed *t* test for unidirectional hypotheses.

Results

User Characteristics and Descriptive Statistics

A cross-sectional online survey was sent to 210 students who were enrolled in the nurse practitioner education program. A total of 146 participants responded to the survey, yielding a response rate of 69.52%; of these, 121 valid responses were used for data analysis. Twenty-five responses were dropped because they only reported demographic information in the surveys. Table 1 summarizes the respondent profiles. The majority were females (72.7%) and employed (full-time workers=54.5%; part-time workers=35.5%). All were well

XSL•FO
RenderX

educated, with an undergraduate or higher degree, and nearly half were of white race (47.5%). Table 1 summarizes the characteristics of the survey participants.

In addition, an adequate sample size is necessary to estimate the PLS path model, which is guided by the 10-times rule and power analysis. On one hand, the 10 times rule dictates the use of "10 times the largest number of formative indicators used to measure a single construct, or 10 times the largest number of structural paths directed at a particular construct in the structural model" [83]. On the other hand, power analysis provides a threshold for the statistical power necessary to detect an effect based on the maximum number of independent variables in the measurement and structural models. In our case, we had five independent variables, three maximum arrows pointing to a latent construct, and thus needed at least 45 observations to achieve a statistical power of 80% and R^2 values of at least 0.25 (with a 5% probability of error) [83]. Thus, our sample size (N=121) was deemed adequate to test the research model.

XSL•FO
RenderX

Table 1. Participant characteristics (N=121).

Demographic variables	Values, n (%)
Gender	
Male	33 (27.3)
Female	88 (72.7)
Age (years)	
18-25	12 (9.9)
26-40	85 (70.2)
41-55	20 (16.5)
56-65	4 (3.3)
Income status (US $)	
25,000-49,999	24 (19.8)
50,000-74,999	48 (39.7)
75,000-99,999	16 (13.2)
≥100,000	13 (10.7)
Prefer not to answer	20 (16.5)
Education	
Bachelor's degree	72 (59.5)
Master's degree	41 (33.9)
PhD	2 (1.7)
Others	6 (5.0)
Occupation[a]	
Working full time	66 (54.5)
Working part time	43 (35.5)
Unemployed	8 (6.6)
Unable to work	1 (0.8)
Other	3 (2.5)
Race[b]	
African American	21 (17.5)
Asian	18 (15.0)
Native Hawaiian or Pacific Islander	1 (0.8)
White	57 (47.5)
Other	19 (15.8)
Prefer not to answer	4 (3.3)

[a]All demographic questions were optional. Four respondents reported their occupation status as either "unable to work" or "other." For clarity, we removed these responses and reran partial least squares analysis, producing the identical results.

[b]N=120.

Nonresponse Bias and Common Method Bias

As our survey was self-reported, we evaluated two possible biases carefully: nonresponse bias and common method bias. Nonresponse bias derives from the differences between participants and nonparticipants in the survey [84,85]. This bias can be assessed by comparing our sample's characteristics with those in the population and by comparing early and late respondents. We compared the early respondents (74.82%) and late respondents (25.18%) on each of the demographic characteristics (age, gender, education, income, and occupational status) and health application experience using a *t* test. There were no significant differences between the early and late respondents in our sample.

Common method bias potentially threatened the veracity of our results, as data for the independent variables and dependent variable were collected in the same survey. Following a previous

XSL•FO
RenderX

study [86], we designed the survey instrument's contents and order carefully. Furthermore, we performed the Harman single factor analysis to assess the bias. The results showed that one factor explained 36.90% of the variance, confirming that no single factor accounted for the majority of covariance. Thus, nonresponse bias and common method bias did not threaten this study's findings.

Measurement Model

To determine each construct's internal reliability, we first examined the item loadings and composite reliabilities. Each item loaded above 0.75 on its respective construct and was significant at $P<.05$. Cronbach alpha was calculated to assess composite reliability and confirmed that all items' values were above 0.7. Convergent validity was established if the average variance extracted (AVE) was above the threshold of 0.5, which suggests that the variance explained by indicators was greater than the unexplained variance. Discriminant validity was tested by assessing the Fornell-Larcker criterion and cross-loadings. To confirm the discriminant validity, the AVE of each construct should be greater than its squared correlations with other constructs [87]. As shown in Tables 2 and 3, internal reliability, convergent validity, and discriminant validity were all confirmed.

Hypothesis Testing: Partial Least Squares Modeling

Following structural model assessment suggestions of de Guinea [69], we evaluated the proposed path model using SmartPLS 3.0. The structural model's quality was assessed by checking multicollinearity (variance inflation factor), path coefficients, R^2 (variance explained), f^2 (effect size), and the Stone-Geisser Q^2 (model's predictive relevance). The variance inflation factor was checked and confirmed to be less than 5, indicating that multicollinearity was not a problem in the study. We report the path coefficients' significance, R^2 and f^2 in the full model results. Effect size (f^2) explains the changes when an exogenous construct of focus is included and when it is omitted from the model. As a rule of thumb, if f^2 is 0.02, 0.15, and 0.35, the effects are considered to be small, medium, and large, respectively [88]. Lastly, to assess the model's predictive power, the Stone-Geisser Q^2 was used to indicate the sample's predictive relevance. A Q^2 value>0 demonstrates that the path model has predictive relevance to a reflective, endogenous latent variable. The Q^2 values for three of the endogenous constructs were >0, indicating exploitative use (Q^2=0.51), exploratory use (Q^2=0.36), and virtual service performance (Q^2=0.54). Thus, the model's predictive relevance was confirmed.

For the path coefficients' significance and the variance explained, R^2, our results demonstrated that exploitative technology use was positively associated with virtual service performance (β=0.76, $P<.001$), while exploratory use did not explain the variation in virtual service performance (β=0.036, $P=.49$). With regard to the effects of individual characteristics, computer self-efficacy was significantly associated with exploitative technology use (β=0.31, $P=.05$). Lastly, habit was associated with both exploitative (β=0.29, $P=.04$) and exploratory use (β=0.41, $P=.006$). Therefore, hypotheses 1, 4a, 5a, and 5b were supported but hypotheses 2, 3a, 3b, 4b, and 5b were not supported. Among the control variables, education (β=–0.12, $P=.01$) was shown to affect virtual service performance negatively, while income level (β=0.11, $P=.008$) was positively associated with virtual service performance, as shown in Table 4 and Figure 3.

XSL•FO
RenderX

Table 2. Internal and convergent validity.

Construct and items	Factor loading	Cronbach alpha	Average variance extracted	Mean (SD)
HAB[a]		0.96	0.89	5.38 (1.22)
HAB1	0.95			
HAB2	0.92			
HAB3	0.94			
HAB4	0.96			
PIT[b]		0.97	0.94	4.67 (0.86)
PIT1	0.97			
PIT2	0.96			
PIT3	0.97			
CSE[c]		0.97	0.92	5.90 (0.96)
CSE1	0.94			
CSE2	0.96			
CSE3	0.96			
CSE4	0.96			
EIU[d]		0.98	0.94	5.92 (1.14)
EIU1	0.98			
EIU2	0.98			
EIU3	0.99			
EIU4	0.92			
ERU[e]		0.96	0.9	4.78 (1.52)
ERU1	0.92			
ERU2	0.94			
ERU3	0.96			
ERU4	0.96			
PERF[f]		0.97	0.9	5.63 (1.22)
PERF1	0.95			
PERF2	0.97			
PERF3	0.97			
PERF4	0.96			
PERF5	0.89			

[a]HAB: habit.

[b]PIT: personal innovativeness with information technology.

[c]CSE: computer self-efficacy.

[d]EIU: exploitative use.

[e]ERU: exploratory use.

[f]PERF: care performance.

XSL•FO
RenderX

Table 3. Discriminant validity. Diagonals represent the value of the average variance extracted, and off-diagonal elements are the squared correlations among construct.

Constructs	1	2	3	4	5	6
1. Care performance	0.95	—[a]	—	—	—	—
2. Computer self-efficacy	0.59	0.96	—	—	—	—
3. Exploitive use	0.80	0.74	0.97	—	—	—
4. Explorative use	0.64	0.60	0.79	0.95	—	—
5. Habit	0.61	0.86	0.74	0.65	0.94	—
6. Personal innovativeness with information technology	0.59	0.86	0.73	0.62	0.87	0.97

[a]Not applicable.

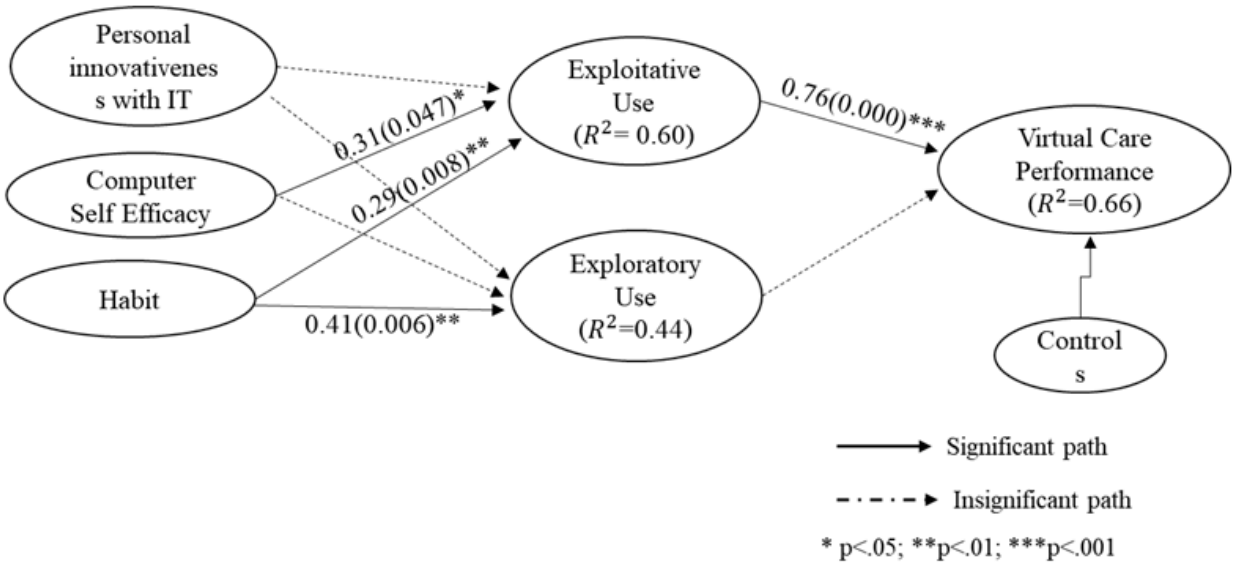
Table 4. Complete results of the hypothesis testing.

Path	β[a]	SD	*t* test	*P* value	f^{2}[b]	Effect size	Hypothesis testing
Virtual service performance							
H[c]1: Exploitive use	0.762	0.075	10.16	<.001	0.483	Large	Supported
H2: Exploratory use	0.036	0.049	0.69	.49	0.001	—[d]	Not supported
Exploitative use							
H3a: Personal innovativeness with IT[e]	0.201	0.134	1.531	0.13	0.019	—	Not supported
H4a: Computer self-efficacy	0.311	0.155	1.991	0.047	0.05	Small	Supported
H5a: Habit	0.293	0.112	2.648	0.008	0.042	Small	Supported
Exploratory use							
H3b: Personal innovativeness with IT	0.168	0.155	1.095	0.27	0.01	—	Not supported
H4b: Computer self-efficacy	0.102	0.149	0.672	0.5	0.004	—	Not supported
H5b: Habit	0.414	0.15	2.781	0.006	0.06	Small	Supported
Control variables							
Age	0.01	0.06	0.11	0.91	0	—	N/A[f]
Education	–0.12	0.05	2.44	0.02	0.031	Small	N/A
Gender	–0.09	0.05	1.67	0.1	0.009	—	N/A
Income	0.11	0.04	2.65	0.01	0.02	Small	N/A

[a]Standard regression coefficient.
[b]Effect size.
[c]H: hypothesis.
[d]Not available.
[e]IT: information technology.
[f]N/A: not applicable.

XSL•FO
RenderX

Figure 3. Structural evaluation of the telehealth adaptive use model. IT: information technology.



Discussion

Principal Findings

The purpose of this study was to contextualize and test a research model that examined the determinants of the adaptive use of telehealth technologies in e-consultation. Using AST at the individual level, we examined the mechanisms by which exploitative and exploratory use of telehealth technologies and three personal traits (ie, personal innovativeness in IT, computer self-efficacy, and habit) influence virtual service performance.

Our results indicate that care providers are willing to use telehealth technologies exploitatively to communicate with specialists in e-consultations and perceive it as a virtual service performance enhancer. With regard to the hypotheses on adaptive use and virtual service performance (hypotheses 1 and 2), exploitative use of telehealth technologies was found to be a strong factor that explains care providers' virtual service performance (β=0.762, *P*<.001), while exploratory use was not (β=0.036, *P*=.49). Moreover, with regard to hypotheses 3 through 5, we found that personal innovativeness in IT is insignificant for explaining adaptive use in our context (*P*=.13 for exploitative use and *P*=.27 for exploratory use, respectively); computer self-efficacy has a significant, positive effect on the exploitative use of telehealth technologies (β=0.311, *P*=.047); and habitual use of nonhealth technologies in daily life is associated with care providers' willingness to engage in both exploitative (β=0.293, *P*=.008) and exploratory adaptive use of new telehealth technology (β=0.414, *P*=.006).

Limitations

Despite the meaningful and practical findings in this study, our results should be interpreted with caution due to some limitations. First, we selected a purposive sample of graduate nursing students for our study. Although this is a legitimate sample, the ability to generalize the findings may be limited because our participants have experienced e-consultation in the education-focused, simulated contexts. Second, while our cross-sectional sample shows adequate responses to estimate our hypothesized path model using PLS-SEM, a larger sample with a panel structure would increase the statistical power of the findings and control for any unobserved confounding factor using panel data analysis. Particularly, time-series cross-sectional data on capturing various technology use among care providers would provide more in-depth understanding of virtual telehealth care services. Third, we did not include task adaptation behavior based on our research context's characteristics. The education program requires individuals to adhere to the standard protocol of patient cases, such that any adaptive task behaviors are evaluated as a failure in the medical setting. Thus, we limited our focus to the participants' adaptive use of telehealth technology only. It would be meaningful to investigate task and technology adaptation behaviors in the health care contexts in which the selection of tasks and technologies are flexible in future research. Fourth, this research considered care providers' perception about their willingness to use telehealth technology either in a traditional way or in an innovative way and its downstream effect on perceived service performance. This is because we viewed telehealth technology as a set of related information and communication technologies and asked participants' adaptation behavior in a general context. It would be worthwhile to revisit our research model to certain care contexts (eg, diabetes) and capture actual measures of virtual service performance. Lastly, future research can be further extended to explore other stakeholders' attitude and behaviors about telehealth technology in e-consultation. For example, our research model can explain how adaptive use of telehealth technology by frontline care providers influences the level of patients' satisfaction and health outcomes as well as those of specialists.

Comparison with Prior Work

This study contributes to both the IS and health care literature on the postadoptive use of telehealth technology. First, this research contributes to the IS literature by exploring adaptive IT use behavior from an individual meeting facilitator's perspective and attempting to identify a contextualized theory

XSL•FO
RenderX

in e-consultation. Specifically, we considered e-consultations between care providers and specialists in a virtual meeting context and proposed that as virtual meeting facilitators, care providers' willingness to use telehealth technologies is an important predictor of virtual service tasks' success. Previous IS literature on online meeting technology use has documented that technology types, environmental factors, and user characteristics are key factors that predict outcomes from a multilevel perspective. From a group perspective, technical support for group users and the fit between tasks and technologies are important determinants of the success of online meeting technology use [89]. In terms of individuals' behavior, extant studies have emphasized the important role that human facilitators' characteristics play in predicting the outcomes of online meetings [43]. Given care providers' unique position as those who relay clinical information and decisions between patients and specialists in health care and the availability of flexible options to select and use telehealth technologies for provider-to-provider communication, there is an urgent need to examine whether, and in what way, such meeting facilitators' adaptive use of IT predicts the success of virtual service performance.

Our study is unique in that we focused explicitly on care providers' virtual role in telehealth communication and their performance under a new definition of telehealth technology (ie, use of telehealth and telehealth-related technologies). Care providers' role differs from that of general online facilitators in online meetings, in that they manage multilateral communication between patients and specialists and although there is a designated telehealth platform for their communications, sometimes, communication with the two different groups involves the use of additional technologies that complement or substitute the existing technology's capabilities. Given that providers [58] and patients [90] were the main user groups of interest in previous studies for predicting telehealth technologies' success, this study contributes to the health care literature by examining the adaptive structuration theory of individual from a care provider's perspective and identifying salient constructs (the exploitative use of IT) in a research model to develop a contextualized theory. Second, the result that habit plays a significant role in adaptive IT use calls for more attention in IS research on how users' habits accumulate from different life domains to affect their postadoptive use of multiple technologies [91]. Previous postadoption studies have established that habit reflects the extent to which people tend to perform behaviors automatically because of past learning [69,92-95]. However, the relationship between habit and the continued use of IT has been mostly tested in a single domain; for example, mobile phone habits predict mobile phone use [93] and mobile internet habits affect mobile internet technology use [95]. In a health care setting, however, our findings emphasize that cross-domain, habitual IT use influences care providers' adaptive technology behaviors. Our interviews with two nurse practitioners in family medicine also reflected these positive effects of the habitual use of technology in nonwork domains. One nurse practitioner stated,

> *My use of personal non-health-related technology gives me hope that I am computer savvy and would be able to learn new computer technology that is used in patient settings.*

Another said:

> *Because of technology, I am able to see who is waiting in patient rooms and who is still waiting in the waiting room. When the numbers are high, it makes […] clinician[s] want to work faster so that they are not too delayed.*

As telehealth medical services include three different communication modalities (video-conferencing, texting, and a hybrid of the two), care providers' existing technology habits can help them select and use these three forms of communication and manage communication with patients and other providers.

Moreover, we found a strong effect of exploitative use (expected use of telehealth technology) on virtual service performance by frontline care providers. These findings are in line with those of a previous study [20], such that two modes of IT adaptation behaviors (exploitative and exploratory use) are differentially salient across research contexts. In other words, exploitative and exploratory adaptation behaviors may not coincide under the same context because users have different coping strategies toward information technologies [96]. For example, explorative use of technology became salient in context of mobile technology in the BYOD (bring your own device) context [20]; in the contexts of enterprise resource planning and product lifecycle management system use, exploratory as well as exploitative adaptation were differentially significant, contingent upon input factors [36]. In our research context, the strong effect of exploitative use of telehealth technology may be due to the contextual characteristics of care process. Actually, health care is a controlled and highly concentrated environment such that care providers expect the technologies to function as expected by supplementing their clinical tasks [5]. Since this study explored technology-related traits as antecedents to technology adaptation behavior in provider-to-provider communication, it is much anticipated that such technology adaptation behavior can vary by different contexts and heterogenous technology users.

Lastly, education and income were shown to differentially influence virtual service performance. Prior studies have documented negative effects of demographic variables on technology use, as less educated participants may have less knowledge, whereas those who earn less income may have less opportunities to access advanced information technology [97,98]. In our research context, we can interpret that the current level of technology education from bachelor's degree may not be on par with specifics of telehealth technology use for communication with specialists. In addition, our graduate nursing students with high income level may have been exposed to various technologies within and outside the education program or care settings. Thus, demographic characteristics of individual users need to be included when exploring technology adaptation behavior.

Practical Implications

This study's findings can be applied to inform health care practitioners and health app designers. Strategic IT management

XSL•FO
RenderX

is necessary for care providers who serve as virtual meeting managers in a telehealth setting [81]. As telehealth medical services have garnered much attention, nurse practitioners have played an increasingly important role in supporting various online care services in which multiple technologies need to be operated appropriately. Our results demonstrate that care providers' adaptive use of technology can help predict telehealth care performance, and therefore, more consideration should be given to the role of intermediary care providers in the care process between patients and specialists. As organizational structures influence both offline and online meeting outcomes [99], health practitioners need to focus on organization-level strategies to enhance care providers' online facilitation by examining the gaps that they have experienced using a variety of technologies across multiple life domains and providing relevant education in the use of focal technology [31,100]. Moreover, this study can be beneficial to telehealth designers and developers in terms of the design of HITs. Prior studies have documented that telehealth apps' design of features, icons, and terminologies is important and that care providers expect all of these to function as intended [14]. As health app developers continue to add new features to stay abreast of rapidly changing health care trends, it is important for them to consider health care consumers' needs and users' familiarity and comfort with the existing features that are evolving across a wide range of technologies and systems [101].

Conclusions

This study investigated frontline care providers' unique role in e-consultation with specialists. By regarding the care providers explicitly as virtual meeting facilitators, we tested the association between their adaptive use of multiple telehealth-related technologies and virtual service performance. Care providers' standard use of telehealth technologies was shown to be a salient factor that predicts success in virtual service, while the innovative use of telehealth technologies remained insignificant. Among their personal characteristics, the habits and computer self-efficacy that care providers acquired and developed in nonwork settings stimulated and enhanced their willingness to use multiple telehealth technologies in standard and creative ways.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Survey scenario.
[PDF File (Adobe PDF File), 117KB - jmir_v21i8e15087_app1.pdf]

Multimedia Appendix 2
Survey instruments.
[PDF File (Adobe PDF File), 138KB - jmir_v21i8e15087_app2.pdf]

References

1. Agarwal S, Perry B, Long A, Labrique B. Evidence on feasibility and effective use of mHealth strategies by frontline health workers in developing countries: systematic review. Trop Med Int Health 2015 Aug;20(8):1003-1014 [FREE Full text] [doi: 10.1111/tmi.12525] [Medline: 25881735]
2. Frontline Health Workers Coalition. 2019. Frontline Health Workers URL: https://www.frontlinehealthworkers.org/frontline-health-workers [accessed 2019-07-24]
3. Daskivich LP, Vasquez C, Martinez C, Tseng C, Mangione CM. Implementation and Evaluation of a Large-Scale Teleretinal Diabetic Retinopathy Screening Program in the Los Angeles County Department of Health Services. JAMA Intern Med 2017 May 01;177(5):642-649 [FREE Full text] [doi: 10.1001/jamainternmed.2017.0204] [Medline: 28346590]
4. Kaufman D, Pevzner J, Rodriguez M, Cimino JJ, Ebner S, Fields L, et al. Understanding workflow in telehealth video visits: observations from the IDEATel project. J Biomed Inform 2009;42(4):581-592 [FREE Full text] [doi: 10.1016/j.jbi.2009.03.012] [Medline: 19358897]
5. MacNeill V, Sanders C, Fitzpatrick R, Hendy J, Barlow J, Knapp M, et al. Experiences of front-line health professionals in the delivery of telehealth: a qualitative study. Br J Gen Pract 2014;64(624):e401-e407 [FREE Full text] [doi: 10.3399/bjgp14X680485] [Medline: 24982492]
6. Magadzire BP, Budden A, Ward K, Jeffery R, Sanders D. Frontline health workers as brokers: provider perceptions, experiences and mitigating strategies to improve access to essential medicines in South Africa. BMC Health Serv Res 2014 Nov 05;14:520. [doi: 10.1186/s12913-014-0520-6] [Medline: 25370799]
7. Joschko J, Keely E, Grant R, Moroz I, Graveline M, Drimer N, et al. Electronic Consultation Services Worldwide: Environmental Scan. J Med Internet Res 2018 Dec 21;20(12):e11112 [FREE Full text] [doi: 10.2196/11112] [Medline: 30578187]
8. Hansen SW, Gogan JL, Baxter RJ, Garfield MJ. Informed collaboration in health care: An embedded-cases study in geriatric telepsychiatry. Info Systems J 2018 Aug 22;29(2):514-547. [doi: 10.1111/isj.12218]

XSL•FO
RenderX

9. Healthit.gov. Telemedicine and Telehealth URL: https://www.healthit.gov/topic/health-it-initiatives/telemedicine-and-telehealth [accessed 2019-07-24]
10. Mobihealthnews. 2017. Enterprise video conferencing company Zoom teams up with Epic to launch configurable telehealth platform URL: https://www.mobihealthnews.com/content/enterprise-video-conferencing-company-zoom-teams-epic-launch-configurable-telehealth [accessed 2019-07-24]
11. Chan WS, Leung AY. Use of Social Network Sites for Communication Among Health Professionals: Systematic Review. J Med Internet Res 2018 Mar 28;20(3):e117. [doi: 10.2196/jmir.8382] [Medline: 29592845]
12. 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 May 20;21(6):323-330. [doi: 10.1177/1357633X15582108]
13. Lee MS, Ray KN, Mehrotra A, Giboney P, Yee HF, Barnett ML. Primary Care Practitioners' Perceptions of Electronic Consult Systems: A Qualitative Analysis. JAMA Intern Med 2018 Dec 01;178(6):782-789 [FREE Full text] [doi: 10.1001/jamainternmed.2018.0738] [Medline: 29801079]
14. Klingberg A, Wallis LA, Hasselberg M, Yen P, Fritzell SC. Teleconsultation Using Mobile Phones for Diagnosis and Acute Care of Burn Injuries Among Emergency Physicians: Mixed-Methods Study. JMIR Mhealth Uhealth 2018 Oct 19;6(10):e11076 [FREE Full text] [doi: 10.2196/11076] [Medline: 30341047]
15. Hensel JM, Yang R, Rai M, Taylor VH. Optimizing Electronic Consultation Between Primary Care Providers and Psychiatrists: Mixed-Methods Study. J Med Internet Res 2018 Dec 06;20(4):e124 [FREE Full text] [doi: 10.2196/jmir.8943] [Medline: 29625949]
16. Buck J, Manges K, Kaboli P. Asynchronous teleneurology: A systematic review of electronic provider-to-provider communications. Neurology 2016;86:3.400.
17. Healthcare IT News. 2018. Nurse burnout? Try telehealth, clinical decision support and analytics tools, expert says URL: https://www.healthcareitnews.com/news/nurse-burnout-try-telehealth-clinical-decision-support-and-analytics-tools-expert-says [accessed 2019-07-24]
18. Brewster L, Mountain G, Wessels B, Kelly C, Hawley M. Factors affecting front line staff acceptance of telehealth technologies: a mixed-method systematic review. J Adv Nurs 2014 Jan;70(1):21-33. [doi: 10.1111/jan.12196] [Medline: 23786584]
19. Schoen C, Osborn R, Squires D, Doty M, Rasmussen P, Pierson R, et al. A survey of primary care doctors in ten countries shows progress in use of health information technology, less in other areas. Health Aff (Millwood) 2012 Dec;31(12):2805-2816. [doi: 10.1377/hlthaff.2012.0884] [Medline: 23154997]
20. Schmitz K, Teng J, Webb K. Capturing the Complexity of Malleable IT Use: Adaptive Structuration Theory for Individuals. MIS Quarterly 2016 Mar 3;40(3):663-686. [doi: 10.25300/MISQ/2016/40.3.07]
21. Venkatesh V, Sykes TA, Zhang X. 'Just what the doctor ordered': a revised UTAUT for EMR system adoption and use by doctors. 2011 Presented at: The 44th Hawaii International Conference on System Sciences; 2011; Hawaii.
22. Dünnebeil S, Sunyaev A, Blohm I, Leimeister JM, Krcmar H. Determinants of physicians' technology acceptance for e-health in ambulatory care. Int J Med Inform 2012;81(11):746-760. [doi: 10.1016/j.ijmedinf.2012.02.002] [Medline: 22397989]
23. Tang Z, Johnson TR, Tindall RD, Zhang J. Applying heuristic evaluation to improve the usability of a telemedicine system. Telemed J E Health 2006;12(1):24-34. [doi: 10.1089/tmj.2006.12.24] [Medline: 16478410]
24. Agnisarman S, Narasimha S, Chalil Madathil K, Welch B, Brinda F, Ashok A, et al. Toward a More Usable Home-Based Video Telemedicine System: A Heuristic Evaluation of the Clinician User Interfaces of Home-Based Video Telemedicine Systems. JMIR Hum Factors 2017;4(2):e11 [FREE Full text] [doi: 10.2196/humanfactors.7293] [Medline: 28438724]
25. Torrent-Sellens J, Díaz-Chao Á, Soler-Ramos I, Saigí-Rubió F. Modeling and Predicting Outcomes of eHealth Usage by European Physicians: Multidimensional Approach from a Survey of 9196 General Practitioners. J Med Internet Res 2018 Oct 22;20(10):e279 [FREE Full text] [doi: 10.2196/jmir.9253] [Medline: 30348628]
26. Rogers E. A prospective and retrospective look at the diffusion model. J Health Commun 2004;9 Suppl 1:13-19. [doi: 10.1080/10810730490271449] [Medline: 14960401]
27. MobihealthNews.com. 2018. In-Depth: Four major telemedicine trends of 2018 URL: https://www.mobihealthnews.com/content/depth-four-major-telemedicine-trends-2018 [accessed 2019-07-24]
28. Health Information Technology Regional Extension Center. 2014. Telehealth start-up and resource guide URL: http://www.gptrac.org/wp-content/uploads/2015/01/TelligenTelehealthGuide-Final-2014.pdf [accessed 2017-07-24]
29. Rutledge C, Kott K, Schweickert PA, Poston R, Fowler C, Haney TS. Telehealth and eHealth in nurse practitioner training: current perspectives. Adv Med Educ Pract 2017;8:399-409 [FREE Full text] [doi: 10.2147/AMEP.S116071] [Medline: 28721113]
30. Hoff T. Deskilling and adaptation among primary care physicians using two work innovations. Health Care Manage Rev 2011;36(4):338-348. [doi: 10.1097/HMR.0b013e31821826a1] [Medline: 21685794]
31. Kvedar J, Coye MJ, Everett W. Connected health: a review of technologies and strategies to improve patient care with telemedicine and telehealth. Health Aff (Millwood) 2014;33(2):194-199. [doi: 10.1377/hlthaff.2013.0992] [Medline: 24493760]

XSL•FO
RenderX

32. DeSanctis G, Poole M. Capturing the Complexity in Advanced Technology Use: Adaptive Structuration Theory. Organization Science 1994 May;5(2):121-147. [doi: 10.1287/orsc.5.2.121]
33. Goh J, Gao G, Agarwal R. Evolving Work Routines: Adaptive Routinization of Information Technology in Healthcare. Information Systems Research 2011 Sep;22(3):565-585. [doi: 10.1287/isre.1110.0365]
34. Kane G, Alavi M. Information Technology and Organizational Learning: An Investigation of Exploration and Exploitation Processes. Organization Science 2007 Oct;18(5):796-812. [doi: 10.1287/orsc.1070.0286]
35. Sun H. Understanding User Revisions When Using Information System Features: Adaptive System Use and Triggers. MIS Quarterly 2012;36(2):453-478. [doi: 10.2307/41703463]
36. Bala H, Venkatesh V. Adaptation to Information Technology: A Holistic Nomological Network from Implementation to Job Outcomes. Management Science 2015 Oct 27;62(1):156-179. [doi: 10.1287/mnsc.2014.2111]
37. Weigel F, Landrum WH, Hall DJ. Human-technology adaptation fit theory for healthcare. 2009 Presented at: The Twelfth Annual Conference of the Southern Association for Information Systems (SAIS), Charleston, SC; 2009; Charleston, SC.
38. de Jong CC, Ros WJ, van Leeuwen M, Schrijvers G. How Professionals Share an E-Care Plan for the Elderly in Primary Care: Evaluating the Use of an E-Communication Tool by Different Combinations of Professionals. J Med Internet Res 2016 Nov 24;18(11):e304 [FREE Full text] [doi: 10.2196/jmir.6332] [Medline: 27884811]
39. Niederman F, Beise CM, Beranek PM. Issues and Concerns about Computer-Supported Meetings: The Facilitator's Perspective. MIS Quarterly 1996 Mar;20(1):1-22. [doi: 10.2307/249540]
40. Niederman F, Volkema R. The Effects of Facilitator Characteristics on Meeting Preparation, Set Up, and Implementation. Small Group Research 2016 Aug 19;30(3):330-360. [doi: 10.1177/1046496499030003004]
41. Griffith T, Fuller MA, Northcraft GB. Facilitator influence in group support systems: Intended and unintended effects. Information Systems Research 1998;9(1):20-36.
42. Kelly G, Bostrom R. A Facilitator's General Model for Managing Socioemotional Issues in Group Support Systems Meeting Environments. Journal of Management Information Systems 2015 Dec 08;14(3):23-44. [doi: 10.1080/07421222.1997.11518173]
43. Kolfschoten G, Niederman F, Briggs R, de Vreede G. Facilitation Roles and Responsibilities for Sustained Collaboration Support in Organizations. Journal of Management Information Systems 2014 Dec 08;28(4):129-162. [doi: 10.2753/MIS0742-1222280406]
44. Delone W. The DeLone and McLean Model of Information Systems Success: A Ten-Year Update. Journal of Management Information Systems 2014 Dec 23;19(4):9-30. [doi: 10.1080/07421222.2003.11045748]
45. Petter S, DeLone W, McLean E. Measuring information systems success: models, dimensions, measures, and interrelationships. European Journal of Information Systems 2017 Dec 19;17(3):236-263. [doi: 10.1057/ejis.2008.15]
46. Devaraj S, Kohli R. Performance Impacts of Information Technology: Is Actual Usage the Missing Link? Management Science 2003 Mar;49(3):273-289. [doi: 10.1287/mnsc.49.3.273.12736]
47. Goodhue DL, Thompson RL. Task-technology fit and individual performance. MIS Quarterly 1995;19:213-236.
48. Junglas I, Abraham C, Watson R. Task-technology fit for mobile locatable information systems. Decision Support Systems 2008 Nov;45(4):1046-1057. [doi: 10.1016/j.dss.2008.02.007]
49. Safran DG. Primary Care Performance in Fee-for-Service and Prepaid Health Care Systems. JAMA 1994 May 25;271(20):1579-1586. [doi: 10.1001/jama.1994.03510440039030]
50. Hoffer Gittell J. Coordinating Mechanisms in Care Provider Groups: Relational Coordination as a Mediator and Input Uncertainty as a Moderator of Performance Effects. Management Science 2002 Nov;48(11):1408-1426. [doi: 10.1287/mnsc.48.11.1408.268]
51. Advisory Board. 2017. You're probably measuring telehealth ROI wrong. Use these 6 metrics to do it right URL: https://www.advisory.com/research/market-innovation-center/the-growth-channel/2017/05/telehealth-roi [accessed 2019-07-24]
52. Spectrum Enterprise. 2017. How telehealth is transforming healthcare URL: https://enterprise.spectrum.com/content/dam/business/omega/enterprise/pdf/landing-pages/webelieve/SE_Telehealth_122217_v3_TS.pdf [accessed 2019-07-24]
53. Schwamm L. Telehealth: seven strategies to successfully implement disruptive technology and transform health care. Health Aff (Millwood) 2014 Feb;33(2):200-206. [doi: 10.1377/hlthaff.2013.1021] [Medline: 24493761]
54. Nittas V, Mütsch M, Ehrler F, Puhan MA. Electronic patient-generated health data to facilitate prevention and health promotion: a scoping review protocol. BMJ Open 2018 Aug 10;8(8):e021245 [FREE Full text] [doi: 10.1136/bmjopen-2017-021245] [Medline: 30099392]
55. Pauleen D, Yoong P. Relationship Building and the use of ICT in Boundary-Crossing Virtual Teams: A Facilitator's Perspective. Journal of Information Technology 2001 Dec;16(4):205-220. [doi: 10.1080/02683960110100391]
56. Chau P, Hu P. Investigating healthcare professionals' decisions to accept telemedicine technology: an empirical test of competing theories. Information & Management 2002 Jan;39(4):297-311. [doi: 10.1016/S0378-7206(01)00098-2]
57. Agarwal R, Prasad J. A Conceptual and Operational Definition of Personal Innovativeness in the Domain of Information Technology. Information Systems Research 1998 Jun;9(2):204-215. [doi: 10.1287/isre.9.2.204]
58. Serrano C, Karahanna E. The Compensatory Interaction Between User Capabilities and Technology Capabilities in Influencing Task Performance: An Empirical Assessment in Telemedicine Consultations. MIS Quarterly 2016 Mar 3;40(3):597-621. [doi: 10.25300/MISQ/2016/40.3.04]

XSL•FO
RenderX

59. 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]
60. Chevalking SL, Ben Allouch S, Brusse-Keizer M, Postel MG, Pieterse ME. Identification of Users for a Smoking Cessation Mobile App: Quantitative Study. J Med Internet Res 2018 Apr 09;20(4):e118 [FREE Full text] [doi: 10.2196/jmir.7606] [Medline: 29631988]
61. Emani S, Yamin CK, Peters E, Karson AS, Lipsitz SR, Wald JS, et al. Patient perceptions of a personal health record: a test of the diffusion of innovation model. J Med Internet Res 2012;14(6):e150 [FREE Full text] [doi: 10.2196/jmir.2278] [Medline: 23128775]
62. Li H, Wu J, Gao Y, Shi Y. Examining individuals' adoption of healthcare wearable devices: An empirical study from privacy calculus perspective. Int J Med Inform 2016;88:8-17. [doi: 10.1016/j.ijmedinf.2015.12.010] [Medline: 26878757]
63. Chow M, Herold DK, Choo T, Chan K. Extending the technology acceptance model to explore the intention to use Second Life for enhancing healthcare education. Computers & Education 2012 Dec;59(4):1136-1144. [doi: 10.1016/j.compedu.2012.05.011]
64. Compeau D, Higgins CA. Computer Self-Efficacy: Development of a Measure and Initial Test. MIS Quarterly 1995 Jun;19(2):189-211. [doi: 10.2307/249688]
65. Ozturk B, Bilgihan A, Nusair K, Okumus F. What keeps the mobile hotel booking users loyal? Investigating the roles of self-efficacy, compatibility, perceived ease of use, and perceived convenience. International Journal of Information Management 2016 Dec;36(6):1350-1359 [FREE Full text] [doi: 10.1016/j.ijinfomgt.2016.04.005]
66. Tams S, Thatcher B, Craig K. How and why trust matters in post-adoptive usage: The mediating roles of internal and external self-efficacy. The Journal of Strategic Information Systems 2018 Jun;27(2):170-190 [FREE Full text] [doi: 10.1016/j.jsis.2017.07.004]
67. Thatcher J, Zimmer J, Gundlach M, McKnight D. Internal and External Dimensions of Computer Self-Efficacy: An Empirical Examination. IEEE Trans Eng Manage 2008 Nov;55(4):628-644. [doi: 10.1109/TEM.2008.927825]
68. Venkatesh V, Thong JYL, Xu X. Consumer Acceptance and Use of Information Technology: Extending the Unified Theory of Acceptance and Use of Technology. MIS Quarterly 2012;36(1):157-178. [doi: 10.2307/41410412]
69. de Guinea A, Markus ML. Why Break the Habit of a Lifetime? Rethinking the Roles of Intention, Habit, and Emotion in Continuing Information Technology Use. MIS Quarterly 2009;33(3):433-444. [doi: 10.2307/20650303]
70. Jasperson J, Carter PE, Zmud RW. A Comprehensive Conceptualization of Post-Adoptive Behaviors Associated with Information Technology Enabled Work Systems. MIS Quarterly 2005;29(3):525-557. [doi: 10.2307/25148694]
71. McGowan B, Wasko M, Vartabedian BS, Miller RS, Freiherr DD, Abdolrasulnia M. Understanding the factors that influence the adoption and meaningful use of social media by physicians to share medical information. J Med Internet Res 2012 Sep 24;14(5):e117 [FREE Full text] [doi: 10.2196/jmir.2138] [Medline: 23006336]
72. Limayem M, Cheung CM. Understanding information systems continuance: The case of Internet-based learning technologies. Information & Management 2008 Jun;45(4):227-232. [doi: 10.1016/j.im.2008.02.005]
73. Limayem M, Hirt S. Force of Habit and Information Systems Usage: Theory and Initial Validation. JAIS 2003;4(1):65-97. [doi: 10.17705/1jais.00030]
74. Nevo S, Nevo D, Pinsonneault A. A Temporally Situated Self-Agency Theory of Information Technology Reinvention. MIS Quarterly 2016 Jan 1;40(1):157-186. [doi: 10.25300/MISQ/2016/40.1.07]
75. Subramani M. How Do Suppliers Benefit from Information Technology Use in Supply Chain Relationships? MIS Quarterly 2004;28(1):45-73. [doi: 10.2307/25148624]
76. Koo C, Chung N, Kim H. Examining explorative and exploitative uses of smartphones: a user competence perspective. Info Technology & People 2015 Mar 02;28(1):133-162. [doi: 10.1108/ITP-04-2013-0063]
77. Huang C, Goo J, Nam K, Yoo CW. Smart tourism technologies in travel planning: The role of exploration and exploitation. Information & Management 2017 Sep;54(6):757-770 [FREE Full text] [doi: 10.1016/j.im.2016.11.010]
78. Health Leaders. 2018. 4 ways telemedicine is changing healthcare URL: https://www.healthleadersmedia.com/innovation/4-ways-telemedicine-changing-healthcare [accessed 2019-07-24]
79. Teddlie C, Yu F. Mixed Methods Sampling. Journal of Mixed Methods Research 2017 Jun 12;1(1):77-100. [doi: 10.1177/1558689806292430]
80. Ringle C, Sarstedt M, Straub DW. A Critical Look at the Use of PLS-SEM in MIS Quarterly. MIS Quarterly 2012;36(1):iii-xiv. [doi: 10.2307/41410402]
81. Ahuja M, Thatcher JB. Moving beyond Intentions and toward the Theory of Trying: Effects of Work Environment and Gender on Post-Adoption Information Technology Use. MIS Quarterly 2005;29(3):427-459. [doi: 10.2307/25148691] [Medline: 10.2307/25148691]
82. Hair Jr JF, Sarstedt M, Hopkins L, Kuppelwieser GV. Partial least squares structural equation modeling (PLS-SEM). European Business Review 2014 Mar 04;26(2):106-121. [doi: 10.1108/EBR-10-2013-0128]
83. Hair JF, Hult GTM, Ringle CM, Sarstedt M. A Primer On Partial Least Squares Structural Equation Modeling (PLS-SEM). CA: Sage Publications, Inc; 2019:328.
84. Armstrong J, Overton TS. Estimating Nonresponse Bias in Mail Surveys. Journal of Marketing Research 1977 Aug;14(3):396-402. [doi: 10.2307/3150783]

XSL•FO
RenderX

85. Pavlou P. Consumer Acceptance of Electronic Commerce: Integrating Trust and Risk with the Technology Acceptance Model. International Journal of Electronic Commerce 2014 Dec 23;7(3):101-134. [doi: 10.1080/10864415.2003.11044275]
86. Podsakoff P, MacKenzie SB, Lee JY, Podsakoff NP. Common method biases in behavioral research: a critical review of the literature and recommended remedies. J Appl Psychol 2003 Oct;88(5):879-903. [doi: 10.1037/0021-9010.88.5.879] [Medline: 14516251]
87. Fornell C, Larcker DF. Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. Journal of Marketing Research 1981 Feb;18(1):39-50. [doi: 10.2307/3151312]
88. Cohen J. Statistical Power Analysis For The Behavioral Sciences (2nd Edition). Mahwah, NJ: Routledge; 1988.
89. Brown S, Dennis AR, Venkatesh V. Predicting Collaboration Technology Use: Integrating Technology Adoption and Collaboration Research. Journal of Management Information Systems 2014 Dec 08;27(2):9-54. [doi: 10.2753/MIS0742-1222270201]
90. Abel E, Shimada SL, Wang K, Ramsey C, Skanderson M, Erdos J, et al. Dual Use of a Patient Portal and Clinical Video Telehealth by Veterans with Mental Health Diagnoses: Retrospective, Cross-Sectional Analysis. J Med Internet Res 2018 Nov 07;20(11):e11350 [FREE Full text] [doi: 10.2196/11350] [Medline: 30404771]
91. Tamilmani K, Rana N, Dwivedi Y. Use of 'Habit' Is not a Habit in Understanding Individual Technology Adoption: A Review of UTAUT2 Based Empirical Studies. In: Smart Working, Living and Organising. TDIT 2018. IFIP Advances in Information and Communication Technology. Cham, Switzerland: Springer; 2018:277-294.
92. Polites G, Karahanna E. Shackled to the Status Quo: The Inhibiting Effects of Incumbent System Habit, Switching Costs, and Inertia on New System Acceptance. MIS Quarterly 2012;36(1):21-41. [doi: 10.2307/41410404]
93. Soror A, Hammer B, Steelman ZR, Davis FD, Limayem MM. Good habits gone bad: Explaining negative consequences associated with the use of mobile phones from a dual-systems perspective. Info Systems J 2015 Feb 24;25(4):403-427. [doi: 10.1111/isj.12065]
94. Venkatesh V, Thong J, Xu X. Unified Theory of Acceptance and Use of Technology: A Synthesis and the Road Ahead. JAIS 2016 May;17(5):328-376. [doi: 10.17705/1jais.00428]
95. Lee YJ, Ha S. Consumer Use of the Internet for Health Management. Journal of Consumer Health on the Internet 2016 May 09;20(1-2):1-18. [doi: 10.1080/15398285.2015.1127114]
96. Beaudry A, Pinsonneault A. Understanding User Responses to Information Technology: A Coping Model of User Adaptation. MIS Quarterly 2005;29(3):493-524. [doi: 10.2307/25148693]
97. Correa T, Hinsley W, de Zúñiga G. Who interacts on the Web?: The intersection of users' personality and social media use. Computers in Human Behavior 2010 Mar;26(2):247-253 [FREE Full text] [doi: 10.1016/j.chb.2009.09.003]
98. Porter C, Donthu N. Using the technology acceptance model to explain how attitudes determine Internet usage: The role of perceived access barriers and demographics. Journal of Business Research 2006 Sep;59(9):999-1007 [FREE Full text] [doi: 10.1016/j.jbusres.2006.06.003]
99. Carney M. Middle manager involvement in strategy development in not-for profit organizations: the director of nursing perspective - how organizational structure impacts on the role. J Nurs Manag 2004 Jan;12(1):13-21 [FREE Full text] [doi: 10.1111/j.1365-2834.2004.00388.x]
100. Hah H, Goldin D. Exploring Care Providers' Perceptions and Current Use of Telehealth Technology at Work, in Daily Life, and in Education: Qualitative and Quantitative Study. JMIR Med Educ 2019 Apr 22;5(1):e13350 [FREE Full text] [doi: 10.2196/13350] [Medline: 31008708]
101. Fathi J, Modin H, Scott J. Nurses advancing telehealth services in the era of healthcare reform. OJIN 2017;22(2):10-10. [doi: 10.3912/OJIN.Vol22No02Man02]

Abbreviations

AST: adaptive structuration theory
AVE: average variance extracted
E-consultation: electronic consultation
HCI: human computer interaction
HIT: health information technology
ICT: information and communication technology
IS: information systems
IT: information technology
PLS: partial least squares
SEM: structural equation modeling

XSL•FO
RenderX

Edited by G Eysenbach; submitted 18.06.19; peer-reviewed by N Archer, PY Yen, S Jalil, G Signorelli; comments to author 16.07.19; revised version received 28.07.19; accepted 28.07.19; published 29.08.19.

Please cite as:
Hah H, Goldin D, Ha S
The Association Between Willingness of Frontline Care Providers' to Adaptively Use Telehealth Technology and Virtual Service Performance in Provider-to-Provider Communication: Quantitative Study
J Med Internet Res 2019;21(8):e15087
URL: http://www.jmir.org/2019/8/e15087/
doi:10.2196/15087
PMID:31469078

©Hyeyoung Hah, Deana Goldin, Sejin Ha. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 29.08.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Identifying Attrition Phases in Survey Data: Applicability and Assessment Study

Camille J Hochheimer[1], PhD; Roy T Sabo[1], PhD; Robert A Perera[1], PhD; Nitai Mukhopadhyay[1], PhD; Alex H Krist[2], MD, MPH

[1]Department of Biostatistics, Virginia Commonwealth University, Richmond, VA, United States

[2]Department of Family Medicine and Population Health, Virginia Commonwealth University, Richmond, VA, United States

Corresponding Author:
Camille J Hochheimer, PhD
Department of Biostatistics
Virginia Commonwealth University
PO Box 980032
Richmond, VA, 23298-0032
United States
Phone: 1 804 828 9824
Fax: 1 804 828 8900
Email: hochheimercj@vcu.edu

Abstract

Background: Although Web-based questionnaires are an efficient, increasingly popular mode of data collection, their utility is often challenged by high participant dropout. Researchers can gain insight into potential causes of high participant dropout by analyzing the dropout patterns.

Objective: This study proposed the application of and assessed the use of user-specified and existing hypothesis testing methods in a novel setting—survey dropout data—to identify phases of higher or lower survey dropout.

Methods: First, we proposed the application of user-specified thresholds to identify abrupt differences in the dropout rate. Second, we proposed the application of 2 existing hypothesis testing methods to detect significant differences in participant dropout. We assessed these methods through a simulation study and through application to a case study, featuring a questionnaire addressing decision-making surrounding cancer screening.

Results: The user-specified method set to a low threshold performed best at accurately detecting phases of high attrition in both the simulation study and test case application, although all proposed methods were too sensitive.

Conclusions: The user-specified method set to a low threshold correctly identified the attrition phases. Hypothesis testing methods, although sensitive at times, were unable to accurately identify the attrition phases. These results strengthen the case for further development of and research surrounding the science of attrition.

(J Med Internet Res 2019;21(8):e12811) doi:10.2196/12811

KEYWORDS

patient dropouts; surveys and questionnaires; survival analyses; statistical models

Introduction

Background

Web-based surveys and questionnaires are an increasingly popular mode of data collection because of factors such as the ease of delivery, cost-effectiveness, automated data management, and ability to reach a large pool of respondents. Compared with participants of paper-based surveys, Web-based survey participants have greater freedom to drop out at any point, especially when they feel that the questionnaire is no longer relevant to them. This results in dropout attrition, where a participant starts but does not complete the survey [1,2]. Consequently, the rate of dropout attrition is often much higher in Web-based surveys than their paper-based counterparts [1,3]. Dropout attrition is different from nonresponse attrition, for which the research is widely established, where participants are solicited but choose not to participate in a survey [1,2,4,5]. For the purposes of this paper, dropout and attrition will both refer to dropout attrition.

XSL•FO
RenderX

Identifying distinct phases of dropout also presents a missing data problem. When dropout is associated with topics addressed by the survey itself, it may be considered either missing at random (MAR) or missing not at random (MNAR) [6]. Ignoring data MAR or MNAR may result in a biased inference when analyzing data, where statistical significance changes depending on whether or not the missing data are addressed [7]. Thus, when attrition phases are present, a researcher may consider using different methods to evaluate the results that account for data MAR or MNAR.

In Eysenbach's original call for a *science of attrition*, he introduced the idea that survey attrition occurs in distinct phases: the *curiosity plateau* at the beginning of a survey where the participation rate remains high while respondents gauge their interest in the survey, the *attrition phase* where participants exit at a higher rate, and the *stable use phase* where most remaining participants are likely to complete the survey [1]. We previously responded to Eysenbach's call with a 3-step process for investigating where and why attrition occurs by first visualizing dropout trends, then confirming statistically significant dropout, and finally exploring factors associated with attrition [8].

Although our previous study suggests that visualization of dropout patterns (particularly plotting the number of dropouts at each question) provides a reasonable estimation of attrition phases, a natural next step would be to directly estimate Eysenbach's phases of attrition using statistical methods. If clear attrition patterns are identified, this suggests that the dropout is because of the content of the questionnaire itself (research suggests that it is the survey content rather than the survey length that drives participant dropout [9-11]) and thus, any further analyses of the survey data should account for both the missing data and the mechanism by which their missingness differs. Identifying the attrition phases not only highlights questions where dropout is high but also classifies the entire series of questions with similar dropout rates; doing so in an empirical manner also reduces researcher reliance on subjective classifications.

Objectives

This paper aimed to describe the application of clinical thresholds and existing statistical methods to the task of identifying Eysenbach's phases of attrition. In the Methods section, we first propose an approach searching for clinically or practically meaningful attrition through user-specified thresholds. We also propose the novel application of 2 existing statistical techniques to identify the phases of attrition in survey data. In the Results section, we conduct a simulation study and apply all 3 approaches to the results of a Web-based survey about cancer screening to demonstrate the performance of these methods for identifying phase transitions. Finally, in the Discussion section, we suggest future directions for this research.

Methods

Methods for Identifying Eysenbach Phases of Attrition

We have proposed an approach motivated by establishing meaningful dropout standards and 2 additional approaches motivated by statistically significant attrition. Statistical significance is not always indicative of a meaningful difference, especially when the sample size is large, as is often the case with Web-based surveys. Thus, these serve as complementary methods to identify clear inflection points of the dropout rate.

User-Specified Attrition Thresholds

For the first proposed method, the researcher specifies the amount of dropout that they consider to be clinically or practically meaningful, allowing them to choose the sensitivity of this method. Specifically, they define thresholds for the start and end of the dropout phase of attrition. The proportion of dropout for each survey question is then compared with both the thresholds. When applying this method, the first question for which dropout exceeds the start threshold was interpreted as the beginning of the dropout phase and the last time that dropout exceeds the end threshold was interpreted as the end of the dropout phase. Note that different numerical values could be used for the start and end thresholds.

Hypothesis Testing Methods

For the following 2 hypothesis testing methods, we fit the model to the entire survey and then apply successive differences contracts to test pairwise differences in the dropout rates between adjacent questions; note that here the outcome is a binary dropout indicator and not the original survey response. The null hypothesis that dropout rates do not change between questions was rejected in favor of the alternative hypothesis that the rates were different between questions when the corresponding *P* value was lower than an appropriately adjusted significance level. We assumed dropout monotonically increased throughout the survey. The first and last instances of statistically significant differences in the dropout rate between questions were interpreted as the start and end of the dropout phase of attrition. *P* values were adjusted for multiple comparisons using Benjamini and Hochberg's false discovery rate correction, with adjusted *P* values evaluated at the 5% level [12].

Generalized Linear Mixed Model

Survey questions are individual, discrete units that are dependent within survey participants. In other words, once a participant drops out of a survey, they cannot answer any further questions. A generalized linear mixed model (GLMM) using the logit link can account for these features while allowing researchers to compare differences in the proportion of respondents between sequential questions. In this model, the binary outcome was whether or not a participant had dropped out of the survey by that question. Question number was included as a time-varying covariate with the number of levels equal to the number of questions and an indicator function identifying the question of interest. A subject-level random effect was included to account for within-subject dependence between response rates, accounting for the fact that a participant cannot reenter the survey once they have dropped out.

Discrete Time Survival Analysis

Discrete time survival analysis (DTSA) is another appropriate method for analyzing survey attrition because questions occur in a discrete order and respondents can only dropout at these distinct points. We propose treating dropout as a time-to-event

XSL•FO
RenderX

outcome, meaning we counted how many questions were answered before the respondent dropped out and used this as a time-to-event outcome. In the survival setting, we inherently assume participants are followed over time, allowing us to eliminate the subject-level random effect. The multivariate, dependent set of outcomes was replaced with a minimally sufficient outcome, meaning no information was lost in this simplified model.

The outcome of this model was whether or not the participant experienced the event of survey dropout. The baseline hazard function was a step function with a dummy variable for each question at which the participant could drop out (which is the last question if they completed the survey); thus, we tested for changes in the hazard of dropping out of the survey [13].

Simulation Study

To compare the performance of these methods in detecting attrition phases, we simulated a variety of dropout patterns. For each pattern, we simulated 10,000 datasets, each with 200 simulated participants answering 20 questions. Respondents had a random chance of dropping out at any point in the survey, including the first question, and a participant could not reenter once they had dropped out.

Simulated dropout patterns corresponded to constant, 2-phase, and 3-phase attrition. Constant attrition could be either the stable use or dropout phase of attrition throughout the survey (see the top-left panel of Figure 1). Two-phase attrition either began with the stable use phase and then transitioned into the dropout phase for the remainder of the survey or began with the dropout phase with a transition at some point during the survey into the stable use phase (see the top-right and middle-left panels of Figure 1). Three-phase attrition followed Eysenbach's proposed pattern in that there were stable use phases at the start and end of the survey with a dropout phase in the middle (see the middle-right and bottom panel of Figure 1). We tested both mild and severe attrition rates for dropout phases, where severe attrition rates demonstrated more pronounced differences in dropout rates between phases, to determine the sensitivity of these methods. The location of the phase transition was varied to see whether these methods better identified phase transitions that occurred near the start, middle, or end of the survey.

The overall null hypothesis being tested was that there are no phases of attrition, represented by the constant attrition patterns in the top-left panel of Figure 1. When the dropout rate was mild, we expected our proposed methods would not detect practically meaningful or statistically significant attrition at any point in the survey. When this rate was severe, we expected the methods to detect the first and last instances of significant attrition at the start and end of the survey, again suggesting constant attrition throughout. In other words, if the dropout rate was constant but significant, we interpreted this result as the dropout phase occurring throughout the survey.

When implementing the user-specified method, we assessed user-specified thresholds of 3%, 5%, and 8%. Although these are arbitrary thresholds, we found that using a sensitive 3% threshold was the only threshold that was able to distinguish attrition patterns in both the simulation study and the test case application (discussed below); thus, only the results of the 3% threshold are presented.

We determined how well these methods achieved the goal of detecting phases of attrition by calculating and comparing the type I error and sensitivity of each method. Here, type I error was defined as when at least 2 phases of attrition were detected when the underlying attrition pattern was constant (ie, phases of attrition were detected when they did not exist). Sensitivity was defined as finding the correct number of phase transitions when they did exist. Ideally, these methods would achieve a type I error of 5% and higher sensitivity.

We also visually inspected how often each question was chosen as the start or end of the attrition phase by plotting these distributions with histograms. The histograms for the user-specified method highlight the first and last question at which the amount of dropout at a question surpassed 3%; histograms for the GLMM and DTSA show the first and last comparisons between questions where there was a significant adjusted *P* value.

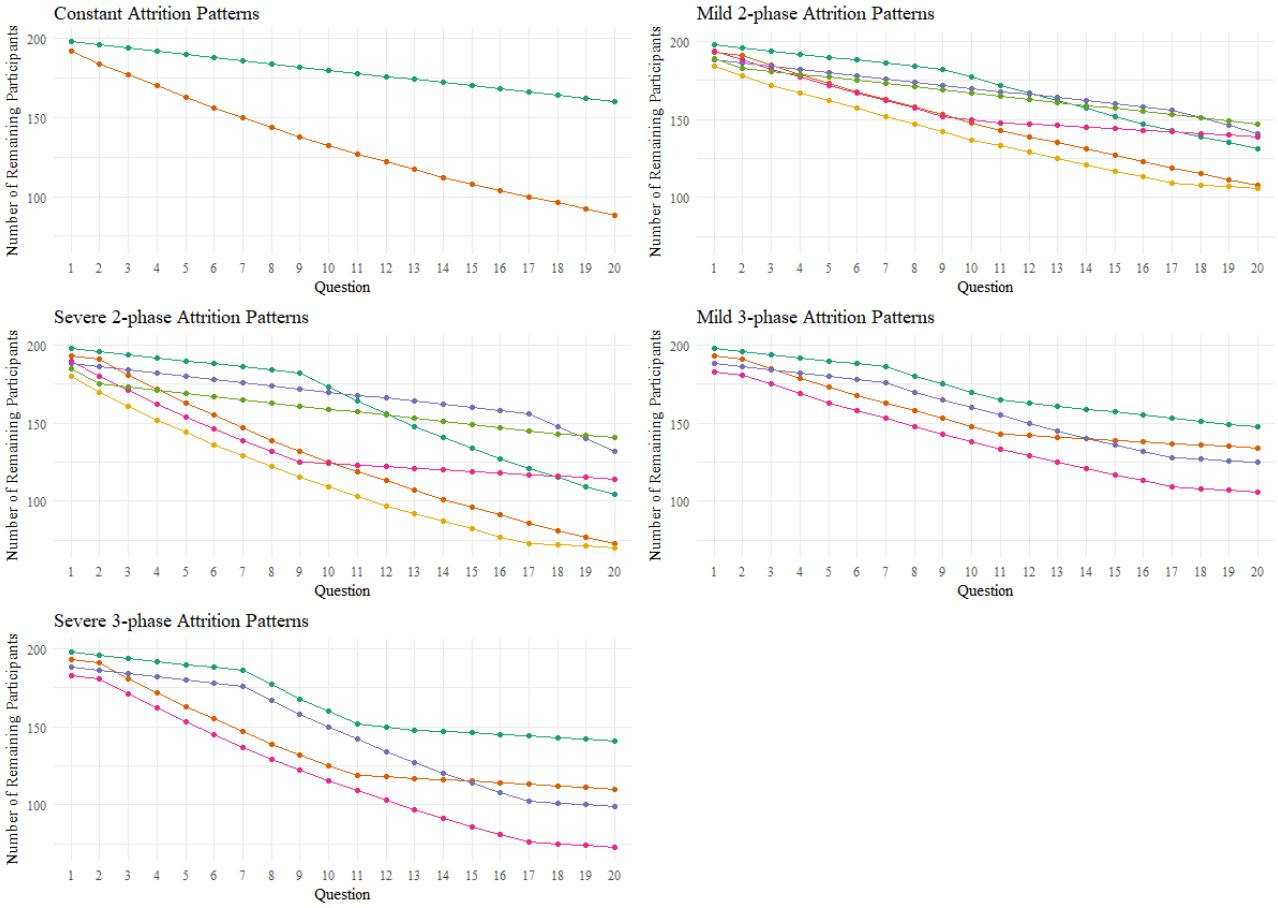
Test Case

Our test case data were from a Web-based survey entitled the informed decision-making (IDM) module. The 17-question survey explored how patients approach decisions regarding screening for breast, colorectal, and prostate cancer. Here we focus exclusively on the results for colorectal cancer, where there were 1249 participants. Questions addressed the awareness of screening eligibility, screening options, primary concerns about cancer screening, and planned next steps [14]. This survey was designed by the Virginia Commonwealth University Department of Family Medicine and Population Health research team and administered from January to August, 2014, in 12 primary care practices throughout northern Virginia through the interactive Web-based patient portal, *MyPreventiveCare* [15-18]. More specific details regarding the survey, including screenshots of the questionnaire itself, can be found in the studies by Hochheimer et al and Woolf et al [8,19].

All simulations and analyses were conducted using the R statistical software version 3.5.0 [20]. When applying the user-specified method, we used a threshold of 3%, 5%, and 8%, although only the results of the 3% threshold are discussed because of the poor performance of the 5% and 8% thresholds. The *survey* package was used to apply DTSA with questions treated as categorical factors of equal weight [21]. A successive differences contrast was applied to the results of both hypothesis testing models using the *multcomp* package to test each pairwise difference in the proportion of participants remaining in the survey and the hazard of dropping out between questions for the GLMM and DTSA, respectively [22]. All figures in this paper were created using the *ggplot2* package [23].

XSL•FO
RenderX

Figure 1. Simulated attrition patterns.



Results

Simulation Study Results

The resulting type I error and sensitivity of each method for each simulated attrition pattern can be found in Table 1. A selection of the histograms displaying the distribution of questions chosen as the start and end of the attrition phase is displayed in Figure 2. In the case of severe constant attrition, we expected to see the first instance of meaningful attrition at question 1 for the user-specified method and the first instance of significant attrition between questions 1 and 2 for the GLMM and DTSA. Then, we hoped to see this attrition phase last throughout the survey, with the last instance of meaningful attrition found at the final question and significant attrition found between the last two questions.

User-Specified Attrition Thresholds

The 3% user-specified threshold had high type I error but also high sensitivity to detect attrition phases, especially in cases of a severe dropout phase and 3 phases of attrition. This threshold achieved higher sensitivity than the other 2 methods when the simulated dropout pattern for 2 phases began with the dropout phase and ended with the stable use phase.

The user-specified method failed to identify the start of the severe constant attrition phase immediately at question 1 (see left column of Figure 2). In the majority of simulations, this method correctly detected severe attrition phases simulated to last from questions 1 to 10 and from questions 10 to 20.

Generalized Linear Mixed Model

The GLMM had a conservative type I error rate in the case of constant mild attrition and was unable to control type I error in the case of constant severe attrition. This method demonstrated low sensitivity to detect 2 phases of attrition in general and 3 phases with a mild dropout phase. The GLMM had higher sensitivity to detect all 3 of Eysenbach's attrition phases in simulation patterns with a severe dropout phase.

The histograms in the middle column of Figure 2 reveal that the GLMM failed to identify the start of the attrition phase immediately at question 1 when the simulated pattern was that of severe constant attrition. These plots also show that when a severe attrition phase was simulated to begin in the middle of the survey and continue until the end of the survey, the GLMM most often detected the start of the attrition phase correctly at question 10 but detected the end of the attrition phase before the end of the survey (resulting in the low sensitivity seen in Table 1). When a severe attrition phase was simulated to begin at the start of the survey and last until the middle of the survey, the GLMM failed to identify an immediate start to the attrition phase but correctly distinguished the phase transition at question 10.

XSL•FO
RenderX

Table 1. Simulation study results.

Metric, simulation type, and attrition severity	Location of phase transition	3% user-specified threshold	GLMM[a]	DTSA[b]
Type I error				
Constant				
Mild	—[c]	0.19	0.01	0.74
Severe	—[c]	0.53	0.98	0.96
Sensitivity				
Two phases				
Mild	Middle	0.47	0.22	0.90
Mild	Start	0.48	0.17	0.94
Mild	End	0.72	0.13	0.87
Mild reverse	Middle	0.44	0.20	0.64
Mild reverse	Start	0.51	0.01	0.69
Mild reverse	End	0.41	0.14	0.58
Severe	Middle	0.82	0.24	0.91
Severe	Start	0.79	0.20	0.97
Severe	End	0.90	0.56	0.88
Severe reverse	Middle	0.88	0.02	0.58
Severe reverse	Start	0.90	0.01	0.69
Severe reverse	End	0.88	0.00	0.50
Three phases				
Mild	Middle	0.67	0.12	0.08
Mild	1 start	0.96	0.32	0.07
Mild	1 end	0.96	0.35	0.05
Mild	Ends	0.96	0.51	0.03
Severe	Middle	0.98	0.84	0.08
Severe	1 start	0.96	0.99	0.07
Severe	1 end	0.96	0.99	0.06
Severe	Ends	0.94	0.99	0.04

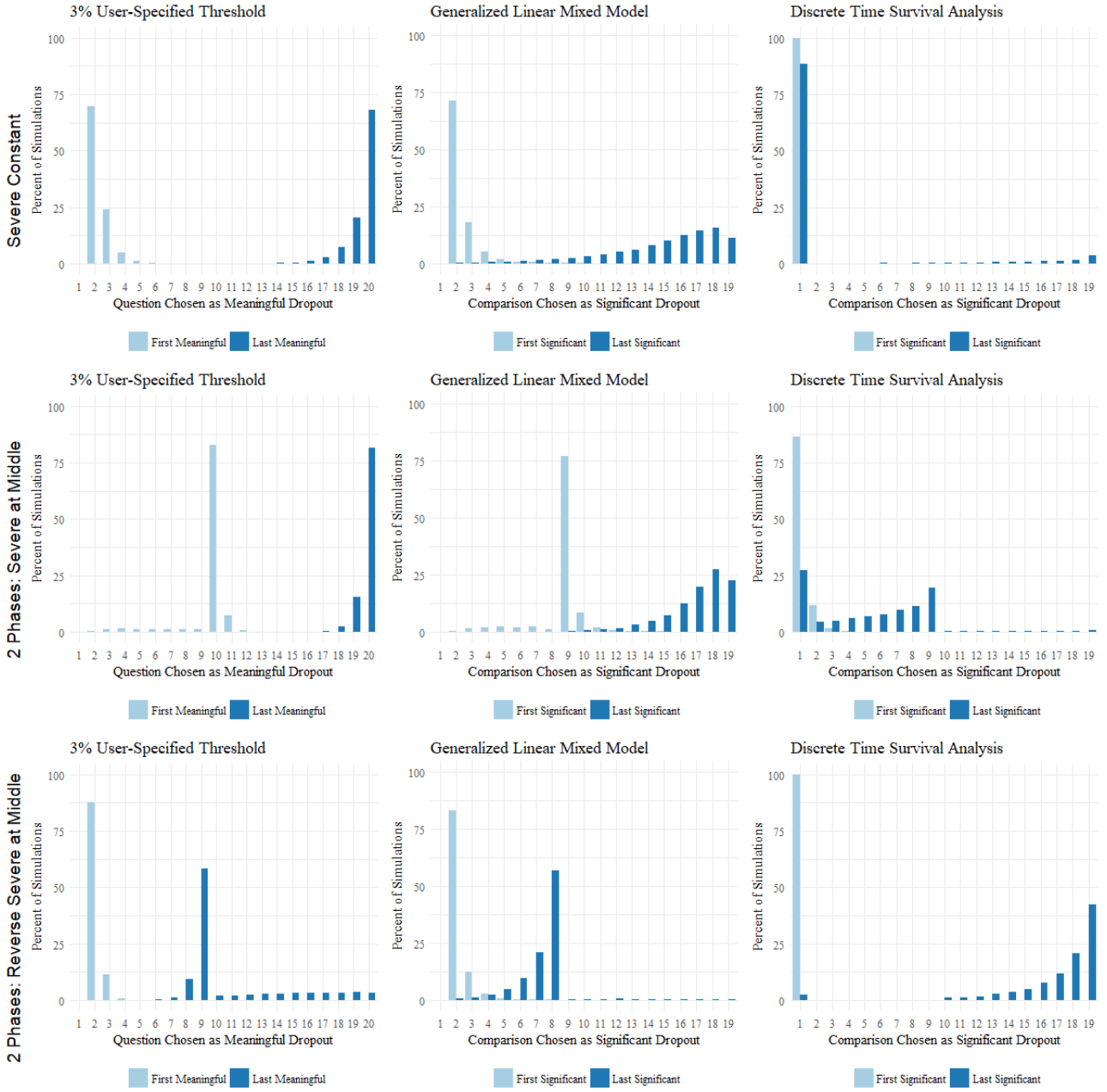
[a]GLMM: generalized linear mixed model.

[b]DTSA: discrete time survival analysis.

[c]Not applicable in the case of constant attrition.

XSL•FO
RenderX

Figure 2. Selected bar charts displaying the percent of simulations where each question was chosen as the first and last instance of meaningful/significant attrition.



Discrete Time Survival Analysis

DTSA did not control type I error and had low sensitivity to detect 3 attrition phases. For simulated patterns with 2 phases, DTSA demonstrated high sensitivity to detect both the phases. DTSA achieved higher sensitivity than the user-specified and GLMM approaches for simulated patterns with a mild change in attrition between phases as well as when the simulated dropout pattern survey began with the stable use phase and transitioned to a severe dropout phase (Table 1). Overall, we observed higher sensitivity for DTSA when the simulated dropout pattern began with the stable use phase and then transitioned to the dropout phase compared with patterns beginning with the dropout phase and ending with the stable use phase. DTSA also consistently had higher sensitivity when the simulated phase transition occurred toward the start of the survey.

Histograms displaying the accuracy of DTSA can be found in the right column of Figure 2. The comparison between questions 1 and 2 was correctly recognized as the first instance of significant attrition but also incorrectly chosen as the last instance when the dropout phase was simulated to last throughout the survey. Although this method detected the correct number of phases in the majority of simulations with 2 phases, it did not choose the correct questions as the start and end of the attrition phase. Specifically, DTSA was unable to detect an abrupt change in dropout rate in the middle of the survey. The histograms suggest a dropout phase at the beginning of the survey when the underlying simulated pattern had a dropout phase in the second half of the survey and suggest constant

XSL•FO
RenderX

attrition when the simulated pattern had a dropout phase in the first half of the survey.

Results of the Informed Decision-Making Module Application

First, we inspected a plot of the number of dropouts at each question of the IDM module for colorectal cancer as suggested in the study by Hochheimer et al [8]. We observed high attrition from questions 3 to 5 with another spike at question 9. We hypothesized that our proposed methods would detect the dropout phase to last from questions 3 to 9 and the 3% user-specified threshold was able to detect exactly that. The GLMM was unable to detect any significant changes in the dropout rate throughout the survey. Finally, DTSA was only able to detect the start of the dropout phase. The results suggested a significant increase in the hazard of dropping out between questions 2 and 3 but also that the dropout phase lasted until the end of the survey. This is inconsistent with the observed dropout pattern, where we saw visual proof of the stable use phase from questions 10 to 17.

Discussion

Principal Findings

In our simulation study, none of the 3 proposed methods consistently detected the correct number of phases while controlling type I error. The 3% user-specified threshold had a high type I error rate but also accurately detected the phases of attrition and had moderate to high sensitivity for all simulated scenarios. Although high sensitivity estimates of DTSA in the case of 2 attrition phases appeared promising, histograms revealed that this method consistently identified the wrong questions as the start and end of the dropout phase. This explains the low sensitivity of DTSA to detect all 3 phases of attrition. DTSA was extremely sensitive, finding a significant difference between the first 2 questions even when the simulated dropout rate was very small (eg, 0.001).

We did not see any distinct patterns in sensitivity when the phase transitions occurred toward the start or end of the survey compared with the middle of the survey. Sensitivity was often higher for the hypothesis testing methods when there was a sudden increase in attrition than when there was a sudden decrease in attrition. This suggests that these methods do not consistently detect a phase transition when dropout starts off at a high rate and then levels off at a certain point in the survey. This issue persisted even when the change was more pronounced, as it was in the severe cases. Although this should limit the ability of the GLMM and DTSA to detect 3 phases of attrition, we actually observed increased sensitivity for the GLMM when 3 phases were present.

We also investigated user-specified thresholds of 5% and 8%. The 5% threshold was unable to control type I error, with 0% type I error in the mild constant case and 94% type I error in the severe constant case. The 8% threshold had low sensitivity (often 0%) to detect any phase transition. Those of the 3% threshold had the best outcomes, both in type I error and sensitivity, suggesting that lower user-specified thresholds perform better at identifying the attrition phases.

When applied to our test case data, the 3% user-specified threshold was the only method able to detect the dropout phase from questions 3 through 9. The GLMM was not sensitive enough to detect a distinct dropout phase and DTSA was able to detect the abrupt increase in the hazard of dropping out between questions 2 and 3 but not the abrupt decrease between questions 9 and 10.

Limitations

One important difference between this study and our previous study is that participants were able to drop out at the first question. Previously, having 100% compliance in the first question limited the number of questions to 10 when applying the GLMM because of convergence issues. By assuming simulated participants could drop out at question 1, we were able to apply the GLMM to the entire survey and compare this method with the other 2 discussed here. The methods discussed in this paper apply specifically to dropout attrition and do not address nonresponse or longitudinal attrition (see the discussion in the study by Hochheimer et al [8]). By assuming dropout monotonically increases throughout the survey, these methods do not account for the functionality to skip questions. Finally, although the greatest region of dropout in the IDM module was from questions 3 to 5, it could be argued that there were 2 dropout phases, with the second occurring around question 9.

Future Studies

Instead of searching for the instance where dropout exceeds a set threshold, the user-specified method could also search for when the difference in the dropout rate between 2 sequential questions exceeds the threshold. Pooled logistic regression is yet another strategy to identify the dropout patterns without modeling between question dependence, with an interpretation closer to that of DTSA [24,25]. Although we are interested in the exact question or questions at which the attrition rate changes, our results suggest that researchers should not search for these inflection points question by question. Future directions will include a search for a method to model overall patterns that in turn reveal significant changes in dropout or hazard rate at particular questions, potentially through change-point modeling.

Conclusions

Our research suggests that when applying practical thresholds and existing statistical methods to the task of identifying Eysenbach's phases of attrition, sensitive user-specified thresholds correctly identify dropout phases at the cost of high type I error, whereas hypothesis testing methods are unable to correctly identify these phases. As we continue to advance the science of attrition, these results strengthen the case for developing new methods to identify attrition phases.

XSL•FO
RenderX

Acknowledgments

The test case data for this research was made possible by the Patient Centered Outcomes Institute (Grant Number IP2PI000516-01) and the National Center for Advancing Translational Sciences Biostatistical Shared Resource (Grant Number ULTR00058). The opinions expressed in this paper are those of the authors and do not necessarily reflect those of the funders.

Conflicts of Interest

None declared.

References

1. 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]
2. Kelders SM, Kok RN, Ossebaard HC, van Gemert-Pijnen JE. Persuasive system design does matter: a systematic review of adherence to web-based interventions. J Med Internet Res 2012 Nov 14;14(6):e152 [FREE Full text] [doi: 10.2196/jmir.2104] [Medline: 23151820]
3. Tortora R. Respondent burden, reduction of. In: Balakrishnan N, Colton T, Everitt B, Piegorsch W, Ruggeri F, Teugels JL, editors. Wiley StatsRef: Statistics Reference Online. Hoboken, New Jersey: John Wiley & Sons; 2014.
4. Edwards P, Roberts I, Clarke M, DiGuiseppi C, Pratap S, Wentz R, et al. Increasing response rates to postal questionnaires: systematic review. Br Med J 2002 May 18;324(7347):1183 [FREE Full text] [doi: 10.1136/bmj.324.7347.1183] [Medline: 12016181]
5. Edwards PJ, Roberts I, Clarke MJ, Diguiseppi C, Wentz R, Kwan I, et al. Methods to increase response to postal and electronic questionnaires. Cochrane Database Syst Rev 2009 Jul 8(3):MR000008. [doi: 10.1002/14651858.MR000008.pub4] [Medline: 19588449]
6. Rubin DB. Inference and missing data. Biometrika 1976 Dec;63(3):581-592. [doi: 10.2307/2335739]
7. Ibrahim JG, Molenberghs G. Missing data methods in longitudinal studies: a review. Test (Madr) 2009 May 1;18(1):1-43 [FREE Full text] [doi: 10.1007/s11749-009-0138-x] [Medline: 21218187]
8. Hochheimer CJ, Sabo RT, Krist AH, Day T, Cyrus J, Woolf SH. Methods for evaluating respondent attrition in web-based surveys. J Med Internet Res 2016 Dec 22;18(11):e301 [FREE Full text] [doi: 10.2196/jmir.6342] [Medline: 27876687]
9. McCambridge J, Kalaitzaki E, White IR, Khadjesari Z, Murray E, Linke S, et al. Impact of length or relevance of questionnaires on attrition in online trials: randomized controlled trial. J Med Internet Res 2011 Nov 18;13(4):e96 [FREE Full text] [doi: 10.2196/jmir.1733] [Medline: 22100793]
10. Cunningham JA, Ansara D, Wild TC, Toneatto T, Koski-Jännes A. What is the price of perfection? The hidden costs of using detailed assessment instruments to measure alcohol consumption. J Stud Alcohol 1999 Nov;60(6):756-758. [doi: 10.15288/jsa.1999.60.756] [Medline: 10606486]
11. Hoerger M. Participant dropout as a function of survey length in internet-mediated university studies: implications for study design and voluntary participation in psychological research. Cyberpsychol Behav Soc Netw 2010 Dec;13(6):697-700 [FREE Full text] [doi: 10.1089/cyber.2009.0445] [Medline: 21142995]
12. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J Royal Stat Soc: Ser B 2018 Dec 5;57(1):289-300. [doi: 10.1111/j.2517-6161.1995.tb02031.x]
13. Steele F, Washbrook E. University of Bristol. 2013. Discrete-Time Event History Analysis URL: http://www.bristol.ac.uk/media-library/sites/cmm/migrated/documents/discrete-time-eha-july2013-combined.pdf [accessed 2018-11-09]
14. Woolf SH, Zimmerman E, Haley A, Krist AH. Authentic engagement of patients and communities can transform research, practice, and policy. Health Aff (Millwood) 2016 Apr;35(4):590-594 [FREE Full text] [doi: 10.1377/hlthaff.2015.1512] [Medline: 27044956]
15. Krist AH, Peele E, Woolf SH, Rothemich SF, Loomis JF, Longo DR, et al. Designing a patient-centered personal health record to promote preventive care. BMC Med Inform Decis Mak 2011 Nov 24;11:73 [FREE Full text] [doi: 10.1186/1472-6947-11-73] [Medline: 22115059]
16. Krist AH, Woolf SH, Rothemich SF, Johnson RE, Peele JE, Cunningham TD, et al. Interactive preventive health record to enhance delivery of recommended care: a randomized trial. Ann Fam Med 2012;10(4):312-319 [FREE Full text] [doi: 10.1370/afm.1383] [Medline: 22778119]
17. Krist AH, Woolf SH, Bello GA, Sabo RT, Longo DR, Kashiri P, et al. Engaging primary care patients to use a patient-centered personal health record. Ann Fam Med 2014;12(5):418-426 [FREE Full text] [doi: 10.1370/afm.1691] [Medline: 25354405]
18. Alpert JM, Krist AH, Aycock RA, Kreps GL. Applying multiple methods to comprehensively evaluate a patient portal's effectiveness to convey information to patients. J Med Internet Res 2016 May 17;18(5):e112 [FREE Full text] [doi: 10.2196/jmir.5451] [Medline: 27188953]
19. Woolf SH, Krist AH, Lafata JE, Jones RM, Lehman RR, Hochheimer CJ, et al. Engaging patients in decisions about cancer screening: exploring the decision journey through the use of a patient portal. Am J Prev Med 2018 Dec;54(2):237-247. [doi: 10.1016/j.amepre.2017.10.027] [Medline: 29241715]
20. The R Project for Statistical Computing. 2016. URL: https://www.R-project.org/

XSL•FO
RenderX

21. Lumley T. Analysis of complex survey samples. J Stat Softw 2004 Jan;9(8):1-19 [FREE Full text] [doi: 10.18637/jss.v009.i08]
22. Hothorn T, Bretz F, Westfall P. Simultaneous inference in general parametric models. Biom J 2008 Jun;50(3):346-363. [doi: 10.1002/bimj.200810425] [Medline: 18481363]
23. Wickham H. Ggplot2: Elegant Graphics for Data Analysis. New York: Springer International Publishing; 2009.
24. D'Agostino RB, Lee ML, Belanger AJ, Cupples LA, Anderson K, Kannel WB. Relation of pooled logistic regression to time dependent Cox regression analysis: the Framingham Heart Study. Stat Med 1990 Dec;9(12):1501-1515. [doi: 10.1002/sim.4780091214] [Medline: 2281238]
25. Cupples LA, D'Agostino RB, Anderson K, Kannel WB. Comparison of baseline and repeated measure covariate techniques in the Framingham Heart Study. Stat Med 1988;7(1-2):205-222. [doi: 10.1002/sim.4780070122] [Medline: 3353604]

Abbreviations

DTSA: discrete time survival analysis
GLMM: generalized linear mixed model
IDM: informed decision-making
MAR: missing at random
MNAR: missing not at random

Edited by G Eysenbach; submitted 16.11.18; peer-reviewed by I White, D Guertler; comments to author 14.04.19; revised version received 07.06.19; accepted 10.06.19; published 23.08.19.

Please cite as:
Hochheimer CJ, Sabo RT, Perera RA, Mukhopadhyay N, Krist AH
Identifying Attrition Phases in Survey Data: Applicability and Assessment Study
J Med Internet Res 2019;21(8):e12811
URL: http://www.jmir.org/2019/8/e12811/
doi:10.2196/12811
PMID:31444875

©Camille J Hochheimer, Roy T Sabo, Robert A Perera, Nitai Mukhopadhyay, Alex H Krist. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 23.08.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Patient and Health Care Provider Experiences With a Recently Introduced Patient Portal in an Academic Hospital in the Netherlands: Mixed Methods Study

Maria M T Vreugdenhil[1*], MSc, MD; Sander Ranke[1*], MSc; Yvonne de Man[1*], MSc; Maaike M Haan[1*], MSc; Rudolf B Kool[1*], MD, PhD

IQ healthcare, Radboudumc, Nijmegen, Netherlands

[*]all authors contributed equally

Corresponding Author:
Maria M T Vreugdenhil, MSc, MD
IQ healthcare
Radboudumc
Postbus 9101
Nijmegen, 6500 HB
Netherlands
Phone: 31 2436153 ext 66868
Email: tjitske.vreugdenhil@radboudumc.nl

Abstract

Background: In the Netherlands, the health care system and related information technology landscape are fragmented. Recently, hospitals have started to launch patient portals. It is not clear how these portals are used by patients and their health care providers (HCPs).

Objective: The objective of this study was to explore the adoption, use, usability, and usefulness of a recently introduced patient portal in an academic hospital to learn lessons for the implementation of patient portals in a fragmented health care system.

Methods: A mixed methods study design was used. In the quantitative study arm, characteristics of patients who used the portal were analyzed, in addition to the utilization of the different functionalities of the portal. In the qualitative study arms, think-aloud observations were made to explore usability. Focus group discussions were conducted among patients and HCPs of the dermatology and ophthalmology outpatient departments. Thematic content analysis of qualitative data was carried out and overarching themes were identified using a framework analysis.

Results: One year after the introduction of the portal, 24,514 patients, 13.49% of all patients who visited the hospital, had logged in to the portal. Adoption of the portal was associated with the age group 45 to 75 years, a higher socioeconomic status, and having at least one medical diagnosis. Overarching themes from the qualitative analyses were (1) usability and user-friendliness of the portal, (2) HCP-patient communication through the portal, (3) usefulness of the information that can be accessed through the portal, (4) integration of the portal in care and work processes, and (5) HCP and patient roles and relationships.

Conclusions: One year after the introduction of the patient portal, patients and HCPs who used the portal recognized the potential of the portal to engage patients in their care processes, facilitate patient-HCP communication, and increase patient convenience. Uncertainties among patients and HCPs about how to use the messaging functionality and limited integration of the portal in care and work processes are likely to have limited portal use and usefulness.

(J Med Internet Res 2019;21(8):e13743) doi:10.2196/13743

KEYWORDS

patient portals; patient access to records; patient participation; professional-patient relations

XSL•FO
RenderX

Introduction

Background

Patient portals are promising tools to support patient involvement and patient-centered care [1-3]. Some of the functionalities of patient portals are primarily aimed at increasing patients' convenience, for example, the functionalities for scheduling appointments or medication reconciliation [4,5]. Other functionalities provide patients with relevant information about their health and offer access to general medical information, such as clinical guidelines, or to personal health information in their medical records [4]. In addition, many patient portals can be used for secure messaging to facilitate communication between patients and health care providers (HCPs) [4].

In the United States, patient portals are generally owned and administered by health institutions such as hospitals or health insurance companies together with their affiliated hospitals and clinics. In contrast, in Australia [1] and various European countries, for example, Denmark, Estonia, France [6], and, more recently, Finland [1] and Sweden [7], national patient portals have been launched. Beside these institution-based and national portals that are meant for the general population, there are also disease-specific patient portals that offer Web-based services for specific groups of patients, for example, for patients with diabetes or chronic kidney disease [8-11].

In the Netherlands, patient portals aimed at the general population are relatively new. In 2011, the development of a national patient portal was called off because of political resistance owing to privacy concerns. Since then, various health care organizations have started to develop institution-based patient portals. Academic hospitals have been leading in the implementation of hospital-based patient portals, using the options provided by the health information systems that they use. General hospitals have been following this trend [12]. Between 2014 and 2016, the percentage of medical specialists that provided their patients Web-based access to test results increased from 6% to 18%. In 2016, approximately 30% of the medical specialists indicated that patients could ask them questions over the internet. However, many patients were not aware of this, and only 5% of the Dutch patients sent a message to their HCP through a Web-based service [13,14].

The current Dutch health care information landscape entails a variety of institution-based patient portals that are not connected with each other, reflecting the fragmented health care system. In the health care system, the general practitioner (GP) has a central position and functions as a gate keeper for most diagnostic services and specialist care. When referred by their GP, patients can visit several diagnostic centers and see HCPs in several hospitals. Test results, diagnoses, care plans, and progression of treatment are generally reported back to the GP. Direct exchange of information between medical specialists from different hospitals is less common.

Objectives

Now that more and more health care organizations are implementing patient portals, patients are likely to come across a number of different patient portals that allow them access to personal health information in their records and offer opportunities to exchange secure messages with their HCPs. Theoretically, this positions them in the center of communication concerning their health, together with their GP. Therefore, the implementation of patient portals creates opportunities for patients to play a more active role in decision making, management of their health conditions, and coordination of their health care [5,15-17]. However, it is not clear how patient portals are used by patients and HCPs in the Dutch setting. To expand our knowledge of the implementation of institution-based patient portals in a fragmented health care system, we explored the adoption, utilization, usability, and usefulness of one of the first patient portals in an academic hospital in the Netherlands, 1 year after its introduction.

Methods

The Patient Portal

The patient portal under study is a version of MyChart, the patient portal of Epic Systems (Epic Systems Corporation). The functionalities of the portal that could be used during the study period were the list of diagnoses, list of prescribed medications, letters addressed to the GP, laboratory results, and functionalities to schedule appointments and send messages to HCPs. When sending a message, patients are informed that they should receive a response within 2 weeks. After triage by a nurse, HCPs receive the messages in their in-basket, which is not integrated with their email inbox. HCPs are instructed to answer the messages within 1 week. During the study period, the portal could also be used for filling out disease-specific, preconsultation questionnaires and patient satisfaction surveys.

The portal was introduced in 2015. Since then, flyers and banners in the waiting rooms of the hospital and the hospital website have been used to promote the portal as an additional hospital service aimed at patient convenience. In addition, all new patients have been invited to sign up for the portal upon registration with the hospital. After face-to-face identification at the hospital registration desk, patients receive an activation code for opening an account with the portal. The hospital website offers basic instructions on how to use the portal. In addition, patients can get in-person support from the help desk during office hours. All hospital staff received a 4-hour instruction for using Epic, which did not include a formal training in using the patient portal.

Study Design

The conceptual framework of the study was based on the *Unified Theory of Acceptance and Use of Technology* (*UTAUT*) [18]. The UTAUT relates adoption of a new technology to performance expectancy of the new technology (perceived usefulness), effort expectancy (perceived usability), social influence, and other facilitating factors [18]. We included the constructs of the UTAUT in a multilevel mixed methods study. At patient level, quantitative data were collected regarding utilization and user characteristics. In addition, qualitative data were collected regarding usability and usefulness, facilitating factors, and social influence. At HCP level, only qualitative data were collected on how the portal was used by HCPs

XSL•FO
RenderX

(facilitating factors and social influence for adoption by patients) and user experiences (usability and usefulness). Patient and hospital staff qualitative data were first analyzed separately and subsequently compared in a framework analysis. Quantitative and qualitative findings were interpreted together, using the qualitative findings to explain the quantitative results. Textbox 1 presents the study design.

Textbox 1. The multilevel mixed methods study design.

Patient level quantitative study arm

- **Outcomes:** user characteristics, portal adoption, use of functionalities
- **Data collection:** hospital information system data, log data
- **Analyses:** descriptive, logistic regression
- **Interpretation:** qualitative findings are used to explain quantitative findings

Patient level qualitative study arm

- **Topics:** perceived usability, perceived usefulness, social influence, facilitators for portal use
- **Data collection:** think aloud observation, focus groups
- **Analyses:** thematic content analysis, framework analysis of patient and health care provider (HCP) data together
- **Interpretation:** qualitative HCP findings are used to explain qualitative patient findings

HCP level quantitative study arm

- **Topics:** perceived usefulness, social influence, facilitators for portal use
- **Data collection:** focus groups
- **Analyses:** thematic content analysis, framework analysis of patient and HCP data together
- **Interpretation:** qualitative patient findings are used to explain qualitative HCP findings

Setting and Participants

The study was conducted at the Radboud University Medical Center, one of the 8 academic hospitals in the Netherlands. This was one of the first hospitals that launched a patient portal in the Netherlands. In 2013, a less advanced portal with fewer functionalities had been introduced, which was replaced by the MyChart portal when Epic was implemented in the hospital in 2015.

The quantitative study arm included patients who were registered on the portal between January 1, 2016, and January 1, 2017, and patients older than 12 years who had visited the hospital at least once in the year 2016 and were alive on December 31, 2016. Think-aloud participants were recruited through referrals from the hospital registration desk. The focus group participants (patients and HCPs) were recruited from 2 outpatient departments: the dermatology and ophthalmology departments. These departments were selected because of the differing portal adoption rates among patients and the divergent attitudes toward the portal among the staff of these departments. The adoption rate of the portal was 25% among patients of the dermatology department and 15% among patients of the ophthalmology department. In an informal qualitative assessment by the information technology department of the hospital, the staff of the dermatology outpatient department was most enthusiastic about the portal, whereas the staff of the ophthalmology department was least enthusiastic of all outpatient departments of the hospital.

For the 2 focus groups with patients, a randomly selected sample of 184 adult patients who had used the portal at least once between September 1, 2016, and February 28, 2017, was invited by their HCPs. For the 2 focus groups with HCPs, hospital staff who had been working at the departments during the last 3 months of 2016 and who had some experience using the portal were invited to participate. To gain a broad perspective on HCPs' experiences with and attitudes toward the portal, we applied a purposeful sampling strategy including all relevant HCPs, that is, doctors, nurses, assistants, administrative staff, and managers. Participants of the think-aloud observations and focus groups were asked to sign for their consent at the beginning of the sessions. Participating patients received a gift card of €20, whereas HCPs were not compensated for their participation. Ethical approval was requested and granted by the Research Ethics Committee of the Radboud University Nijmegen Medical Centre under number 2016-3091.

Data Collection

For the quantitative study at patient level, log data were collected from all portal users in the year 2016, including the number of times they had logged on to the portal and the functionalities they had used. In addition, we collected data from the hospital information system on all patients older than 12 years (portal users and nonusers) who visited the hospital at least once during 2016, including sex, age, number of open medical specialist trajectories, and number of open diagnoses. For socioeconomic status (SES), we used the status scores for postal code zones provided by the Netherlands Institute for Social Research for the year 2014. These scores are based on income, employment, and education levels of the inhabitants in the postal code zones [19]. Higher scores refer to higher SES.

XSL•FO
RenderX

We used think-aloud observation and inquiry to assess usability and user-friendliness of the portal while avoiding recall bias [20,21]. Patients without previous experience with the portal were asked to perform a number of predefined tasks, including logging in to the portal, booking a fictitious appointment, checking test results and new messages, and navigating to questionnaires. They were instructed to *think aloud* or verbalize their thoughts while performing these tasks. After completing the tasks, they were asked to reflect on their experiences. To gain further insight into the patients' reasons to use the portal, experiences with the portal, and their opinions concerning its usefulness, potential benefits, and drawbacks, we organized 2 focus group discussions with patients who were familiar with the portal. In addition, we conducted 2 focus groups with hospital staff to explore their experiences and opinions as well. All focus groups were guided by 2 moderators (SR, MH, YdM, or RBK) who used predefined topic lists, based on the constructs of the UTAUT. The lists included reasons to use the portal, experiences with the portal in general and with the different functionalities, and opinions about the usability and usefulness of the portal and how the portal might be improved. In addition, participants were invited to deliberate about facilitating factors for adoption of the portal and the specific functionalities and the role the portal had or could have in the care processes. The potential impact of the portal was also included as a topic. All think-aloud sessions and all focus group discussions were audiotaped.

Analysis

Analysis of Quantitative Data

Descriptive statistics were used to analyze portal users' gender, age, number of diagnosis-treatment combinations, number of treatment programs, and SES. We compared the characteristics of portal users and nonusers using a 2-sample *t* test (mean age), a chi-square test (sex), and the nonparametric Kolmogorov-Smirnov test (number of open diagnoses, SES score, and number of medical specialist treatment trajectories). Logistic regression was used to analyze which characteristics were associated with having logged in to the portal at least once, stratified for sex. The covariates included in the logistic regression analyses were the number of medical specialist trajectories (dichotomous: <1, ≥1), number of open diagnoses (dichotomous: <1, ≥1), SES score (continuous), and age (categorical: <45 years, 45-75 years, and >75 years). Age was included as a categorical variable as we expected a nonlinear relation between age and portal use, based on the distribution of age among users, which demonstrated 2 peaks, one around the age of 30 years and another around the age of 60 years, and a steep down slope at the age of 75 years. We used descriptive statistics for the analysis of the frequency of use of the portal in general and the specific functionalities. All analyses were performed using SAS Enterprise Guide.

Analysis of Qualitative Data

All think-aloud observations and focus group discussions were transcribed verbatim. We performed a thematic content analysis primarily using an inductive approach and secondarily using a deductive approach with a focus on the constructs of the UTAUT (usability, usefulness, social influence, and other facilitating factors) and the HCP-patient relation and roles. Overall, 3 researchers (SR, MH, and YdM) coded the transcripts independently using ATLAS.ti. After having coded the first think-aloud session and the first focus group discussion, the team discussed codes until consensus was reached and the code book was composed. This was used for coding the other sessions. After all sessions were coded, subthemes were identified for the focus group discussions among HCPs and patients and the think-aloud observations. After discussion with the complete research team, the subthemes were entered in a matrix to enable a thematic content analysis using framework methodology [22]. The matrix displayed the subthemes for HCP focus groups, patient focus groups, and think-aloud observations, together with the codes, and for each code, the most representative quote (see Multimedia Appendix 1). The matrix was used to identify overarching themes by 3 researchers (TV, SR, and RBK), combining the findings from the 3 qualitative study arms.

Results

Quantitative Results

In 2016, 24,514 out of 181,679 patients, 13.49% of all patients who visited the hospital, used the portal at least once. Users and nonusers of the portal did not differ with respect to gender, 54.80% of users were female versus 54.49% of nonusers (P=.29). The mean age of portal users (50.8 years) was slightly higher than the mean age of nonusers (50.3 years). Users had a higher median SES score than nonusers (0.26 vs 0.18; P<.001). More open diagnoses were registered for portal users than nonusers (median: 2.0 vs 1.0; P<.001; Table 1).

In 2016, a total of 28,419 patients logged in to the portal at least once. The mean number of times that portal users logged in to the portal during that year was 13 (interquartile range 3-16, mode 1). Logistic regression analyses demonstrated that in both females and males, portal usage was associated with belonging to the age group 45 to 75 years, having a higher SES, having at least one medical specialist trajectory, and having at least one open diagnosis (Table 2). For females, the effect of having 1 or more open diagnoses on portal use was stronger in the youngest and oldest age groups than in the age group 45 to 75 years. For males, there was an interaction between the effects of having a trajectory with a medical specialist and having 1 or more diagnoses.

Most patients who logged in to the portal checked their laboratory results (89.7%), the incoming messages (88.9%), the letters to their GP (87.5%), and their appointments in the calendar (87.4%). The medication list was viewed by 42.6% of the portal users. Fewer portal users sent a message through the portal (20.5%) or used the portal to book an appointment (1.8%).

XSL•FO
RenderX

Table 1. Characteristics of portal users versus nonusers.

Characteristics	Patients who visited the hospital in 2016 (N=181,679)	Users of the portal (N=24,514)	Nonusers of the portal (N=157,165)	*P* value
Females, %	55.90	54.80	54.49	.29[a]
Age (years), mean (SD)	50.4 (19.6)	50.8 (16.8)	50.3 (20.0)	<.001[b]
Number of medical specialist trajectories, median (range)	1 (1-16)	1 (1-15)	1 (1-16)	<.001[c]
Number of open diagnoses, median (range)	1 (0-28)	2 (0-23)	1 (0-28)	<.001[c]
Socioeconomic status score, median (range)	0.19 (−5.38 to 3.02)	0.26 (−5.38 to 3.02)	0.18 (−5.38 to 3.02)	<.001[c]

[a]χ^2 test.

[b]2-sample *t* test.

[c]Kolmogorov-Smirnov test.

Table 2. Logistic regression analysis for females and males with dependent variable use of the portal.

Covariates	Model A			Model B		
	Regression coefficient (SE)	OR[a] (95% CI)	*P* value	Regression coefficient (SE)	OR (95% CI)	*P* value
Females						
Age group 45-75 years[b]	0.5320 (0.0193)	1.13 (1.09-1.17)	<.001	0.5489 (0.0199)	1.73 (1.66-1.80)	<.001
Age group >75 years[b]	−0.9427 (0.0338)	0.26 (0.23-0.29)	<.001	−0.9384 (0.0350)	0.39 (0.20-0.77)	<.001
≥1 diagnosis	0.4851 (0.0163)	2.64 (2.48-2.81)	<.001	0.4847(0.0353)	1.62 (1.51-1.74)	<.001
≥1 medical specialist trajectory	0.1062 (0.0165)	1.24 (1.16-1.32)	<.001	0.1071 (0.0347)[c]	1.11 (1.04-1.19)	.001
SES[c]	0.1055 (0.0099)	1.11 (1.09-1.13)	<.001	0.1065 (0.0099)	1.11 (1.09-1.13)	<.001
≥1 medical specialist trajectory, age group 45-75 years	—[d]	—	—	0.1487 (0.0376)	1.16 (1.08-1.25)	<.001
≥1 medical specialist trajectory, age group >75 years	—	—	—	−0.0455 (0.0667)	0.96 (0.84-1.09)	.47
≥1 diagnosis, age group 45-75 years	—	—	—	−0.1966 (0.0381)	0.82 (0.76-0.88)	<.001
≥1 diagnosis, age group >75 years	—	—	—	0.0242 (0.0679)	1.02 (0.89-1.17)	.65
Males						
Age group 45-75 years[b]	0.5068 (0.0171)	1.72 (1.64-1.81)	<.001	0.5072 (0.0276)	1.72 (1.64-1.81)	<.001
Age group >75 years[b]	−0.4723 (0.0273)	0.65 (0.59-0.71)	<.001	−0.4732 (0.0171)	0.65 (0.59-0.70)	<.001
≥1 diagnosis	0.3344 (0.0204)	1.95 (1.80-2.11)	<.001	0.3748 (0.0025)	1.45 (1.44-1.46)	<.001
≥1 medical specialist trajectory	0.1958 (0.0205)	1.48 (1.37-1.60)	<.001	0.1558 (0.0250)	1.12 (1.07-1.18)	<.001
SES	0.1053 (0.0112)	1.11 (1.09-1.14)	<.001	0.1053 (0.0112)	1.11 (1.09-1.14)	<.001
≥1 medical specialist trajectory, ≥1 diagnosis	—	—	—	0.0788 (0.0250)[c]	1.08 (1.03-1.13)	<.001

[a]OR: odds ratio.

[b]Reference age group <45 years.

[c]SES: socioeconomic status.

[d]Not included in model A.

Qualitative Results

A total of 8 patients were recruited to participate in the think-aloud observations, 5 males and 3 females. The ages ranged from 21 to 71 years, with a median of 59 years. They differed in education level and in self-reported digital skills (Table 3). One of the participants got confused when trying to log in to the portal and failed to log in even with encouragement and assistance from the researcher. The think-aloud observations lasted approximately 20 min.

XSL•FO
RenderX

Table 3. Characteristics of all participants of the qualitative study arms.

Characteristics	Think-aloud observation participants (N=8)	Focus groups patients (N=12)	Focus groups hospital staff (N=17)
Sex, n (%)			
Male	5 (63)	7 (58)	5 (29)
Female	3 (38)	5 (42)	12 (71)
Age (years), median (range)	59 (21-71)	63 (34-79)	46 (23-64)
Education level, n (%)			
Low	None	None	—[a]
Medium	4 (50)	5 (42)	—
High	3 (38)	6 (50)	—
Unknown	1 (13)	1 (8)	—
Self-reported digital skills (1: very bad, 10: excellent), median (range)	7 (5-10)	7 (5-10)	8 (7-9)
Position (hospital staff), n (%)			
Medical specialist	—	—	3 (17)
Medical specialist in training	—	—	4 (24)
Nurse	—	—	4 (24)
Doctor's assistant	—	—	1 (6)
Administrative employee	—	—	3 (18)
Manager	—	—	2 (12)

[a]Not applicable.

Overall, 7 male and 5 female patients aged between 34 and 79 years participated in the focus groups. Their education levels were medium or high and their self-reported digital skills ranged from mediocre to excellent (Table 3). Furthermore, 3 medical specialists, 4 medical specialists in training, 4 nurses, 3 administrative assistants, 1 doctor's assistant, and 2 managers of the outpatient departments participated in the focus group discussions that were conducted with hospital staff. The characteristics of all participants of the qualitative study arms are summarized in Table 3.

The subthemes that we distinguished from the think-aloud observations and all 4 focus groups were adoption of the portal by patients and HCPs; stimulating use of the portal; learning to use the portal; available support for using the portal; procedure to log on; understandability of the information in the portal for patients; functionalities, benefits, patient engagement, and control; patient-HCP relationship; work process; and care process (see Multimedia Appendix 1). Analysis of the matrix resulted in the identification of 5 overarching themes: (1) usability and user-friendliness of the portal, (2) HCP-patient communication through the portal, (3) usefulness of the information that can be accessed through the portal, (4) integration of the portal in care and work processes, and (5) patient and doctor roles and relationships.

Overarching Theme 1: Usability and User-Friendliness of the Portal

The think-aloud observations and the focus group discussions revealed some difficulties concerning the log in procedure. However, once logged in to the portal, the layout and the drop-down menus were appreciated, and there were no difficulties to navigate to the different pages:

> *I think if you have looked at mijnRadboud a couple of times, you know how it works.* [Patient 7]

Although the think-aloud participants succeeded in booking a fictitious appointment, some focus group patients mentioned that they were not able to do so.

In the focus groups of the HCPs, concerns were raised about the understandability of the information that patients could access through the portal:

> *If I write an ophthalmologic report, even other medical specialists ask me: what do you mean. Therefore, my reporting cannot directly be translated into information that patients can understand.* [HCP 3]

However, in the patient focus groups, this was discussed as a surmountable problem. Patients pointed out how they found explanations for what they did not understand on the internet or with their family or friends:

> *Well, sometimes you need to search on the internet (...) sometimes you see terminology and you think: What are they talking about? But usually it is easy to find out.* [Patient 2]

Patients appreciated that test results were provided with reference values as it helped them to interpret the results. Regarding the full histological reports, it was mentioned that

XSL•FO
RenderX

these were difficult to understand and that the conclusion would be sufficient:

> *Just report the final result. What needs to be done.* [Patient 5]

Overarching Theme 2: Health Care Provider–Patient Communication Through the Portal

Both HCPs and patients expressed uncertainties about the new opportunities for communication that the portal provides. There was disagreement among the HCPs who attended the focus group discussions on how the messaging functionality should be used. Nurses and administrative employees reported that they advised patients to ask their questions through the portal. In contrast, some doctors commented that they preferred not to use the messaging functionality of the portal to answer questions, especially when it concerned complex problems:

> *Sometimes the messages are so complex that a telephone call should be scheduled. However, we are not so well organized yet, and not all doctors agree on this.* [HCP 16]

> *We try to limit the number messages, because it takes so much time, or it may cause misunderstandings or conflicts.* [HCP 1]

Patients appreciated that they could ask questions in between consultations through the portal:

> *That you can ask your questions. That you do not need to call, that you do not need to wait. That you get an answer within a certain period of time. I find that great.* [Patient 10]

Uncertainties about sending messages were also mentioned in the focus groups with patients. It was not always clear who would read the messages:

> *(It says) You can ask questions to your team. Then I think: Who are all these people? To whom do I send this question?* [Patient 11]

In addition, there were uncertainties about the kind of questions and the number of questions that might be included in a message:

> *I felt hesitant because I had very many questions and because it was new to me.* [Patient 9]

Some patients also mentioned feeling reluctant sending messages because of concerns about the workload of the doctors:

> *I do not want to bother my doctor with this, he is much too busy.* [Patient 11]

Among the HCPs, there were reservations concerning patients' access to the letters to the GP. On one hand, it was recognized as an additional way to communicate with their patients:

> *It helps me sometimes as well, if I know that a patient reads it (the letter to the GP). It gives me another opportunity to show how thorough I am. That I can say: We have done this and that, we have excluded that, it is not cancer.* [HCP 6]

On the other hand, there were concerns that this might limit the original function of the letters to report relevant medical information to the GP:

> *I also notice it when writing the letters to the GP. Then I think, the patient reads this as well. Then I formulate more cautiously and I just hope that the GP will still understand what I mean.* [HCP 3]

The experiences of HCPs with the questionnaire functionality differed. On one hand, they were considered a useful time-saving tool:

> *The alternative is that someone walks in with a (paper) questionnaire and that I have to go over it in 5 or 10 minutes. While now, it (the filled-out questionnaire) is just there, with one press of the button. Then I check it and it is done.* [HCP 1]

On the other hand, there were concerns that the information from the questionnaires was not accurate:

> *If the patient fills something out, and does not understand the question correctly, then you don't know. Then he fills something out that is not correct.* [HCP 2]

In the focus groups with patients, the questionnaire functionality was not discussed.

Overarching Theme 3: Usefulness of the Information That Can Be Accessed Through the Portal

Some patients were satisfied with the amount and type of information they could access through the portal, whereas others would like to be able to view more information. Patients appreciated being able to view their test results, especially for monitoring their conditions:

> *Another advantage is that you can compare the lab results with those from a few months ago. You can see whether they have gone up or down, are they better or not.* [Patient 5]

Although patients appreciated to have timely access to their results, some reservations about receiving potentially sensitive information through the portal were put forward:

> *Something I did not like was when I had had a biopsy. Then you have the results before you have seen the doctor. I googled it and found out what it meant. Not exactly of course, because I am not a doctor. I think that the doctor should discuss the results before you can see them through the portal.* [Patient 5]

The HCPs agreed with this, emphasizing that informing patients about test results is their responsibility and that they did not want their patients to worry unnecessarily about their results when viewing them without explanations:

> *You are medically responsible and you prefer the patient not to see the results before you, the professional, sees them.* [HCP 6]

Notwithstanding the uncertainties about how to use the secure messaging functionality, there was appreciation among patients for this functionality as an opportunity to formulate their questions more accurately and ask them in between visits:

XSL•FO
RenderX

> *I can ask my questions when they come up. (...) I can take time to formulate my questions, re-read them to check if I put it right and then send them. During visits, you are not sure whether you ask the right question. Afterwards you think: I should have asked something else.* [Patient 6]

Patients also appreciated being able to check the calendar as a reminder for upcoming appointments.

The medication list was considered less useful as this was not always up to date. In contrast, access to the letters of GPs was considered useful by patients as it provided them with a summary of what was discussed:

> *For me the great advantage is that you can read what was decided. You do not need to remember it, it is just there.* [Patient 3]

Overarching Theme 4: Integration of the Portal in Care and Work Processes

Patients noticed that the use of the portal for care processes varies among HCPs:

> *There are some (HCPs) that really use it. For example, the nephrologist. He says, check your blood pressure and send the results. That stimulates. There are other departments where the portal is hardly mentioned at all.* [Patient 7]

HCPs mentioned that some colleagues do not want to use the portal:

> *There is a group (of medical specialists) who is fundamentally against these sort systems. Also, against electronic health records and all sorts of innovations in this hospital.* [HCP 5]

Patients reported using the portal for preparation of their consultations as a reminder of what was discussed during visits and of upcoming appointments. In addition, they used the portal to contact their HCP in between visits, thus using it as a continuation of their contact with their HCPs and regular care:

> *If you have read it beforehand, you can ask the right questions. (...) That is much more efficient, also for the doctor.* [Patient 6]

In addition, HCPs felt that the consultations might improve when patients use the portal:

> *They are better informed about their medical history, which makes it easier to ask questions during consultations (...) I think this may help during consultations.* [HCP 17]

According to the HCPs, the portal was not firmly integrated in the work process of HCPs yet:

> *What we don't do very well yet, I suppose because the impact is not substantial yet, is that we do not adjust our work processes to the portal.* [HCP 1]

Answering the messages was not formally incorporated in the daily work process of doctors:

> *There is an extra work load because of the in-basket messages, which seem to be sent with little reluctance. (...) You need to plan extra time to process these messages.* [HCP 3]

Some HCPs mentioned that they answered messages in their spare time.

Overarching Theme 5: Doctor and Patient Roles and Relationships

Some patients reported feeling more engaged in their health care and having more control by using the portal:

> *I feel that I have a little more control. Before I depended on the GP or the medical specialist. Now I can monitor the results myself and I like to be able to do that.* [Patient 4]

Others did not think that their relationship with the HCP had changed after they had started to use the portal:

> *No, it has not changed the relationship because there is not enough information in it. Only appointments and letters to the GP. So, you can't refer to it, like "Doctor I read this and that...*" [Patient 10]

HCPs noted changes in the role patients played in their care, using the portal for checking their test results before visits:

> *I had a patient who said, the ALAT (alanine aminotransferase) has increased, is that a problem doctor? That was a nice question (...) I liked the conversation. Before, you did not have these types of discussions (...) it changes the dynamic.* [HCP 5]

They also mentioned changes in their role related to the use of the portal:

> *It feels a little uncomfortable. It feels as if you lose your autonomy as a doctor, because you have your partner, the patient in this case, sitting next to you. Personally, I think this is a good development; I suppose we have to get used to it.* [HCP 5]

HCPs also felt that they lost some control over the health care and communication processes by allowing patients to reschedule appointments through the portal and by providing patients the possibility to send 24×7 messages without any restrictions.

Discussion

Principal Findings

One year after the implementation of the patient portal, 1 out of 7 patients who visited the hospital logged on to the portal. Predictors for adoption were having at least one diagnosis or a medical specialist trajectory open in the hospital, having a higher SES, and belonging to the age group 45 to 75 years. Patients mainly used the portal to view their laboratory results, incoming messages, and letters to their GP. Qualitative analyses revealed how these functionalities could be useful for patients to increase their engagement in their care, that is, to monitor their condition, to remember what was decided during visits, and to prepare for consultations. In addition, patients described how the portal might make their interactions with the hospital more convenient, for example, using the calendar and asking questions in between visits. Regarding the use of the portal for patient-HCP communication, uncertainties were raised by patients and HCPs

XSL•FO
RenderX

about what type of questions and how many questions might be asked or answered through the portal. These uncertainties might explain the rather low utilization of secure messaging functionality. Patients and HCPs noted that the use of the portal was not integrated in care and work processes in all the departments of the hospital. Regarding the potential impact of the portal on patient and HCPs roles, potential loss of autonomy and control over care and communication were brought up by HCPs. Although some patients mentioned the feeling of having more control, others did not find that the portal had changed the relationship with their HCP substantially.

Relation With Findings From Other Studies

Adoption of the Portal

To our knowledge, this study is the first to report an adoption rate of a patient portal in a real-world setting in the Netherlands. When compared with international studies, the adoption rate of 13.5% corresponds to the mean adoption rate of 23% (95% CI 13%-33%) for patient portals in real-world studies reported in a systematic review and meta-analysis on portal adoption [23].

Similar to the results of some reviews on patient portal adoption [24,25], we found an association between portal adoption and belonging to the middle-aged group and having a higher SES. However, in our study, the older age group, being over 75 years, was older than the older age group in most of the studies that were included in these reviews. Lower adoption rates among older patients and patients with lower SES have been explained by limited access to the internet, limited skills to use the internet, limited health literacy, and concerns about security and privacy among older and less educated patients [24-27]. Our study did not include these variables, but limited access to the internet is likely to have contributed to lower adoption among patients older than 75 years in our study. In the Netherlands, in 2016, the year this study was conducted, access to the internet was much lower among people older than 75 years as opposed to the younger age groups, 60% versus 90% to 99% [28].

The other predictors of portal adoption that we found, having at least one open medical diagnosis or trajectory with a medical specialist rather than being a 1-time visitor to the hospital, may point at a higher perceived relevance of the portal for patients with a (chronic) disease and frequent users of health services, as suggested by some reviews [2,23,25]. However, frequent visitors of the hospital may also be more aware of the portal as they are more likely to come across the banners and the flyers about the portal.

Usability of the Portal

Limited user-friendliness or usability problems have been described as barriers for portal adoption [29]. We found that, apart from the challenging procedure to log in, usability problems do not seem to have hampered adoption and continuing use. Patients found the portal easy to navigate and also explained that they understood most of the information, including medical terminology, if necessary, with help from their family, friends, or the internet. This corresponds with the results of the OpenNotes study that demonstrated that few patients reported not being able to understand the information or reported being confused after reading their visit notes [30]. In addition, a study that investigated the terminologies that patients used in Web-based patient-patient communication suggests that patients are more familiar with the medical terminologies concerning their own health problems than HCPs might be aware of [31]. However, both the OpenNotes study and this study on patient-patient communication did not investigate the patients' actual understanding of the terminologies and information. In our study, HCPs had concerns about patients not being able to understand the letters to the GP or their test results. These concerns have also been reported in other studies [32], which is not surprising as before the introduction of patient portals and Web-based access to records, the information in the medical records was aimed at professional use. Some HCPs in our study mentioned that they adjusted their reporting to avoid patients getting confused or anxious. Adjusted reporting may not be necessary as patients find explanations from other sources and they may also like to read the medical terminology to learn from this and bring the communication to a level playing field [33,34]. Communicating sensitive information through the portal seemed to be a different issue, not so much related to understanding the terminology, but to the interpretation of the results and putting them in the perspective of the next steps that need to be undertaken regarding further testing or treatment. Among both HCPs and patients, a preference for face-to-face or telephone contact for sensitive results was brought up. In line with this, at the time of our study, there was a delay in presenting test results to provide HCPs the possibility to communicate results to their patients first personally.

Usefulness of the Portal

The relevance and usefulness of a patient portal for patients to manage their conditions and health care have been suggested to determine portal adoption and use [2,23]. Our quantitative results that demonstrated an association between portal use and having at least one open diagnosis or at least one medical specialist trajectory, also point at this. Our qualitative results demonstrated that both HCPs and patients who use the portal feel that the portal can help to engage patients in their care, add convenience, and provide a new way of communication between patients and HCP. These issues correspond with the mechanisms through which patient portals have been found to produce effects in a realist review, that is, insight into information and activation of information, patient convenience, and continuity of interpersonal care [4].

Concerning the new channel for communication between patients and HCPs that the portal offers, patients explained how the secure messaging function enabled them to interact with their HCPs in between visits and to reflect on their questions. It has been suggested that patients feel more confident asking questions through secure messaging than during face-to-face contacts and that secure messaging enables patients to set the agenda [35]. Therefore, secure messaging can contribute to patient empowerment and patient-centered communication and eventually have an effect on the power relation between HCP and patients [36]. Our study differs from other studies on portal adoption in the relatively low utilization of the secure messaging functionality by patients [5]. This may be explained by the uncertainties that patients and HCPs have about how to use this functionality and concerns about the workload of HCPs. Similar

XSL•FO
RenderX

barriers for using secure messaging have been described by Sieck et al [37]. They emphasize that secure messaging is a new way of communication and therefore requires new rules. They propose to define new *rules of engagement* for patient-HCP communication using secure messaging. In addition, training of patients and HCPs on how to use secure messaging in the care processes have been proposed to overcome these barriers [38-40].

Social Influence: Health Care Provider Endorsement and Integration in Care and Work Processes

HCP endorsement of portal use and integration of the portal in care and administrative processes have been found to have a positive effect on portal adoption by patients [23,25]. In terms of the constructs of the UTAUT, HCP endorsement and integration in care processes can be considered as social influence. In the setting of our study, the portal was introduced as an additional service in flyers and on the hospital website. We found that the portal was not embraced equally by all HCPs in the hospital and some HCPs encouraged patients to use the portal, whereas others did not. Furthermore, we found that the portal was not embedded in the care processes in the ophthalmology and dermatology departments. Integration of the portal in care processes will engage patients more with the portal and will help them to use it as a tool for their health care rather than just as an additional service. Integration of the portal in care processes and in the workflow of HCPs is also likely to engage HCPs with the portal, as has been described for the implementation of new technologies [29,41,42]. In addition, it has been argued that in the digital era, HCPs have to become *e-physicians* and need to be empowered to be able to benefit from digital technologies in their work [43].

Potential Impact of the Portal on the Patient–Health Care Provider Relationship

Our study was not designed to evaluate the impact of the portal on health outcomes or care processes. The previously mentioned realist review linked the mechanisms through which patient portals have been found to work (insight into information and activation of information, convenience, and continuity of interpersonal care) to the effects on health outcomes, patient empowerment, adherence, patient satisfaction, communication, and the patient-HCP relationship [4]. Of these potential effects, the patient-HCP relationship was discussed in the focus groups. Some HCPs felt changes in their control over care processes and also patients mentioned the feeling of having more control. Neither mentioned substantial changes in their roles or in relationship, in the sense of this getting better or worse or more equal. This may be because it was still early to observe the effects on the HCP-patient relationship. However, it may also have to do with the low use of the secure messaging function and the limited integration in care processes or the limited information that can be accessed through the portal. Furthermore, some patients and HCPs pointed out that the portal plays only a small role in their interactions.

Strengths and Limitations

A strength of this study is the multilevel mixed methods design. Mixing qualitative think-aloud observations and focus group discussions among patients and HCPs with quantitative log data and user characteristics helped to deepen our insight into portal utilization. However, the obtained insight may not be complete as we did not include patients who did not use the portal in the focus group discussions. Therefore, we do not have information on other barriers for adoption of the portal, other than the challenging procedure to log in to the portal. In addition, we did not collect individual data on health literacy and education level, variables that are likely to be of more importance for predicting portal adoption than SES score, which is a score at neighborhood level. Another limitation is the selection of the departments from which we recruited patients and HCPs for the focus group discussions. As we expected, the HCPs in the dermatology and ophthalmology departments differed in their utilization of the portal, and thus provided us a broad range of experiences and opinions. However, this choice may have limited the range of patient experiences that we were able to identify. To obtain more variation in patient experiences and opinions, especially regarding engagement in care and self-management, it would have been interesting to include a department where more patients with chronic diseases were treated, for example, the nephrology or oncology department.

Implications for Clinical Practice

This study suggests that the adoption and use of a patient portal might improve if patients and HCPs are informed about how the functionalities of the portal should be used. In addition, use and usefulness of patient portals may improve if HCPs incorporate the portal in their care practices. They can, for example, explain patients how to adjust medication after certain test results or ask patients to prepare for consultations by looking at their test results before a face-to-face consultation. They also might encourage patients to share information that they can access through the portal with their HCPs from other health care organizations and stimulate them to play an active role in the coordination of their care in the fragmented health care system. Hence, HCPs can contribute to a portal being more than just a new service and make it a valuable instrument to improve health care. Regarding the secure messaging functionality, the hospital and individual HCPs should be clear about what sort of questions can be asked through the portal and who will read the messages. In addition, HCPs should be aware that elderly patients and patients with lower SES scores are less likely to use the portal. Furthermore, the HCPs should be allowed to reserve time in their schedules for answering the messages.

Further Research

We have used the constructs of the UTAUT, perceived effort, perceived usefulness, and social influence to investigate the adoption and use of the portal. We included HCP endorsement and integration in care and work processes in our study as social influence, which provided us some insight into the complexity of portal adoption. However, the UTAUT may not be the best model to investigate the complex interplay among the HCPs, between HCPs and patients, HCPs and the hospital, and patients and the hospital related to portal adoption and nonadoption. Recently, Greenhalgh et al proposed a framework for theorizing and evaluating nonadoption, abandonment, scale-up, spread, and sustainability of health and care technologies (NASSS)

XSL•FO
RenderX

[44]. The NASSS framework includes the following 7 domains: the condition, the technology, the value proposition, the adopter system (HCPs, patients, and caregivers), the health organization, the wider institutional and societal context, and the adaptations over time. The NASSS framework overlaps partly with the UTAUT but uses a broader and less linear perspective on adoption. Therefore, it may be useful to use this framework instead of the UTAUT in future studies.

We found that portal adoption may have been hampered by limited integration in the care and work processes. Therefore, implementation research addressing how to embed portal use in the care and work processes is called for. Furthermore, as portals provide new ways for HCP-patient communication, research is needed on how to use these new communication tools in relation to other communication tools such as WhatsApp and the conventional telephone calls and face-to-face consultations.

Conclusions

This study identified some factors that are associated with the adoption of a recently introduced patient portal: age between 45 and 75 years, higher SES, and having an open diagnosis or open trajectory with a medical specialist. It demonstrated that patients and HCPs recognize the potential of a patient portal to engage patients in their care processes, facilitate patient-HCP communication, and increase patient convenience. Limited integration of a patient portal in care is likely to limit its adoption by patients and its usefulness for patient engagement. Vague and inconsistent information about how to use the new communication opportunities of a portal are likely to hinder the utilization of the communication functionality. Instructions on how to use the functionalities of a portal and integration of a portal in care and work processes may improve the utilization of a patient portal and are likely to contribute to a portal becoming a more valuable tool for improving patient engagement, HCP-patient communication, and patient convenience.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Matrix qualitative data.
[PDF File (Adobe PDF File), 127 KB - jmir_v21i8e13743_app1.pdf]

References

1. Rigby M, Georgiou A, Hyppönen H, Ammenwerth E, de Keizer N, Magrabi F, et al. Patient portals as a means of information and communication technology support to patient-centric care coordination - the missing evidence and the challenges of evaluation. A joint contribution of IMIA WG EVAL and EFMI WQ EVAL. Yearb Med Inform 2015 Aug 13;10(1):148-159 [FREE Full text] [doi: 10.15265/IY-2015-007] [Medline: 26123909]
2. 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]
3. Shah SD, Liebovitz D. It takes two to tango: engaging patients and providers with portals. PM R 2017 May;9(5S):S85-S97. [doi: 10.1016/j.pmrj.2017.02.005] [Medline: 28527507]
4. Otte-Trojel T, de Bont A, Rundall TG, van de Klundert J. How outcomes are achieved through patient portals: a realist review. J Am Med Inform Assoc 2014;21(4):751-757 [FREE Full text] [doi: 10.1136/amiajnl-2013-002501] [Medline: 24503882]
5. 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 Sep 8;4(9):e006021 [FREE Full text] [doi: 10.1136/bmjopen-2014-006021] [Medline: 25200561]
6. de Lusignan S, Ross P, Shifrin M, Hercigonja-Szekeres M, Seroussi B. A comparison of approaches to providing patients access to summary care records across old and new Europe: an exploration of facilitators and barriers to implementation. Stud Health Technol Inform 2013;192:397-401. [doi: 10.3233/978-1-61499-289-9-397] [Medline: 23920584]
7. Armstrong S. Patient access to health records: striving for the Swedish ideal. Br Med J 2017 Dec 2;357:j2069. [doi: 10.1136/bmj.j2069] [Medline: 28465316]
8. Bartlett C, Simpson K, Turner AN. Patient access to complex chronic disease records on the internet. BMC Med Inform Decis Mak 2012 Aug 6;12:87 [FREE Full text] [doi: 10.1186/1472-6947-12-87] [Medline: 22867441]
9. Amante DJ, Hogan TP, Pagoto SL, English TM. A systematic review of electronic portal usage among patients with diabetes. Diabetes Technol Ther 2014 Nov;16(11):784-793. [doi: 10.1089/dia.2014.0078] [Medline: 24999599]
10. Nijland N, van Gemert-Pijnen JE, Kelders SM, Brandenburg BJ, Seydel ER. Factors influencing the use of a web-based application for supporting the self-care of patients with type 2 diabetes: a longitudinal study. J Med Internet Res 2011 Sep 30;13(3):e71 [FREE Full text] [doi: 10.2196/jmir.1603] [Medline: 21959968]
11. Urowitz S, Wiljer D, Dupak K, Kuehner Z, Leonard K, Lovrics E, et al. Improving diabetes management with a patient portal: a qualitative study of diabetes self-management portal. J Med Internet Res 2012 Nov 30;14(6):e158 [FREE Full text] [doi: 10.2196/jmir.2265] [Medline: 23195925]

XSL•FO
RenderX

12. Hoe online is jouw ziekenhuis?. URL: https://www.hoeonlineisjouwziekenhuis.nl/
13. Krijgsman J, Swinkels I, van Lettow B, de Jong J, Out K, Friele R, et al. Meer Dan Techniek: eHealth-Monitor 2016. The Hague and Utrecht, Netherlands: Nictiz and NIVEL; 2016.
14. Wouters M, Swinkels I, Sinnige J, de Jong J, Brabers A, van Lettow B, et al. Kies Bewust Voor eHealth: eHealth-monitor 2017. The Hague and Utrecht, Netherlands: Nictiz and NIVEL; 2017.
15. Jilka SR, Callahan R, Sevdalis N, Mayer EK, Darzi A. 'Nothing about me without me': an interpretative review of patient accessible electronic health records. J Med Internet Res 2015 Jun 29;17(6):e161 [FREE Full text] [doi: 10.2196/jmir.4446] [Medline: 26123476]
16. Kaelber DC, Jha AK, Johnston D, Middleton B, Bates DW. A research agenda for personal health records (PHRs). J Am Med Inform Assoc 2008;15(6):729-736 [FREE Full text] [doi: 10.1197/jamia.M2547] [Medline: 18756002]
17. Kruse CS, Bolton K, Freriks G. The effect of patient portals on quality outcomes and its implications to meaningful use: a systematic review. J Med Internet Res 2015 Feb 10;17(2):e44 [FREE Full text] [doi: 10.2196/jmir.3171] [Medline: 25669240]
18. Venkatesh V, Morris M, Davis GF. User acceptance of information technology: toward a unified view. MIS Q 2003;27(3):425-478. [doi: 10.2307/30036540]
19. Sociaal en Cultureel Planbureau. URL: https://www.scp.nl/Onderzoek/Lopend_onderzoek/A_Z_alle_lopende_onderzoeken/Statusscores
20. Boren T, Ramey J. Thinking aloud: reconciling theory and practice. IEEE Trans Profess Commun 2000;43(3):261-278. [doi: 10.1109/47.867942]
21. Jaspers MW, Steen T, van den Bos C, Geenen M. The think aloud method: a guide to user interface design. Int J Med Inform 2004 Nov;73(11-12):781-795. [doi: 10.1016/j.ijmedinf.2004.08.003] [Medline: 15491929]
22. Gale NK, Heath G, Cameron E, Rashid S, Redwood S. Using the framework method for the analysis of qualitative data in multi-disciplinary health research. BMC Med Res Methodol 2013 Sep 18;13:117 [FREE Full text] [doi: 10.1186/1471-2288-13-117] [Medline: 24047204]
23. Fraccaro P, Vigo M, Balatsoukas P, Buchan IE, Peek N, van der Veer SN. Patient portal adoption rates: a systematic literature review and meta-analysis. Stud Health Technol Inform 2017;245:79-83. [doi: 10.3233/978-1-61499-830-3-79] [Medline: 29295056]
24. Wildenbos GA, Peute L, Jaspers M. Facilitators and barriers of electronic health record patient portal adoption by older adults: a literature study. Stud Health Technol Inform 2017;235:308-312. [doi: 10.3233/978-1-61499-753-5-308] [Medline: 28423804]
25. Irizarry T, DeVito DA, Curran CR. Patient portals and patient engagement: a state of the science review. J Med Internet Res 2015 Jun 23;17(6):e148 [FREE Full text] [doi: 10.2196/jmir.4255] [Medline: 26104044]
26. Lyles CR, Allen JY, Poole D, Tieu L, Kanter MH, Garrido T. 'I want to keep the personal relationship with my doctor': understanding barriers to portal use among African Americans and Latinos. J Med Internet Res 2016 Dec 3;18(10):e263 [FREE Full text] [doi: 10.2196/jmir.5910] [Medline: 27697748]
27. Sakaguchi-Tang DK, Bosold AL, Choi YK, Turner AM. Patient portal use and experience among older adults: systematic review. JMIR Med Inform 2017 Oct 16;5(4):e38 [FREE Full text] [doi: 10.2196/medinform.8092] [Medline: 29038093]
28. StatLine - CBS. URL: https://opendata.cbs.nl/statline/#/CBS/nl/dataset/83429NED/table?ts=1560359792018
29. Kruse CS, Argueta DA, Lopez L, Nair A. Patient and provider attitudes toward the use of patient portals for the management of chronic disease: a systematic review. J Med Internet Res 2015 Feb 20;17(2):e40 [FREE Full text] [doi: 10.2196/jmir.3703] [Medline: 25707035]
30. 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]
31. Fage-Butler AM, Nisbeth JM. Medical terminology in online patient-patient communication: evidence of high health literacy? Health Expect 2016 Dec;19(3):643-653 [FREE Full text] [doi: 10.1111/hex.12395] [Medline: 26287945]
32. Walker J, Leveille SG, Ngo L, Vodicka E, Darer JD, Dhanireddy S, et al. Inviting patients to read their doctors' notes: patients and doctors look ahead: patient and physician surveys. Ann Intern Med 2011 Dec 20;155(12):811-819 [FREE Full text] [doi: 10.7326/0003-4819-155-12-201112200-00003] [Medline: 22184688]
33. Britto MT, Hesse EA, Kamdar OJ, Munafo JK. Parents' perceptions of a patient portal for managing their child's chronic illness. J Pediatr 2013 Jul;163(1):280-1.e1. [doi: 10.1016/j.jpeds.2013.02.041] [Medline: 23541773]
34. Earnest MA, Ross SE, Wittevrongel L, Moore LA, Lin CT. Use of a patient-accessible electronic medical record in a practice for congestive heart failure: patient and physician experiences. J Am Med Inform Assoc 2004;11(5):410-417 [FREE Full text] [doi: 10.1197/jamia.M1479] [Medline: 15187074]
35. Houston TK, Sands DZ, Jenckes MW, Ford DE. Experiences of patients who were early adopters of electronic communication with their physician: satisfaction, benefits, and concerns. Am J Manag Care 2004 Sep;10(9):601-608 [FREE Full text] [Medline: 15515992]
36. Fage-Butler AM, Jensen MN. The relevance of existing health communication models in the email age: an integrative literature review. Commun Med 2015;12(2-3):117-128. [doi: 10.1558/cam.18399] [Medline: 29048139]

XSL•FO
RenderX

37. Sieck CJ, Hefner JL, Schnierle J, Florian H, Agarwal A, Rundell K, et al. The rules of engagement: perspectives on secure messaging from experienced ambulatory patient portal users. JMIR Med Inform 2017 Jul 4;5(3):e13 [FREE Full text] [doi: 10.2196/medinform.7516] [Medline: 28676467]
38. Haun JN, Lind JD, Shimada SL, Martin TL, Gosline RM, Antinori N, et al. Evaluating user experiences of the secure messaging tool on the veterans affairs' patient portal system. J Med Internet Res 2014 Mar 6;16(3):e75 [FREE Full text] [doi: 10.2196/jmir.2976] [Medline: 24610454]
39. 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 Dec 21;17(12):e282 [FREE Full text] [doi: 10.2196/jmir.5152] [Medline: 26690761]
40. Nazi KM. The personal health record paradox: health care professionals' perspectives and the information ecology of personal health record systems in organizational and clinical settings. J Med Internet Res 2013 Apr 4;15(4):e70 [FREE Full text] [doi: 10.2196/jmir.2443] [Medline: 23557596]
41. Miller Jr DP, Latulipe C, Melius KA, Quandt SA, Arcury TA. Primary care providers' views of patient portals: interview study of perceived benefits and consequences. J Med Internet Res 2016 Jan 15;18(1):e8 [FREE Full text] [doi: 10.2196/jmir.4953] [Medline: 26772771]
42. Granja C, Janssen W, Johansen MA. Factors determining the success and failure of ehealth interventions: systematic review of the literature. J Med Internet Res 2018 Dec 1;20(5):e10235 [FREE Full text] [doi: 10.2196/10235] [Medline: 29716883]
43. Mesko B, Győrffy Z. The rise of the empowered physician in the digital health era: viewpoint. J Med Internet Res 2019 Mar 26;21(3):e12490 [FREE Full text] [doi: 10.2196/12490] [Medline: 30912758]
44. Greenhalgh T, Wherton J, Papoutsi C, Lynch J, Hughes G, A'Court C, et al. Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. J Med Internet Res 2017 Dec 1;19(11):e367 [FREE Full text] [doi: 10.2196/jmir.8775] [Medline: 29092808]

Abbreviations

GP: general practitioner
HCP: health care provider
NASSS: nonadoption, abandonment, scale-up, spread, and sustainability of health and care technologies
SES: socioeconomic status
UTAUT: Unified Theory of Acceptance and Use of Technology

Edited by G Eysenbach; submitted 18.02.19; peer-reviewed by T Irizarry, I Riippa; comments to author 28.03.19; revised version received 14.06.19; accepted 29.06.19; published 20.08.19.

Please cite as:
Vreugdenhil MMT, Ranke S, de Man Y, Haan MM, Kool RB
Patient and Health Care Provider Experiences With a Recently Introduced Patient Portal in an Academic Hospital in the Netherlands: Mixed Methods Study
J Med Internet Res 2019;21(8):e13743
URL: https://www.jmir.org/2019/8/e13743
doi:10.2196/13743
PMID:31432782

©Maria M T Vreugdenhil, Sander Ranke, Yvonne de Man, Maaike M. Haan, Rudolf B. Kool. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 20.08.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Comparing Characteristics of Patients Who Connect Their iPhones to an Electronic Health Records System Versus Patients Who Connect Without Personal Devices: Cohort Study

William J Gordon[1,2,3], MD, MBI; David W Bates[1,2], MSc, MD; Daniel Fuchs[3], BSc; John Pappas[3], AS; Sara Silacci[4], BSc; Adam Landman[2,3,5], MS, MD, MIS, MHS

[1]Department of Medicine, Brigham and Women's Hospital, Boston, MA, United States

[2]Harvard Medical School, Boston, MA, United States

[3]Partners HealthCare, Somerville, MA, United States

[4]Massachusetts General Hospital, Boston, MA, United States

[5]Department of Emergency Medicine, Brigham and Women's Hospital, Boston, MA, United States

Corresponding Author:
William J Gordon, MD, MBI
Department of Medicine, Brigham and Women's Hospital
75 Francis St
Boston, MA, 02115
United States
Phone: 1 (617) 732 5500
Email: WJGORDON@partners.org

Abstract

Background: While individual access to health records has traditionally been through paper and other physical media, there has been a recent push toward digitizing this process. Direct patient access to health data through application programming interfaces (APIs) is an important part of current United States policy initiatives, and Apple has created the product “Health Records on iPhone” to leverage APIs for this purpose.

Objective: The objective of this study was to examine the characteristics of patients at our institution who connected their personal iPhone devices to our electronic health records (EHRs) system through “Health Records on iPhone”, as compared to patients at our institution who used our patient portal but did not connect a personal device to our system.

Methods: We examined adult patients at our institution who had authorized an iPhone device to download their health data from the Partners HealthCare EHR via APIs through “Health Records on iPhone” from February 18, 2018 (the date this feature was enabled at our health system) until February 17, 2019. We compared these patients to adult patients who used our portal at least once during this period but did not authorize an iPhone device to download their data via APIs.

Results: Variables associated with an increased likelihood of using “Health Records on iPhone” included male gender (adjusted OR 3.36; 95% CI 3.11-3.62; *P*<.001) and younger age, particularly below 50 years of age. With each decade of age over 50, people were less likely to be “Health Records on iPhone” product users. Asian patients were more likely to use the product than Caucasian patients (adjusted OR 1.32; 95% CI 1.16-1.51; *P*<.001), though there was no significant difference between African Americans and Caucasians (adjusted OR 1.15; 95% CI 0.94-1.41; *P*=.17). Patients who resided in higher ZIP code income quartiles were more likely to be users than those in the lowest quartile.

Conclusions: Early results from the implementation of patient-facing APIs at a single institution suggest that there are opportunities for expanding these technologies to ensure all patients are aware of, and have access to, their health data on their personal devices. More work is needed on expanding these technologies to different patient populations.

(J Med Internet Res 2019;21(8):e14871) doi:10.2196/14871

KEYWORDS

health information interoperability; patient participation; information technology; mobile health

XSL•FO
RenderX

Introduction

Giving patients access to their own health data is widely felt to be beneficial for numerous reasons, including better patient engagement, enhanced care coordination, and improved patient safety [1-4]. In fact, defining an individual's right to access their own health records represents an important component of the Health Insurance Portability and Accountability Act's (HIPAA) Privacy Rule [5]. Personal health records (PHRs) are electronic applications, designed to be used by individuals, that allow for accessing, managing, and capturing health data about an individual. PHRs stand in contrast to electronic health records (EHRs), which are primarily used by clinicians and health care institutions, though EHRs may offer patients an online portal where they can access a subset of EHR information and enter additional information about their health [6]. PHRs have existed for decades, with many different implementations, from vendor products directly linked to EHRs to industry efforts led by companies like Google and Microsoft [1,6,7].

Notwithstanding the potential benefits of patients accessing their own health data electronically, PHR usage is often low, and as a result, many patients do not have electronic access to their health data [8-10]. Therefore, individual access to health records remains largely through paper and other physical media [11], even though physical records can be challenging to access. For example, in the United States, under HIPAA, hospitals can charge a fee for release of medical records which is often higher than recommended [12,13]. Regulatory noncompliance, procedural hurdles, and convenience issues with managing paper and physical media are additional barriers to patients accessing their health data [12,13].

Because of these concerns, the United States Centers for Medicare and Medicaid Services (CMS) Promoting Interoperability program (formerly Meaningful Use) has required various forms of electronic access to health data for patients for many years, with the most recent requirement that certified EHRs include functionality for patients to connect third-party applications via application programming interface (API) technology (a mechanism for applications to communicate directly with EHRs) [14]. Additionally, the 21st Century Cures Act (21CCA) requires that certified EHRs have published APIs available for patients [15] and the United States Department of Health and Human Services recently proposed a new rule that would implement these data sharing provisions from 21CCA and expand the number of required data elements to be shared [16].

With these regulations in mind, in 2018 Apple announced a new "Health Records on iPhone" feature that would enable patients to directly connect their iPhone through APIs to EHRs using a direct connection and would allow them to download, aggregate, and view their records (medications, allergies, results, etc). Patients could additionally choose to allow third-party apps to access these data [17]. The number of participating healthcare organizations has expanded substantially since the original release [18], and Apple recently expanded availability of this functionality to any US healthcare system with a compatible EHR [19]. Though this technology is limited to iPhone devices only, there are an estimated 193 million iPhone units in the United States [20], which presents a major opportunity for improving patient access to health data.

Despite the growing availability of third-party applications that connect to EHRs via APIs, little is known about the patients who have begun using them. Therefore, we examined the characteristics of patients at our institution who connected their personal Apple devices to our EHR through "Health Records on iPhone", as compared to patients at our institution who used our patient portal but did not connect a personal device to our EHR.

Methods

We identified adult patients at our institution who had authorized an iPhone device to download their health data via APIs with "Health Records on iPhone," from the Partners HealthCare EHR, from February 18, 2018 (the date this feature was enabled at our health system) until February 17, 2019. For the purposes of this study, using the API was defined as authorizing the iPhone product "Health Records on iPhone" to download health data at least once during the study period. Authorizations were retrieved from an internal audit log database. Our healthcare system (Partners HealthCare) uses the Epic EHR software (Epic Systems Corporation, Verona, Wisconsin). During the study period, our portal was a custom-developed product that utilized native Epic MyChart functionality for many components, including the API functionality. Our control group consisted of adult patients who used our portal at least once during this period but did not authorize an iPhone device to download their data via APIs. The control group was not limited to patients with iPhone devices. Due to data use agreements, we were unable to report the total number of "Health Records on iPhone" users, so we instead took a random sample of the total from each population.

We calculated descriptive statistics and performed a multivariable logistic regression to compute odds ratios (ORs) with 95% CI for the odds that a patient would be in our case group ("Health Records on iPhone" users). Covariates included gender, age (split into ranges of 18-40, 41-50, 51-60, 61-70, and >80), race and ethnicity (using United States Census Bureau groupings), primary language, and the United States census median household income quartile of their primary ZIP code, which we obtained from the United States Census Bureau [21]. Patient characteristics data were obtained from an internal clinical reporting system. Data analysis was conducted using R statistical software version 3.5.1 (R Project for Statistical Computing, Vienna, Austria). The Partners HealthCare institutional review board approved this study.

Results

We randomly sampled 3000 "Health Records on iPhone" users and compared them to 100,000 randomly sampled patient portal users who did not use the feature (Table 1).

XSL•FO
RenderX

Table 1. Association between patient characteristics and usage of "Health Records on iPhone".

Characteristic	Non-"Health Records on iPhone" users (n=100,000), n (%)	"Health Records on iPhone" users (n=3000), n (%)	Adjusted OR[a] (95% CI)	*P* value
Gender			3.36 (3.11-3.62)	<.001
Female	62,813 (62.8)	1069 (35.6)	—[b]	—
Male	37,187 (37.2)	1931 (64.4)	—	—
Primary language				
English	95,251 (95.3)	2897 (96.6)	—	—
Spanish	681 (0.68)	17 (0.57)	0.66 (0.40-1.11)	.12
Other	1163 (1.16)	24 (0.80)	0.71 (0.47-1.07)	.11
Not available	2905 (2.90)	62 (2.07)	0.83 (0.63-1.09)	.19
Race				
Caucasian	83,215 (83.2)	2408 (80.3)	—	—
Asian	5596 (5.60)	262 (8.73)	1.32 (1.16-1.51)	<.001
African American	3196 (3.20)	109 (3.63)	1.15 (0.94-1.41)	.17
Other	3390 (3.39)	124 (4.13)	0.91 (0.71-1.15)	.43
Not available	4603 (4.60)	97 (3.23)	0.66 (0.52-0.83)	<.001
Ethnicity				
Hispanic or Latino	3309 (3.31)	137 (4.57)	—	—
Not Hispanic or Latino	88,092 (88.1)	2598 (86.6)	0.67 (0.53-0.85)	.001
Not available	8599 (8.60)	265 (8.83)	0.77 (0.59-0.99)	.05
Age range				
18-40	31,377 (31.4)	1233 (41.1)	—	—
41-50	16,255 (16.3)	690 (23.0)	1.00 (0.90-1.10)	.94
51-60	19,573 (19.6)	513 (17.1)	0.61 (0.54-0.67)	<.001
61-70	18,844 (18.8)	363 (12.1)	0.43 (0.38-0.49)	<.001
71-80	10,827 (10.8)	177 (5.90)	0.35 (0.39-0.41)	<.001
>80	3124 (3.12)	24 (0.80)	0.16 (0.11-0.24)	<.001
Median household income quartile, by ZIP code (US $)				
4836-41,406.50	24,844 (24.8)	669 (22.3)	—	—
41,406.50-51,897	24,787 (24.8)	725 (24.2)	1.12 (1.01-1.25)	.05
51,897-65,903.50	24,735 (24.7)	778 (25.9)	1.15 (1.03-1.28)	.02
65,903.50-244,671	24,703 (24.7)	809 (27.0)	1.21 (1.09-1.35)	<.001
Non-US, invalid, or missing	931 (0.93)	19 (0.63)	0.91 (0.57-1.45)	.71

[a]OR: odds ratio.

[b]Not applicable.

In a multivariable analysis, characteristics associated with an increased likelihood of using "Health Records on iPhone" included male gender (adjusted OR 3.36; 95% CI 3.11-3.62; *P*<.001) and younger age, particularly below 50 years of age. With each decade over 50 years of age, people were less likely to be "Health Records on iPhone" users. Additionally, Asian patients were more likely to use the app than Caucasians (adjusted OR 1.32; 95% CI 1.16-1.51; *P*<.001), though there was no significant difference between African Americans and Caucasians (adjusted OR 1.15; 95% CI 0.94-1.41; *P*=.17). Spanish as a primary language was not associated with "Health Records on iPhone" usage, as compared to English as a primary language (adjusted OR 0.66; 95% CI 0.40-1.11; *P*=.12). Hispanic ethnicity was more associated with "Health Records on iPhone" usage than non-Hispanic or non-Latino ethnicity (adjusted OR 0.67; 95% CI 0.53-0.85; *P*=.001). Finally, patients who resided in higher ZIP code income quartiles were more likely to be users than those in the lowest quartile, with comparisons to quartile 2 (adjusted OR 1.12; 95% CI 1.01-1.25; *P*=.05), to quartile 3 (adjusted OR 1.15; 95% CI 1.03-1.28;

XSL•FO
RenderX

P=.02) and to quartile 4 (adjusted OR 1.21; 95% CI 1.09-1.35; *P*<.001). Full results are listed in Table 1. Unadjusted results did not substantially differ from the adjusted results and can be found in Multimedia Appendix 1.

Discussion

In this single center study of the characteristics of patients who used "Health Records on iPhone", patients that used the product differed in important ways from patients that also used our online patient portal but did not use the app. Initial users were more likely to be male and reside in a ZIP code with a higher median household income than patients who used our portal but did not connect their personal device to our EHR. Additionally, users of this technology were more likely to be younger than 50 compared to non-API portal users. Patients of Asian race used "Health Records on iPhone" more than patients of Caucasian race, but no other racial or ethnic differences were observed beyond non-Hispanic or non-Latino ethnicity using it less than Hispanic or Latino ethnicity.

In the United States, patients have the right to access their health information, as the HIPAA Privacy Rule requires that covered entities give individuals access to their data, upon request, with exceptions for some items like psychotherapy notes [22]. PHRs and patient portals have existed for decades in various forms, though there has been significant expansion of these technologies in the last decade driven by overall technological advancements and US federal regulatory requirements in the form of Meaningful Use (now called Promoting Interoperability). Prior work has shown that, overall, patient portal usage seems to be increasing [23,24] but remains low, with typically less than 50% of patients using online portals [9,10,25]. Additionally, prior work has shown that patient portal use may reflect or exacerbate a digital divide between sociodemographic patient groups. For example, Perzynski et al [25] showed that racial and ethnic minorities and patients of lower socioeconomic status were less likely to use a patient portal, along with those without broadband internet access in their neighborhood. Pho et al and Gerber et al [23,24] showed that in an oncology population, certain characteristics were more associated with portal usage, such as younger age, Caucasian race, and Spanish-speaking patients. Lockwood et al [9] also demonstrated important sociodemographic differences in portal usage in a pre and post-kidney transplant population. These, and our findings, exist within a broader literature describing a very real digital divide between patients who do not have ready access to the internet due to literacy, cost, or other barriers, and those that do have immediate access [26-31].

APIs provide an opportunity for deeper, more seamless integration with health data and EHRs than is possible through patient portals, and their availability is now required as part of a certified EHR. Through APIs an entire ecosystem of PHRs can form, with patients free to choose among multiple solutions depending on their needs. APIs are particularly enabling for mobile devices to retrieve health data, with these devices becoming increasingly common. For example, smartphone usage in the United States has dramatically increased over the past decade, from 33% in 2011 to 84% in 2019 [32]. More Americans now have a smartphone than a desktop or laptop computer [32]. "Health Records on iPhone" is one of the first major products to take advantage of these functionalities. In the United States, the CMS program Promoting Interoperability requires certified EHRs to include patient-facing API technology as of 2019. Because this technology is new, little is known about what types of patients are connecting their personal devices to the EHR to retrieve data. Initial reports suggest patients are receptive to these technologies [33], but we are not aware of any other work examining the characteristics of patients using APIs. Our results are an important first look at which patients are connecting their personal devices to the EHR system through APIs.

Our study has several limitations. First, we looked only at patient demographics. More work is needed to look at the clinical characteristics of these patients. Additionally, our study is a single site analysis and looked only at one product ("Health Records on iPhone") on one type of smartphone operating system (Apple, iOS). "Health Records on iPhone" is by far the largest current implementation of these technologies. As these technologies expand, we expect more product usage across different types of personal devices, and it will be essential that these analyses are replicated for difference devices and populations. Third, we did not have access to the specific model of iPhone device used by the patients. Finally, we were unable to account for unmeasured clustering of patients using these technologies (eg, driven by specific provider groups that encouraged usage). However, we are not aware of any such behaviors at our institution.

In summary, we report here a first look at the characteristics of the initial cohort of patients who used "Health Records on iPhone" and show that these initial patients differ in important ways from patients who did not use this product but still used our portal. More work is needed to understand how to expand this technology to other members of the community and how policies can be modified to improve patient access to data broadly.

Acknowledgments

We would like to acknowledge Cindy Bero for helping think through study design and data availability.

Conflicts of Interest

AL reported receiving personal fees from Abbott Medical Device Cybersecurity Council outside the submitted work. DWB reported consulting for EarlySense, which makes patient safety monitoring systems. He receives cash compensation from CDI-Negev Ltd, which is a not-for-profit incubator for health information technology startups. He receives equity from ValeraHealth,

XSL•FO
RenderX

which makes software to help patients with chronic diseases, from Clew, which makes software to support clinical decision-making in intensive care, and from MDClone, which produces deidentified versions of clinical data.

Multimedia Appendix 1
Unadjusted patient characteristics of health records on iPhone users compared to patient portal users who did not use health records on iPhone.
[PDF File (Adobe PDF File), 226KB - jmir_v21i8e14871_app1.pdf]

References

1. Halamka JD, Mandl KD, Tang PC. Early experiences with personal health records. J Am Med Inform Assoc 2008;15(1):1-7 [FREE Full text] [doi: 10.1197/jamia.M2562] [Medline: 17947615]
2. Kim J, Jung H, Bates DW. History and Trends of "Personal Health Record" Research in PubMed. Healthc Inform Res 2011 Mar;17(1):3-17 [FREE Full text] [doi: 10.4258/hir.2011.17.1.3] [Medline: 21818452]
3. Fossa AJ, Bell SK, DesRoches C. OpenNotes and shared decision making: a growing practice in clinical transparency and how it can support patient-centered care. J Am Med Inform Assoc 2018 Sep 01;25(9):1153-1159. [doi: 10.1093/jamia/ocy083] [Medline: 29982659]
4. Irizarry T, DeVito DA, Curran CR. Patient Portals and Patient Engagement: A State of the Science Review. J Med Internet Res 2015;17(6):e148 [FREE Full text] [doi: 10.2196/jmir.4255] [Medline: 26104044]
5. Health Insurance Portability Accountability Act of 1996. Sect 45 CFR 164308 (5). URL: https://www.federalregister.gov/documents/2013/01/25/2013-01073/modifications-to-the-hipaa-privacy-security-enforcement-and-breach-notification-rules-under-the [accessed 2019-08-15]
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. Roehrs A, da Costa CA, Righi RD, de Oliveira KS. Personal Health Records: A Systematic Literature Review. J Med Internet Res 2017 Jan 06;19(1):e13 [FREE Full text] [doi: 10.2196/jmir.5876] [Medline: 28062391]
8. Dumitrascu AG, Burton MC, Dawson NL, Thomas CS, Nordan LM, Greig HE, et al. Patient portal use and hospital outcomes. J Am Med Inform Assoc 2018 Apr 01;25(4):447-453. [doi: 10.1093/jamia/ocx149] [Medline: 29300961]
9. Lockwood M, Dunn-Lopez K, Pauls H, Burke L, Shah S, Saunders M. If you build it, they may not come: modifiable barriers to patient portal use among pre- and post-kidney transplant patients. JAMIA Open 2018 Oct 01;1(2):264. [doi: 10.1093/jamiaopen/ooy024]
10. Hoogenbosch B, Postma J, de Man-van Ginkel JM, Tiemessen NA, van Delden JJ, van Os-Medendorp H. Use and the Users of a Patient Portal: Cross-Sectional Study. J Med Internet Res 2018 Sep 17;20(9):e262 [FREE Full text] [doi: 10.2196/jmir.9418] [Medline: 30224334]
11. Murphy-Abdouch K. Patient Access to Personal Health Information: Regulation vs. Reality. Perspect Health Inf Manag 2015;12:1c [FREE Full text] [Medline: 26807075]
12. Jaspers AW, Cox JL, Krumholz HM. Copy Fees and Limitation of Patients' Access to Their Own Medical Records. JAMA Intern Med 2017 Apr 01;177(4):457-458. [doi: 10.1001/jamainternmed.2016.8560] [Medline: 28135350]
13. Lye CT, Forman HP, Gao R, Daniel JG, Hsiao AL, Mann MK, et al. Assessment of US Hospital Compliance With Regulations for Patients' Requests for Medical Records. JAMA Netw Open 2018 Oct 05;1(6):e183014. [doi: 10.1001/jamanetworkopen.2018.3014]
14. Department of Health and Human Services. Fed Regist 83 No 160. 2018 Aug 17. Medicare Program; Hospital Inpatient Prospective Payment Systems for Acute Care Hospitals and the Long-Term Care Hospital Prospective Payment System and Policy Changes and Fiscal Year 2019 Rates; Quality Reporting Requirements for Specific Providers; Medicare and Medicaid Electronic Health Record (EHR) Incentive Programs (Promoting Interoperability Programs) Requirements for Eligible Hospitals, Critical Access Hospitals, and Eligible Professionals; Medicare Cost Reporting Requirements; and Physician Certification and Recertification of Claims URL: https://www.govinfo.gov/content/pkg/FR-2018-08-17/pdf/2018-16766.pdf [accessed 2019-08-15]
15. 114th Congress (2015-2016). H.R.34 - 21st Century Cures Act. 2016. 21st Century Cures Act URL: https://www.congress.gov/bill/114th-congress/house-bill/34/ [accessed 2019-08-15]
16. Department of Health and Human Services. Fed Regist 84 No 42. 2019 Mar 4. 21st Century Cures Act: Interoperability, Information Blocking, and the ONC Health IT Certification Program URL: https://www.govinfo.gov/content/pkg/FR-2019-03-04/pdf/2019-02224.pdf [accessed 2019-05-22]
17. Apple Inc. 2018. Apple announces effortless solution bringing health records to iPhone URL: https://www.apple.com/newsroom/2018/01/apple-announces-effortless-solution-bringing-health-records-to-iPhone/ [accessed 2019-05-22]
18. Apple Inc. Institutions that support health records on iPhone (beta) URL: https://support.apple.com/en-us/HT208647 [accessed 2019-05-22]

XSL•FO
RenderX

19. Terry K. Medscape. Apple Opens iPhone EHR Feature to All Healthcare Organizations URL: https://www.medscape.com/viewarticle/915208 [accessed 2019-07-11]
20. MacRumors. Apple's U.S. iPhone User Base Sees Slowing Growth in Q1 2019 URL: https://www.macrumors.com/2019/05/16/apple-us-iphone-user-base-growth-slows-q1-2019/ [accessed 2019-08-15]
21. United States Census Bureau. URL: https://www.census.gov/en.html [accessed 2019-05-22]
22. US Department of Health and Human Services. 45 CFR § 164.524. 2014 Feb 06. Access of individuals to protected health information URL: https://www.hhs.gov/hipaa/for-professionals/privacy/guidance/access/index.html [accessed 2019-08-15]
23. Pho K, Lu R, Gates S, Xie Y, Lee SJC, Gerber DE. Characteristics of Patients Using Patient Portals in Oncology. JAMA Oncol 2018 Mar 01;4(3):416-418 [FREE Full text] [doi: 10.1001/jamaoncol.2017.5257] [Medline: 29372242]
24. Gerber DE, Laccetti AL, Chen B, Yan J, Cai J, Gates S, et al. Predictors and intensity of online access to electronic medical records among patients with cancer. J Oncol Pract 2014 Sep;10(5):e307-e312 [FREE Full text] [doi: 10.1200/JOP.2013.001347] [Medline: 25006222]
25. Perzynski AT, Roach MJ, Shick S, Callahan B, Gunzler D, Cebul R, et al. Patient portals and broadband internet inequality. J Am Med Inform Assoc 2017 Sep 01;24(5):927-932 [FREE Full text] [doi: 10.1093/jamia/ocx020] [Medline: 28371853]
26. Ohno-Machado L. Understanding and mitigating the digital divide in health care. J Am Med Inform Assoc 2017 Sep 01;24(5):881. [doi: 10.1093/jamia/ocx082]
27. Jha AK, DesRoches CM, Shields AE, Miralles PD, Zheng J, Rosenbaum S, et al. Evidence of an emerging digital divide among hospitals that care for the poor. Health Aff (Millwood) 2009 Jan 01;28(6):w1160-w1170. [doi: 10.1377/hlthaff.28.6.w1160] [Medline: 19858142]
28. Din HN, McDaniels-Davidson C, Nodora J, Madanat H. Profiles of a Health Information-Seeking Population and the Current Digital Divide: Cross-Sectional Analysis of the 2015-2016 California Health Interview Survey. J Med Internet Res 2019 May 14;21(5):e11931 [FREE Full text] [doi: 10.2196/11931] [Medline: 31094350]
29. Woods SS, Forsberg CW, Schwartz EC, Nazi KM, Hibbard JH, Houston TK, et al. The Association of Patient Factors, Digital Access, and Online Behavior on Sustained Patient Portal Use: A Prospective Cohort of Enrolled Users. J Med Internet Res 2017 Oct 17;19(10):e345 [FREE Full text] [doi: 10.2196/jmir.7895] [Medline: 29042345]
30. López L, Green AR, Tan-McGrory A, King RS, Betancourt JR. Bridging the Digital Divide in Health Care: The Role of Health Information Technology in Addressing Racial and Ethnic Disparities. The Joint Commission Journal on Quality and Patient Safety 2011 Oct;37(10):437-445. [doi: 10.1016/S1553-7250(11)37055-9]
31. Mackert M, Mabry-Flynn A, Champlin S, Donovan EE, Pounders K. Health Literacy and Health Information Technology Adoption: The Potential for a New Digital Divide. J Med Internet Res 2016 Oct 04;18(10):e264 [FREE Full text] [doi: 10.2196/jmir.6349] [Medline: 27702738]
32. Pew Research Center. 2019 Core Trends Survey URL: https://www.pewresearch.org/wp-content/uploads/2019/05/FT_19.05.23_DigitalDivideCommType_methodology_topline.pdf [accessed 2019-08-15]
33. Dameff C, Clay B, Longhurst CA. Personal Health Records: More Promising in the Smartphone Era? JAMA 2019 Jan 29;321(4):339-340. [doi: 10.1001/jama.2018.20434] [Medline: 30633300]

Abbreviations

21CCA: 21st Century Cures Act
API: application programming interface
CMS: Centers for Medicare and Medicaid Services
EHR: electronic health record
HIPAA: Health Insurance Portability and Accountability Act
OR: odds ratio
PHR: personal health record

Edited by G Eysenbach; submitted 30.05.19; peer-reviewed by A Hernández, A Nguyen, B Smaradottir, S Lim, H Mehdizadeh; comments to author 01.07.19; revised version received 21.07.19; accepted 13.08.19; published 22.08.19.

Please cite as:
Gordon WJ, Bates DW, Fuchs D, Pappas J, Silacci S, Landman A
Comparing Characteristics of Patients Who Connect Their iPhones to an Electronic Health Records System Versus Patients Who Connect Without Personal Devices: Cohort Study
J Med Internet Res 2019;21(8):e14871
URL: http://www.jmir.org/2019/8/e14871/
doi:10.2196/14871
PMID:31441430

XSL•FO
RenderX

©William J Gordon, David W Bates, Daniel Fuchs, John Pappas, Sara Silacci, Adam Landman. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 22.08.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Translation of the eHealth Impact Questionnaire for a Population of Dutch Electronic Health Users: Validation Study

Koen Ilja Neijenhuijs[1,2], BA, MSc; Anja van der Hout[1,2], BA, MSc; Evalien Veldhuijzen[2,3], BA, MSc; Gwendolijne G M Scholten-Peeters[4,5], PhD; Cornelia F van Uden-Kraan[1,2], PhD; Pim Cuijpers[1], PhD; Irma M Verdonck-de Leeuw[1,2,3], PhD

[1]Department of Clinical, Neuro- and Developmental Psychology, Amsterdam Public Health Research Institute, Vrije Universiteit Amsterdam, Amsterdam, Netherlands

[2]Cancer Center Amsterdam, Amsterdam UMC, Amsterdam, Netherlands

[3]Department of Otolaryngology-Head and Neck Surgery, Amsterdam Public Health Research Institute, Amsterdam UMC, Amsterdam, Netherlands

[4]Department of Human Movement Sciences, Amsterdam Movement Sciences, Vrije Universiteit Amsterdam, Amsterdam, Netherlands

[5]Department of Orthopedics and Research, Kliniek ViaSana, Mill, Netherlands

Corresponding Author:
Irma M Verdonck-de Leeuw, PhD
Department of Clinical, Neuro- and Developmental Psychology
Amsterdam Public Health Research Institute
Vrije Universiteit Amsterdam
Van Der Boechorststaat 7
Amsterdam, 1081 BT
Netherlands
Phone: 31 20 444 0931
Fax: 31 20 444 3688
Email: im.verdonck@amsterdamumc.nl

Abstract

Background: The eHealth Impact Questionnaire (eHIQ) provides a standardized method to measure attitudes of electronic health (eHealth) users toward eHealth. It has previously been validated in a population of eHealth users in the United Kingdom and consists of 2 parts and 5 subscales. Part 1 measures attitudes toward eHealth in general and consists of the subscales attitudes towards online health information (5 items) and attitudes towards sharing health experiences online (6 items). Part 2 measures the attitude toward a particular eHealth application and consists of the subscales confidence and identification (9 items), information and presentation (8 items), and understand and motivation (9 items).

Objective: This study aimed to translate and validate the eHIQ in a Dutch population of eHealth users.

Methods: The eHIQ was translated and validated in accordance with the COnsensus-based Standards for the selection of health status Measurement INstruments criteria. The validation comprised 3 study samples, with a total of 1287 participants. Structural validity was assessed using confirmatory factor analyses and exploratory factor analyses (EFAs; all 3 samples). Internal consistency was assessed using hierarchical omega (all 3 samples). Test-retest reliability was assessed after 2 weeks, using 2-way intraclass correlation coefficients (sample 1). Measurement error was assessed by calculating the smallest detectable change (sample 1). Convergent and divergent validity were assessed using correlations with the remaining measures (all 3 samples). A graded response model was fit, and item information curves were plotted to describe the information provided by items across item trait levels (all 3 samples).

Results: The original factor structure showed a bad fit in all 3 study samples. EFAs showed a good fit for a modified factor structure in the first study sample. This factor structure was subsequently tested in samples 2 and 3 and showed acceptable to good fits. Internal consistency, test-retest reliability, convergent validity, and divergent validity were acceptable to good for both the original as the modified factor structure, except for test-retest reliability of one of the original subscales and the 2 derivative subscales in the modified factor structure. The graded response model showed that some items underperformed in both the original and modified factor structure.

Conclusions: The Dutch version of the eHIQ (eHIQ-NL) shows a different factor structure compared with the original English version. Part 1 of the eHIQ-NL consists of 3 subscales: attitudes towards online health information (5 items), comfort with sharing

XSL•FO
RenderX

health experiences online (3 items), and usefulness of sharing health experiences online (3 items). Part 2 of the eHIQ-NL consists of 3 subscales: motivation and confidence to act (10 items), information and presentation (13 items), and identification (3 items).

(J Med Internet Res 2019;21(8):e13408) doi:10.2196/13408

KEYWORDS

eHealth; evaluation; e-Health Impact Questionnaire; psychometrics

Introduction

Currently, patients and care providers are encouraged to use electronic health (eHealth) apps to improve health care, including self-management [1,2]. A standardized measure to evaluate eHealth apps throughout the development process is needed. In the Netherlands, more than 98% of the population has access to the internet [3], and the use of eHealth apps is stimulated by both government and health care organizations. Internationally, the access to the internet is also growing rapidly. A standardized measure to evaluate eHealth apps is therefore much needed. However, evaluating eHealth apps is difficult because of a number of factors, including the difficulty of creating controlled experiments and confounding variables such as proficiency with the internet [4], and the continued development of eHealth app in comparison with more traditional forms of health care. Currently, evaluation of eHealth apps usually consists of 2 components: testing efficacy using randomized controlled trials (RCTs) and in-depth evaluation of the content of the app using structured and unstructured interviews. These methods require a large investment of time and resources. Given the rapid development of technology, this creates a state of *playing catch-up* for eHealth developers. A standardized way of evaluating eHealth apps can be invaluable in the process of constant development and evaluation. Although some such standardized measures exist (eg, the system usability, which measures the usability of software apps), they do not offer similar insight into the user experience through interviews.

In 2013, Kelly et al [5] developed the eHealth Impact Questionnaire (eHIQ) to measure the self-reported impact of eHealth on its users. On the basis of 5 themes, which were identified from interviews, the questionnaire consists of 2 parts. The first part (11 items) measures the overall attitude of eHealth users regarding eHealth, consisting of 2 subscales: *attitudes towards online health information* (5 items) and *attitudes towards sharing health experiences online* (6 items). The second part (26 items) measures the attitude of eHealth users regarding a specific eHealth app, consisting of 3 subscales: *confidence and identification* (9 items), *information and presentation* (8 items), and *understand and motivation* (9 items). This questionnaire was validated in 2015 for the British eHealth users [6].

The goal of this study was to translate and validate the eHIQ in a Dutch population of eHealth users—resulting in the Dutch version called eHIQ-NL—according to the COnsensus-based Standards for the selection of health status Measurement INstruments (COSMIN) criteria [7]. These criteria provide a systematic roadmap for appropriate analyses and interpretation of different types of validity and reliability. To our knowledge, the eHIQ has not been previously translated and/or validated outside of the original development and validation [5,6].

In the first study (the main study), Dutch users of the website Kanker.nl (an eHealth website for Dutch cancer patients) completed both parts of the eHIQ twice. In the second study, the first part of the eHIQ was completed by Dutch cancer survivors who were invited to participate in a survey on supportive cancer care, which was part of a randomized controlled trial to evaluate the efficacy of Oncokompas (an eHealth self-management app that supports Dutch cancer survivors in finding and obtaining optimal supportive care) [8]. In the third study, the second part of the eHIQ was completed by Dutch patients who had undergone orthopedic surgery and were participants in a pilot study of an app providing health information regarding pre- and postoperative care.

Methods

Translation

The questionnaire was translated from English into Dutch by 2 independent translators: 1 eHealth expert and 1 language expert who is a Dutch native and fluent in English. These translations were combined into a single Dutch questionnaire by 2 independent reviewers. In case of discrepancies, the final translation was decided by consensus. The Dutch translation was then translated back into English by 2 independent experts in language who are English natives and fluent in Dutch. The back-translated questionnaire was compared with the original English version by 2 independent reviewers. Discrepancies between the back-translated and the original English questionnaire were discussed, and some final changes were made. An example copy of the final translated questionnaire can be found in Multimedia Appendix 1.

Recruitment and Procedure

Due to the results of the main study (study sample 1), the eHIQ was subsequently presented to 2 other samples of (prospective) eHealth users (study samples 2 and 3).

Study Sample 1

Dutch users of the national website Kanker.nl (an eHealth website for cancer patients) who had signed to participate in scientific research were asked to fill in both parts of the eHIQ-NL twice, with an interval of 2 weeks. On the second measurement, they were also asked to answer 2 questions designed to gauge attitudes to eHealth apps; 1 question asked them to grade their satisfaction with Kanker.nl on a 10-point scale (overall satisfaction), whereas the other question asked how likely they were to recommend Kanker.nl to a fellow cancer patient (the Net Promoter Score, NPS). They were further asked

XSL•FO
RenderX

to fill in the 5-level EuroQol-5D version (EQ-5D-5L), which measures self-reported health-related quality of life [9].

Study Sample 2

A random sample of cancer survivors (breast cancer, colorectal cancer, head and neck cancer, or lymphoma) was drawn from the Netherlands Cancer Registry and invited to complete a survey on supportive cancer care, which was part of an RCT investigating the efficacy of Oncokompas (an eHealth self-management app that supports Dutch cancer survivors in finding and obtaining optimal supportive care) [8]. Patients were excluded who had severe cognitive impairment, insufficient mastery of the Dutch language, physical inability to complete a questionnaire, or received palliative care. Participants with internet access filled in the first part of the eHIQ-NL during the survey on supportive care. They were also asked to fill in the Functional, Communicative and Critical Health Literacy (FCCHL) scales (Cronbach α was .94 in the current sample), which measures the capacity of individuals to access, understand, and use health information [10], and the European Organisation for Research and Treatment of Cancer core quality of life questionnaire, version 3.0 Cronbach α was .98 in the current sample), which measures cancer-related quality of life [11]. Medical ethical approval was provided by the Medical Ethics Review Board of the VU Medical Center in Amsterdam, the Netherlands (reference number 2015.523).

Study Sample 3

Patients were recruited from a single clinic (ViaSana, Mill, The Netherlands) to participate in a pilot study of an app providing health information regarding pre- and postoperative care. Patients were eligible when aged older than 18 years and had undergone orthopedic surgery. Patients were excluded if they were not accessible by e-mail. Participants filled in the second part of the eHIQ-NL up to 2 weeks after using the app. Participants also filled in the System Usability Scale (SUS; Cronbach α was .90 in the current sample), which measures the usability of software apps [12], and 2 questions designed to gauge attitudes to eHealth apps, 1 question asked them to grade the app on a 10-point scale, whereas the other question asked how likely they were to recommend the app to a fellow patient. Medical ethical approval was provided by Medical Ethics Review Board of the Elisabeth Hospital in Tilburg, the Netherlands (reference number METC-T2012-11).

Statistical Analysis

All analyses were performed in R version 3.3.3 [13]. Measurement properties were assessed in accordance with the COSMIN criteria [7].

Study Sample 1

First, structural validity was assessed with a combination of confirmatory factor analyses (CFAs) and exploratory factor analyses (EFAs). All CFAs were run using the cfa function of the lavaan package version 0.6-3 [14], whereas all EFAs were run using the efaUnrotate function of the semTools package version 0.5-1 [15], and Oblimin rotation was applied using the oblqRotate function of the semTools package version 0.5-1 [15].

Second, internal consistency was assessed by calculating hierarchical omega [16] using the reliability function of the semTools package version 0.5-1 [15]. Third, test-retest reliability was assessed by calculating a 2-way intraclass correlation coefficient (ICCs) between the 2 measurement times, using the icc function of the irr package version 0.84.1 [17]. Fourth, measurement error was assessed by calculating the standard error of measurement using the SE.Meas function of the psychometric package version 2.2 [18]. The smallest detectable change (SDC) was calculated by hand using the standard error of measurement.

Fifth, convergent validity and divergent validity were tested by correlating the subscales of the eHIQ-NL with the questions concerning satisfaction with Kanker.nl and the NPS (a positive correlation was hypothesized) and the EQ-5D-5L of which the items for *daily activities* and *anxiety or depression* were assumed to show a positive correlation. No correlation was hypothesized to exist between the eHIQ-NL and the remaining EQ-5D-5L items. Correlations were calculated using the rcorr function of the Hmisc package [19].

Sixth and last, a graded response model was fit using the grm function of the ltm package [20]. Item information curves were plotted for each subscale to describe the information provided by items across the item trait level (ie, the construct measured by the subscale).

Study Sample 2

Structural validity was assessed with a combination of CFAs and EFAs. Internal consistency was assessed with hierarchical omega. Divergent validity was tested by correlating the subscales of the eHIQ-NL with the FCCHL, as no correlation was hypothesized to exist. Finally, a graded response model was fit. All analyses were performed using the same functions and R packages as in study sample 1.

Study Sample 3

Structural validity was assessed with a combination of CFAs and EFAs. Internal consistency was assessed with hierarchical omega. Convergent validity was tested by correlating the subscales of the eHIQ-NL with the SUS and the questions concerning the grade and likelihood of recommending the app, as positive correlations were hypothesized. Finally, a graded response model was fit. All analyses were performed using the same functions and R packages as in study sample 1.

Results

Study Population

Table 1 shows the demographic and clinical characteristics of the 3 study samples. In study sample 1, 304 cancer survivors participated with a mean age of 58.12 years (SD=11.26), and 177 were female (58.2%, 177/304). The study sample consisted of more than 17 cancer diagnoses; most were diagnosed with breast cancer (27.1%, 82/340) or prostate cancer (13.8%, 42/340). The feasibility of the eHIQ-NL was good: of the 304 participants who started the first measurement, 288 (94.7%, 288/304) completed the eHIQ-NL. A total of 242 (79.6%,

XSL•FO
RenderX

242/304) participants started the second measurement, of which 217 (71.4%, 217/304) completed all questionnaires.

In study sample 2, 566 cancer survivors completed the first part of the eHIQ-NL with a mean age of 64.18 years (SD=10.65), and 351 (62.1%, 351/565) were female. The study sample consisted of 4 cancer diagnoses: breast cancer (39.2%, 222/566), colorectal cancer (29.7%, 168/566), head and neck cancer (19.1%, 108/566), and lymphoma (12.0%, 68/566).

In study sample 3526 orthopedic patients completed the second part of the eHIQ-NL with a median age of 59.00 years (interquartile range=50-66), and 267 were female (50.7%, 267/526). The study sample consisted of patients who underwent various orthopedic surgeries; the main group had undergone a total knee arthroplasty (31.1%, 164/526).

Table 1. Descriptive statistics of the study population.

Study sample, characteristic	Study sample 1 (N=288)	Study sample 2 (N=566)	Study sample 3 (N=526)
Age, mean (SD)	58.12 (11.26)	64.18 (10.65)	59.00 (50-66)[a]
Gender, n (%)			
Male	126	214	259
Female	177	351	267
Diagnosis, n (%)			
Breast cancer	82 (26.9)	222 (39.2)	__[b]
Miscellaneous cancer	47 (15.4)	—	—
Prostate cancer	42 (13.8)	—	—
Lymphoma	20 (6.5)	68 (12.0)	—
Colon cancer	18 (5.9)	—	—
Skin cancer	18 (5.9)	—	—
Lung cancer	17 (5.5)	—	—
Bladder and kidney cancer	17 (5.5)	—	—
Rectal cancer	14 (4.6)	—	—
Head and neck cancer	12 (3.9)	108 (19.1)	—
Esophageal cancer	10 (3.2)	—	—
Leukemia	10 (3.2)	—	—
Other	33 (8.8)	—	—
Colorectal cancer	—	168 (29.7)	—
Total knee arthroplasty	—	—	164 (31.2)
Total hip arthroplasty	—	—	89 (16.9)
Anterior cruciate ligament reconstruction	—	—	56 (10.6)
Knee arthroscopy	—	—	47 (8.9)
Cuff repair	—	—	30 (5.7)
High tibial osteotomy	—	—	23 (4.4)
Lumbar discectomy	—	—	17 (3.2)
Acromionplasty	—	—	14 (2.7)
Remaining group	—	—	86 (16.3)

[a]Median (interquartile range).

[b]Not assessed in this study.

[c]Remaining group: shoulder arthroplasty, femoral osteotomy, patella stabilization (medial patellofemeral ligament), mortons neurom, hallux valgus/rigidus, exostosis, and talocrual arthrodesis.

XSL•FO
RenderX

Study Sample 1

Structural Validity

A CFA was run on a 2-level hierarchical model, with the specified subscales as first-order factors, and the 2 different sections (general attitude and specific attitude) as second-order factors. This model had a bad fit (minimum discrepancy per degree of freedom [CMIN]=2.61, adjusted goodness-of-fit index [AGFI]=0.719, Comparative Fit Index [CFI]=0.752, Tucker-Lewis index [TLI]=0.753, standardized root mean square residual [SRMR]=0.076, and root mean square error of approximation [RMSEA]=0.075 [0.070-0.079]). Inspecting the modification indices revealed cross-loadings of items on the second-order factors. Such cross-loadings made sense when looking at the content of the items (eg, items on information on the specific eHealth tool showing cross-loadings with general attitude toward health information); however, shifting items from one section to another made no theoretical or practical sense. Therefore, 2 CFAs were run separately for each section, removing the second-order factor from the analysis.

The fit for the first part of the questionnaire was better than the first model fit, but not yet acceptable (CMIN=5.14, AGFI=0.796, CFI=0.847, TLI=0.804, SRMR=0.074, and RMSEA=0.118 [0.103-0.134]). A 3-factor EFA using Oblimin rotation was run to investigate an alternative to the original factor structure. This model showed a good fit (CMIN=3.16, AGFI=0.989, CFI=0.954, TLI=0.898, SRMR=0.032, and RMSEA=0.085 [0.065-0.107]). The 3 factors were interpretable (Multimedia Appendix 2), with the subscale *attitudes towards sharing health experiences online* being split into the 2 factors *comfort with sharing health experiences online* and *usefulness of sharing health experiences online*. The third factor was identical to the original factor of *attitudes towards online health information*.

The fit for the second part of the questionnaire was also better than the first model fit, but not yet acceptable (CMIN=3.20, AGFI=0.747, CFI=0.755, TLI=0.731, SRMR=0.082, and RMSEA=0.087 [0.081-0.094]). A 4-factor EFA using Oblimin rotation was run to investigate an alternative to the original factor structure. The model showed a good fit (CMIN=2.01, AGFI=0.988, CFI=0.914, TLI=0.876, SRMR=0.037, and RMSEA=0.059 [0.051-0.067]), but the factor structure was not clearly interpretable, many items had double loadings, and the fourth factor had very low factor loadings. A 5-factor EFA using Oblimin rotation showed a similar fit (CMIN=1.93, AGFI=0.988, CFI=0.928, TLI=0.886, SRMR=0.033, and RMSEA=0.057 [0.048-0.065]). Although the double loadings were mostly taken care of, the loadings on the fourth and fifth factors were very low.

A 3-factor EFA using Oblimin rotation was run to investigate problematic items. Items 10, 8, 16, 4, 17, and 11 showed double loadings and no clear distinction to any one factor. Removing these items and performing a CFA on the original factor structure resulted in a bad fit (CMIN=3.39, AGFI=0.779, CFI=0.786, TLI=0.757, SRMR=0.084, and RMSEA=0.091 [0.083-0.099]). Running an EFA using Oblimin rotation on the same subset of items resulted in a good fit (CMIN=2.22, AGFI=0.990, CFI=0.920, TLI=0.886, SRMR=0.041, and RMSEA=0.062 [0.052-0.072]), but with a different factor structure than theorized (Multimedia Appendix 2): the first factor being a combination of items from the subscales *confidence and identification* and *Understanding and motivation* and interpretable as *motivation and confidence to act*; the second factor being identical to the original subscale *information and presentation* with the addition of item 2; and the third factor consisting of 3 items from the subscale *confidence and identification* and interpretable as *identification*.

Internal Consistency

Multimedia Appendix 3 shows the results on internal consistency of the original factor structure and the modified factor structure, respectively. All values were acceptable (omega>0.70), and the values of the original first part and the modified first part were comparable. The values of the modified second part were better than of the original second part.

Test-Retest Reliability

Table 2 shows the results on test-retest reliability of the original factor structure and the modified factor structure. All original subscales, except for *attitudes towards sharing health experiences online* (ICC=0.63) showed acceptable ICCs (ICC>0.70). All modified subscales, except for *comfort with sharing health experiences online* (ICC=0.62) and *usefulness of sharing health experiences online* (ICC=0.53) showed acceptable ICCs (ICC>0.70). The ICCs for the original factor structure and the modified factor structure were comparable.

XSL•FO
RenderX

Table 2. Test-retest reliability.

Structure, subscale	ICC[a]	CI
Original factor structure		
Attitudes towards online health information	0.71	0.64-0.77
Attitudes towards sharing health experiences online	0.63	0.54-0.7
Confidence and identification	0.73	0.66-0.78
Information and presentation	0.72	0.64-0.78
Understanding and motivation	0.74	0.67-0.8
Modified factor structure		
Attitudes towards online health information	0.71	0.64-0.77
Comfort with sharing health experiences online	0.62	0.53-0.69
Usefulness of sharing health experiences online	0.53	0.43-0.62
Motivation and confidence to act	0.76	0.7-0.81
Information and presentation	0.73	0.66-0.79

[a]ICC: intraclass correation coefficient.

Measurement Error

Table 3 shows the results of the measurement error of the original factor structure and the modified factor structure. For the original factor structure, the SDC ranged between 15.77 and 26.18, which represents a measurement error of 15%-26% of the 100 subscale range. Consequently, we can be 95% certain that a change score larger than 15% to 26% of the subscale range is not an artifact of measurement error. For the modified factor structure, the SDC ranged between 15.05 and 34.81, which represents a measurement error of 15% to 35% of the 100 subscale range. The highest SDCs were reported for the Part 1 *attitudes towards sharing health experiences online* (34.81) and *comfort with sharing health experiences online* (28.91) subscales. This makes sense, as both subscales only consisted of 3 items, and small scales are susceptible to high measurement error.

Convergent and Divergent Validity

All subscales correlated significantly with both the overall satisfaction and the NPS. The correlations between the subscales of the first part of the eHIQ-NL and the overall satisfaction and the NPS were small (r<0.30). There were either no significant or very small (r<0.20) correlations with the EQ-5D questions on daily activities and anxiety and depression. The 3 remaining EQ-5D items did not correlate significantly with any of the eHIQ-NL subscales (Table 4)

Graded Response Model

Figure 1 shows the item information curves for the original subscales. A number of items of part 1 did not provide much extra information to the subscale: items 1, 2, 8, 9, and 11. Notably, most items in the subscale attitudes *towards sharing health experiences online* provided information at the same item trait levels. A number of items of Part 2 also did not provide much extra information to the subscale: items 2, 10, 11, 13, 16, 23, and 25. Notably, items 10, 11, and 16 were items that fit poorly in the factor analysis.

Figure 2 shows the item information curves of the modified subscales. Of part 1, the information of the subscale *comfort with sharing health experiences online* was rather low across the entire latent trait spectrum. For the subscale *usefulness of sharing health experiences online* information was high on certain points of the latent trait spectrum, but all 3 items overlap almost completely. Of part 2, the subscale *motivation and confidence to act* showed a good range of information across latent trait levels. However, 3 items hardly contributed information (items 1, 7, and 13). The subscale *information and presentation* still suffered from multiple items adding little information, as well as a lot of overlap. Finally, the subscale *identification* showed a good range of information as well as high peaks for all 3 items, but still a lot of overlap between items on information range.

XSL•FO
RenderX

Table 3. Measurement error.

Structure, subscale Original factor structure	SEM[a]	SDC[b]
Original factor structure		
Attitudes towards online health information	9.14	25.32
Attitudes towards sharing health experiences online	9.44	26.18
Confidence and identification	6.79	18.83
Information and presentation	5.69	15.77
Understanding and motivation	6.33	17.54
Modified factor structure		
Attitudes towards online health information	9.14	25.32
Comfort with sharing health experiences online	12.56	34.81
Usefulness of sharing health experiences online	10.43	28.90
Motivation and confidence to act	7.19	19.93
Information and presentation	5.43	15.05

[a]SEM: standard error of measurement.

[b]SDC: smallest detectable change.

XSL•FO
RenderX

Table 4. Convergent and divergent validity.

Study sample	Original factor structure					Modified factor structure					
	OHI[a]	SHEO[b]	C&I[c]	I&P[d]	U&M[e]	OHI	CSHEO[f]	USHEO[g]	M&CA[h]	I&P	ID[i]
A: Study sample 1—convergent											
Grade	0.17[j]	0.28[k]	0.50[k]	0.47[k]	0.49[k]	0.17[j]	0.23[k]	0.27[k]	0.46[k]	0.49[k]	0.35[k]
NPS[l]	0.21[m]	0.25[k]	0.48[k]	0.44[k]	0.44[k]	0.21[m]	0.19[m]	0.26[k]	0.43[k]	0.45[k]	0.37[k]
EQ-5D[n]—daily activities	−0.01	0.07	−0.05	0.00	−0.05	−0.01	0.08	0.04	−0.05	0.00	0.01
EQ-5D—anxiety/depression	−0.07	0.00	−0.13	−0.18[m]	−0.14[j]	−0.07	0.06	−0.07	−0.15[j]	−0.18[m]	−0.02
B: Study sample 1—divergent											
EQ-5D—mobility	−0.04	0.02	−0.04	−0.05	−0.02	−0.04	0.01	0.02	−0.01	−0.03	0.01
EQ-5D—selfcare	0.02	0.07	−0.04	−0.03	0.03	0.02	0.06	0.07	0.01	−0.02	−0.04
EQ-5D—pain	−0.04	0.00	−0.08	−0.06	−0.08	−0.04	0.03	−0.04	−0.09	−0.06	−0.03
C: Study sample 2—divergent											
FCCHL[o]	0.14[m]	0.03	—[p]	—	—	0.14[m]	−0.01	0.06	—	—	—
EORTC QLQ-C30[q]	0.01	0.01	—	—	—	0.01	0.00	0.02	—	—	—
D: Study sample 3—convergent											
System Usability Scale	—	—	0.29[k]	0.53[k]	0.39[k]	—	—	—	0.30[k]	0.55[k]	0.12[m]
NPS	—	—	0.46[k]	0.42[k]	0.52[k]	—	—	—	0.48[k]	0.45[k]	0.31[k]
Grade[r]			0.59[k]	0.53[k]	0.62[k]				0.58[k]	0.55[k]	0.43[k]

[a]OHI: attitudes towards online health information.

[b]SHEO: attitudes towards sharing health experiences online.

[c]C&I: confidence and identification.

[d]I&P: information and presentation.

[e]U&M: understanding and motivation.

[f]CSHEO: comfort with sharing health experiences online.

[g]USHEO: usefulness of sharing health experiences online.

[h]M&CA: motivation and confidence to act.

[i]ID: identification.

[j]$P<.05$.

[k]$P<.01$.

[l]NPS: Net Promoter Score.

[m]$P<.001$.

[n]EQ-5D: EuroQol-5D.

[o]FCCHL: Functional, Communicative and Critical Health Literacy scale.

[p]Not applicable.

[q]EORTC QLQ-C30: European Organisation for Research and Treatment of Cancer core quality of life questionnaire, version 3.0.

[r]Grade: overall satisfaction.

XSL•FO
RenderX

Figure 1. Study sample 1: item information curves (IICs) of original subscales.

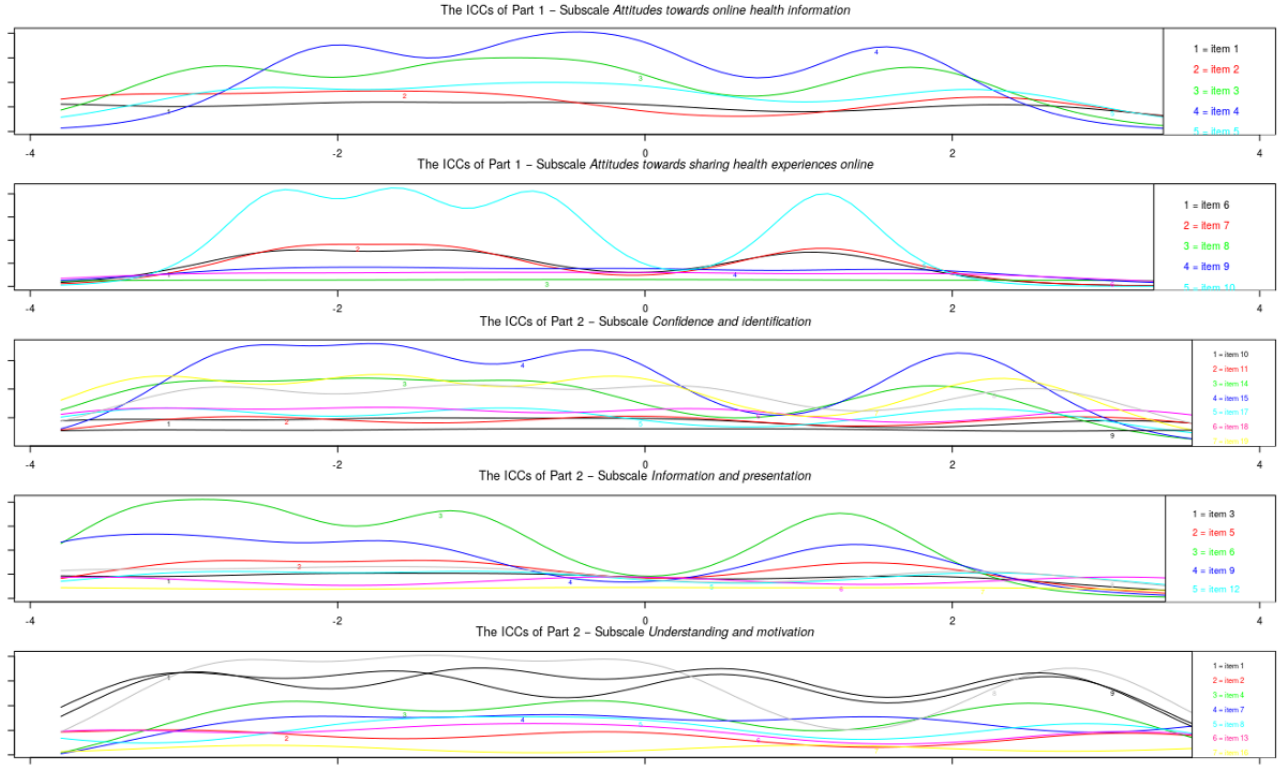
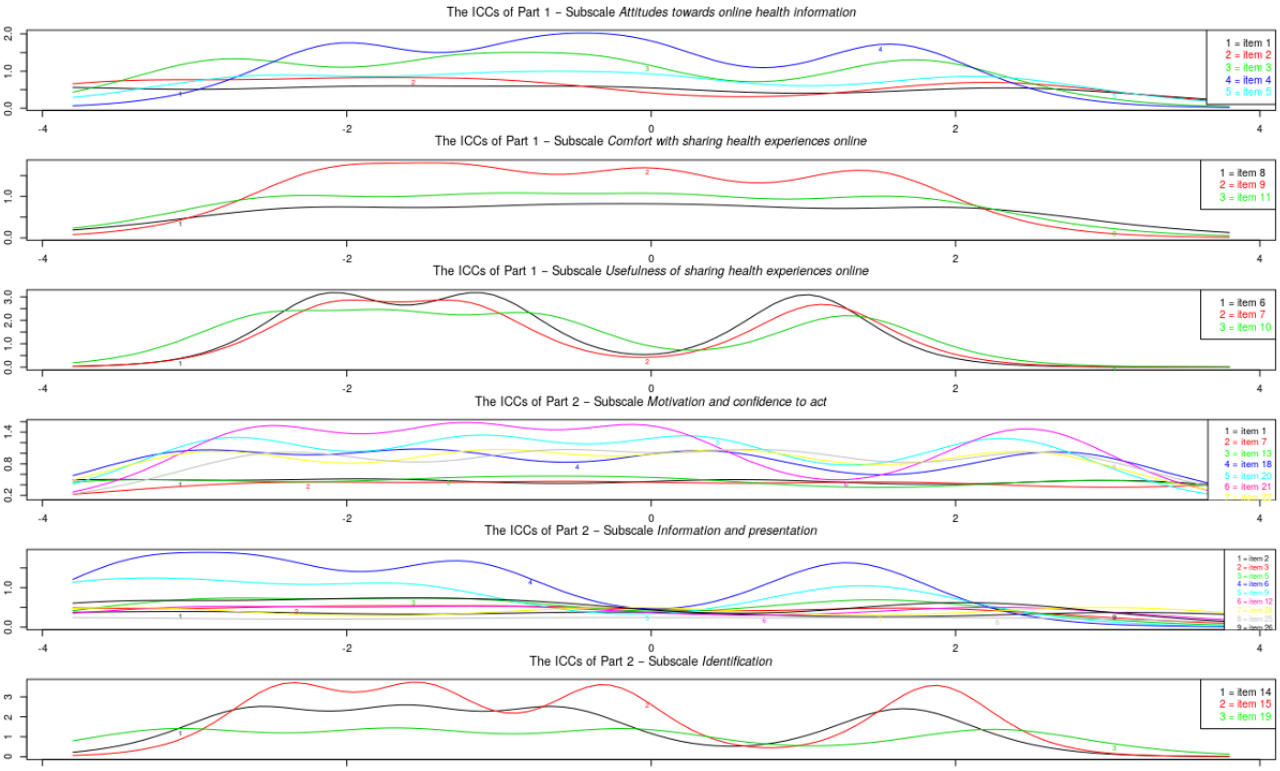


Figure 2. Study sample 1: item information curves (IICs) of the modified subscales.



Study Sample 2

Structural Validity

A CFA was run with the 2 original subscales as first-order factors. This model had a bad fit (CMIN=8.13, AGFI=0.829, CFI=0.893, TLI=0.863, SRMR=0.054, and RMSEA=0.113 [0.102-0.124]). A second CFA was run with the factor structure found in study 1. This model had a barely acceptable fit (CMIN=7.37, AGFI=0.849, CFI=0.909, TLI=0.878, SRMR=0.049, and RMSEA=0.107 [0.096-0.118]). An EFA using Oblimin rotation was run with 3 factors to determine possible deviations from the 3 subscales found in study 1. This model had a good fit (CMIN=4.86, AGFI=0.978, CFI=0.966,

XSL•FO
RenderX

TLI=0.926, SRMR=0.025, and RMSEA=0.084 [0.069-0.098]). The 3 factors (Multimedia Appendix 2) were identical to the subscales found in study 1, except for item 8 loading on both subscales concerning the sharing of health experiences online.

Internal Consistency

Multimedia Appendix 3 shows the internal consistency of the original factor structure and the modified factor structure. All values were acceptable (omega>0.70) and comparable between both factor structures.

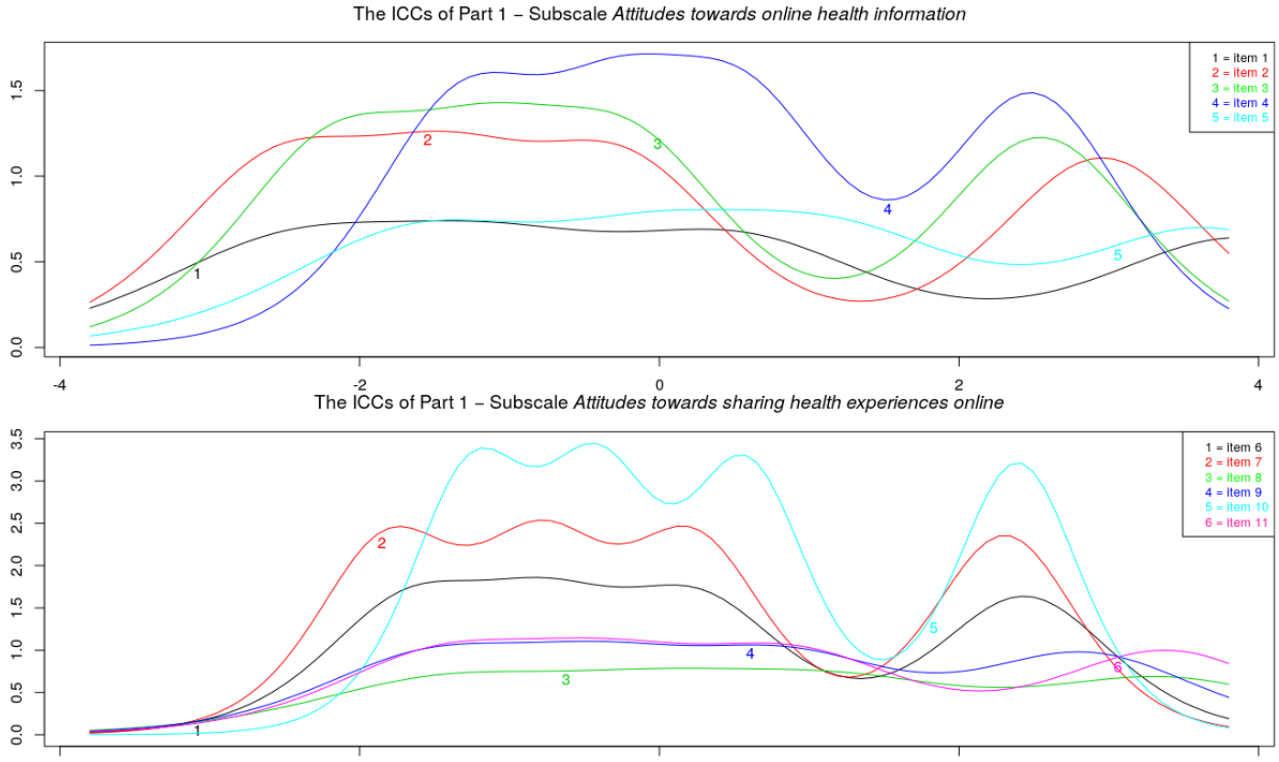
Divergent Validity

For both the original and the modified factor structure, only the subscale *attitudes toward online health information* showed a significant correlation with the FCCHL (Table 4). However, this correlation is small enough to be acceptable for divergent validity (r<0.15).

Graded Response Model

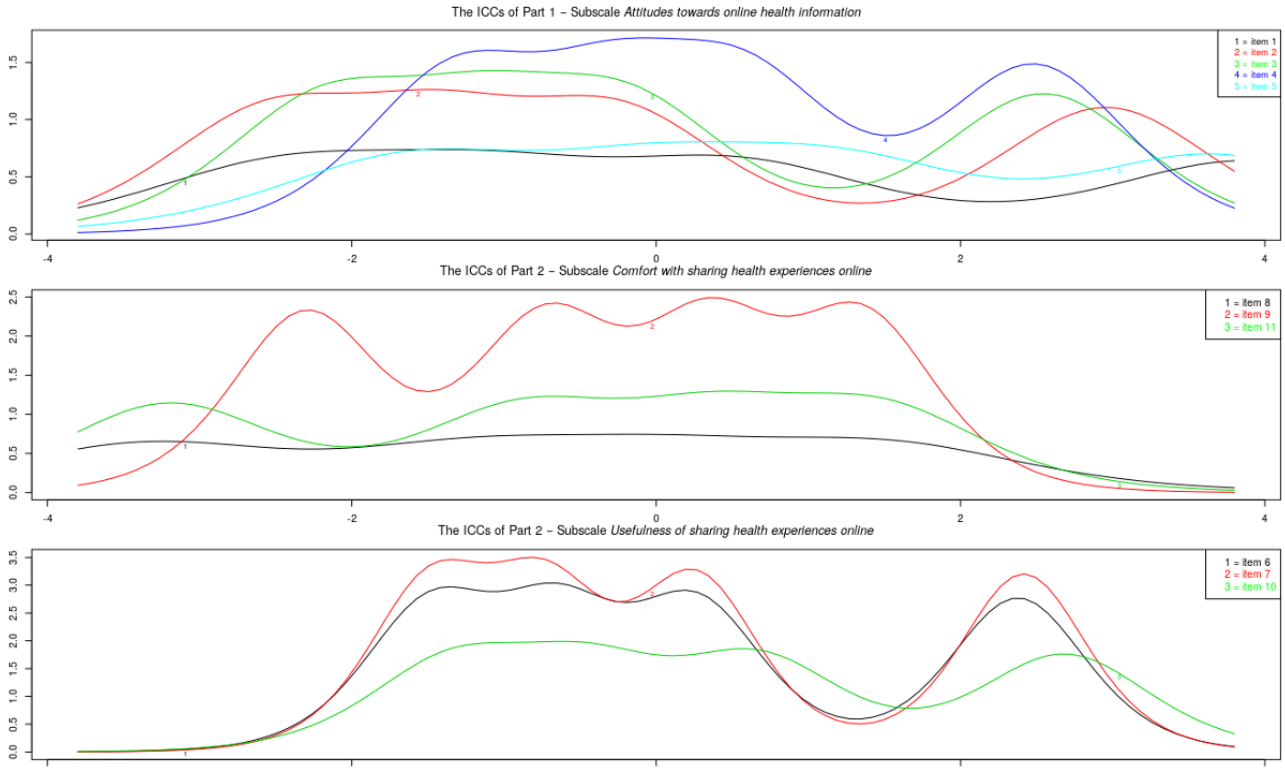
Figure 3 shows the item information curves for the original subscales of part 1. A number of items do not provide much extra information over the others: items 5, 8, and 9. Figure 4 shows the item information curves of the modified subscales of part 1. The information of the subscale *comfort with sharing health experiences online* showed large dips on certain levels of ability. For the subscale *usefulness of sharing health experiences online* information was high on certain points of the latent trait spectrum, but the items overlap a great deal.

Figure 3. Study sample 2: item information curves (IICs) of original subscales.



XSL•FO
RenderX

Figure 4. Study sample 2: item information curves (IICs) of modified subscales.



Study Sample 3

Structural Validity

A CFA was run with the 3 original subscales as first-order factors. This model had a slightly below acceptable fit (CMIN=5.568, AGFI=0.717, CFI=0.811, TLI=0.792, SRMR=0.092, and RMSEA=0.093 [0.089-0.098]). A second CFA was run with the 3 subscales found in study sample 1. This model had an acceptable fit (CMIN=4.447, AGFI=0.828, CFI=0.889, TLI=0.873, SRMR=0.075, and RMSEA=0.081 [0.075-0.087]). An EFA using Oblimin rotation was run with 3 factors and including the items that were deemed problematic in study sample 1 to determine whether including them would result in a better fit. This model had a good fit (CMIN=2.496, AGFI=0.990, CFI=0.948, TLI=0.932, SRMR=0.029, and RMSEA=0.053 [0.048-0.059]).

In the EFA, items 8, 11, and 17 showed no problematic cross-loadings. Items 4, 10, and 16 did show problematic cross-loadings, but not as extreme as in study sample 1 (Multimedia Appendix 2). Items 8 and 10 were found to load most highly on the factor representing *motivation and confidence to act*. Items 4, 11, 16, and 17 were found to load most highly on the factor representing *information and presentation*. Beyond the problematic items, only 1 item loaded differently than in study sample 1: item 20 loaded as highly on the factor representing *motivation and confidence to act* (on which it loaded in study sample 1) as it did on the factor representing *identification*.

Internal Consistency

Multimedia Appendix 3 shows the internal consistency of the original factor structure, the modified factor structure without previously problematic items, and the modified factor structure with previously problematic items, respectively. The internal consistency of the modified factor structure with previously problematic items is represented by Cronbach alpha instead of Omega, as Omega is based on factor variance and unsuitable for factor structures fit based on EFAs. All values, except for the original subscale *information and presentation* (omega=0.65), were acceptable and comparable between the 3-factor structures.

Convergent Validity

Both the original and modified subscales correlated significantly with the SUS, NPS, and *grade* questions (Table 4). All correlations were acceptable for convergent validity (r>0.30), except for the original subscale *confidence and identification* with the SUS (r=0.29) and the modified subscale *identification* with the SUS (r=0.12).

Graded Response Model

Figure 5 shows the item information curves for the original subscales of part 2. With a large number of items per scale, there was a good range of information across latent trait levels. Some items did not add much to the information provided by other items: items 3, 8, 10, 11, 17, 21, 24, 25, and 26. Notably, items 8, 10, 11, and 17 were items that were judged problematic in study 1. Figure 6 shows the item information curves of the modified subscales of part 2. The subscale *motivation and confidence to act* showed a good range of information across latent trait levels. However, 3 items hardly contributed information: items 1, 7, and 13. The subscale *information and presentation* still suffered from multiple items adding little information, as well as a lot of overlap. Finally, the subscale

XSL•FO
RenderX

identification showed a good range of information as well as high peaks for all 3 items, but still a lot of overlap.

Figure 5. Study sample 3: item information curves (IICs) of original subscales.

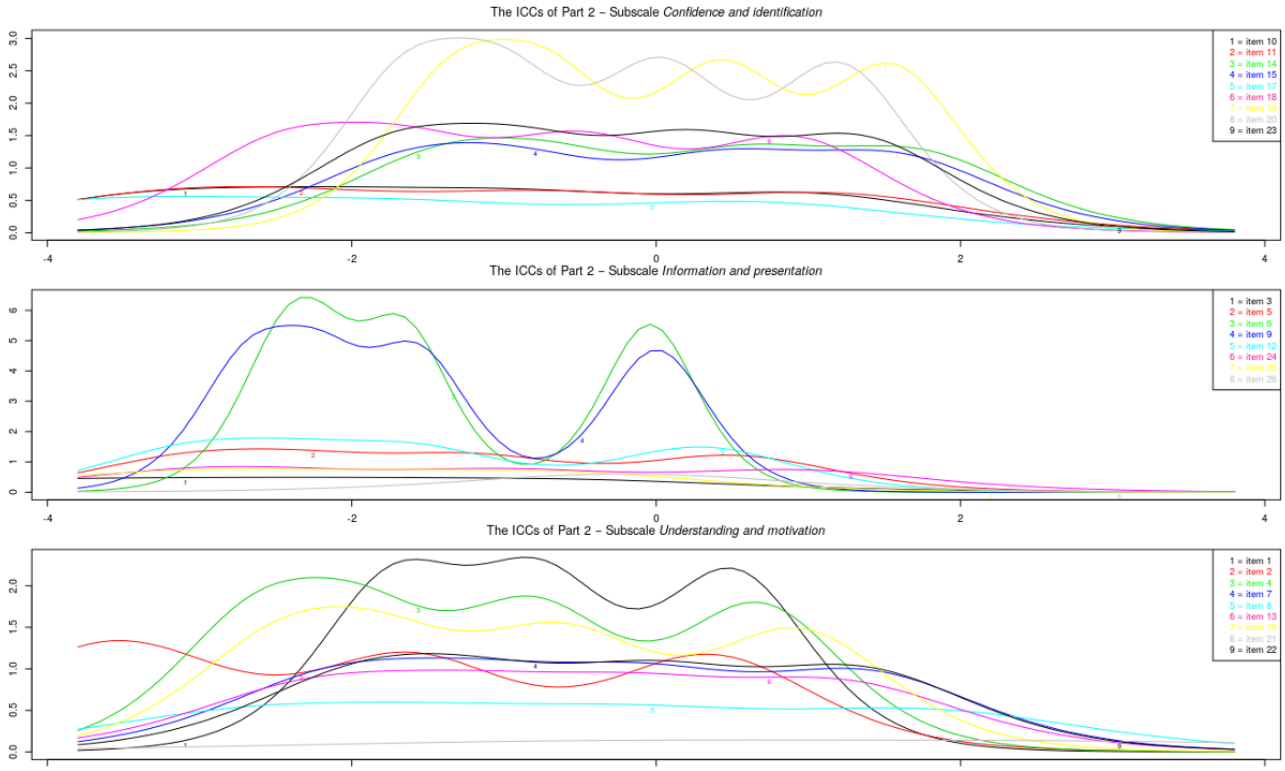
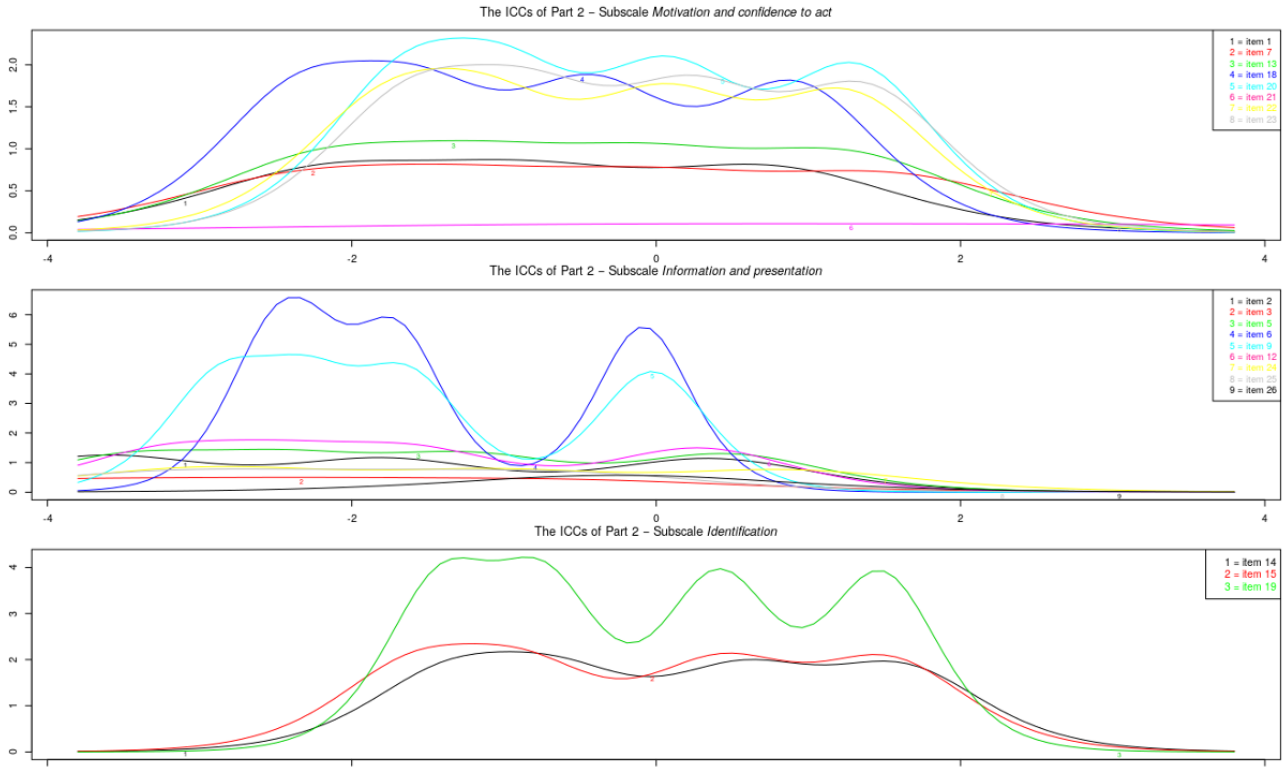


Figure 6. Study sample 3: item information curves (IICs) of modified subscales.



XSL•FO
RenderX

Discussion

Principal Findings

In this study, the eHIQ was translated into Dutch, and the measurement properties were investigated. Feasibility was good: more than 94% of participants in the main study completed the eHIQ-NL. The eHIQ-NL showed a different factor structure compared with the original English version. Part 1 of the eHIQ-NL consists of 3 subscales: *attitudes towards online health information* (5 items), *comfort with sharing health experiences online* (3 items), and *usefulness of sharing health experiences online* (3 items). Part 2 of the eHIQ-NL consists of 3 subscales: *motivation and confidence to act* (10 items), *information and presentation* (13 items), and *identification* (3 items). These factor structures were replicated in subsequent samples and altogether showed acceptable to good internal consistency, test-retest reliability, and construct validity.

Limitations

Limitations of this study are some underperforming measurement properties of the modified factor structure. In particular, test-retest reliability for *comfort with sharing health experiences online* and *usefulness of sharing health experiences online* (ICC=0.62 and ICC=0.53, respectively) was below acceptable threshold. Notably, the original subscale comprised these 2 subscales *attitudes towards sharing health experiences online* also underperformed on test-retest reliability (ICC=0.63).

Furthermore, the correlations testing convergent validity were small in the main study (r<0.30), as well as some smaller correlations in study sample 3 for the subscales *confidence and identification* (r=0.29), and the modified subscale *identification* (r=0.12). We recognize that this may be because of subpar a priori hypotheses in regard to the EQ-5D (study 1) and the SUS (study 3). The reasoning for these hypotheses was somewhat tenuous. For the first sample, we expected the specific eHealth app Kanker.nl to provide useful information for patients with issues regarding daily activities and anxiety/depression resulting in a correlation between a higher score on these issues and eHIQ scores. For the third sample, we expected a higher usability score to be correlated to higher eHIQ scores, but we recognize that the subscales *confidence and identification* and *identification* may be theoretically unrelated to usability. Further research is necessary to further investigate test-retest reliability and construct validity of the eHIQ-NL. Future validations in different nationalities and different patient populations may shed more light on these measurement properties.

Comparison With Prior Work

The findings of this study do not entirely match the findings of the original validation of the eHIQ for the British population [6]. The differences may be the results of a number of differences between the current and previous validation studies. The first explanation is that in the translation of the questionnaire, the meaning of some items may have changed. Although we followed a strict protocol for the translation, this explanation cannot be ruled out.

The second explanation can be found in the use of a different study populations. The original validation study presented the eHIQ to a range of health groups, who were not necessarily eHealth users at the time of the study. The participants in the original validation study were invited to the laboratory and were briefly (at least 15 min) acquainted with an eHealth app relevant to their personal health situation [6]. This study presented the eHIQ-NL to eHealth users who were familiar with the app under investigation (study samples 1 and 3) and noncurrent eHealth users. Furthermore, the current validation study presented the eHIQ-NL only to cancer patients (study samples 1 and 2) and patients with musculoskeletal disorders (study sample 3). As such, the populations differ quite a bit beyond nationality.

We realize that the results of this study present complexities to which subscales should be adhered to in case the user of the eHIQ-NL aims to compare their results with international samples. In such cases, we recommend that, besides the results using the subscales as presented in this study, the results using the original subscales are also reported. The caveat is that one cannot be sure of the structural validity using this method, and we recommend factor analysis to back up any such interpretation.

Conclusions

Nevertheless the limitations specified above, the eHIQ-NL shows a consistent factor structure, sufficient internal consistency, and mostly sufficient test-retest reliability and construct validity. The eHIQ-NL is a valid and reliable tool for measuring attitudes of eHealth users. Interested users can contact Oxford Innovations (healthoutcomes@innovation.ox.ac.uk) for a license to use the eHIQ.

Acknowledgments

This research was funded through the Citrien fund, which was funded by ZonMw, grant number: ZonMw 80-83920-98-201.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Example copy of Dutch version of the eHealth Impact Questionnaire.
[PDF File (Adobe PDF File), 153KB - jmir_v21i8e13408_app1.pdf]

XSL•FO
RenderX

Multimedia Appendix 2
Structural validity: exploratory factor analysis factor loadings.
[PDF File (Adobe PDF File), 95KB - jmir_v21i8e13408_app2.pdf]

Multimedia Appendix 3
Internal consistency.
[PDF File (Adobe PDF File), 78KB - jmir_v21i8e13408_app3.pdf]

References

1. Elbert NJ, van Os-Medendorp H, van Renselaar W, Ekeland AG, Hakkaart-van Roijen L, Raat H, et al. Effectiveness and cost-effectiveness of ehealth interventions in somatic diseases: a systematic review of systematic reviews and meta-analyses. J Med Internet Res 2014 Apr 16;16(4):e110 [FREE Full text] [doi: 10.2196/jmir.2790] [Medline: 24739471]
2. Slev VN, Mistiaen P, Pasman HR, Verdonck-de Leeuw IM, van Uden-Kraan CF, Francke AL. Effects of ehealth for patients and informal caregivers confronted with cancer: a meta-review. Int J Med Inform 2016 Mar;87:54-67. [doi: 10.1016/j.ijmedinf.2015.12.013] [Medline: 26806712]
3. European Commission. 2019. Level of Internet Access - Households URL: https://ec.europa.eu/eurostat/web/products-datasets/-/tin00134 [accessed 2019-08-08]
4. Greenhalgh T, Russell J. Why do evaluations of ehealth programs fail? An alternative set of guiding principles. PLoS Med 2010 Nov 2;7(11):e1000360 [FREE Full text] [doi: 10.1371/journal.pmed.1000360] [Medline: 21072245]
5. Kelly L, Jenkinson C, Ziebland S. Measuring the effects of online health information for patients: item generation for an e-health impact questionnaire. Patient Educ Couns 2013 Dec;93(3):433-438 [FREE Full text] [doi: 10.1016/j.pec.2013.03.012] [Medline: 23598293]
6. Kelly L, Ziebland S, Jenkinson C. Measuring the effects of online health information: scale validation for the e-health impact questionnaire. Patient Educ Couns 2015 Nov;98(11):1418-1424 [FREE Full text] [doi: 10.1016/j.pec.2015.06.008] [Medline: 26162953]
7. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. COSMIN - Improving the selection of outcome measurement instruments. 2012. COSMIN Checklist Manual URL: http://www.cosmin.nl/COSMIN%20checklist.html8 [accessed 2019-08-08]
8. van der Hout A, van Uden-Kraan CF, Witte BI, Coupé VM, Jansen F, Leemans CR, et al. Efficacy, cost-utility and reach of an eHealth self-management application 'Oncokompas' that helps cancer survivors to obtain optimal supportive care: study protocol for a randomised controlled trial. Trials 2017 May 22;18(1):228 [FREE Full text] [doi: 10.1186/s13063-017-1952-1] [Medline: 28532439]
9. Herdman M, Gudex C, Lloyd A, Janssen MF, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res 2011 Dec;20(10):1727-1736 [FREE Full text] [doi: 10.1007/s11136-011-9903-x] [Medline: 21479777]
10. 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]
11. Bjordal K, de Graeff A, Fayers PM, Hammerlid E, van Pottelsberghe C, Curran D, et al. A 12 country field study of the EORTC QLQ-C30 (version 3.0) and the head and neck cancer specific module (EORTC QLQ-H&N35) in head and neck patients. EORTC Quality of Life Group. Eur J Cancer 2000 Sep;36(14):1796-1807. [doi: 10.1016/s0959-8049(00)00186-6] [Medline: 10974628]
12. Brooke J. SUS: A 'quick and dirty' usability scale. In: Jordan PW, Thomas B, McClelland IL, Weerdmeester B, editors. Usability Evaluation In Industry. London, UK: Taylor and Francis; 1996:189-194.
13. R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2017. URL: https://www.R-project.org/
14. Rosseel Y. Lavaan: an R package for structural equation modeling. J Stat Softw 2012;48(2):1-36. [doi: 10.18637/jss.v048.i02]
15. Jorgensen TD, Pornprasertmanit S, Schoemann AM, Rosseel Y. The Comprehensive R Archive Network. 2018. SemTools: Useful Tools for Structural Equation Modeling URL: https://cran.r-project.org/web/packages/semTools/index.html [accessed 2019-08-08]
16. McDonald RP. Test Theory: A Unified Treatment. New Jersey, United States: Lawrence Erlbaum Associates; 1999.
17. Gamer M, Lemon J, Singh IF. The Comprehensive R Archive Network. 2019. Irr: Various Coefficients of Interrater Reliability and Agreement URL: https://cran.r-project.org/web/packages/irr/index.html [accessed 2019-08-08]
18. Fletcher TD. The Comprehensive R Archive Network. 2010. Psychometric: Applied Psychometric Theory URL: https://cran.r-project.org/web/packages/psychometric/index.html [accessed 2019-08-08]
19. Harrell Jr FE, Dupont C. The Comprehensive R Archive Network. 2019. Hmisc: Harrell Miscellaneous URL: https://cran.r-project.org/web/packages/Hmisc/index.html [accessed 2019-08-08]
20. Rizopoulos D. Ltm: An R package for latent variable modelling and item response theory analyses. J Stat Softw 2006 Nov;17(5):1-25. [doi: 10.18637/jss.v017.i05]

XSL•FO
RenderX

Abbreviations

AGFI: adjusted goodness-of-fit index
CFA: confirmatory factor analyses
CFI: Comparative Fit Index
CMIN: minimum discrepancy per degree of freedom
COSMIN: Consensus-based Standards for the selection of health status Measurement INstruments
EFA: exploratory factor analyses
eHealth: electronic health
eHIQ: eHealth Impact Questionnaire
eHIQ-NL: Dutch version of the eHealth Impact Questionnaire
EQ-5D-5L: 5-level EuroQol-5D version
FCCHL: Functional, Communicative and Critical Health Literacy
ICC: intraclass correlation coefficient
NPS: Net Promoter Score
RCT: randomized controlled trialRMSEA: root mean square error of approximation
SDC: smallest detectable change
SRMR: standardized root mean square residual
SUS: System Usability Scale
TLI: Tucker-Lewis Index

Edited by G Eysenbach; submitted 14.02.19; peer-reviewed by E McElroy, G Pérez; comments to author 11.04.19; revised version received 20.05.19; accepted 25.05.19; published 26.08.19.

Please cite as:
Neijenhuijs KI, van der Hout A, Veldhuijzen E, Scholten-Peeters GGM, van Uden-Kraan CF, Cuijpers P, Verdonck-de Leeuw IM
Translation of the eHealth Impact Questionnaire for a Population of Dutch Electronic Health Users: Validation Study
J Med Internet Res 2019;21(8):e13408
URL: http://www.jmir.org/2019/8/e13408/
doi:10.2196/13408
PMID:31452516

©Koen Ilja Neijenhuijs, Anja van der Hout, Evalien Veldhuijzen, Gwendolijne G M Scholten-Peeters, Cornelia F van Uden-Kraan, Pim Cuijpers, Irma M Verdonck-de Leeuw. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 26.08.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Development of an Instructional Design Evaluation Survey for Postgraduate Medical E-Learning: Content Validation Study

Robert Adrianus de Leeuw[1,2], MSc, MD, PhD; Michiel Westerman[3], MD, PhD; Kieran Walsh[4], MD, PhD; Fedde Scheele[1], MD, PhD

[1]Athena Institute for Trans-Disciplinary Research, VU University, Amsterdam, Netherlands

[2]Amsterdam University Medical Center, Amsterdam, Netherlands

[3]Department of Internal Medicine, Franciscus Gasthuis en Vlietland Hospital, Rotterdam, Netherlands

[4]British Medical Journal Learning, British Medical Association House, London, United Kingdom

Corresponding Author:
Robert Adrianus de Leeuw, MSc, MD, PhD
Athena Institute for Trans-Disciplinary Research
VU University
Boelelaan 1118
Amsterdam,
Netherlands
Phone: 31 618 39 0269
Email: r.deleeuw@amsterdamumc.nl

Abstract

Background: E-Learning has taken a firm place in postgraduate medical education. Whereas 10 years ago it was promising, it now has a definite niche and is clearly here to stay. However, evaluating the effect of postgraduate medical e-learning (PGMeL) and improving upon it can be complicated. While the learning aims of e-learning are evaluated, there are no instruments to evaluate the instructional design of PGMeL. Such an evaluation instrument may be developed by following the Association for Medical Education in Europe (AMEE) 7-step process. The first 5 steps of this process were previously performed by literature reviews, focus group discussion, and an international Delphi study.

Objective: This study will continue with steps 6 and 7 and answer the research question: Is a content-validated PGMeL evaluation survey useful, understandable, and of added value for creators of e-learning?

Methods: There are five phases in this study: creating a survey from 37 items (phase A); testing readability and question interpretation (phase B); adjusting, rewriting, and translating surveys (phase C); gathering completed surveys from three PGMeL modules (phase D); and holding focus group discussions with the e-learning authors (phase E). Phase E was carried out by presenting the results of the evaluations from phase D, followed by a group discussion. There are four groups of participants in this study. Groups A and B are experienced end users of PGMeL and participated in phase B. Group C are users who undertook e-learning and were asked to complete the survey in phase D. Group D are the authors of the e-learning modules described above.

Results: From a list of 36 items, we developed a postgraduate Medical E-Learning Evaluation Survey (MEES). Seven residents participated in the phase B group discussion: 4 items were interpreted differently, 3 were not readable, and 2 items were double. The items from phase B were rewritten and, after adjustment, understood correctly. The MEES was translated into Dutch and again pilot-tested. All items were clear and were understood correctly. The MEES version used for the evaluation contained 3 positive domains (motivation, learning enhancers, and real-world translation) and 2 negative domains (barriers and learning discouragers), with 36 items in those domains, 5 Likert scale questions of 1 to 10, and 5 open questions asking participants to give their own comments in each domain. Three e-learning modules were evaluated from July to November 2018. There were a total of 158 responses from a Dutch module, a European OB/GYN (obstetrics and gynecology) module, and a surgical module offered worldwide. Finally, 3 focus group discussions took place with a total of 10 participants. Usefulness was much appreciated, understandability was good, and added value was high. Four items needed additional explanation by the authors, and a Creators' Manual was written at their request.

Conclusions: The MEES is the first survey to evaluate the instructional design of PGMeL and was constructed following all 7 steps of the AMEE. This study completes the design of the survey and shows its usefulness and added value to the authors. It finishes with a final, publicly available survey that includes a Creators' Manual. We briefly discuss the number of responses

XSL•FO
RenderX

needed and conclude that more is better; in the end, however, one has to work with what is available. The next steps would be to see whether improvement can be measured by using the MEES and continue to work on the end understandability in different languages and cultural groups.

(J Med Internet Res 2019;21(8):e13921) doi:10.2196/13921

KEYWORDS

postgraduate medical education; continuing medical education; e-learning; distance education; survey; evaluation

Introduction

Background

E-Learning and distance education are a growing part of postgraduate and continuous medical education. The cost effectiveness and logistical benefits have previously been shown [1] and, whereas 10 years ago e-learning was promising, it is now part of mainstream medical education [2]. However, the overall effectiveness and added value of e-learning over conventional education such as face-to-face learning is debatable, and results in the literature are diverse [3]. One of the problems in evaluating e-learning is the lack of a proper evaluation tool [4].

The effectiveness of e-learning can be separated into two parts: effect of the learning aim and instructional design of the e-learning module. By learning aim, we mean the ability of an e-learning module to achieve the learning goals, usually either new knowledge, skills, or attitude/behavior [5]. For example, the learning aim may be to tie a laparoscopic knot, which can be evaluated by an objective skill assessment tool. By instructional design, we mean the functionalities (affordances) and their design [6]. In the case of e-learning, these are the design of the digital medium and affordances used to achieve the learning aim—for example, the virtual reality program with interactivity, feedback, and gamification to practice the knot tying.

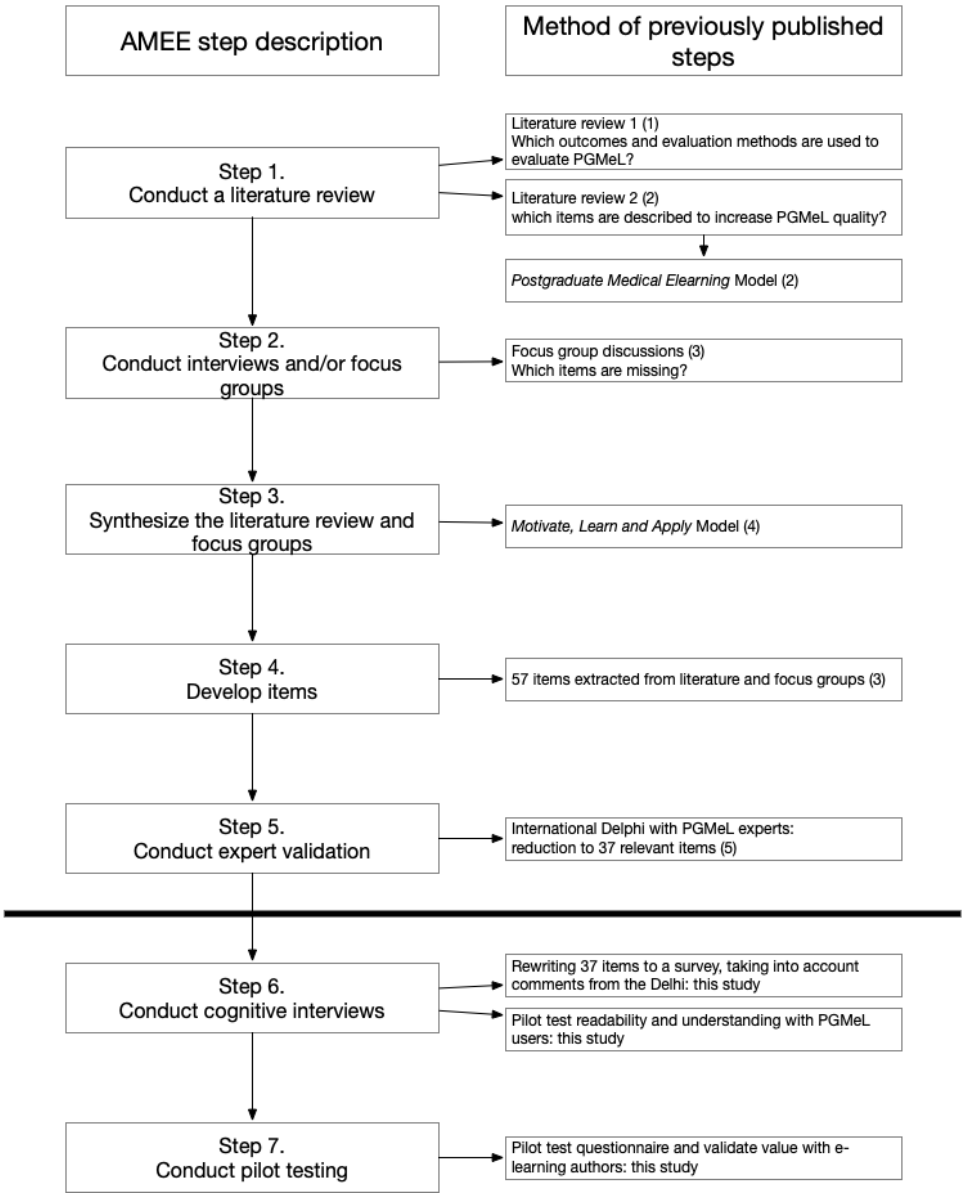
Despite the methodological limitations, e-learning is often evaluated by comparing test results before and after it is used. Usually, the learning aim is used as the primary outcome, and even when the design is also evaluated, this is generally done using instruments that are not aimed at postgraduate medical e-learning (PGMeL) [7]. When the design is not included in the e-learning evaluation study, however, the question arises about how to ascertain whether the e-learning modus is suited to the learning aim (eg, virtual reality is not suited for learning to tie a knot) or if there were essential flaws in the e-learning itself (eg, the virtual reality box was poorly designed). We believe that the evaluation of the learning aim should always go together with the evaluation of the design because they are interwoven in the final outcome. To properly evaluate the design, we need an instrument that has proper content validation and is aimed at the right target audience—in our case, PGMeL.

Development of a Survey in Medical Education

The development of an evaluation instrument is complex and involves many steps [8]. In 2014, the Association for Medical Education in Europe (AMEE) published a 7-step design process for developing surveys for medical education [9]. Steps 1 through 5 of the design were previously published in two reviews [7,10], focus group discussions [11], and an international Delphi study [4] (Figure 1). The aim of this study is to proceed with steps 6 and 7 and evaluate the results of the survey with the creators of a PGMeL. We want to know if the creators find the results helpful to improve the e-learning (usefulness), if they can understand the indicators that are used in the context of instructional design (understandability), and, finally, if the survey is offering them additional information over existing evaluation methods (added value). This leads to the following research question that this study will try to answer: Is a content-validated evaluation survey for PGMeL useful, understandable, and of added value for the creators of PGMeL?

XSL•FO
RenderX

Figure 1. Association for Medical Education in Europe (AMEE) 7-step design process for developing surveys for medical education. PGMeL: postgraduate medical e-learning.



Methods

Study Design

To conclude the content validation, this study collected evidence of response process validity to assess how participants interpreted the items (AMEE step 6) and conducted pilot testing (AMEE step 7). To prevent confusion between the AMEE steps and the methodology of this study, we call the steps of this study phases. To answer the research question, this study has five phases, A through E (Figure 2).

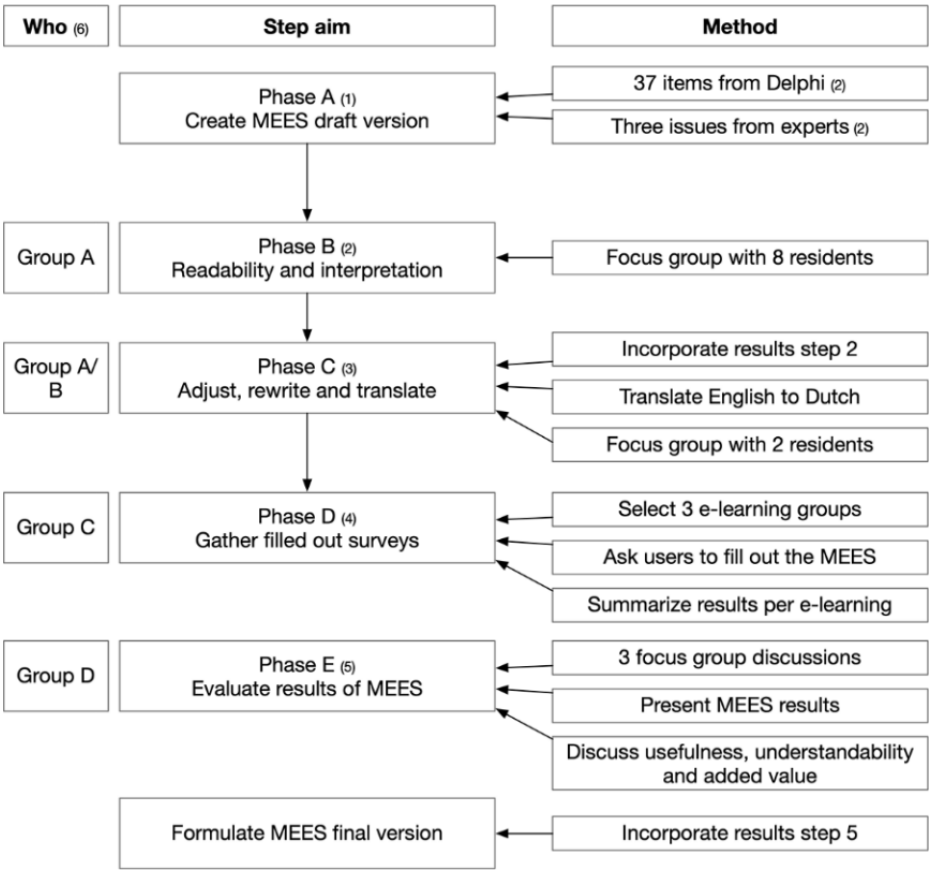
- Create a survey draft based on 37 items of the previous Delphi study [4], and address three concerns of experts in this Delphi study. First, the term e-learning can be confusing; second, the added value of another survey might be limited; and third, the indicators may be too general for the evaluated e-learning module [4]. We called the survey the postgraduate Medical E-Learning Evaluation Survey (MEES).
- Determine readability and understandability with experienced postgraduate students by use of a focus group discussion.
- Adjust the survey draft according to the feedback in phase B. The English survey was also needed in Dutch (for phase D); therefore, the rewritten English survey was translated and again pilot tested for readability. Because less discussion was expected, only two (native Dutch) residents (not part of the first group of residents) were asked to read the Dutch survey and provide feedback.
- Use the survey by evaluating three PGMeL modules. Contacts of several European PGMeL groups were emailed and asked to participate, and the first three agreed. They were sent the participant information and the MEES. After agreeing to use the survey, they were asked to add it to the standard evaluation survey they might already have. Users participated voluntarily. Anonymized results of the surveys were sent to RDL.

XSL•FO
RenderX

- Perform focus group discussions with the creators of the e-learning modules about the survey results. The results were presented per domain. For each domain, the minimum, maximum, and average score; all items; and free text comments were discussed (Multimedia Appendix 1). Finally, a strength-weakness analysis with scores, summary of recognized items, and summary of free texts per domain were carried out. The discussion guide and short demographics questionnaire are in Multimedia Appendix 2.

We chose focus group discussions as the main methodology because they are an appropriate method to investigate attitudes and beliefs and generate new ideas [12].

Figure 2. Five phases to address steps 6 and 7 of the Association for Medical Education in Europe design process and evaluate the usefulness, understandability, and added value of the Medical e-Learning Evaluation Survey (MEES).



Study Participants

There were three groups of participants in this study. Group A members were experienced end users of PGMeL who participated in phase 1. These were OB/GYN (obstetrics and gynecology) residents in their fourth and fifth year at the Amsterdam University Medical Center. Group B members (also experienced residents) were asked to participate with the translation from English to Dutch. Residents were invited by RDL to participate by email and could decline without any consequences or repercussions.

The second group (group C) comprised users who undertook an e-learning module and were asked to complete the survey in phase C. This group was asked to evaluate the e-learning module they had just taken as part of the usual evaluation process. The e-module had to be (1) an approach aiming to teach and learn the adoption of new knowledge, skills, or attitude/behavior representing all or part of an educational model and (2) based on the use of electronic media and devices as tools to improve training access, communication, and interaction. The users had to be postgraduates in medicine and able to read and write in English or Dutch.

The third group (group D) were creators or authors of the e-learning modules described above. For phase D, we asked representatives tasked with the usual evaluation and improvement responsibilities of each evaluated module. The Dutch Association of Medical Education Research gave ethical consent (ID 2018.5.1).

E-Learning Groups

This study evaluated three PGMeL modules. The aim was to gather survey outcomes to determine the usefulness, understandability, and added value of the modules with the creators. The aim was not to evaluate the modules themselves.

Module 1 was aimed at new doctors in a big teaching hospital in Amsterdam, the Netherlands. The e-learning aim was to train them to use the local electronic patient records. It was mandatory for all new doctors to complete the module, after which they were asked to complete the evaluation. The author group did not add extra items to the survey. A total of 160 participants were asked to fill out the survey from June to October 2018.

Module 2 was aimed at surgical residents and offered globally. The platform offered different surgical modules for a variety

XSL•FO
RenderX

of specialties focusing on anatomy, surgical steps, and pitfalls during surgery. The author group added 19 extra items. After finishing at least three modules, users were asked to voluntarily complete the survey. From August to November 2018, 395 participants were asked to evaluate the e-learning modules.

Module 3 was aimed at OB/GYN residents practicing minimally invasive surgery, mainly in Europe. The e-learning module is part of a certification with face-to-face and hands-on training. The author group added 8 extra items. From August until the end of October 2018, about 2400 participants were asked to complete the survey. Of these, most were older users, some with email addresses that no longer worked. An estimated 1600 participants had recently used the e-learning module and were reached by email.

Data Collection and Procedure

Data were collected in phases B and C by focus group discussion with experienced users in the comfort of their university environment. Data from phase D were collected by providing the MEES as an online survey with a short introductory text (Multimedia Appendix 3). Data for phase E were collected by audio recordings after written consent was given. These focus group discussions were facilitated by RDL at the main offices of the e-learning groups.

Data Analysis

No analysis was undertaken of the first four phases as the data were used during the phases themselves. The focus group evaluations in phase E were analyzed. All interviews were transcribed verbatim and a thematic analysis was performed as per Braun et al [13]. The transcribing of the interviews was completed by RDL to enable the author to familiarize himself with the data. We used ATLAS.ti version 8.0 (ATLAS.ti Scientific Software Development GmbH) for the initial coding. Thematic analysis has been shown to usefully summarize key features and generate unanticipated insights [12,13]. To perform the analysis, we ordered the codes by the predefined themes from our research question: usefulness, understandability, and added value. We then reviewed the content of the themes and, if needed, redefined them as per Braun et al [13].

Results

The results are described per phase in Figure 2. The initial draft and a change log of the MEES can be found online at www.MotivateLearnApply.com. The final version of the MEES is attached as Multimedia Appendix 3 and a creators' manual is attached as Multimedia Appendix 4.

Phase A: Postgraduate Medical E-Learning Evaluation Survey

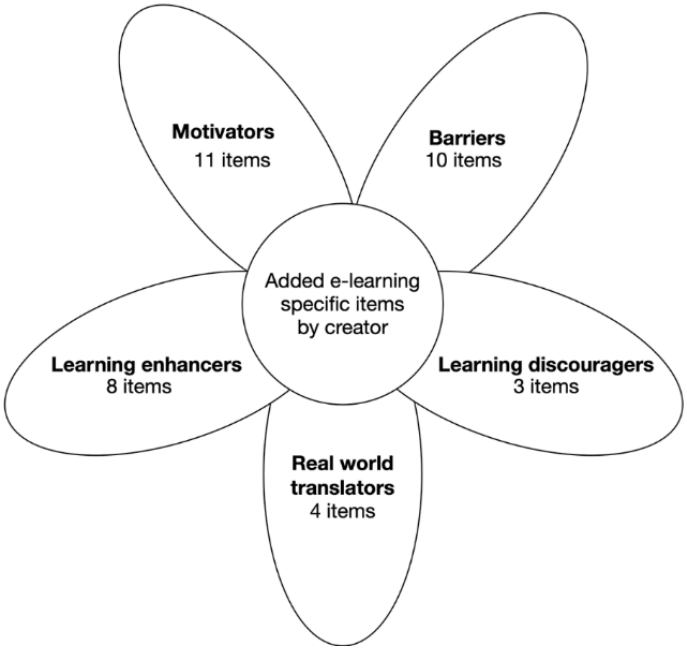
To address expert feedback, we added an explanation to the survey and the option for e-learning creators to add additional items. The term e-learning was defined in a previous Delphi study and has been used for all studies so far: "an approach to teaching and learning, representing all or part of the educational model applied, that is based on the use of electronic media and devices as tools for improving access to training, communication, and interaction and that facilitates the adoption of new knowledge, skills, and/or behavior/attitude" [14]. This means that all forms of electronic learning based on an educational model are e-learning.

To address the generalizability of the indicators, they are used as examples. There are, for example, many ways to be motivated. The previous phases provided nine ways to achieve this, but the creators of the e-learning module might have used other strategies as well. Before the survey started, the creators of the modules were asked to add those indicators to each domain as examples. The domains were thus questioned using the general indicators from the literature but also items that might be unique to that one module.

The MEES contains questions in five domains, each of which starts with a Likert scale of 1 to 10. These are followed by indicators from the previous phases and those added by the e-learning creators depending on the aims of the specific module. Finally, there is an open question about the domain. The domains are motivators, barriers, learning enhancers, learning discouragers, and real-world translators. The MEES contains 10 questions, 36 examples, and five open questions. Figure 3 shows the relationship between the domains, and Table 1 lists all 36 examples with a short explanation of the purpose.

XSL•FO
RenderX

Figure 3. The domains and structure of the Medical e-Learning Evaluation Survey.



XSL•FO
RenderX

Table 1. The 36 items from the final Medical E-Learning Evaluation Survey with a short explanation of their meaning from the creators' points of view.

Domain and original item	Short explanation
Motivators	
I felt this module was important.	Creating a feeling of importance is very important for the user. The challenge is to convey to your users that the learning aims are important for their work and personal development.
I felt it was my responsibility to undertake this module.	Along with importance, your user needs to feel responsible for the learning aim as well. This can be done by emphasis on the importance, but also by, for example, rewarding or giving responsibility for an outcome.
I had enough time to complete the module.	Proving time to do the module seems contradictory to "anytime, anywhere" learning, but it does give the learning the feeling of priority from a management level.
I had a good understanding of the general purpose of the module.	The general purpose is the learning aim: knowledge, skills, or attitude/behavior. It should be very clear to the user what they gain from finishing the module.
The e-learning objectives (for each educational section) were clear to me.	When a module is separated into different sections/chapters, make sure you communicate what the learning objectives are for each section.
There was a clear overview of all content.	Providing an overview of all lessons, objectives, and options gives the user the possibility to manage expectations and, if possible, create their own learning process.
I knew how to navigate to the content.	Navigation is an important part of the user interface and should be very clear for the user so they can find content easily and go back and forward through the content.
I felt comfortable with the quality and truthfulness of the content.	Trust is important when learning. If the user has doubts about the truthfulness or quality, it will limit the working memory used for learning. Trust can be gained by the transparency of the creators, referring to recent literature, etc.
I was able to undertake this module without being forced.	Forcing a user to undertake a module is the opposite of motivating them. If force or even blackmail is needed, the user will feel resentment, which kills motivation.
I felt taken seriously as an adult learner.	Taking the learning seriously means avoiding childish illustrations or examples and aiming at the level of experience means that you take into account what the user already knows to prevent repetition of basic knowledge.
The module was aimed at my level of experience.	Making your module too easy will decrease motivation, and making it too complicated will make users learn less. This is why knowing the background knowledge of the target audience is of great importance.
Barriers	
I was not able to create my own learning path to my own needs.	This questions the difference between synchronized and asynchronized learning paths. Creating your own learning path means the option to test and skip already known sections or to go from A to C and then to B.
The module was not easily accessible at my location or with my device.	Accessing the module should ideally be possible from every device and location, so consider, for example, internet speeds in foreign countries. If access is not possible, consider helping your users get the right device.
The navigation did not make sense to me.	Good navigation is helpful but poor navigation will not only limit a module but make it impossible to finish. Make sure your users can follow all steps without using their cognitive load for navigation.
The layout of the module was too complicated.	Navigation and layout are both important aspects of the user interface. The less cognitive energy is used for the learning environment, the more can be used for the learning itself.
There was no instrument to help me navigate the module (eg, a sitemap).	Even if the navigation is of a high standard, it is still very helpful to have an instrument that gives an overview of all content and helps direct users where they want to be.
I had concerns about the security and safety of the module regarding my personal information.	Worries about security and privacy are relevant in many countries and may even have a legal aspect.
The module was slow and took too long to load.	Fast and logical use of the module is also an important aspect of the user experience. Waiting on affordances or loading frustrates and distracts and should be minimized.
I did not know which devices the module was compatible with, and I might have used the wrong one.	If your module has specific needs (eg, a specific operating system such as iOS) you need to clearly state that at the beginning. Try to prevent users from experiencing your module in a wholly different way than planned because they use the wrong device.
The module was too long.	The duration should have been specified. Duration of videos, sections, and the module overall are taken together as one item. If there are, for example, longer videos, their duration can be added as a separate item.
The module did not divide the content into proper sections.	Learning and memory theories suggest that learning has a limited time span. Sectioning or chunking is a very effective way to help users through a bigger module.

XSL•FO
RenderX

Domain and original item	Short explanation
Learning enhancers	
I could personalize the module (eg, by saving and continuing, filling out questionnaires, and getting my personal score).	Personalizing a learning experience allows the user to know how they are doing and follow a preferred method and path. The more personal and specific such things as feedback are, the more the user will gain. This is a very important motivator as well.
I could create my own learning path and was not forced to follow the directed path (eg, by skipping parts or returning to previous sections if needed).	This questioned the difference between synchronized and asynchronized learning paths. Creating your own learning path means the option to test and skip already known sections or to go from A to C and then to B.
I had an idea of the progress I had made and what was left to do (eg, by a progress bar).	When learning, it's important to manage expectations. Knowing what is already done and what is left to do is an important affordance of, for example, a book, and should preferably be available in a module as well.
I had access to technical support if needed.	To minimize the effort spent on technical aspects rather than learning, providing support as fast as possible will prevent users from stopping learning.
The module provided summaries where needed.	Learning theory suggests that summaries support learning by offering repetition of content in a new format and allowing chunking of the bigger picture.
The module provided feedback on my answers.	Learning theory also suggests that learning is more effective when based on previous experience and knowledge, and providing feedback helps the user to make connections between new knowledge and their mistaken or correct assumptions.
There were exercises and/or assignments in the module.	Learning theory suggests that actively using new knowledge will help it to go from working memory to long-term memory. Therefore, exercises or assignments help the transfer of the learning aim to long term memory.
I could interact with the content of the module (eg, questions, exercises, or other interactivities).	Interaction is another example of actively using the content, helping users learn more efficiently.
Learning discouragers	
I got stressed or frustrated by the module for whatever reason.	Stress can be caused by many things but will always distract from learning. Stress can come from failing hardware, deadlines, the consequences of failing, etc.
The content was not able to adapt to my device when needed (eg, the module should work on a mobile device, but the icons were much too small for that).	Nonadaptable content can cause frustration and degrade the user experience, again moving energy away from learning and toward technical aspects.
The e-learning design and visuals were too distracting for me.	Multimedia learning provides a theory and guidelines for how to use the combination of visuals and auditory stimuli effectively. Distraction should always be prevented.
Real-world translators	
The e-learning content and examples are translatable to my daily real-world work.	Adult learning theory suggests that adults prefer learning in a professional environment, if they can use the lessons learned in daily practice. Providing content and examples that are relatable will help.
The module seems up to date and properly maintained.	When a user thinks they are learning old material, it might not seem applicable to their daily work anymore. This will kill motivation and minimize the effort the user is willing to put in.
The module provided sources for the information that were also accessible after finishing it.	Health care professionals in particular might want to undertake further reading in a relevant topic or refresh their memory after finishing the module. Providing this and letting the user know this is possible will increase motivation.
Besides this questionnaire, the module was evaluated on topics like user experience, effectiveness, usability, and/or costs.	The literature suggests that evaluation is an important step. This question is an oxymoron because by asking it, you are already evaluating. Therefore, the question is: Are OTHER evaluation instruments ADDED to this evaluation—for example, focus group discussions?

Phase B: Readability and Item Interpretation

Eight residents were asked to participate and all agreed, but one was unable to come to the discussion on time. The discussion lasted 65 minutes and took place in May 2017 at the Amsterdam University Medical Center, the Netherlands. Four items were interpreted differently than intended, three items were not readable, and two items were in two domains. Details can be found in the change log online. Overall, domains and questions were well understood.

Phase C: Adjust, Rewrite, and Translate

The items from phase A were rewritten. After they had been adjusted, they were understood correctly. After finishing the

XSL•FO
RenderX

English MEES, RDL translated the survey into Dutch. Two native Dutch residents from the Amsterdam University Medical Center read the survey in May 2017. All items were clear and were understood correctly. No other changes were made. The MEES version used for the evaluation contained three positive domains (motivation, learning enhancers, and real-world translation) and two negative domains (barriers and learning discouragers), with 36 items in those domains, 5 Likert scale questions of 1 to 10, and 5 open questions asking the participants their own comments in each domain. Figure 3 provides an overview of the domains and number of questions per domain in the MEES survey.

Phase D: Gathering Completed Surveys

Details of the three evaluated modules are summarized in Table 2, and the scores per domain are in Table 3. All evaluations took place between July and the end of November 2018, after which we concluded the evaluation.

The higher the score in the positive domains (motivators, learning enhancers, and real-world translators), the better. In the negative domains barriers and learning discouragers, a lower score is better. Note the discrimination between the positive and negative domains.

In total, there were 77 free text comments (see Multimedia Appendix 1 for all comments). The main positive comments concerned the availability and added value of the module to local education.

> *Unfortunately, surgical skills at my university are not well taught due to the large number of residents. There is also lack of standardization in teaching. I was happy to find a fun way to learn the best, standard way to perform common gynecological procedures and no longer rely on sketchy YouTube videos.*

On the negative side, users complained mostly about technical barriers such as long loading times, log-in problems, software crashes, and nonfunctioning affordances such as search functions.

> *It sometimes took too long to load despite good internet connection, and I have often been forced to abandon a procedure due to this.*
>
> *Video streaming is a serious limitation. There is a need for video downloads.*

Another frequent complaint was about the language barriers such as poor English or videos with hard-to-understand speakers.

> *Difficulte de langue. Je ferais effort d'aprendre.*

Table 2. Summary details of the three evaluated modules. For more details, see Multimedia Appendix 1.

Module	Target audience	Location	Additional items	Survey status	
				Invited, n	Completed, n (%)
1	Medical staff	Netherlands	0	160	16 (10.0)
2	Residents	Worldwide	19	395	36 (9.1)
3	OB/GYN[a] residents	Worldwide	8	1600	106 (6.6)

[a]OB/GYN: obstetrics and gynecology.

Table 3. Domain scores (range 1-10) per domain of the three evaluated modules.

Module	Motivator, median (IQR[a])	Barrier, median (IQR)	Learning enhancer, median (IQR)	Learning discourager, median (IQR)	Real-world translator, median (IQR)
1	7.5 (2.0)	3.0 (3.8)	7.0 (2.0)	4.0 (2.0)	8.0 (3.0)
2	9.0 (2.3)	3.0 (3.0)	9.0 (3.0)	3.0 (3.0)	8.5 (3.0)
3	10.0 (2.0)	2.0 (5.0)	9.0 (2.0)	1.0 (4.0)	9.0 (2.0)

[a]IQR: interquartile range.

Phase E: Focus Group Discussions With the E-Learning Creators

Three focus group discussions took place, one with each e-learning creator group at their main office in November and December 2018. The average age of the 10 participants was 51 years with participants having 0 to 5 years' experience in creating e-learning, and there were content, didactic, and technical experts at the interviews. While 80% (8/10) had experience with previous evaluations, only 10% (1/10) had used any formal evaluation methods. The three subjects of discussion are now described.

Usefulness: all participants described the results as very useful.

> *Grouping the results into positive (domains) and negative (domains) resulted in a clear overview of what we need to keep and what to improve.* [1B]

All groups said that the option to add items specific to a module increases the usefulness. It provides feedback on the additional items that are considered important for the creator group. The first group regretted not adding any items themselves and, seeing the results now, said they would have added them.

Understandability: going over the items one by one, some were not clear to the creators. Even though they understood the question, they did not know how to interpret it from a creator's point of view. These items (10, 12, and 36) are in Multimedia

XSL•FO
RenderX

Appendix 4. The general advice was to have an explanation manual for the creators of each item that could be consulted in the event of misunderstanding.

> *It would be nice to have a short explanation per item.* [2C]

There were also worries about the form in which the results were presented. All the completed surveys were presented by RDL and summarized by him as well.

> *How much time did it take you to formulate the results like this, and can this be done by us as well?* [3A]

Added value: although the participants had experience evaluating e-learning modules, only one had used a formal evaluation method (although it is unknown which one). There were three subjects that added value to the MEES, the first of which was the domains. Using the domains gave the creators a structure that they did not have with other, informal, evaluation methods. Second, the items provided concrete examples of do's and don'ts, which inspired the changes needed in an update. Third, using a formal method gave the creators a feeling of importance and allowed them to formalize the needs for improvement. The creators believed that using a formal method would allow them to more easily convince management and increase the commercial benefits of the module.

> *This really helps to improve the commercial value of our e-learning.* [2A]

The five phases of this study provided a first draft of the MEES, an adjusted version (Multimedia Appendix 3), evaluation data from three PGMeL modules (Multimedia Appendix 1), and the thematic analysis of three focus group discussions with the creators of these modules.

Discussion

Principal Findings

To our knowledge, the MEES is the first survey designed to evaluate PGMeL. Content validation has been completed, and this study completes all seven steps described by the AMEE. We set out to investigate the usefulness, understandability, and added value of this survey with the creators of three PGMeL modules. This study shows that the MEES is very useful, that understandability is clear with the help of a creators' manual (Multimedia Appendix 4), and that the MEES is of added value in connection with the structured domains and validated nature of the survey. Although the MEES has reached a second stage, two questions remain.

First, how do we know the items have been correctly understood by all the participants? The pilot evaluation was carried out with Dutch participants who seemed to understand the questions in a group session. That might be generalizable to Dutch residents. However, the second and third evaluated e-learning modules involved European and worldwide users who might not have interpreted the questions correctly. To then interpret an evaluation, it is important to at least consider the content integration and evaluate the equity pedagogy [15]. Content integration means using content that will illustrate the same examples in a variety of cultures. Equity pedagogy exists when the curriculum is modified to accommodate and facilitate the academic achievements of a diverse group. Although debatable, using the cultural dimensions of Hofstede [16] can also help to place extreme high or extreme low scores in a cultural light. For example, in a culture with high power distance (in other words, where power and inequality are fundamental facets of the society), users might be tempted to express gratitude by providing extremely positive feedback.

Second, how many completed surveys are needed for a proper evaluation? Ideally, you would ask participants to keep completing the survey until a theoretical data saturation is met. This would require data analysis, after, for example, every tenth survey. But as seen above, the average reply was around 8%; thus, 92% of users are missed when reply is made voluntary. Besides the practical problem of determining data saturation, replies stopped whether you need more or not. This can be increased by, for example, asking respondents to fill out a survey before supplying a needed certificate. This might raise an ethical dilemma in a research setting like this. For this study, the response rate was less relevant. We were not aiming to evaluate the e-learning modules, we were aiming to evaluate the outcome with the creators of the modules. When using this survey in practice, the authors should consider making it mandatory to get a higher response rate. Other ways to assemble a representative group include purposive sampling [17], but the question remains as to how many is enough. The conclusion is that the more respondents, the better the data, but the reality is that researchers can only work with the data made available to them.

On the other hand, although missing 92% of users might seem to indicate a failed evaluation, the question of who is most likely to complete the survey is a valid one. Hu et al [18] show that users who provide feedback are most likely to be either very satisfied or very disappointed and thus may be exactly the respondents required; that is, it may be that the middle group yields less information of use to the evaluation.

Strengths and Limitations

Following all steps of the AMEE guide and peer publishing them might be the biggest strength of the MEES. Questionnaire validation is a complex and diverse field of expertise, and this study takes an important next step. An international Delphi study in 2010 (Consensus-Based Standards for the Selection of Health Status Measurement Instruments) helped to structure the diverse terminology within this field by providing this definition: "validity is the degree to which an instrument truly measures the constructs it proposes to measure" [8]. Many types of validity can be evaluated, and validation is a continuous process. A short list, partly based on Tsang et al [19], of forms of validation is given in Table 4 with an explanation of their applicability to the MEES. It can be seen that face and content validity are completed and that construct, criterion, and predictive validity leave room for future research. Because structural and concurrent validity are not applicable, factor analysis was deliberately not performed.

XSL•FO
RenderX

Table 4. Forms of validation of a survey in relation to the Medical E-Learning Evaluation Survey.

Type of validation	Explanation	Relationship to MEES[a]
Face	Whether the instrument is understandable and relevant	Checked in this study by asking the PGMeL[b] creators about understandability and added value
Content	Whether the instrument measures the most important aspects of a concept that it is designed to evaluate	Checked in three previous studies by means of a review [10], focus groups [11], and an international Delphi study [4]
Construct	The degree to which the instrument's scores relate to other measures in a manner that is consistent with an a priori hypothesis concerning the concepts being measured	Awaiting future validation: the construct of the MEES is predicting efficiency and effectiveness by evaluating the experience of affordances to determine the quality of the instructional design
Criterion	How well one measure predicts the outcome of another	Awaiting future validation: MEES might predict the satisfaction or learning aim transference of the PGMeL
Predictive	The instrument's ability to predict future test results	Awaiting future validation: as regards the MEES, this is in line with the criterion validity and can be checked in the future
Structural	Whether all items in a scale or subscale measure the same concept of the dimensionality of the instrument	Not applicable; this can be done by factor analysis but assumes that a scale or subscale is highly correlated, which might not be the case in the MEES
Known-group	The ability to be sensitive to differences between groups of users that may be expected to score differently in the predicted direction	Not applicable; this can be done by comparing PGMeL designed for certain cultural groups, but the assumptions are too complex to use this in practice
Concurrent	The association of the instrument with accepted standards	Not applicable; there is no gold standard of evaluating PGMeL

[a]MEES: Medical E-Learning Evaluation Survey.

[b]PGMeL: postgraduate medical e-learning.

Another strength lies in the involvement of end users. By including experienced residents in the focus group discussions, Delphi surveys, and pilot evaluation, we believe the MEES is truly a user-centered method of evaluation as it evaluates not only theoretical items but also subjects that matter to the user. The numbers from the Likert scale questions do not offer enough insight, but the free texts fields do add knowledge about the users' needs and wishes. Most of the comments concerned e-learning affordances and technical execution (eg, interactive videos that load too slow). This emphasis shows the importance of evaluating not only the content of the module but the instructional design as a whole.

The biggest limitation of the MEES is also addressed above. It is not possible to know what feedback has been missed from those users who did not fill out the form. This might be reduced by making an evaluation mandatory, but it is impossible to predict how motivated users will be to provide useful feedback. We therefore believe that the MEES should always be accompanied by in-depth focus group evaluations with the users. Not only will these provide missed feedback, but they will also allow researchers to find out why some items are recognized or missed. Proper evaluation should never contain only an online survey.

Future Research

Validation is a never-ending story, and it is necessary to continue collecting validity evidence [9]. We believe three steps should be taken. First, as part of the construct and criterion validation, it would be interesting to see if the MEES can be used to measure improvement by taking a PGMeL, evaluating it with the MEES and a focus group discussion, adjusting it accordingly, and reevaluating to see if the second evaluation is better. This would provide insight into the actual benefit for future learning. The second step would be to evaluate readability and understandability in different languages. To this end, the survey should be translated into other languages and those new translations pilot-tested for readability and understandability.

The third step is to evaluate the understandability and reliability of the survey within different subcultures. Evaluating how different cultural groups interpret digital evaluation and the questions can provide insight into the way the creators should use the results of the survey.

Conclusion

This study provides the first instructional design evaluation survey for postgraduate medical e-learning. Content validation has been completed, and this study completes all seven steps described by the AMEE for the development of an evaluation instrument for medical education. The survey was experienced as useful, understandable, and of added value. Future research can continue the validation process and follow up in the daily practice of evaluating e-learning.

Conflicts of Interest

None declared.

XSL•FO
RenderX

Multimedia Appendix 1
E-Learning survey outcomes.
[DOCX File, 34KB - jmir_v21i8e13921_app1.docx]

Multimedia Appendix 2
Focus group discussion guide.
[DOCX File, 16KB - jmir_v21i8e13921_app2.docx]

Multimedia Appendix 3
Final version of the Medical E-Learning Evaluation Survey.
[DOCX File, 86KB - jmir_v21i8e13921_app3.docx]

Multimedia Appendix 4
Creators' manual.
[DOCX File, 25KB - jmir_v21i8e13921_app4.docx]

References

1. Taroco A, Valente T, Carbogim C. Distance learning for updating health professionals in palliative care: a systematic review. BMJ Support Palliat Care 2017 Jun;7(2):205-211. [doi: 10.1136/bmjspcare-2015-001042] [Medline: 28062410]
2. O'Doherty D, Dromey M, Lougheed J, Hannigan A, Last J, McGrath D. Barriers and solutions to online learning in medical education: an integrative review. BMC Med Educ 2018 Jun 07;18(1):130 [FREE Full text] [doi: 10.1186/s12909-018-1240-0] [Medline: 29880045]
3. Vaona A, Banzi R, Kwag KH, Rigon G, Cereda D, Pecoraro V, et al. E-learning for health professionals. Cochrane Database Syst Rev 2018 Jan 21;1:CD011736. [doi: 10.1002/14651858.CD011736.pub2] [Medline: 29355907]
4. de Leeuw RA, Walsh K, Westerman M, Scheele F. Consensus on quality indicators of postgraduate medical e-learning: Delphi study. JMIR Med Educ 2018 Dec 26;4(1):e13 [FREE Full text] [doi: 10.2196/mededu.9365] [Medline: 29699970]
5. Taveira-Gomes T, Ferreira P, Taveira-Gomes I, Severo M, Ferreira MA. What are we looking for in computer-based learning interventions in medical education? A systematic review. J Med Internet Res 2016 Aug 01;18(8):e204 [FREE Full text] [doi: 10.2196/jmir.5461] [Medline: 27480053]
6. Gagné R, Golas KC, Kaller JM. Principles of Instructional Design, 5th Edition. Belmont: Wadsworth; 2005.
7. de Leeuw R, de Soet A, van der Horst S, Walsh K, Westerman M, Scheele F. How we evaluate postgraduate medical e-learning: systematic review. JMIR Med Educ 2019 Apr 05;5(1):e13128 [FREE Full text] [doi: 10.2196/13128] [Medline: 30950805]
8. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. J Clin Epidemiol 2010 Jul;63(7):737-745. [doi: 10.1016/j.jclinepi.2010.02.006] [Medline: 20494804]
9. Artino AR, La Rochelle JS, Dezee KJ, Gehlbach H. Developing questionnaires for educational research: AMEE Guide No. 87. Med Teach 2014 Jun;36(6):463-474 [FREE Full text] [doi: 10.3109/0142159X.2014.889814] [Medline: 24661014]
10. de Leeuw RA, Westerman M, Nelson E, Ket JCF, Scheele F. Quality specifications in postgraduate medical e-learning: an integrative literature review leading to a postgraduate medical e-learning model. BMC Med Educ 2016 Dec 08;16:168 [FREE Full text] [doi: 10.1186/s12909-016-0700-7] [Medline: 27390843]
11. de Leeuw RA, Westerman M, Scheele F. Quality indicators for learner-centered postgraduate medical e-learning. Int J Med Educ 2017 Apr 27;8:153-162 [FREE Full text] [doi: 10.5116/ijme.58ce.60aa] [Medline: 28456781]
12. Wong LP. Focus group discussion: a tool for health and medical research. Singapore Med J 2008 Mar;49(3):256-260 [FREE Full text] [Medline: 18363011]
13. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
14. Sangrà A, Vlachopoulos D, Cabrera N. Building an inclusive definition of e-learning: an approach to the conceptual framework. Int Rev Res Open Dist Learn 2012;13(2):146-159 [FREE Full text]
15. Bank JA. Cultural Diversity and Education. 6th Edition. New York: Routledge; 2016.
16. Hofstede G. Dimensionalizing cultures: the Hofstede Model in context. Online Read Psychol Cult 2011 Dec 01;2(1):1-26. [doi: 10.9707/2307-0919.1014]
17. Palinkas LA, Horwitz SM, Green CA, Wisdom JP, Duan N, Hoagwood K. Purposeful sampling for qualitative data collection and analysis in mixed method implementation research. Adm Policy Ment Health 2015 Sep;42(5):533-544 [FREE Full text] [doi: 10.1007/s10488-013-0528-y] [Medline: 24193818]
18. Hu N, Pavlou P, Zhang J. On self-selection biases in online product reviews. MISQ 2017 Feb 1;41(2):449-471. [doi: 10.25300/MISQ/2017/41.2.06]

XSL•FO
RenderX

19. Tsang S, Royse CF, Terkawi AS. Guidelines for developing, translating, and validating a questionnaire in perioperative and pain medicine. Saudi J Anaesth 2017 May;11(Suppl 1):S80-S89 [FREE Full text] [doi: 10.4103/sja.SJA_203_17] [Medline: 28616007]

Abbreviations

PGMeL: postgraduate medical e-learning
AMEE: Association for Medical Education in Europe
IQR: interquartile range
MEES: Medical E-Learning Evaluation Survey
OB/GYN: obstetrics and gynecology

Edited by N Zary; submitted 06.03.19; peer-reviewed by J Last, T Taveira-Gomes; comments to author 02.04.19; revised version received 03.04.19; accepted 17.04.19; published 09.08.19.

Please cite as:
de Leeuw RA, Westerman M, Walsh K, Scheele F
Development of an Instructional Design Evaluation Survey for Postgraduate Medical E-Learning: Content Validation Study
J Med Internet Res 2019;21(8):e13921
URL: https://www.jmir.org/2019/8/e13921/
doi:10.2196/13921
PMID:31400102

©Robert Adrianus de Leeuw, Michiel Westerman, Kieran Walsh, Fedde Scheele. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 09.08.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Evaluation of a Technology-Based Peer-Support Intervention Program for Preventing Postnatal Depression (Part 1): Randomized Controlled Trial

Shefaly Shorey[1], PhD; Cornelia Yin Ing Chee[2], MBBS, MMed; Esperanza Debby Ng[1], BA; Ying Lau[1], PhD; Cindy-Lee Dennis[3], PhD; Yiong Huak Chan[1], PhD

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

[2]Department of Psychological Medicine, University Medicine Cluster, National University Hospital, Singapore, Singapore

[3]Division of Medical Oncology and Hematology, Department of Medicine, University of Toronto, Toronto, ON, Canada

Corresponding Author:
Shefaly Shorey, PhD
Alice Lee Centre for Nursing Studies
Yong Loo Lin School Of Medicine
National University of Singapore
Clinical Research Centre, MD 11
Level 2, 10 Medical Drive
Singapore, 117597
Singapore
Phone: 65 6601 1294
Fax: 65 6776 7135
Email: nurssh@nus.edu.sg

Abstract

Background: The frenzy of postbirth events often takes a toll on mothers' mental well-being, leaving them susceptible to postpartum psychological disorders such as postnatal depression (PND). Social support has been found to be effective in restoring the emotional well-being of new mothers. Therefore, mothers need to be supported during the crucial postpartum period to buffer the negative after effects of childbirth and to promote healthier maternal well-being.

Objective: This study aimed to evaluate the effectiveness of a technology-based peer-support intervention program (PIP) on maternal outcomes during the early postpartum period.

Methods: A randomized, parallel-armed controlled trial was conducted. The study recruited 138 mothers (69 in intervention group, 69 in control group) at risk of PND from a tertiary hospital in Singapore. To support these mothers, 20 peer volunteers were recruited by word of mouth and trained by a psychiatrist in social support skills before the intervention commenced. The 4-week–long intervention included a weekly follow-up with a peer volunteer through phone calls or text messages. The intervention group received peer support in addition to the standard care offered by the hospital. The control group only received postnatal standard care. Maternal outcomes (PND, postnatal anxiety [PNA], loneliness, and perceived social support) were measured with reliable and valid instruments. Data were collected immediately postpartum, at 1 month postpartum and at 3 months postpartum. The general linear model was used to compare the groups for postpartum percentage changes in the outcome variables at first and third months, and the linear mixed model was used to compare the trend over the study period.

Results: There was a statistically significant difference in Edinburgh Postnatal Depression Scale scores (d=–2.11; 95% CI −4.0 to −0.3; P=.03) between the intervention and control groups at 3 months postpartum after adjusting for covariates. The intervention group had a significant change over time compared with the control group.

Conclusions: The technology-based PIP was found to be effective in reducing the risk of PND among new mothers and showed a generally positive trend in reducing PNA and loneliness and increasing perceived social support. This study highlights the importance of training paraprofessionals to provide needed support for new mothers postpartum. A further long-term evaluation of the PIP on maternal and family outcomes and its cost-effectiveness is needed to inform clinical practices.

Trial Registration: ISRCTN Registry ISRCTN14864807; https://www.isrctn.com/ISRCTN14864807

International Registered Report Identifier (IRRID): RR2-10.2196/resprot.9416

XSL•FO
RenderX

(J Med Internet Res 2019;21(8):e12410) doi:10.2196/12410

KEYWORDS

anxiety; loneliness; postpartum depression; social support; technology; digital health; peer support; peer-to-peer support; online support groups; internet

Introduction

Postnatal Depression

In a recent effort to improve women's well-being globally, the World Health Organization has stepped up preventive efforts to reduce maternal morbidity and mortality [1] . As one of the leading causes of maternal morbidity [2], postnatal depression (PND) has an approximate global prevalence of 10% to 15% [3], with less than 1% of this population diagnosed with postpartum psychosis [4]. Common symptoms of PND include changes in appetite, insomnia, higher irritability, mood swings, anxiety [5], and even suicidal ideation in severe cases [6]. Women have also reported feelings of inadequacy [7,8], role conflicts [8,9], disconnection from others [6,9], loneliness [9], and dissatisfaction with life [7].

PND has garnered much attention because of its potential contribution to maternal mortality and its ripple effects on the family unit. According to a study by Goodman et al [10], maternal PND is the biggest risk factor for paternal PND, affecting 24% to 50% of all fathers, often jeopardizing marital relationships. Furthermore, PND adversely affects the quality of mother-child interaction and bonding [11] as mothers who suffer from PND tend to be more negligent, less tolerant, and hostile toward their children [12-14]. These tendencies not only impair the cognitive, behavioral, social-emotional development, and physical health of the child [14-17], but also increase their attachment anxiety [14], proneness to violence [18], and risks of psychopathology [14]. Therefore, early detections and preventions of PND are necessary to mitigating detrimental consequences at the individual and societal levels.

Despite its unspecified causes and suggested multifactorial etiologies [5,9], many studies have identified high-risk predictors of PND, including demographic, biological, psychological, obstetric, social, and lifestyle risk factors [9,19,20]. An in-depth analysis has also further revealed other risk factors that were derived from underlying unmet needs and social deficiencies experienced by postpartum mothers, such as the need for close, nonjudgmental confidants who empathize with them [9,21,22] and initiated support from others [5,23]. This indicates the vital role of social support during the postpartum period in reducing the risk of maternal PND.

Importance of Social Support

Social support has long been proven to buffer stress and promote healthy psychological well-being [24]. This is especially crucial for new mothers during the stressful postpartum period. Although professional advice and informational support were much preferred by mothers [23], social support from partners and family members was shown to sustain mothers' quality of life after childbirth and serve effectively as a buffer against PND [24]. However, Dennis et al [21] also stressed on the importance of support from other paraprofessionals such as experienced mothers. Other studies discovered that the sharing of experiences among mothers helps to develop a tight-knit community, which promotes a sense of belongingness, improves one's sense of self-worth, boosts parenting confidence, and prevents PND [6,7,25]. This suggests that a support system involving sharing with another experienced mother who has undergone similar situations can potentially meet mothers' needs in terms of empathy and having a nonjudgmental listening ear.

Existing Technology-Based Interventions

Numerous studies identified help-seeking barriers among women at risk of psychological issues, namely lack of knowledge, practical barriers (eg, financial difficulties and work), and attitudinal barriers (eg, stigma) [21,26,27]. In a conservative multiracial country such as Singapore, traditional views and homebound confinement practices serve as additional help-seeking barriers. Therefore, technology-based interventions are an ideal alternative to increase local women's accessibility to professional help and improve maternal outcomes [28,29]. With other advantages such as improved health care accessibility, flexibility, individualized care, and privacy [30], many randomized controlled trials have begun adopting technology-based supportive interventions [31-34].

Most of the existing literature has established the effectiveness of various technology-based interventions on maternal outcomes [31-33]. A Web-based study consisting of weekly Web educational sessions and phone calls from a coach was shown to decrease the risk of PND in 90% of the mothers at 6 months postpartum [32]. Another recent study utilizing telephone-based support provided by midwives was found to be effective in reducing the risk of PND in at-risk women at 8 and 12 weeks postpartum [33]. Similarly, a Canadian-based study [35] involving weekly telephone-based peer support was also found to reduce the risk of PND and postnatal anxiety (PNA) among at-risk mothers at 12 and 24 weeks postpartum. Despite encouraging results on maternal outcomes, these studies were mainly conducted in Western countries [34-36], required a health care professional [34], did not sample at-risk mothers [34,36], or did not have their interventions administered immediately postpartum [35,36]. Additionally, a study by Sjoberg et al [37] revealed that the new generation of mothers preferred online peer support over face-to-face or online consultations with health care professionals. Therefore, there is a need to adopt a technology-based approach and paraprofessional peer support to effectively meet the desires of new generation mothers in Singapore.

Aim and Hypotheses

According to a recent review [30], an effective technology-based PND prevention intervention should be short term, be conducted immediately postpartum at an individual level, and target at-risk women instead of the general population. By incorporating all these elements, this study aims to examine the effectiveness of

XSL•FO
RenderX

a technology-based peer-support intervention program (PIP) among mothers at risk of PND during the early postpartum period (3 months postpartum). The secondary maternal outcomes examined were PNA, loneliness, and perceived social support.

The hypothesis is that compared with the control group, mothers in the intervention group will report significantly lower scores for PND, PNA, and loneliness and higher scores for perceived social support at 3 months postpartum.

Methods

Study Design

The protocol of this study has been published [38]. The study was conducted from May 2017 to May 2018 at a local tertiary hospital, National University Hospital, in Singapore. This study adopted a randomized controlled, single-blinded 2-group pretest and posttest design. The research assistant who was responsible for data collection was blinded to the group allocation of the participants. Participants were randomized to the intervention and control groups using opaque envelopes containing nonduplicated numbers (1-138). A set of 69 numbers was generated from a research randomizer [39] to determine the allocation of the intervention group. Specific details on the randomization process can be found in the study protocol [38].

Participants

Two samples of participants were recruited: (1) peer volunteers to facilitate the intervention program and (2) postnatal mothers at risk of PND. Peer volunteers were recruited through a blasting of emails to the study venue's working community and by word of mouth based on the following inclusion criteria: (1) mothers who were aged at least 21 years, (2) proficient in verbal and written English, (3) delivered a healthy baby in the past, (4) had a self-reported history of and recovery from PND, (5) had a mobile phone and were willing to share their number and call needy mothers as instructed by the research team, and (6) planned to stay in Singapore for the next 6 months after recruitment to administer the peer-support intervention. Peer volunteers were excluded if they had any physical or mental conditions that interfered with their ability to participate in the study.

Mothers at risk of PND were recruited from the postnatal wards of a local tertiary hospital immediately postbirth based on the following inclusion criteria: (1) were aged at least 21 years, (2) could read and speak English, (3) owned a mobile phone and were willing to share their number, (4) planned to stay in Singapore for 3 months postbirth, (5) delivered a healthy baby without birth defects and/or medical complications, and (6) had a baseline Edinburgh Postnatal Depression Scale (EPDS) score of more than or equal to 9. Mothers were excluded if (1) they had a history of existing psychiatric illness, cognitive impairment, and/or major medical conditions that could interfere with their abilities to participate in the study and/or (2) had a vacuum- or forceps-assisted delivery with a fourth-degree perineal tear.

Sample Size Calculation

On the basis of an independent sample *t* test to examine the differences between the control and intervention groups, assuming that there would be at least a medium Cohen effect size of 0.6 with 80% power and 0.05 significance level (2-sided), 47 participants were required in each group [40]. Factoring an attrition rate of 30%, a total of 138 participants (69 in each group) were recruited. A specification of the estimation of the effect size and the attrition rate is reported in the study protocol [38]. On the basis of a previous similar intervention study [35], 20 peer volunteers were recruited.

Intervention

Mothers in the control group received standard routine postnatal care by the hospital, which included in-hospital care by an obstetrician, nurses, and a lactation consultant. Posthospital discharge, the only continuity of care provided, was in the form of appointments with obstetricians or neonatologists and breastfeeding hotline numbers. In addition to this standard postnatal care by the hospital, mothers in the intervention group received a technology-based peer-support program for 4 weeks postpartum. Before the recruitment of postnatal mothers, the peer volunteers underwent a half-a-day training session by a psychiatrist. The training session inculcated roleplaying and strategizing to hone skills required in administering successful technology-based peer support. Volunteers were also taught to conduct appropriate referrals to health care professionals, should the need arise. A training booklet was prepared and given to each peer volunteer for future references. The PIP intervention involved correspondence with a trained peer volunteer at least once a week (for 4 weeks) via phone calls, emails, or mobile communication applications (eg, WhatsApp), depending on each mother's preference and convenience. During the introductory phone session, both sides shared their experience regarding emotional distress during the early postpartum period and extra efforts were made by the peer volunteer to build a strong relationship with the mother. Mothers were also informed that health care professionals would be notified if the mothers became too stressed during the correspondence. Subsequent sessions were individualized based on the unique needs of the mothers (eg, how to seek help from the family members and sharing one's feelings with their partners). Peer volunteers were encouraged to keep a free text journal of their conversations, and the intensity and duration of each correspondence were recorded in an activity log. More specification on the peer volunteer training and intervention process can be found in the research protocol [38].

Outcome Measures

The demographic data of the mothers were collected at the baseline using a self-reported questionnaire. Symptom scores for PND (primary outcome), PNA, loneliness score, and scores for perceived social support (secondary outcomes) were measured using a self-reported face-to-face questionnaire at the baseline and via Web-based questionnaires at the 4th and 12th week postpartum. The internal consistency of each instrument was measured using Cronbach alpha.

XSL•FO
RenderX

PND symptom score was measured using the 10-item EPDS [41]. The total score ranges from 0 to 30, with a higher score indicating a higher risk of PND. On the basis of previous trials [35,42], a recommended cut-off score of 9 was used to screen mothers at risk of PND and a score of more than 12 as a probable diagnosis for PND. The internal consistencies at baseline, 1 month postpartum, and 3 months postpartum were 0.59, 0.87, and 0.86, respectively.

The 9-item Patient Health Questionnaire (PHQ-9) [43] was extracted from the full PHQ used to diagnose and measure the severity of major depression. The total score ranges from 0 to 27, with a higher score indicating a higher severity of PND. The Cronbach alpha values for this study were 0.83, 0.86, and 0.92 for baseline, 1 month postpartum, and 3 months postpartum, respectively.

The State-Trait Anxiety Inventory (STAI) [44], a 40-item questionnaire using a 4-point Likert scale, was used to measure maternal anxiety. The total score ranges from 40 to 160, with a higher score suggesting a higher severity of anxiety. The STAI had high internal consistencies of 0.96, 0.97, and 0.98 for baseline, 1 month postpartum, and 3 months postpartum, respectively.

Loneliness score was measured using the 10-item University of California, Los Angeles Loneliness Scale (ULS) [45]. Items are rated on a 4-point Likert scale, with the total score ranging from 10 to 40. A higher score represents a higher level of loneliness. The ULS had high internal consistencies of 0.96, 0.97, and 0.97 at baseline, 1 month postpartum, and 3 months postpartum, respectively, in this study.

The Perceived Social Support for Parenting (PSSP) instrument developed by Leerkes and Crockenberg [46] was used to measure maternal satisfaction of the social support received from partners and others during the postpartum period. The instrument had a 5-point Likert scale and 2 4-item subparts: (1) social support received from the partner and (2) social support received from others. The total score ranges from 5 to 40, with a higher score implying a higher level of satisfaction of the received social support. The Cronbach alpha values for baseline, 1 month postpartum, and 3 months postpartum were 0.93, 0.89, and 0.92, respectively. Detailed descriptions of the instruments can be found in the study protocol [38].

Data Collection

Nurses in the postnatal wards of the study hospital assisted in identifying suitable healthy mothers for the study. After being screened for eligibility, participants were given a thorough briefing on the study's purpose and details. Mothers each signed a consent form upon agreeing to participate and proceeded to complete the EPDS questionnaire as a baseline measure. Mothers with EPDS scores of 9 and above were then randomly assigned to either the intervention or control group. Mothers in the intervention group were matched to a peer volunteer who contacted these mothers 2 to 3 days after their discharge from the hospital. Correspondence between peer volunteers and their paired mothers occurred at least once a week for 4 weeks via phone calls, emails, and mobile communication applications. Mothers in both the control and intervention groups concurrently proceeded with their standard hospital care and postnatal follow-ups. Weekly text reminders were sent by the research assistant nearing the 4th and 12th weeks to remind participants to complete the upcoming Web-based questionnaires. After 4 and 12 weeks, a research assistant who was blinded to the allocation of the participants forwarded a text message containing a Web link to the follow-up questionnaires to the participants and requested them to complete the questionnaires as soon as possible. An elaboration on the data collection process can be found in the research protocol [38].

Data Analysis

All analyses were conducted using the IBM SPSS version 24.0 software (International Business Machines Corporation) with the statistical significance set at $P<.05$. The analysis was performed on the intention-to-treat population. Descriptive statistics were presented as mean (SD) and n (%) for continuous and categorical variables, respectively. A repeated measures analysis using a linear mixed model was used to assess the trend of the examined outcomes over the period of 3 months. The random effects were on subjects, and the fixed effects were on the following factors: age, marital status, antenatal class attendance, gender of infant, confinement period, and group and time interaction. The interaction effect of group and time was used to assess the differences over time. A general linear model was performed to compare the difference between intervention and control groups at 1 month and 3 months postpartum for the outcomes of EPDS, PSSP, ULS, and STAI by adjusting for the baseline. On the basis of previous literature [5,9,19,2], the following demographic factors were also adjusted in the general linear model for all outcomes: age, marital status, antenatal class attendance, gender of infant, and confinement period.

Ethical Considerations

Ethics approval from the National Health Group Domain Specific Review Board (Ref number: NHG DSRB: 2017/00185) was obtained before the commencement of the study. All participants were briefed in detail on the research process before their written consents were obtained. Participation was strictly voluntary, and the participants were guaranteed anonymity and informed of their rights to withdraw at any time without consequences.

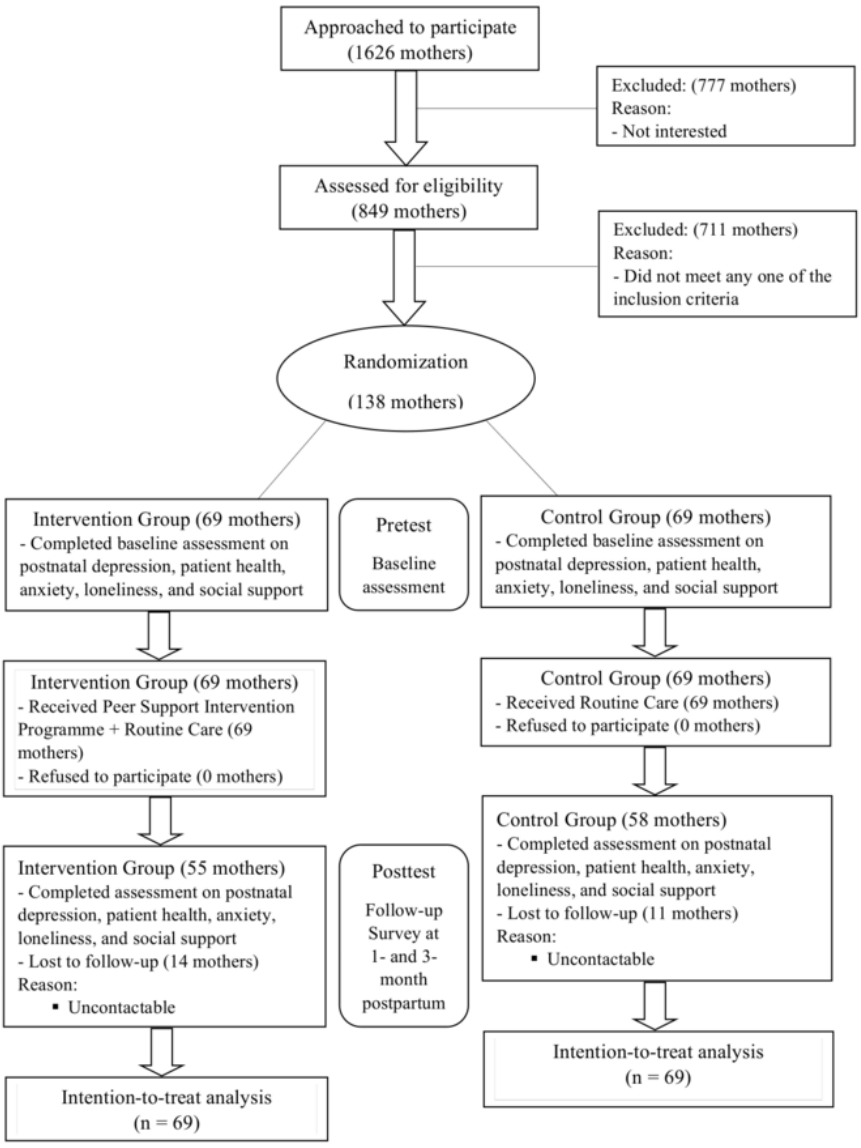
Results

Participants Data

Figure 1 shows the Consolidated Standards of Reporting Trials flowchart of the study. A total of 138 mothers were recruited and randomized to the control (n=69) and intervention (n=69) groups. The baseline demographic data of all the participants are presented in Table 1. The participants had a mean age of 32.1 years (SD 4.35, range 23-43). Forty-two percent (53/138) of the participants were Chinese 96.4% (133/138) were married, 60.1% (83/138) had a university degree, and 67.6% (92/138) had a monthly household income of more than SGD $3000.

XSL•FO
RenderX

Figure 1. Consolidated Standards of Reporting Trials (CONSORT) flowchart of the study.



XSL•FO
RenderX

Table 1. Comparison of demographic characteristics between the 2 groups (N=138).

Characteristics	Total	Intervention group (n=69)	Control group (n=69)
Age (years), mean (SD); range	32.05 (4.35); 23-43	32.26 (3.90); 24-43	31.84 (4.77); 23-41
Ethnicity, n (%)			
Chinese	58 (42.0)	30 (43)	28 (41)
Malay	47 (34.1)	26 (38)	21 (30)
Indian	16 (11.6)	5 (7)	11 (16)
Others	17 (12.3)	8 (12)	9 (13)
Marital status, n (%)			
Married	133 (96.4)	66 (96)	67 (97)
Not married	5 (3.6)	3 (4)	2 (3)
Highest education level, n (%)			
Secondary and below	15 (10.9)	9 (13)	6 (9)
Preuniversity	40 (29.0)	17 (25)	23 (33)
University	83 (60.1)	43 (62)	40 (58)
Monthly household income (SGD $), n (%)			
<3000	44 (32.4)	24 (35)	20 (29)
>$3000	92 (67.6)	44 (65)	48 (71)
Attendance for prenatal course, n (%)			
Yes	40 (29.0)	26 (38)	14 (20)
No	98 (71.0)	43 (62)	55 (80)
Type of birth, n (%)			
Normal vaginal delivery	75 (54.7)	30 (43)	45 (66)
Assisted delivery	14 (10.2)	11 (16)	3 (4)
Cesarean section	48 (35.0)	28 (41)	20 (30)
Baby's gender, n (%)			
Twins	2 (1.4)	1 (1)	1 (1)
Male	70 (50.8)	31 (45)	39 (57)
Female	66 (47.8)	37 (54)	29 (42)
Baby's birth order, n (%)			
First	81 (58.7)	46 (67)	35 (51)
Second	32 (23.2)	12 (17)	20 (29)
Others	25 (18.1)	11 (16)	14 (20)
First-time mother, n (%)			
Yes	81 (58.7)	46 (67)	35 (51)
No	57 (41.3)	23 (33)	34 (49)
Maternity leave, n (%)			
Yes	96 (69.6)	53 (77)	43 (62)
No	42 (30.4)	16 (23)	26 (38)
Confinement period, n (%)			
Yes	94 (87.0)	48 (89)	46 (85)
No	14 (13.0)	6 (11)	8 (15)
Baby's feeding method, n (%)			
Breastfeeding	46 (42.2)	21 (39)	25 (45)

XSL•FO
RenderX

Characteristics	Total	Intervention group (n=69)	Control group (n=69)
Bottle feeding	6 (5.5)	3 (6)	3 (6)
Both	57 (41.3)	30 (55)	27 (49)

Mothers mostly did not attend antenatal classes, had a normal vaginal delivery, followed a confinement period, and were first-time mothers. Follow-up assessments at 1 month postpartum were completed for all mothers in both the control (n=69) and intervention (n=69) groups. At 3 months postpartum, follow-up assessments were completed for 55 mothers from the intervention group (79%, 55/69) and 58 mothers from the control group (84%, 58/69). The overall attrition rate was 18.1%. As the intention-to-treat analysis was used, the outcome data for all 138 mothers were analyzed. Mothers who dropped out or did not provide outcome data at 3 months were still included in the main analysis (as the linear mixed model models data points rather than subjects). For sensitivity of the effect of the missing values on the results, a best-case and worst-case scenario was performed.

On the basis of previous literature, the outcome measures were adjusted for age, antenatal class attendance, gender of infant, and confinement period. Symptom scores for PND were measured using the EPDS and the PHQ.

The total scores for EPDS were lower in the intervention than the control group at both 1 month and 3 months postpartum. However, the difference of scores between groups was not statistically significant at 1 month even after adjusting for covariates (unadjusted: difference [*d*]=–0.91; 95% CI −2.5 to 0.6; *P*=.25; adjusted: *d*=−1.02; 95% CI −2.7 to 0.6; *P*=.23, but it was statistically significant at 3 months postpartum before and after adjustment (unadjusted: *d*=−1.77; 95% CI −3.5 to 0.0; *P*=.04; adjusted: *d*=−2.11; 95% CI −4.0 to −0.3; *P*=.03). On the basis of the linear mix model, there was also a statistically significant difference in the change of the total adjusted EPDS scores from baseline to 3 months postpartum for the intervention over the control group (*d*=−1.16; 95% CI −2.0 to −0.4; *P*=.004).

The total PHQ scores for the intervention group were lower than those of the control group at both 1 month and 3 months postpartum. At 1 month and 3 months, the unadjusted difference in scores between groups were statistically significant (first month: *d*=−1.80; 95% CI −3.3 to −0.3; *P*=.02; third month: *d*=−1.9; 95% CI −3.7 to −0.1; *P*=.04). However, the difference of scores between groups at both time points was no longer statistically significant after adjusting for covariates (first month: *d*=−1.59; 95% CI −3.3 to 0.1; *P*=.06; third month: *d*=−1.49; 95% CI −3.4 to 0.4; *P*=.11). Additionally, the difference in change of the total adjusted PHQ score for the intervention group from baseline to 3 months postpartum was also statistically significant (*d*=−1.00; 95% CI −1.9 to −0.1; *P*=.03).

At both 1 month and 3 months postpartum, the total STAI scores were lower in the intervention than the control group. The difference of scores between groups at 1 month and 3 months was not statistically significant even after adjusting for covariates (first month unadjusted: *d*=−3.63; 95% CI −10.7 to 3.5; *P*=.31; first month adjusted: *d*=−2.45; 95% CI −9.9 to 5.0; *P*=.52; 3 months unadjusted: *d*=−8.61; 95% CI −17.2 to 0.0; *P*=.05; 3 months adjusted: *d*=−7.89; 95% CI −16.4 to 0.7; *P*=.07). However, the difference in change of adjusted STAI scores for the intervention group was statistically significant across the 3 months (*d*=−4.16; 95% CI −7.9 to −0.4; *P*=.03).

The total scores for ULS were higher in the control group than in the intervention group at both 1 month and 3 months postpartum. At both time points, the unadjusted and adjusted mean differences of scores between groups were not statistically significant (first month unadjusted: *d*=−2.14; 95% CI −6.3 to 2.1; *P*=.32; first month adjusted: *d*=−2.45; 95% CI −7.0 to 2.1; *P*=.29; 3 months unadjusted: *d*=−3.90; 95% CI −8.2 to 0.4; *P*=.08; 3 months adjusted: *d*=−3.43; 95% CI −8.0 to 1.1; *P*=.14). There was also no statistically significant difference in change of ULS scores for the intervention group across 3 months even after adjustment (*d*=−2.16; 95% CI −4.4 to 0.0; *P*=.06).

At 1 month and 3 months postpartum, the control group had slightly higher PSSP scores than the control group. In addition, the difference of scores between groups was not statistically significant at both 1 month and 3 months even after adjustment (first month unadjusted: *d*=−2.11; 95% CI −5.2 to 1.0; *P*=.18; first month adjusted: *d*=−0.86; 95% CI −3.9 to 2.1; *P*=.57; 3 months unadjusted: *d*=1.09; 95% CI −2.9 to 5.1; *P*=.59; 3 months adjusted: *d*=−0.53; 95% CI −4.7 to 3.6; *P*=.80). The difference in change of total PSSP scores of the intervention group from baseline to 3 months postpartum was also not statistically significant (*d*=0.86; 95% CI −0.9 to 2.6; *P*=.33).

To account for potential type 1 error for multiple outcomes, the resultant *P* values were inflated by the number of outcomes analyzed, which is a factor of 5. Upon doing so, only the adjusted change of EPDS scores from baseline to the third month remained significant.

Table 2 shows the changes in scores for all outcome variables, at each time point, including scores that were not adjusted for, whereas Table 3 shows the differences in change of outcome scores in the intervention group (over the control group) across 3 months. Although a change in ULS and PSSP scores between groups was not statistically significant across 3 months postpartum, Figure 2 shows good overall trending for all maternal outcomes from the baseline to 3 months postpartum. On the basis of the graph, mothers who received the intervention had better maternal outcome scores than the control group by the end of the third month.

A sensitivity analysis was also performed using the best and worst scores to replace missing values. *P* values from the best-case and worst-case analysis were then compared with the *P* values in the main analysis. The best-case analysis showed similar significant results as the main analysis. Sensitivity results are attached in Multimedia Appendices 1 and 2.

XSL•FO
RenderX

Table 2. Change in outcome scores between the intervention (I) and control (C) groups among mothers at 1 month and 3 months postpartum based on the general linear model.

Outcome variable	1 month								3 months							
	I, mean (SD), range	n[a]	C, mean (SD), range	n[a]	Unadjusted		Adjusted[b]		I, mean (SD), range	n[a]	C, mean (SD), range	n[a]	Unadjusted		Adjusted[b]	
					I–C (95% CI)	*P*	I–C (95% CI)	*P*					I–C (95% CI)	*P*	I–C (95% CI)	*P*
Postpartum depression (EPDS[c])	11.4 (2.0), 7.5 to 15.3	56	12.4 (2.1), 8.4 to 16.5	58	−0.91 (−2.5 to 0.6)	.25	−1.02 (−2.7 to 0.6)	.23	9.8 (2.2), 5.5 to 14.3	54	12.0 (2.3), 7.5 to 16.5	57	−1.77 (−3.5 to 0.0)	.04[d]	−2.11 (−4.0 to −0.3)	.03[d]
Postpartum depression (PHQ[e])	4.6 (2.0), 0.6 to 8.7	56	6.2 (2.1), 2.1 to 10.4	58	−1.80 (−3.3 to −0.3)	.02[d]	−1.59 (−3.3 to 0.1)	.06	5.5 (2.2), 1.0 to 10.0	54	7.0 (2.3), 2.4 to 11.6	57	−1.90 (−3.7 to −0.1)	.04[d]	−1.49 (−3.4 to 0.4)	.11
Postpartum anxiety (STAI[f])	85.2 (9.9), 65.4 to 104.9	52	87.6 (9.9), 67.8 to 107.4	53	−3.63 (−10.7 to 3.5)	.31	−2.45 (−9.9 to 5.0)	.52	79.8 (11.3), 57.4 to 102.3	50	87.7 (11.3), 65.3 to 110.2	52	−8.61 (−17.2 to 0.0)	.05	−7.89 (−16.4 to 0.7)	.07
Loneliness (ULS[g])	41.6 (5.5), 30.7 to 52.4	54	44.0 (5.6), 32.8 to 55.2	57	−2.14 (−6.3 to 2.1)	.31	−2.45 (−7.0 to 2.1)	.29	39.5 (5.4), 28.7 to 50.3	53	42.9 (5.6), 31.8 to 54.1	56	−3.90 (−8.2 to 0.4)	.08	−3.43 (−8.0 to 1.1)	.14
Perceived social support (PSSP[h])	32.3 (2.6), 27.2 to 37.5	20	33.2 (2.6), 28.0 to 38.4	31	−2.11 (−5.2 to 1.0)	.18	−0.86 (−3.9 to 2.1)	.57	35.4 (4.4), 26.5 to 44.3	20	35.9 (4.3), 27.2 to 44.6	27	1.09 (−2.9 to 5.1)	.59	−0.53 (−4.7 to 3.6)	.78

[a]n values for adjusted analysis.

[b]Adjusted estimates were obtained from general linear models after being adjusted for baseline, age, marital status, antenatal class attendance, baby's gender, and confinement period.

[c]EPDS: Edinburgh Postnatal Depression Scale.

[d]Significant *P* value <.05.

[e]PHQ: Patient Health Questionnaire.

[f]STAI: State-Trait Anxiety Inventory.

[g]ULS: University of California, Los Angeles Loneliness Scale.

[h]PSSP: Perceived Social Support for Parenting.

Table 3. Differences in the change of outcome scores across 3 months between intervention and control groups based on a linear mixed model.

Outcome variable	Trend difference (reference control group)			
	Unadjusted estimate (95% CI)	*P* value	Adjusted estimate[a] (95% CI)	*P* value
Postpartum depression (EPDS[b])	−0.90 (−1.7 to −0.6)	.02[c]	−1.16 (−2.0 to −0.4)	.004[b]
Postpartum depression (PHQ[d])	−1.02 (−1.9 to −0.2)	.02[c]	−1.00 (−1.9 to −0.1)	.03[b]
Postpartum anxiety (STAI[e])	−3.17 (−6.9 to 0.6)	.09	−4.16 (−7.9 to −0.4)	.03[b]
Loneliness (ULS[f])	−2.19 (−4.3 to −0.1)	.04	−2.16 (−4.4 to 0.0)	.05
Perceived social support (PSSP[g])	0.47 (−1.2 to 2.2)	.58	0.86 (−0.9 to 2.6)	.33

[a]Adjusted estimates were obtained from linear mixed models after being adjusted for baseline, age, marital status, antenatal class attendance, baby's gender, and confinement period.

[b]EPDS: Edinburgh Postnatal Depression Scale.

[c]Significant *P* value <.05.

[d]PHQ: Patient Health Questionnaire.

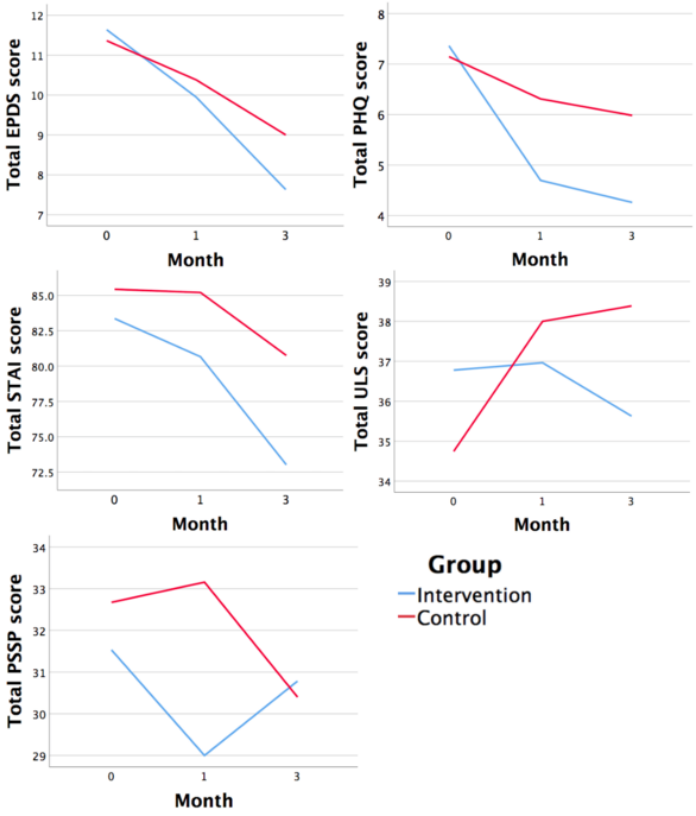
[e]STAI: State-Trait Anxiety Inventory.

[f]ULS: University of California, Los Angeles Loneliness Scale.

[g]PSSP: Perceived Social Support for Parenting.

XSL•FO
RenderX

Figure 2. Trend comparison of mean outcome scores between groups across 3 months postpartum. EPDS: Edinburgh Postnatal Depression Scale; PHQ: Patient Health Questionnaire; PSSP: Perceived Social Support for Parenting; STAI: State-Trait Anxiety Inventory; ULS: University of California, Los Angeles Loneliness Scale.



Discussion

Evaluation of Findings

This study examined the effectiveness of a technology-based PIP among mothers at risk of PND. According to the score trend, the intervention group scored better than the control group for all maternal outcomes at both 1 month and 3 months postpartum, but only the difference in the EPDS scores between groups was shown to be statistically significant. This suggests that compared with mothers who only received routine hospital care, the PIP was generally effective in reducing the risks of PND, PNA, and loneliness and in increasing perceived social support received by the end of 3 months postpartum. Mothers further expressed their satisfaction with the intervention in a separate qualitative interview [47].

Although scores between groups were only significant for the EPDS, the EPDS and PHQ scores of the intervention group were still observably lower than those of the control group at 3 months postpartum, indicating the effectiveness of the peer-support program in reducing PND among at-risk mothers. The results were similar to previous studies [35,36] in which telephone-based peer support was found to reduce PND among depressed and at-risk mothers at 3 months postpartum. Our results are also constant with studies linking higher risks of PND to reduce perceived social support, few social networks, and close relationships [20,42,48]. Additionally, the study's results correspond with large-scale reviews [30,49] that reported how short-term, individually based, technology-based interventions targeting at-risk women that were conducted immediately postpartum were the most effective in promoting positive maternal outcomes.

Many studies also established the high correlation and comorbidity between PND and PNA [50,51]; therefore, improvements in PND scores may inevitably lead to improvements in PNA scores. This corresponds with our results, which showed a significant change of EPDS, PHQ, and STAI scores across 3 months between groups. Similar results were found in Dennis et al's peer-support study [35] where a positive trend in favor of the intervention group was found for maternal PNA. However, little empirical attention has been given to maternal anxiety as a standalone disorder [52]; hence, further research is required to examine effective preventive measures against PNA.

From our results, another notable observation was the sudden drop in PNA scores for both groups between 1 month and 3 months postpartum compared with the first month. Despite the cessation of the PIP at 1 month, it is possible that, by then, parents would have adapted to their new roles and gained sufficient parenting confidence and enhanced self-efficacy, which, in turn, reduced the risks of PNA. This is supported by a Finnish study [53], which reported that having high maternal parenting efficacy scores at 1 week postpartum reduces PNA at 1 month postpartum, and Kohlhoff's study [54], which revealed an inverse correlation between parenting self-efficacy and PNA and PND. Another study [55] reporting the long-term sustainability of parental self-efficacy after the termination of

XSL•FO
RenderX

a 6-week intervention could also be a plausible explanation to the sudden improvement in PNA scores in this study. Owing to the lack of studies examining the direct effects of paraprofessional peer support on parenting self-efficacy, further research is needed to validate these findings.

In terms of loneliness, there was an increase in loneliness scores for both groups from baseline to 1 month postpartum, with the control group having a steeper increase than the intervention group. Although loneliness scores continue to increase for mothers in the control group from 1 month to 3 months, loneliness scores for mothers in the intervention group decreased. This is evident that although the PIP was not able to fully relieve the sense of loneliness among mothers during the postpartum period, it still buffered mothers against loneliness compared with those who did not receive the PIP. A similar result was also noted in Dennis et al's study [35], where although no positive trend was noted for loneliness at 3 months postpartum, the change in PND scores was significant. This conflicts with most literature that reported the predictive effects of loneliness on PND [3,20]. The trend observation was supported by a study that adopted a video conference method to examine loneliness [34]. Mothers reported that online face-to-face interactions were almost equivalent to having a physical presence and it facilitated rapport building [34]. This feature was lacking in our study. Furthermore, other face-to-face peer-support intervention studies also reported reduced feelings of loneliness among mothers [6,56]. This indicates the importance of a face-to-face element in technology-based support programs, which can better facilitate the sharing of experiences and alleviate feelings of loneliness. Another plausible cause is that at the end of the confinement period at 1 month, mothers were no longer physically isolated and could actively seek out family and friends for accompaniment [57]. Although the effectiveness of the PIP in mitigating loneliness might not be equivalent to the physical presence of a family member or friend, it can prevent the escalation of loneliness in mothers during the confinement period [35].

Mothers who did not receive peer support perceived a slight increase in social support at 1 month and a drastic decline in social support at 3 months postpartum, whereas mothers who received peer support only perceived a decrease in social support upon the termination of the intervention at 1 month to 3 months postpartum. This indicates the long-term effectiveness of the PIP in providing mothers with social support and that the termination of the program is a loss of an important source of social support to them. Over reliance on the PIP may result in a sudden drop in scores at 1 month, but the PIP also equipped them with help-seeking skills that might have caused an increase in scores at the third month. These findings correspond with other peer-support studies that reported an increased sense of perceived support in the intervention group [31,58] and studies linking the lack of social support to increased depressive symptoms [5,9,59]. However, our results conflict with other studies that reported an inverse correlation between social support and loneliness [60-62]. High social support scores of mothers in the control group for the first month could be due to the availability of instrumental and emotional support by family members, a partner, or a confinement nanny during the confinement period, whereas the decrease in scores after one month can be attributed to the end of paternal leave and the confinement period. During this time, mothers lose instrumental support from their partners and are abruptly entrusted with infant care responsibilities, which may result in an overwhelming sense of loss [63]. Social support is a three-dimensional construct consisting of emotional, informational, and instrumental support [19]. Peer-support interventions may fulfill the emotional aspect and, to a certain extent, informational support, but mothers reported higher needs for instrumental support during the postpartum period [23]. Therefore, for optimal outcomes, the PIP should be administered concurrently with instrumental help to provide well-rounded social support to mothers.

Limitations and Recommendation for Future Studies

To the best of our knowledge, this is the first technology-based peer-support study based in Asia that showed a preventive effect against PND. Therefore, it is a valuable contribution to ongoing region-specific research on the prevention of PND. However, a major limitation of this study is that it was a single-site study targeting only English-speaking mothers. Future studies can consider integrating more languages to cater to minority groups. Another limitation is that the intervention was only administered during the postpartum period. Considering that antenatal depression is a main predictor of PND, future studies should examine the effectiveness of such interventions during the perinatal period. Additionally, maternal outcomes included in this study were limited and infant outcomes were lacking. Given that most maternal outcomes are interrelated, other outcomes such as parenting self-efficacy and parenting satisfaction as well as an evaluation of the effectiveness of the program on infant development can be included in subsequent studies. Finally, an evaluation of the cost-effectiveness of the PIP will provide a holistic view on the effectiveness of this intervention.

Conclusions

This randomized controlled trial demonstrated the effectiveness of a technology-based peer-support program in reducing maternal PND. Besides receiving standard postnatal care, additional participation in the PIP was shown to improve the general well-being of mothers at the end of 3 months postpartum. This study adds value to the use of technology and trained paraprofessionals in combating PND among new mothers. Although future rigorous trials are needed to evaluate the effectiveness of the PIP further, health care professionals can involve paraprofessionals such as family members in supporting new mothers during the stressful postpartum period. This may enhance not only maternal outcomes but also the future well-being of the family, thus creating positive childbirth experiences for mothers.

XSL•FO
RenderX

Acknowledgments

The authors are thankful to the National University Health System Collaborative Clinician Research grant (ref #T1- NUHS O-CRG 2016, Oct 22) for funding this study. The authors also send their special thanks to the National University Health System, Medical Publications Support Unit, for editing the language and format of this paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Sensitivity analysis (best-case and worst-case scenario) using the general linear model.
[DOCX File, 18KB - jmir_v21i8e12410_app1.docx]

Multimedia Appendix 2
Sensitivity analysis (best-case and worst-case scenario) using the linear mixed model.
[DOCX File, 14KB - jmir_v21i8e12410_app2.docx]

Multimedia Appendix 3
CONSORT - EHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 929KB - jmir_v21i8e12410_app3.pdf]

References

1. Firoz T, Chou D, von Dadelszen P, Agrawal P, Vanderkruik R, Tunçalp O, Maternal Morbidity Working Group. Measuring maternal health: focus on maternal morbidity. Bull World Health Organ 2013 Oct 1;91(10):794-796 [FREE Full text] [doi: 10.2471/BLT.13.117564] [Medline: 24115804]
2. Dennis CL. Detection, prevention, and treatment of postpartum depression. In: Stewart DE, Robertson E, Dennis CL, Grace SL, Wallington T, editors. Postpartum Depression: Literature Review of Risk Factors and Interventions. Geneva. Switzerland: World Health Organisation; 2003.
3. O'hara MW, Swain AM. Rates and risk of postpartum depression—a meta-analysis. Int Rev Psychiatry 2009 Jul 11;8(1):37-54. [doi: 10.3109/09540269609037816]
4. VanderKruik R, Barreix M, Chou D, Allen T, Say L, Cohen LS, Maternal Morbidity Working Group. The global prevalence of postpartum psychosis: a systematic review. BMC Psychiatry 2017 Jul 28;17(1):272 [FREE Full text] [doi: 10.1186/s12888-017-1427-7] [Medline: 28754094]
5. O'Hara MW. Postpartum Depression: Causes and Consequences. First Edition. New York: Springer; 1995.
6. McLeish J, Redshaw M. Mothers' accounts of the impact on emotional wellbeing of organised peer support in pregnancy and early parenthood: a qualitative study. BMC Pregnancy Childbirth 2017 Jan 13;17(1):28 [FREE Full text] [doi: 10.1186/s12884-017-1220-0] [Medline: 28086827]
7. Holopainen D. The experience of seeking help for postnatal depression. Aust J Adv Nurs 2002;19(3):39-44. [Medline: 12002628]
8. Choi P, Henshaw C, Baker S, Tree J. Supermum, superwife, supereverything: performing femininity in the transition to motherhood. J Reprod Infant Psyc 2005 May;23(2):167-180. [doi: 10.1080/02646830500129487]
9. Habel C, Feeley N, Hayton B, Bell L, Zelkowitz P. Causes of women's postpartum depression symptoms: men's and women's perceptions. Midwifery 2015 Jul;31(7):728-734. [doi: 10.1016/j.midw.2015.03.007] [Medline: 25921442]
10. Goodman JH. Paternal postpartum depression, its relationship to maternal postpartum depression, and implications for family health. J Adv Nurs 2004 Jan;45(1):26-35. [doi: 10.1046/j.1365-2648.2003.02857.x] [Medline: 14675298]
11. Logsdon MC, Wisner KL, Pinto-Foltz MD. The impact of postpartum depression on mothering. J Obstet Gynecol Neonatal Nurs 2006;35(5):652-658. [doi: 10.1111/j.1552-6909.2006.00087.x] [Medline: 16958723]
12. Lovejoy MC, Graczyk PA, O'Hare E, Neuman G. Maternal depression and parenting behavior: a meta-analytic review. Clin Psychol Rev 2000 Aug;20(5):561-592. [doi: 10.1016/S0272-7358(98)00100-7] [Medline: 10860167]
13. Murray L, Fiori-Cowley A, Hooper R, Cooper P. The impact of postnatal depression and associated adversity on early mother-infant interactions and later infant outcome. Child Dev 1996 Oct;67(5):2512-2526. [doi: 10.2307/1131637] [Medline: 9022253]
14. Stein A, Pearson RM, Goodman SH, Rapa E, Rahman A, McCallum M, et al. Effects of perinatal mental disorders on the fetus and child. Lancet 2014 Nov 15;384(9956):1800-1819. [doi: 10.1016/S0140-6736(14)61277-0] [Medline: 25455250]
15. Letourneau NL, Dennis CL, Benzies K, Duffett-Leger L, Stewart M, Tryphonopoulos PD, et al. Postpartum depression is a family affair: addressing the impact on mothers, fathers, and children. Issues Ment Health Nurs 2012 Jul;33(7):445-457. [doi: 10.3109/01612840.2012.673054] [Medline: 22757597]

XSL•FO
RenderX

16. Soe NN, Wen DJ, Poh JS, Li Y, Broekman BF, Chen H, et al. Pre- and post-natal maternal depressive symptoms in relation with infant frontal function, connectivity, and behaviors. PLoS One 2016;11(4):e0152991 [FREE Full text] [doi: 10.1371/journal.pone.0152991] [Medline: 27073881]
17. Beck CT. The effects of postpartum depression on child development: a meta-analysis. Arch Psychiatr Nurs 1998 Feb;12(1):12-20. [doi: 10.1007/s00737-003-0024-6] [Medline: 9489170]
18. Oberlander TF, Reebye P, Misri S, Papsdorf M, Kim J, Grunau RE. Externalizing and attentional behaviors in children of depressed mothers treated with a selective serotonin reuptake inhibitor antidepressant during pregnancy. Arch Pediatr Adolesc Med 2007 Jan;161(1):22-29. [doi: 10.1001/archpedi.161.1.22] [Medline: 17199063]
19. Chojenta C, Loxton D, Lucke J. How do previous mental health, social support, and stressful life events contribute to postnatal depression in a representative sample of Australian women? J Midwifery Womens Health 2012;57(2):145-150. [doi: 10.1111/j.1542-2011.2011.00140.x] [Medline: 22432486]
20. Beck CT. Predictors of postpartum depression: an update. Nurs Res 2001;50(5):275-285. [doi: 10.1097/00006199-200109000-00004] [Medline: 11570712]
21. Dennis CL, Chung-Lee L. Postpartum depression help-seeking barriers and maternal treatment preferences: a qualitative systematic review. Birth 2006 Dec;33(4):323-331. [doi: 10.1111/j.1523-536X.2006.00130.x] [Medline: 17150072]
22. Mauthner NS. Postnatal depression: how can midwives help? Midwifery 1997 Dec;13(4):163-171. [doi: 10.1016/S0266-6138(97)80002-2] [Medline: 9511683]
23. Negron R, Martin A, Almog M, Balbierz A, Howell EA. Social support during the postpartum period: mothers' views on needs, expectations, and mobilization of support. Matern Child Health J 2013 May;17(4):616-623 [FREE Full text] [doi: 10.1007/s10995-012-1037-4] [Medline: 22581378]
24. Cohen S, Wills TA. Stress, social support, and the buffering hypothesis. Psychol Bull 1985 Sep;98(2):310-357. [doi: 10.1037/0033-2909.98.2.310] [Medline: 3901065]
25. Cohen S, Underwood LG, Gottlieb BH, Fetzer I, editors. Social Support Measurement and Intervention: A Guide for Health and Social Scientists. New York: Oxford University Press; 2000.
26. Fonseca A, Gorayeb R, Canavarro MC. Women s help-seeking behaviours for depressive symptoms during the perinatal period: socio-demographic and clinical correlates and perceived barriers to seeking professional help. Midwifery 2015 Dec;31(12):1177-1185. [doi: 10.1016/j.midw.2015.09.002] [Medline: 26433622]
27. O'Mahen HA, Flynn HA. Preferences and perceived barriers to treatment for depression during the perinatal period. J Womens Health (Larchmt) 2008 Oct;17(8):1301-1309. [doi: 10.1089/jwh.2007.0631] [Medline: 18816202]
28. Fotheringham MJ, Owies D, Leslie E, Owen N. Interactive health communication in preventive medicine: internet-based strategies in teaching and research. Am J Prev Med 2000 Aug;19(2):113-120. [doi: 10.1016/S0749-3797(00)00188-4] [Medline: 10913902]
29. Barrera AZ, Nichols AD. Depression help-seeking attitudes and behaviors among an internet-based sample of Spanish-speaking perinatal women. Rev Panam Salud Publica 2015 Mar;37(3):148-153. [Medline: 25988251]
30. Fonseca A, Pereira M, Araújo-Pedrosa A, Gorayeb R, Ramos MM, Canavarro MC. Be a mom: formative evaluation of a web-based psychological intervention to prevent postpartum depression. Cogn Behav Pract 2018 Nov;25(4):473-495. [doi: 10.1016/j.cbpra.2018.02.002]
31. Shorey S, Lau YY, Dennis CL, Chan YS, Tam WW, Chan YH. A randomized-controlled trial to examine the effectiveness of the 'Home-but not Alone' mobile-health application educational programme on parental outcomes. J Adv Nurs 2017 Sep;73(9):2103-2117. [doi: 10.1111/jan.13293] [Medline: 28276086]
32. Danaher BG, Milgrom J, Seeley JR, Stuart S, Schembri C, Tyler MS, et al. MomMoodBooster web-based intervention for postpartum depression: feasibility trial results. J Med Internet Res 2013 Nov 4;15(11):e242 [FREE Full text] [doi: 10.2196/jmir.2876] [Medline: 24191345]
33. Posmontier B, Neugebauer R, Stuart S, Chittams J, Shaughnessy R. Telephone-administered interpersonal psychotherapy by nurse-midwives for postpartum depression. J Midwifery Womens Health 2016 Jul;61(4):456-466. [doi: 10.1111/jmwh.12411] [Medline: 26970401]
34. Lindberg I, Christensson K, Ohrling K. Parents' experiences of using videoconferencing as a support in early discharge after childbirth. Midwifery 2009 Aug;25(4):357-365. [doi: 10.1016/j.midw.2007.06.002] [Medline: 17935844]
35. Dennis CL, Hodnett E, Kenton L, Weston J, Zupancic J, Stewart DE, et al. Effect of peer support on prevention of postnatal depression among high risk women: multisite randomised controlled trial. Br Med J 2009 Jan 15;338:3064 [FREE Full text] [doi: 10.1136/bmj.a3064] [Medline: 19147637]
36. Letourneau N, Secco L, Colpitts J, Aldous S, Stewart M, Dennis CL. Quasi-experimental evaluation of a telephone-based peer support intervention for maternal depression. J Adv Nurs 2015 Jul;71(7):1587-1599. [doi: 10.1111/jan.12622] [Medline: 25705786]
37. Sjöberg M, Lindgren S. Challenging the roles of 'skilled' professionals and 'risky' young mothers: peer support, expertise, and relational patterns in Facebook groups. J Technol Hum Serv 2017 Sep 13;35(3):247-270. [doi: 10.1080/15228835.2017.1367350]

XSL•FO
RenderX

38. Shorey S, Chee C, Chong YS, Ng ED, Lau Y, Dennis CL. Evaluation of technology-based peer support intervention program for preventing postnatal depression: protocol for a randomized controlled trial. JMIR Res Protoc 2018 Mar 14;7(3):e81 [FREE Full text] [doi: 10.2196/resprot.9416] [Medline: 29540338]
39. Research Randomizer. 2018. URL: https://www.randomizer.org [accessed 2018-06-10] [WebCite Cache ID 6yjd5HGA9]
40. Cohen J. Statistical power analysis. Curr Dir Psychol Sci 2016 Jun 24;1(3):98-101. [doi: 10.1111/1467-8721.ep10768783]
41. Cox JL, Holden JM, Sagovsky R. Detection of postnatal depression. Development of the 10-item Edinburgh postnatal depression scale. Br J Psychiatry 1987 Jun;150:782-786. [doi: 10.1192/bjp.150.6.782] [Medline: 3651732]
42. Kawachi I, Berkman LF. Social ties and mental health. J Urban Health 2001 Sep;78(3):458-467 [FREE Full text] [doi: 10.1093/jurban/78.3.458] [Medline: 11564849]
43. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613 [FREE Full text] [doi: 10.1046/j.1525-1497.2001.016009606.x] [Medline: 11556941]
44. Spielberger CD. Manual for the State-Trait Anxiety Inventory. Palo Alto, CA: Consulting Psychologists Press; 1983.
45. Russell DW. UCLA Loneliness Scale (Version 3): reliability, validity, and factor structure. J Pers Assess 1996 Feb;66(1):20-40. [doi: 10.1207/s15327752jpa6601_2] [Medline: 8576833]
46. Leerkes EM, Crockenberg SC. The development of maternal self-efficacy and its impact on maternal behavior. Infancy 2002 Apr 1;3(2):227-247. [doi: 10.1207/S15327078IN0302_7]
47. Shorey S, Ng ED. Evaluation of a technology-based peer-support intervention program for preventing postnatal depression (part 2): qualitative study. J Med Internet Res 2019;21(8):e12915 [FREE Full text] [doi: 10.2196/12915]
48. Barnett PA, Gotlib IH. Psychosocial functioning and depression: distinguishing among antecedents, concomitants, and consequences. Psychol Bull 1988 Jul;104(1):97-126. [doi: 10.1037/0033-2909.104.1.97] [Medline: 3043529]
49. Dennis CL. Psychosocial and psychological interventions for prevention of postnatal depression: systematic review. Br Med J 2005 Jul 2;331(7507):15-18 [FREE Full text] [doi: 10.1136/bmj.331.7507.15] [Medline: 15994688]
50. Sockol LE, Epperson CN, Barber JP. The relationship between maternal attitudes and symptoms of depression and anxiety among pregnant and postpartum first-time mothers. Arch Womens Ment Health 2014 Jun;17(3):199-212 [FREE Full text] [doi: 10.1007/s00737-014-0424-9] [Medline: 24643422]
51. Falah-Hassani K, Shiri R, Dennis CL. Prevalence and risk factors for comorbid postpartum depressive symptomatology and anxiety. J Affect Disord 2016 Jul 1;198:142-147. [doi: 10.1016/j.jad.2016.03.010] [Medline: 27016657]
52. Ross LE, McLean LM. Anxiety disorders during pregnancy and the postpartum period: a systematic review. J Clin Psychiatry 2006 Aug;67(8):1285-1298. [doi: 10.4088/JCP.v67n0818] [Medline: 16965210]
53. Salonen AH, Kaunonen M, Astedt-Kurki P, Järvenpää AL, Isoaho H, Tarkka MT. Parenting self-efficacy after childbirth. J Adv Nurs 2009 Nov;65(11):2324-2336. [doi: 10.1111/j.1365-2648.2009.05113.x] [Medline: 19761456]
54. Kohlhoff J, Barnett B. Parenting self-efficacy: links with maternal depression, infant behaviour and adult attachment. Early Hum Dev 2013 Apr;89(4):249-256. [doi: 10.1016/j.earlhumdev.2013.01.008] [Medline: 23398731]
55. Abarashi Z, Tahmassian K, Mazaheri MA, Panaghi L, Mansoori N. Parental self-efficacy as a determining factor in healthy mother-child interaction: a pilot study in Iran. Iran J Psychiatry Behav Sci 2014;8(1):19-25 [FREE Full text] [Medline: 24995026]
56. Khan S, Ion A, Alyass A, Greene S, Kwaramba G, Smith S, et al. Loneliness and perceived social support in pregnancy and early postpartum of mothers living with HIV in Ontario, Canada. AIDS Care 2019 Mar;31(3):318-325. [doi: 10.1080/09540121.2018.1515469] [Medline: 30157684]
57. Wong J, Fisher J. The role of traditional confinement practices in determining postpartum depression in women in Chinese cultures: a systematic review of the English language evidence. J Affect Disord 2009 Aug;116(3):161-169. [doi: 10.1016/j.jad.2008.11.002] [Medline: 19135261]
58. Dennis CL, Hodnett E, Gallop R, Chalmers B. The effect of peer support on breast-feeding duration among primiparous women: a randomized controlled trial. Can Med Assoc J 2002 Jan 8;166(1):21-28 [FREE Full text] [doi: 10.1186/s13006-017-0118-9] [Medline: 11800243]
59. Haslam DM, Pakenham KI, Smith A. Social support and postpartum depressive symptomatology: the mediating role of maternal self-efficacy. Infant Ment Health J 2006 May;27(3):276-291. [doi: 10.1002/imhj.20092] [Medline: 28640472]
60. Beck CT. Postpartum depression: a metasynthesis. Qual Health Res 2002 Apr;12(4):453-472. [doi: 10.1177/104973202129120016] [Medline: 11939248]
61. Cutrona CE, Troutman BR. Social support, infant temperament, and parenting self-efficacy: a mediational model of postpartum depression. Child Dev 1986 Dec;57(6):1507-1518. [doi: 10.2307/1130428] [Medline: 3802975]
62. Hudson DB, Elek SM, Campbell-Grossman C. Depression, self-esteem, loneliness, and social support among adolescent mothers participating in the new parents project. Adolescence 2000;35(139):445-453. [Medline: 11130590]
63. Leung S, Arthur DG, Martinson I. Stress in women with postpartum depression: a phenomenological study. J Adv Nurs 2005 Aug;51(4):353-360. [doi: 10.1111/j.1365-2648.2005.03506.x] [Medline: 16086804]

Abbreviations

EPDS: Edinburgh Postnatal Depression Scale

XSL•FO
RenderX

PHQ: Patient Health Questionnaire
PIP: peer-support intervention program
PNA: postnatal anxiety
PND: postnatal depression
PSSP: Perceived Social Support for Parenting
STAI: State-Trait Anxiety Inventory
ULS: University of California, Los Angeles Loneliness Scale

Edited by G Eysenbach; submitted 05.10.18; peer-reviewed by T Rashid Soron, J Ciolino, A Rahman; comments to author 03.02.19; revised version received 19.03.19; accepted 28.07.19; published 29.08.19.

Please cite as:
Shorey S, Chee CYI, Ng ED, Lau Y, Dennis CL, Chan YH
Evaluation of a Technology-Based Peer-Support Intervention Program for Preventing Postnatal Depression (Part 1): Randomized Controlled Trial
J Med Internet Res 2019;21(8):e12410
URL: http://www.jmir.org/2019/8/e12410/
doi:10.2196/12410
PMID:31469084

©Shefaly Shorey, Cornelia Yin Ing Chee, Esperanza Debby Ng, Ying Lau, Cindy-Lee Dennis, Yiong Huak Chan. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 29.08.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Evaluation of a Technology-Based Peer-Support Intervention Program for Preventing Postnatal Depression (Part 2): Qualitative Study

Shefaly Shorey[1], PhD; Esperanza Debby Ng[1], BA

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

Corresponding Author:
Shefaly Shorey, PhD
Alice Lee Centre for Nursing Studies
Yong Loo Lin School of Medicine
National University of Singapore
Clinical Research Centre, MD 11
Level 2, 10 Medical Drive
Singapore, 117597
Singapore
Phone: 65 6601 1294
Email: nurssh@nus.edu.sg

Abstract

Background: Social support is known to reduce risks of postnatal depression (PND) and improve maternal emotional well-being. However, the Asian cultural context is often neglected when appraising maternal needs and mothers' preferences for social support. While many preventive efforts have experimented with technology, professionals, and paraprofessionals in providing social support to mothers in need, most studies determined the effectiveness of their interventions through quantitative measurements of maternal outcomes. Experiences and feedback from both participants and administrators are rarely discussed, especially in an Asian setting.

Objective: The goal of the research was to evaluate the postnatal experiences of Asian mothers at risk of PND and the perceptions of peer volunteers regarding a technology-based peer-support intervention program (PIP).

Methods: A qualitative semistructured interview was conducted with 20 Asian mothers at risk of depression (10 from the control group and 10 from the intervention group) and 19 peer volunteers from a randomized controlled trial. The PIP included weekly correspondence between peer volunteers and mothers through any telecommunication means over 4 weeks. All interviews were approximately 30 to 60 minutes long, audiotaped, transcribed verbatim, and analyzed using thematic analysis. Study findings were reported according to the Consolidated Standards of Reporting Trials checklist.

Results: Two overarching themes comprising five subthemes were generated: postnatal experience (a bouncy ride, a way forward) and evaluation of the PIP (valuable, flexible, and supportive program; building blocks of a good relationship; and lessons learned and the road ahead). Mothers from both the control and interventions groups were generally satisfied with hospital care and the support received from family. They also shared similar breastfeeding challenges and needs for more informed decisions and follow-up support from the hospital. However, mothers who received the PIP tended to have more positive outlooks of their birth experiences. Overall, peer volunteers and mothers involved in the PIP found the PIP useful and expressed satisfaction with the program's flexibility. They also shared their personal takeaways, the qualities of their friendships, and the need for extended correspondence time and recommended outreach to non–at-risk mothers.

Conclusions: The positive endorsement of the PIP by peer volunteers and mothers suggests the success of the PIP in maintaining positive maternal emotional well-being during the postpartum period. With the help of technology, hospitals can easily provide additional peer support to at-risk mothers in addition to existing standard care offered to these mothers.

Trial Registration: ISRCTN Registry ISRCTN14864807; http://www.isrctn.com/ISRCTN14864807

International Registered Report Identifier (IRRID): RR2-10.2196/resprot.9416

(J Med Internet Res 2019;21(8):e12915) doi:10.2196/12915

XSL•FO
RenderX

KEYWORDS

depression; mothers; postpartum; qualitative; social support; telecommunication; digital health; peer support; peer-to-peer support; online support groups; internet

Introduction

Background on Postnatal Depression

As one of the leading causes of maternal morbidity and a potential contributor to maternal mortality, postnatal depression (PND) has garnered much attention from researchers studying women's health over the years [1]. Apart from affecting the psychological well-being of mothers [2-4], maternal PND can potentially increase risks of depression in their partners [5] and severely hinder their children's neurocognitive, psychological, and social development [6-8]. In consideration of long-term adverse societal consequences, various preventive measures have been implemented to battle PND, but targeting the root cause remains a challenge due to its complex multifactorial etiology (ie, biological, psychological, and social). The contribution of psychosocial factors to PND outweighs that of biological factors, with chronic social adversity in women playing a substantial role [9,10]. This is especially so in Asian countries, where polygamous marriages, conflicting influences of culture, and conflict between mothers- and daughters-in-law are unique risk factors of PND [11-13]. According to a review by Evagorou and colleagues [14], Asian and Western mothers share similar PND risk factors and have similar manifestations of physical symptoms. Differences lie in prevalence rates and cultural factors (beliefs, values, and environment). A local study has revealed a prevalence of 6.8% for PND in Singapore, which is much lower than the global average of 16% [15]. This corresponds with findings from another review [16], but the prevalence of PND varies across Asia itself as well, and South Asia collectively has higher PND rates than Western countries [17-18]. These reviews [16-18] have also highlighted cultural differences that contributed to differences in regional prevalence rates, such as Asian's higher emphasis on family and social relationships, hence mothers' higher dependency on their partners, their own mothers, and extended family. The lack of female empowerment, the Asian view of women as the weaker sex, and other traditional rituals such as practicing the confinement period are additional stressors for Asian mothers. Due to the collective nature of Asian societies, social support from mothers' family members and partners plays a crucial role in mitigating PND. Therefore, positive social relationships and increased social support are exceptionally important in ensuring the healthy psychological well-being of Asian mothers [18].

Types of and Need for Social Support

Defined as an interpersonal process whereby the provider communicates to the recipient that he or she is cared for in this reciprocal relationship [19], social support is multidimensional and varies in terms of range of support network, type, source, and quality [20]. Types of support typically include (1) emotional support, involving expressions of trust, care, and empathy; (2) instrumental support, involving tangible aid; (3) informational support, such as advice, suggestions, and information; and (4) appraisal support, including advice that allows self-evaluation [21]. Additionally, preferred types of support are largely influenced by cultural differences. In an individualistic Western culture, direct forms of support involving verbal and emotional expressions are often sought explicitly from others, whereas in a collectivistic Asian culture, indirect support through companionship and attentiveness from others is preferred [22]. Given that the impact of social support on health and recovery highly depends on the match between support provided and needed, it is necessary to consider cultural contexts when providing postpartum support to mothers in need [23].

Social support from family and partners is known to be an effective stress buffer that reduces risks of maternal PND and maintains quality of life of mothers in the postpartum period [19]. The practice of a 1-month confinement period in the Asian culture also provides most mothers with additional instrumental support in terms of infant care and household chores from confinement nannies or in-laws [24,25]. However, mothers are often dissatisfied with social support, especially due to the lack of emotional support received [24,26,27]. This can be attributed to the conservative nature of Asian societies in which direct emotional expressions are often discouraged [28,29] and the availability of emotion-focused support is rare, even from those close [22]. This highlights an unspoken need for more emotional support for Asian mothers. Mothers often mentioned a need for a close, nonjudgmental confidante to initiate support and empathize with them [4,30]. Hence, peer support is a viable option as it creates a sense of belonging, boosts parenting confidence and self-esteem among mothers, and is effective in preventing PND [3,31]. Moreover, peer support provided by peer volunteers is found to ensure continuity of support during the postpartum period and fill in for health care professionals in between clinical follow-ups [32]. Although most mothers still prioritize the need for informational support from health care professionals [33], Sjoberg and colleagues [34] revealed that new-generation mothers preferred online peer support over face-to-face or online consultations with health care professionals. Using technology to support mothers will not only increase the accessibility, availability, and affordability of maternal health care but will also remove help-seeking barriers in conservative cultures [35-37]. Additionally, Singapore has been undergoing the Smart Nation initiative since 2014 [38], which encourages health care sectors to shift to telehealth in order to optimize resources and overcome manpower constraints. Previous local studies have also seen successful results with supportive technology-based interventions involving the use of online forums and mobile apps [39,40]. Therefore, technology-based supportive interventions in this context are not considered foreign and may be more ideally suited to the needs of new-generation mothers.

Review of Current Literature

While there are a substantial amount of randomized controlled trials (RCTs) that adopted a technology-based supportive approach, the effectiveness of these RCTs is usually determined

XSL•FO
RenderX

by a quantitative analysis of maternal outcomes [41-45]. In-depth qualitative analyses of user experiences and feedback are limited. Technology-based supportive interventions are effective in reducing PND in women, and study designs primarily involved weekly telephone calls or video calls by paraprofessionals or peer volunteers [41-43]. A qualitative analysis provides a descriptive and exploratory insight of user needs that can inform the further refinement and tailoring process of an intervention program to better meet user needs. The limited number of qualitative studies and the heterogeneity of available studies in terms of support source, delivery method, and frequency warrant a need to perform a qualitative study on a technology-based peer support program in an Asian setting [46-50].

Therefore, this study aimed to evaluate the postnatal experiences of mothers at risk of PND and their perceptions of a technology-based peer support intervention program (PIP). Feedback from peer volunteers was also examined in order to gain a well-rounded perspective from both recipients and administrators.

Methods

Design and Setting

This descriptive qualitative study is a follow-up of an RCT [51] that examined the effectiveness of a technology-based PIP on maternal outcomes including PND, postnatal anxiety, loneliness, and perceived social support received. The original study was conducted at a tertiary hospital in Singapore where 136 mothers at risk of PND were recruited through purposive sampling and randomized into either the intervention (n=69) or control (n=69) group. Mothers at risk were identified through an initial screening using the Edinburgh Postnatal Depression Scale. Only those who scored 9 and above were recruited for the RCT. Twenty peer volunteers, mothers who had recovered from PND, were recruited through a blasting of emails to the study venue's working community and by word of mouth to facilitate the support program. Peer volunteers then underwent a training session by a psychiatrist to learn required skills (via role play and discussions) for successful administration of technology-based peer support, which covered topics on alleviating depression, anxiety, loneliness, and how to conduct appropriate referrals to a health care professional when necessary. Details on peer volunteer training can be found in the published protocol [52]. Each peer volunteer was matched with at least three mothers. The fidelity of the PIP intervention was maintained using the strategies proposed by Bellg et al [53] and Eaton et al [54]: (1) the protocol for intervention was developed and has been published [52]; (2) to confirm dosage of the intervention, log sheets were maintained on the time and duration of contacts made with peer volunteers; (3) to ensure consistency, peer volunteers were trained by the same psychiatrist; and (4) a peer volunteer manual was developed for a standardized delivery of information.

Participants in the control and intervention groups received standard hospital postnatal care such as lactation support, parent craft teaching, and postpartum follow-up appointments with an obstetrician. In addition, participants in the intervention group had support from peer volunteers for at least 1 month postpartum involving a minimum of once-a-week correspondence between the mothers and the volunteers through any technology-based means (ie, phone calls, text messages, and WhatsApp). Frequency and duration were tailored to maternal needs. Further details have been published elsewhere [52]. The experiences of mothers in the control group were included in this follow-up to provide different perspectives and form a basis of comparison of needs and received support during the perinatal period. Prior to this, all participants were informed of this optional extended qualitative component of the study. Qualitative feedback from peer volunteers on delivery of the intervention was made mandatory upon the initial consent for the original study.

Recruitment

Purposive sampling was used to achieve an equal number of participants from the control and intervention groups. Recruitment was conducted at 1-month postpartum through the blasting of emails to the 136 participants in the original study until data saturation was reached. Thirty-six mothers volunteered for the interviews. However, data saturation was reached at the eighth participant for both groups when no new findings emerged. Two additional mothers from each group were interviewed to confirm the findings, resulting in a total of 20 interviewed mothers (10 from the control group and 10 from the intervention group). The remaining 16 mothers consented to be excluded from the interviews. One peer volunteer had withdrawn from the original study due to time commitment issues, leaving nineteen peer volunteers in this study. Mothers and peer volunteers were informed of the approximate duration of each interview (30 to 60 minutes) and that all interviews would be audio-recorded for research purposes only.

Data Collection

When participants were between 4 and 12 weeks postpartum, a female research assistant trained in qualitative interviewing techniques conducted individual face-to-face interviews at a time and location to each mother's convenience, typically at the mother's home. Participants were given pseudonyms during the interviews to protect their actual identities. Peer volunteers were only interviewed after the conclusion of the intervention for all 69 mothers, and the Peer Volunteer Activity Logs, which were used by peer volunteers to record the frequency, duration, and key points of the correspondences, were submitted to the research assistant during the interviews. Three versions of a semistructured interview guide specifically tailored to the control group, the intervention group, and peer volunteers were developed and piloted to attain a comprehensive view of the delivery and recipient of the PIP and the postnatal experiences of mothers (Multimedia Appendix 1). The interviews lasted an average 40 minutes and were subsequently transcribed verbatim. Field notes were taken during the interviews to note nonverbal cues that were used to supplement the transcripts.

Data Analysis

A thematic analysis was conducted according to the six phases of analysis described in the research of Braun and Clark [55]. The authors read the 39 transcribed interviews multiple times to gain familiarity with the data and subsequently adopted a

XSL•FO
RenderX

manual color-coding method to highlight different concepts and generate the initial codes independently. These excerpts were then extracted and put into a tabular format, entailing themes and subthemes within a new document. Data source triangulation from all three groups of participants (intervention and control groups and peer volunteers) was performed to look for common themes [56]. The themes were reviewed comprehensively for homogeneity by both authors before overarching themes were decided. To achieve confirmability and objectivity in the analysis of data, the various themes and subthemes were discussed extensively in several meetings between the two authors. Any discrepancies were discussed and clarified between the two authors until a consensus was achieved. Field notes were also constantly referred to as supplementary materials, and constant comparative analyses were performed. Themes constituting frequently reported overlapping data were selected from the authors' independent analyses, renamed, and included in the final analysis.

Ethical Considerations

Ethics approval was obtained from the National Health Group Domain Specific Review Board (reference number: NHG DSRB 2017/00185) of the participating hospital. Prior to obtaining the participants' written informed consents, the participants were briefed thoroughly on the purpose of the research and were informed of their rights to withdraw at any time during the study. Study participation was strictly voluntary, and confidentiality was adhered to.

Results

Overview

Study findings were reported according to the Consolidated Criteria for Reporting Qualitative Research checklist (COREQ, Multimedia Appendix 2) [57]. According to our RCT results [51], the PIP was successful in significantly reducing PND symptoms at the end of three months postpartum. Although not statistically significant, there were also observable decreases in postnatal anxiety and loneliness and an increase in perceived social support at the end of three months.

Participant Characteristics

A total of 39 interviews were completed with 20 Asian mothers (10 from the control group and 10 from the intervention group) and 19 peer volunteers. Mothers had an age range of 25 to 40 years while peer volunteers had an age range of 25 to 54 years. The majority of the mothers were Chinese (n=10), followed by Malay (n=9) and Indian (n=1), which was likewise for the peer volunteers (Chinese, n=17; Malay, n=1; Indian, n=1). All 10 of the mother-volunteer dyads corresponded (regarding emotional needs and available resources) through WhatsApp messages at least once a week; only five dyads attempted to schedule a weekly phone call. Further details of the participants and peer volunteers can be found in Table 1.

Two overarching themes comprising five subthemes were generated: postnatal experience (a bouncy ride, a way forward) and evaluation of the PIP (valuable, flexible, and supportive program; building blocks of a good relationship; and lessons learned and the road ahead). A detailed breakdown of the themes and subthemes is shown in Figure 1. Exact quotes from the participants were used to represent their emic views, and their identities were coded to allow for anonymity, with mothers in the control group labeled with C, mothers in the intervention group labeled with T, and peer volunteers labeled with P.

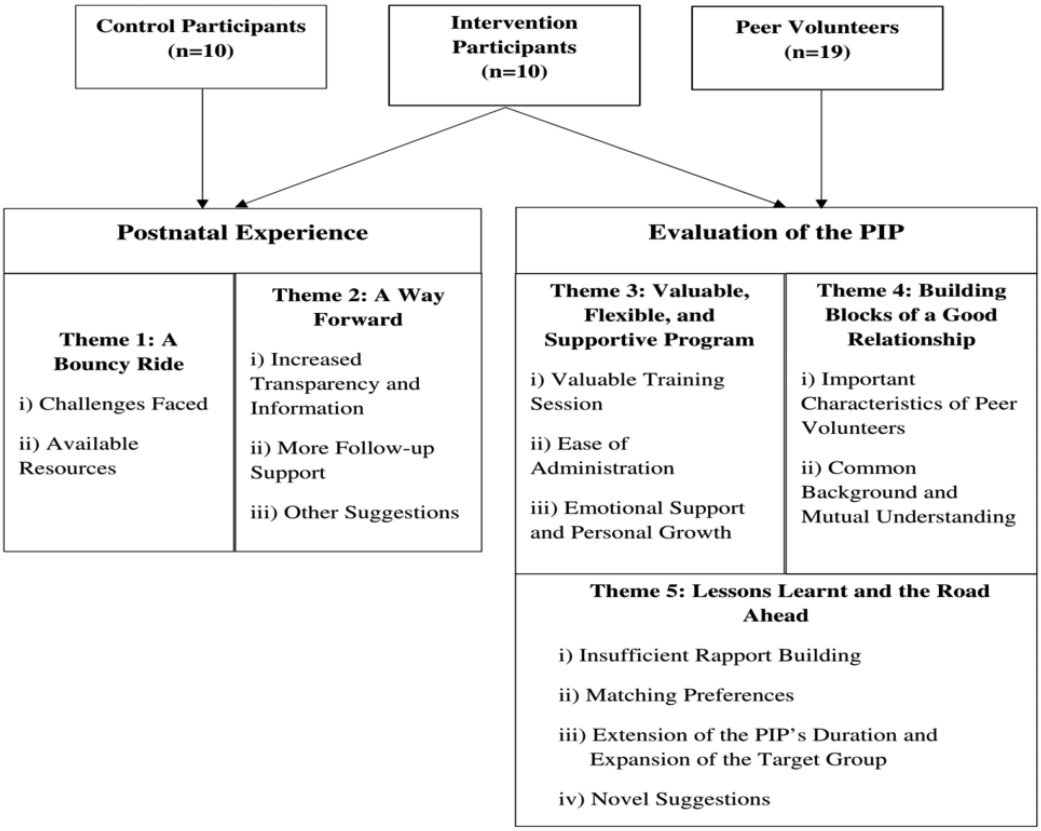
XSL•FO
RenderX

Table 1. Characteristics of study participants (control n=10, intervention n=10, volunteers n=19).

Characteristics	Control	Intervention	Peer volunteers
Age (years), mean (SD), range	32.0 (5.0), 25-40	31.5 (4.7), 25-40	40.4 (7.9), 25-54
Ethnicity, n (%)			
Chinese	6 (60)	4 (40)	17 (90)
Malay	3 (30)	6 (60)	1 (5)
Indian	1 (10)	0 (0)	1 (5)
Marital status, n (%)			
Married	10 (100)	9 (90)	17 (90)
Single	0 (0)	1 (10)	2 (11)
University graduates, n (%)	7 (70)	9 (90)	13 (68)
Monthly household income (S$), n (%)			
<3000	3 (30)	1 (10)	3 (16)
3000-5000	3 (30)	3 (30)	4 (21)
>5000	4 (40)	6 (60)	9 (47)
Attended antenatal class, n (%)	6 (60)	5 (50)	7 (37)
Mode of delivery, n (%)			
Normal vaginal delivery or water birth	8 (80)	5 (50)	8 (42)
Instrument delivery	0 (0)	3 (30)	5 (26)
Cesarean delivery	2 (20)	2 (20)	6 (32)
Number of children, n (%)			
One	8 (80)	8 (80)	5 (26)
Two	2 (20)	0 (0)	8 (42)
Three or more	0 (0)	2 (20)	6 (32)
Breastfeeding, n (%)	9 (90)	9 (90)	17 (90)

XSL•FO
RenderX

Figure 1. Overview of themes and subthemes. PIP: peer support intervention program.



Theme 1: A Bouncy Ride

This theme describes the bumpy postnatal experiences of mothers including the challenges of breastfeeding, their physical and emotional states, and the types of support received.

Challenges Faced

Mothers from both the intervention and control groups frequently described breastfeeding as stressful and difficult. Heavy infant care responsibilities also left mothers feeling exhausted, tired, and overwhelmed. Receiving support from family members served as a double-edged sword with a few mothers reporting conflicts in parenting styles and opinions due to generation gaps causing additional stress.

> *Different sides of [the] family have different cultures and different upbringings, so sometimes, it's like you have too many inputs, you get stressed...* [T9]

A combination of the above factors resulted in many first-time mothers feeling lost, worried, and scared.

Available Resources

During the short hospitalization period, all mothers were satisfied with the support received from health care professionals, often citing them as caring, encouraging, assuring, and patient when guiding them in breastfeeding and answering their queries.

The majority of mothers from both groups were grateful for the high instrumental support provided by their family members, especially their partners. With family members helping to shoulder most household responsibilities and infant care tasks, mothers had more *me time* to catch a breather and rest.

> *I feel much more relieved because I know the baby is [being] taken care [of] by my husband...It helped me to get back my sleep...* [T9]

Baby care information was often sought from friends who had recently given birth, otherwise mothers were embarrassed to seek hands-on support from them. Facebook groups, forums, and websites were also commonly referred to as information sources and support, but some mothers remained skeptical of the credibility of online information.

Theme 2: A Way Forward

In general, mothers indicated a need for more informed decisions and social support to improve their birth experiences.

Increased Transparency and Information

Mothers from both groups identified a high need for hospitals to provide more information to facilitate informed decision-making in terms of medical costs and medical procedures and to increase the awareness of postnatal blues and the availability of support sources.

> *I'm unaware of how this system is like...We didn't do a lot of research...We thought that they would inform us...* [T9]

More Follow-Up Support

In consideration of the 1-month confinement period adhered to by most mothers, there were mixed preferences for home visits

XSL•FO
RenderX

and increasing the frequency of follow-up phone calls by health care professionals.

> *It's difficult to tell to explain the situation over the phone...You really have to, like, watch and observe the problem...* [C7]

> *There could be...more follow-ups...Actually, a phone call is good enough...At least you know that you're being taken care of...* [T3]

Mothers in the control group also expressed a need for peer volunteers as an additional support.

> *My friends are married...but they don't have [a] baby, so there's too little information I can get from them...If there's a peer volunteer for me, I'll definitely find it very good...* [C10]

Other mothers suggested having a breastfeeding buddy or befriender in the hospital to stand in nurses' place and guide them in breastfeeding techniques.

> *Some volunteer that could at least help out a bit...rather than have nobody, cause the nurses, they're busy...I think it will make the journey a bit less stressful...* [T4]

Other Suggestions

Although many mothers referred to the UK-based website babycentre for baby care information, they highly preferred a baby mobile app that could provide more locally relevant information, such as on confinement practices. They also mentioned that they would find the mobile app more reputable and reliable if it were moderated by health care professionals from local hospitals.

> *Google will always come up with...from [the] United States, [the] UK, that kind...Their advice is very westernized, may not be what Singaporeans follow...especially confinement...* [T4]

> *Baby Center was the most popular thing...If there was something with the hospital...we will think it's more reputable...* [C9]

Other suggestions included mobile app features such as having baby food recipes, instant messaging with health care professionals, and an online portal with the hospital to check and update baby appointments.

Theme 3: Valuable, Flexible, and Supportive Program

This theme summarizes the feedback of peer volunteers and mothers who received support (the intervention group only) on the usefulness of the PIP. Overall, the peer volunteers were satisfied with the informative training session. The flexibility of the program was also mentioned by both mothers and peer volunteers, which made them feel supported.

Valuable Training Session

Peer volunteers found the training by the psychiatrist very useful, informative, and helpful in setting their expectations. It also enhanced their self-confidence, sensitivity, and awareness toward at-risk mothers.

> *It kind of gives us the assurance...the confidence level that we can do it...* [P8]

> *It was helpful in creating more awareness in me...It helps me to be a bit more sensitive...* [P3]

A few peer volunteers even mentioned how the training session established a support system for them and created a sense of belonging since they shared a common past and common goal in this study.

> *The sense of belonging is like...cause we all felt the need to support these mothers...Either it's because of our own experiences or there's a sense of same heart in supporting these women...* [P16]

However, there were also suggestions for more culture-specific information and simulated roleplays to increase preparedness and their abilities to cater to mothers from different cultures.

Ease of Administration

Both peer volunteers and mothers who received the PIP praised its flexibility as the mode of contact was to their convenience and preferences, which was mostly through text messages.

> *It's very hard for me to talk on the phone because of the baby and my other kid, so WhatsApp is the best mode cause I can reply as and when I am free...* [T8]

> *WhatsApp is convenient, like, anytime that they reply, you can just reply at work...Very flexible...You can accommodate both working moms and stay-at-home moms...* [P19]

Although a few pairings attempted to converse through phone calls, both parties admitted it was hard to match their busy schedules.

> *The baby is a bit difficult to, you know, find the right time...It's a bit difficult to arrange...* [T1]

> *I don't want to call without warning...When I try to set a time, they don't commit, then it's not very nice also if I just call unannounced...* [P4]

Emotional Support and Personal Growth

The majority of the mothers were appreciative of the extra listening ear and felt they had another friend to talk to and they were not alone. They also felt more comforted and reassured and had reduced negative feelings after receiving continuous peer support.

> *Whenever I have issues, she will comfort me...that I'm not the only one going through it, and it's only normal...It was very nice to have somebody additional apart from my own family.* [T8]

Almost all of the peer volunteers found the PIP experience meaningful, enriching, and fulfilling and felt happy that their previous experiences could benefit other mothers. Some even took it as a chance to self-reflect and for self-improvement.

> *I can learn something new...I can improve myself...The more I give, then the more I feel happy. I just want to make sure that they didn't get depression or stressed.* [P9]

XSL•FO
RenderX

Theme 4: Building Blocks of a Good Relationship

This theme describes the essential traits of a befriender and the basic underlying foundation of a good relationship.

Important Characteristics of Peer Volunteers

Most recipients of the peer support reported that the friendliness, sincerity, proactiveness, positivity, and commitment shown by the peer volunteers allowed them to gradually open up and share their problems.

> *She took [the] initiative to check on me when I was having my confinement...She was also quite cheerful...After that, she still asked me occasionally how I was...Despite her having so many kids, she still bothered to remember this peer volunteer friend...It was a two-way thing…Quite assuring to know that somebody could help me when I needed [it].* [T4]

Many peer volunteers shared that being understanding and open-minded was important in establishing a good relationship. Being perceptive was also vital in gauging the responses of mothers and knowing when to not be overimposing. Additionally, a strong mentality and awareness of their role limitations also helped them to cope with the lack of responses and not to take it personally.

> *I would also remind myself not to take it personally...The whole point of this is to be open-minded and to help them...* [P4]

Common Background and Mutual Understanding

Mothers and peer volunteers who reportedly had closer friendships found it easier to relate with one another due to various common factors (ie, marital status and working background), and they possessed a mutual understanding of each other's busy schedule and that delayed replies were inevitable.

> *I'm a single parent...and the lady is also a single parent, so, like, there are some common factors...We kind of support each other...so it's more like a two-way thing...which is why I say she is more like a friend.* [T3]

Theme 5: Lessons Learned and the Road Ahead

Insufficient Rapport Building

A few mother-peer volunteer pairs faced initial discomfort in sharing their thoughts with a stranger. The lack of connection resulted in awkwardness and one-sided, superficial relationships.

> *Sometimes, I don't feel very connected to the person that I call...so, sometimes, it gets awkward during the phone conversation.* [T2]

A lack of response from mothers also added stress and uncertainty for some peer volunteers as they were unable to identify the mother's needs, hence disabling their support efforts. Peer volunteers were also unsure if they had provided adequate help or whether they were disturbing the mothers.

> *It's a bit discouraging, so...I know they may not have the time to read the message or reply to me...I can't do anything...I cannot help much. I don't know what's going on.* [P2]

Matching Preferences

In order to increase relevance to self, mothers generally preferred to be matched with a volunteer of similar age, same ethnicity, employment status, marital status, recency of childbirth, and similar ages of children. One mother expressed concern over the age gap with her peer volunteer.

> *She doesn't quite understand the current fast pace of a working environment and young parents like us who have to juggle so many things...* [T6]

However, in order to provide relevant and effective advice, peer volunteers preferred to be matched based on delivery mode, mode of feeding, and socioeconomic and employment status.

Extension of the Peer Support Intervention Program's Duration and Expansion of the Target Group

Both mothers and peer volunteers agreed that the PIP should be extended for at least 2 to 3 months postpartum since the onset of PND is unpredictable.

> *It will help, especially the second month...when confinement lady is gone...when my husband goes back to work, then you're left alone with the kids...You need this added support.* [T5]

Since PND is hard to detect in the early stages, the PIP was highly recommended to be made available to all new mothers as a safety net. Most participants acknowledged that having additional support would benefit all mothers emotionally, regardless of their risks of depression

> *Some of them, they are strong enough, but sometimes, they just need a listening ear to release their stress...* [P15]

Novel Suggestions

Mothers and peer volunteers were highly in favor of at least one session of face-to-face meet-up, which would allow easier rapport building and let peer volunteers provide more substantial instrumental support.

> *Maybe one [session] could be like a visit, to see them, say hello...Maybe they will feel better...It will be easier to build the rapport.* [P19]

Developing a mobile app to locate and contact nearby befrienders who could help during emergencies was also suggested. This would increase convenience, reduce the hassle of pairing peer volunteers, and facilitate the natural occurrence of friendships.

Discussion

Principal Findings

Overall, postnatal experiences were highly similar in both groups of mothers, such as common breastfeeding issues and physical fatigue. A large-scale review of the effectiveness of various types of breastfeeding support concluded that regularly scheduled face-to-face support by health care professionals has

XSL•FO
RenderX

a higher success rate in teaching mothers to breastfeed [58]. Findings from another local study that used educational parenting videos emphasized mothers' need for both theory and hands-on learning through visual means [55]. As the PIP did not focus on the physical or breastfeeding needs of mothers, this could be the reason that mothers from the intervention group were equally distressed about their physical and breastfeeding needs.

In terms of emotional experience, mothers who received the PIP had reduced negative feelings and better understanding and acceptability of their emotional situations; hence, they were better able to enjoy their birth experiences. Mothers without additional support felt lost, anxious, and alone, which has been commonly reported in other studies [55-60]. The perceived effectiveness of the PIP in providing adequate emotional support for mothers is also supported by the study findings of Dennis et al [46], in which mothers benefitted from the emotional support received from peer volunteers.

Although mothers in both groups were highly satisfied with the quality and competency of health care providers during their hospitalization stays, they were disappointed with the lack of transparency and insufficient information from health care professionals, which hindered their informed decision making processes. Being involved and making well-informed child care decisions gives mothers a sense of empowerment, responsibility, and preparedness, which promotes better health outcomes and healthier physical and psychological well-being and even affects the long-term health and well-being of their children [61]. Mothers also voiced the need for more hospital follow-ups and continuity of care in terms of phone calls and home visits. Studies have shown that continuity of care is often lacking due to resource constraints, such as a lack of manpower [62], but patients who received continuity after hospital discharge had better health care outcomes, higher satisfaction rates, and more cost-effective health care [59,63]. Therefore, in order to promote positive postpartum experiences, a multidimensional approach to improve current hospital care in the form of follow-up multimodal educational programs needs to be considered.

Interestingly, when describing the support received from family and their partners, mothers from both groups only mentioned their satisfaction with the instrumental support received from their parents, in-laws, and partners. This is supported in a study by Chen and colleagues [64] that revealed that Asians tend to provide more problem-focused support than European Americans who tend to provide more emotion-focused support. This is mainly due to the conservative, collectivist Asian culture whereby indirect support, which avoids mention of the stressor, is preferred and includes displays of attentiveness and companionship from their loved ones [22]. However, conflicts with in-laws in terms of parenting styles and differences in opinions due to generation gaps were commonly reported, adding emotional stress for mothers. As respect for elders is highly observed in the Asian culture, new mothers often have to suppress their emotions, and having a poor relationship with one's mother-in-law potentially jeopardizes her marital relationship and increases risks of PND [11,26,27,65]. Chen and colleagues [64] concluded that the most effective type of social support is a delicate balance between autonomy and dependence. Considering that negative childbirth experiences can cause partner and family strain, an external support in the form of peer-volunteers might be more ideal when providing a nonjudgmental listening ear [66].

Peer Support Intervention Program

Similar to other peer support studies [47,67,68], our peer volunteers underwent mandatory training to learn adequate support provision for at-risk mothers, which most found useful in setting personal expectations and gaining confidence and preparedness. Additionally, since peer volunteers were mothers with a history of PND, training sessions also served as a platform for mothers with similar pasts to connect and form a community of support [47].

In an effort to increase the accessibility and flexibility of support with minimal disruption to the mothers' lives, communication through technology-based platforms was adopted. This aspect was praised by both recipients and peer volunteers as correspondence was flexible and suited to their own levels of comfort, whether it was through text messages or phone calls. Due to unpredictability and difficulty in arranging fixed calling times, most mothers and volunteers preferred correspondence through text messages. However, a few reported the lack of a personal touch and difficulty in gauging mothers' responses over text messages, with some mothers also suggesting at least one home visit or face-to-face meet-up. A review by Shaw [60] acknowledged the effectiveness of peer-support in reducing PND, but maternal satisfaction was higher for home visits. Videoconferencing was not used in this study, but it has been shown to improve mental health outcomes [69] and increase confidence and was deemed to be almost equivalent to the physical presence of a peer volunteer [43]. Parents in a study by Linberg [43] also felt that the midwife understood their situations better with added nonverbal communication through videoconferencing. Hence, the option of using videoconferencing can be considered in future local studies to support mothers at risk of PND.

Initiating a socially meaningful relationship of trust involves certain key characteristics of peer volunteers: positivity, determination, openness, proactiveness, sincerity, and respect. These qualities are critical to successful program delivery and enable peer volunteers to overcome challenges and stress and establish effective functional relationships [68]. Mothers and volunteers shared that the mutual understanding and identification that drew on the shared experiences of parenthood and social circumstances increased feelings of connectedness and ease of sharing. Despite the differences in personality, the sharing of experiences with someone of similar backgrounds, attitudes, and experiences facilitates trust-building and easier communication [68,70]. The matching of peer volunteers to mothers is an important component of successful peer-mentoring as the provision of support should be targeted to fit the nature and situation of the family [46,50,71].

This PIP was also seen as a two-way support system. While most mothers were appreciative of the extra listening ear and support provided that improved their overall emotional well-being, peer volunteers were also glad for the opportunity to impact other lives with their experiences and felt happier

XSL•FO
RenderX

after volunteering. Expressing satisfaction with their roles, many volunteers felt empowered and took this chance to learn new things to enhance their personal growth and self-discovery. Similarly, in other studies peer volunteers felt emotionally rewarded and enriched by their interpersonal relationships; given the chance, they would volunteer again [47,66].

However, challenges and stressful moments arose for some peer volunteers when communication was only one-sided. Mothers and volunteers attributed it to insufficient rapport-building and the discomfort of sharing with a stranger. A few volunteers perceived some relationships as superficial and obligatory, while some mothers felt that they did not need the additional support. This is common in the Asian culture as the stigma of mental illness is related to the loss of face beyond the individual level [72,73]. In a culture that emphasizes emotional restraint, avoidance of shame, and saving face, it may take a longer time for mothers to warm up and share their problems with volunteers [74].

Implication for Future Research and Clinical Practice

Given the easy accessibility and wide availability of technological devices today, technology-based peer support is a potentially time-saving and cost-effective way to extend social support to mothers in need. Being able to provide mothers with a personalized continuity of care post–hospital discharge should incentivize health care sectors to dedicate more resources to train peer volunteers and implement the technology-based peer support to a wider population.

In consideration of the collectivistic culture and strong interdependence on family, future studies on technology-based peer support in Asia can involve partners and family members, which might reduce the emotional strain between mothers and family members during the stressful postpartum period. Future research can also consider including videoconferencing or developing a peer volunteer mobile app with detailed profiles to provide more targeted forms of support. Given that breastfeeding remains a key stressor to mothers, future interventions can adopt a holistic approach of providing both instrumental and emotional support to new mothers. Additionally, the introduction of peer volunteers to mothers should ideally take place during the third trimester, when mothers still have the time to build rapport. Last, the practice of the confinement period may see mothers receiving more help from family, in-laws, or confinement nannies during the immediate postpartum period for one month [24]. Instead, the second month postpartum, when all added help is depleted, may be more crucial as mothers resume work and the full responsibility of infant care. Therefore, the extension of the PIP's duration will be beneficial to mothers during this period.

Strengths and Limitations

Unique to this study is the valuable insight of cultural influences on preferred types of social support and the help-seeking behaviors of Asian mothers in the postpartum period. The inclusion of both mothers' and peer volunteers' perspectives also provided a well-rounded overview of the PIP. To our knowledge, this is the first technology-based peer support intervention in Asia. Therefore, the study findings are vital in informing future research on PND prevention using technology-based peer support, especially in an Asian context.

However, the cross-sectional nature of this study disallows the comparison of the pretest and posttest perceptions of mothers, making it difficult to determine the effectiveness of the intervention. Additionally, due to previous childbirth experiences, the sample inclusion of non–first-time mothers may bias the findings positively and provide inaccurate information on the usefulness of the PIP. Hence, future interviews should consider interviewing an equal number of first-time and experienced mothers and analyzing if there is any uniqueness in their individual experiences of receiving the PIP.

Conclusion

This study not only provided insight into the postnatal experiences and needs of at-risk Asian mothers but also evaluated the perceptions of both recipients and administrators of the technology-based PIP. The positive endorsement by mothers and volunteers suggests the PIP's success and usefulness in meeting mothers' postpartum needs. The technology-based PIP has the potential to connect with and give support to vulnerable mothers and enable them to access services in ways that complement the work of health professionals. However, given the suggestions for improvement by stakeholders, more research and rigorous testing are needed to further refine this approach before its implementation to a wider community.

Acknowledgments

The authors are thankful for the National University Health System Collaborative Clinician Research grant (reference number T1-NUHS O-CRG 2016 Oct 22) that funded this study. The authors also send special thanks to the National University Health System, Medical Publications Support Unit, for editing the language and format of this paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Semistructured interview guide for participants and peer volunteers.
[DOCX File, 20KB - jmir_v21i8e12915_app1.docx]

XSL•FO
RenderX

Multimedia Appendix 2
Consolidated Criteria for Reporting Qualitative Research (COREQ) checklist.
[PDF File (Adobe PDF File), 490KB - jmir_v21i8e12915_app2.pdf]

References

1. Dennis C. Detection, prevention, and treatment of postpartum depression. In: Postpartum Depression: Literature Review of Risk Factors and Interventions. Toronto, Canada: University Health Network; 2003:71-196.
2. O'Hara M. Postpartum Depression: Causes and Consequences. 1st Edition. New York: Springer; 1995.
3. McLeish J, Redshaw M. Mothers' accounts of the impact on emotional wellbeing of organised peer support in pregnancy and early parenthood: a qualitative study. BMC Pregnancy Childbirth 2017 Dec 13;17(1):28 [FREE Full text] [doi: 10.1186/s12884-017-1220-0] [Medline: 28086827]
4. Habel C, Feeley N, Hayton B, Bell L, Zelkowitz P. Causes of women's postpartum depression symptoms: men's and women's perceptions. Midwifery 2015 Jul;31(7):728-734. [doi: 10.1016/j.midw.2015.03.007] [Medline: 25921442]
5. Goodman JH. Paternal postpartum depression, its relationship to maternal postpartum depression, and implications for family health. J Adv Nurs 2004 Jan;45(1):26-35. [Medline: 14675298]
6. Soe NN, Wen DJ, Poh JS, Li Y, Broekman BFP, Chen H, et al. Pre- and post-natal maternal depressive symptoms in relation with infant frontal function, connectivity, and behaviors. PLoS One 2016;11(4):e0152991 [FREE Full text] [doi: 10.1371/journal.pone.0152991] [Medline: 27073881]
7. Stein A, Pearson RM, Goodman SH, Rapa E, Rahman A, McCallum M, et al. Effects of perinatal mental disorders on the fetus and child. Lancet 2014 Nov 15;384(9956):1800-1819. [doi: 10.1016/S0140-6736(14)61277-0] [Medline: 25455250]
8. Letourneau NL, Dennis C, Benzies K, Duffett-Leger L, Stewart M, Tryphonopoulos PD, et al. Postpartum depression is a family affair: addressing the impact on mothers, fathers, and children. Issues Ment Health Nurs 2012 Jul;33(7):445-457. [doi: 10.3109/01612840.2012.673054] [Medline: 22757597]
9. Fisher JR, de Mello MC, Izutsu T, Tran T. The Ha Noi Expert Statement: recognition of maternal mental health in resource-constrained settings is essential for achieving the Millennium Development Goals. Int J Ment Health Syst 2011 Jan 07;5(1):2 [FREE Full text] [doi: 10.1186/1752-4458-5-2] [Medline: 21214891]
10. Stewart D, Robertson E, Dennis C, Grace S, Wallington T. Postpartum depression: literature review of risk factors and interventions. 2003. URL: https://www.who.int/mental_health/prevention/suicide/lit_review_postpartum_depression.pdf [accessed 2019-06-13]
11. Mehta S, Mehta N. An overview of risk factors associated to post-partum depression in Asia. Ment Illn 2014 Mar 04;6(1):5370 [FREE Full text] [doi: 10.4081/mi.2014.5370] [Medline: 25478140]
12. Ho-Yen SD, Bondevik GT, Eberhard-Gran M, Bjorvatn B. Factors associated with depressive symptoms among postnatal women in Nepal. Acta Obstet Gynecol Scand 2007;86(3):291-297. [doi: 10.1080/00016340601110812] [Medline: 17364302]
13. Chandran M, Tharyan P, Muliyil J, Abraham S. Post-partum depression in a cohort of women from a rural area of Tamil Nadu, India: incidence and risk factors. Br J Psychiatry 2002 Dec;181:499-504. [Medline: 12456520]
14. Evagorou O, Arvaniti A, Samakouri M. Cross-cultural approach of postpartum depression: manifestation, practices applied, risk factors and therapeutic interventions. Psychiatr Q 2016 Mar;87(1):129-154. [doi: 10.1007/s11126-015-9367-1] [Medline: 25986531]
15. Chee CYI, Lee DTS, Chong YS, Tan LK, Ng TP, Fones CSL. Confinement and other psychosocial factors in perinatal depression: a transcultural study in Singapore. J Affect Disord 2005 Dec;89(1-3):157-166. [doi: 10.1016/j.jad.2005.09.004] [Medline: 16257451]
16. Abdollahi F, Lye M, Md Zain A, Shariff Ghazali S, Zarghami M. Postnatal depression and its associated factors in women from different cultures. Iran J Psychiatry Behav Sci 2011;5(2):5-11 [FREE Full text] [Medline: 24644441]
17. Shorey S, Chee CYI, Ng ED, Chan YH, Tam WWS, Chong YS. Prevalence and incidence of postpartum depression among healthy mothers: a systematic review and meta-analysis. J Psychiatr Res 2018 Sep;104:235-248. [doi: 10.1016/j.jpsychires.2018.08.001] [Medline: 30114665]
18. Jones E, Coast E. Social relationships and postpartum depression in South Asia: a systematic review. Int J Soc Psychiatry 2013 Nov;59(7):690-700. [doi: 10.1177/0020764012453675] [Medline: 22851133]
19. Cohen S, Wills TA. Stress, social support, and the buffering hypothesis. Psychol Bull 1985 Sep;98(2):310-357. [Medline: 3901065]
20. House JS, Kahn RL, McLeod JD, Williams D. Measures and concepts of social support. In: Cohen S, Syme SL, editors. Social Support and Health. San Diego: Academic Press; 1985:83-108.
21. Heaney C, Israel B. Social networks and social support. In: Glanz K, Rimer BK, Viswanath K, editors. Health Behavior and Health Education: Theory, Research, Practice. San Francisco: Jossey-Bass; 2008:189-210.
22. Chen JM, Kim HS, Sherman DK, Hashimoto T. Cultural differences in support provision: the importance of relationship quality. Pers Soc Psychol Bull 2015 Nov;41(11):1575-1589. [doi: 10.1177/0146167215602224] [Medline: 26351274]
23. Nurullah A. Received and provided social support: a review of current evidence and future directions. Am J Health Stud 2012;27(3):173-188 [FREE Full text]

XSL•FO
RenderX

24. Heh SS, Fu YY, Chin YL. Postpartum social support experience while "doing the month" in Taiwanese women. J Nurs Res 2001 Jun;9(3):13-24. [Medline: 12688569]
25. Liu YQ, Petrini M, Maloni JA. “Doing the month”: postpartum practices in Chinese women. Nurs Health Sci 2015 Mar;17(1):5-14. [doi: 10.1111/nhs.12146] [Medline: 24798890]
26. Rodrigues M, Patel V, Jaswal S, de Souza N. Listening to mothers: qualitative studies on motherhood and depression from Goa, India. Soc Sci Med 2003 Nov;57(10):1797-1806. [Medline: 14499506]
27. Lee JY, Hwang JY. A study on postpartum symptoms and their related factors in Korea. Taiwan J Obstet Gynecol 2015 Aug;54(4):355-363 [FREE Full text] [doi: 10.1016/j.tjog.2014.04.030] [Medline: 26384050]
28. Kim HS, Sherman DK. “Express yourself”: culture and the effect of self-expression on choice. J Pers Soc Psychol 2007 Jan;92(1):1-11. [doi: 10.1037/0022-3514.92.1.1] [Medline: 17201538]
29. Matsumoto D, Yoo SH, Fontaine J. Mapping expressive differences around the world: the relationship between emotional display rules and individualism versus collectivism. J Cross-Cultur Psychol 2008 Jan;39(1):55-74. [doi: 10.1177/0022022107311854]
30. Dennis C, Chung-Lee L. Postpartum depression help-seeking barriers and maternal treatment preferences: a qualitative systematic review. Birth 2006 Dec;33(4):323-331. [doi: 10.1111/j.1523-536X.2006.00130.x] [Medline: 17150072]
31. Holopainen D. The experience of seeking help for postnatal depression. Aust J Adv Nurs 2002;19(3):39-44. [Medline: 12002628]
32. Stevens J, Dahlen H, Peters K, Jackson D. Midwives' and doulas' perspectives of the role of the doula in Australia: a qualitative study. Midwifery 2011 Aug;27(4):509-516. [doi: 10.1016/j.midw.2010.04.002] [Medline: 20889246]
33. Negron R, Martin A, Almog M, Balbierz A, Howell EA. Social support during the postpartum period: mothers' views on needs, expectations, and mobilization of support. Matern Child Health J 2013 May;17(4):616-623 [FREE Full text] [doi: 10.1007/s10995-012-1037-4] [Medline: 22581378]
34. Sjöberg M, Lindgren S. Challenging the roles of “skilled” professionals and “risky” young mothers: peer support, expertise, and relational patterns in Facebook groups. J Technol Human Serv 2017 Sep 13;35(3):247-270. [doi: 10.1080/15228835.2017.1367350]
35. Slomian J, Emonts P, Vigneron L, Acconcia A, Reginster J, Oumourgh M, et al. Meeting the needs of mothers during the postpartum period: using co-creation workshops to find technological solutions. JMIR Res Protoc 2017 May 03;6(5):e76 [FREE Full text] [doi: 10.2196/resprot.6831] [Medline: 28468746]
36. Warren JR, Kvasny L, Hecht ML, Burgess D, Ahluwalia JS, Okuyemi KS. Barriers, control and identity in health information seeking among African American women. J Health Dispar Res Pract 2010;3:68-90 [FREE Full text]
37. Field M. Telemedicine: A Guide to Assessing Telecommunications for Health Care. Washington: National Academies Press; 1996.
38. TeleHealth. Singapore: Smart Nation and Digital Government Office; 2018. URL: https://www.smartnation.sg/what-is-smart-nation/initiatives/Health/telehealth [accessed 2019-06-13]
39. Shorey S, Ng YPM, Ng ED, Siew AL, Mörelius E, Yoong J, et al. Effectiveness of a technology-based supportive educational parenting program on parental outcomes (part 1): randomized controlled trial. J Med Internet Res 2019 Feb 13;21(2):e10816 [FREE Full text] [doi: 10.2196/10816] [Medline: 30758289]
40. Shorey S, Yang YY, Dennis C. A mobile health app-based postnatal educational program (Home-But Not Alone): descriptive qualitative study. J Med Internet Res 2018 Apr 19;20(4):e119 [FREE Full text] [doi: 10.2196/jmir.9188] [Medline: 29674314]
41. Dennis C, Hodnett E, Kenton L, Weston J, Zupancic J, Stewart DE, et al. Effect of peer support on prevention of postnatal depression among high risk women: multisite randomised controlled trial. BMJ 2009 Jan 15;338(jan15 2):a3064. [doi: 10.1136/bmj.a3064]
42. Letourneau N, Secco L, Colpitts J, Aldous S, Stewart M, Dennis C. Quasi-experimental evaluation of a telephone-based peer support intervention for maternal depression. J Adv Nurs 2015 Jul;71(7):1587-1599. [doi: 10.1111/jan.12622] [Medline: 25705786]
43. Lindberg I, Christensson K, Ohrling K. Parents' experiences of using videoconferencing as a support in early discharge after childbirth. Midwifery 2009 Aug;25(4):357-365. [doi: 10.1016/j.midw.2007.06.002] [Medline: 17935844]
44. Matthey S. Telephone based peer support can reduce postnatal depression in women at high risk. Evid Based Ment Health 2009 Aug;12(3):82. [doi: 10.1136/ebmh.12.3.82] [Medline: 19633250]
45. Shamshiri MH, Azargashb E, Beyraghi N, Defaie S, Asbaghi T. Effect of telephone-based support on postpartum depression: a randomized controlled trial. Int J Fertil Steril 2015;9(2):247-253 [FREE Full text] [Medline: 26246884]
46. Dennis C. Postpartum depression peer support: maternal perceptions from a randomized controlled trial. Int J Nurs Stud 2010 May;47(5):560-568. [doi: 10.1016/j.ijnurstu.2009.10.015] [Medline: 19962699]
47. Dennis C. Peer support for postpartum depression: volunteers' perceptions, recruitment strategies and training from a randomized controlled trial. Health Promot Int 2013 Jun;28(2):187-196 [FREE Full text] [doi: 10.1093/heapro/das003] [Medline: 22388589]
48. Gjerdingen DK, McGovern P, Pratt R, Johnson L, Crow S. Postpartum doula and peer telephone support for postpartum depression: a pilot randomized controlled trial. J Prim Care Community Health 2013 Jan;4(1):36-43. [doi: 10.1177/2150131912451598] [Medline: 23799688]

XSL•FO
RenderX

49. Letourneau N, Leger J, Weaver K. Gaining perspective: Peer mentors' experiences with an intervention project for mothers with postpartum depression. Clin Nurs Stud 2017;3(1):29. [doi: 10.5430/cns.v3n1p29]
50. Murphy CA, Cupples ME, Percy A, Halliday HL, Stewart MC. Peer-mentoring for first-time mothers from areas of socio-economic disadvantage: a qualitative study within a randomised controlled trial. BMC Health Serv Res 2008 Feb 27;8:46 [FREE Full text] [doi: 10.1186/1472-6963-8-46] [Medline: 18304334]
51. Shorey S, Chee C, Ng E, Lau Y, Dennis C, Chan YHC. Evaluation of a technology-based peer-support intervention program for preventing postnatal depression (part 1): randomized controlled trial. J Med Internet Res 2019;21(8) [FREE Full text] [doi: 10.2196/12410]
52. Shorey S, Chee C, Chong Y, Ng ED, Lau Y, Dennis C. Evaluation of technology-based peer support intervention program for preventing postnatal depression: protocol for a randomized controlled trial. JMIR Res Protoc 2018 Mar 14;7(3):e81 [FREE Full text] [doi: 10.2196/resprot.9416] [Medline: 29540338]
53. Bellg AJ, Borrelli B, Resnick B, Hecht J, Minicucci DS, Ory M, et al. 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]
54. Eaton L, Doorenbos AZ, Schmitz KL, Carpenter KM, McGregor BA. Establishing treatment fidelity in a web-based behavioral intervention study. Nurs Res 2011;60:6. [Medline: 22048559]
55. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
56. Carter N, Bryant-Lukosius D, DiCenso A, Blythe J, Neville AJ. The use of triangulation in qualitative research. Oncol Nurs Forum 2014 Sep;41(5):545-547. [doi: 10.1188/14.ONF.545-547] [Medline: 25158659]
57. Moher D, Hopewell S, Schulz K, Montori V, Gøtzsche P, Devereaux P. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. BMJ 2010:340. [doi: 10.1136/bmj.c869]
58. McFadden A, Gavine A, Renfrew M, Wade A, Buchanan P, Taylor J. Support for healthy breastfeeding mothers with healthy term babies. Cochrane Database Syst Rev 2017;2:CD001141.
59. Dahlberg U, Aune I. The woman's birth experience: the effect of interpersonal relationships and continuity of care. Midwifery 2013 Apr;29(4):407-415 [FREE Full text] [doi: 10.1016/j.midw.2012.09.006] [Medline: 23399319]
60. Shaw E, Levitt C, Wong S, Kaczorowski J, McMaster University Postpartum Research Group. Systematic review of the literature on postpartum care: effectiveness of postpartum support to improve maternal parenting, mental health, quality of life, and physical health. Birth 2006 Sep;33(3):210-220. [doi: 10.1111/j.1523-536X.2006.00106.x] [Medline: 16948721]
61. Goldberg H. Informed decision making in maternity care. J Perinat Educ 2009;18(1):32-40 [FREE Full text] [doi: 10.1624/105812409X396219] [Medline: 19436598]
62. Jeffers H, Baker M. Continuity of care: still important in modern-day general practice. Br J Gen Pract 2016 Aug;66(649):396-397 [FREE Full text] [doi: 10.3399/bjgp16X686185] [Medline: 27481958]
63. Haggerty JL, Reid RJ, Freeman GK, Starfield BH, Adair CE, McKendry R. Continuity of care: a multidisciplinary review. BMJ 2003 Nov 22;327(7425):1219-1221 [FREE Full text] [doi: 10.1136/bmj.327.7425.1219] [Medline: 14630762]
64. Chen JM, Kim HS, Mojaverian T, Morling B. Culture and social support provision: who gives what and why. Pers Soc Psychol Bull 2012 Jan;38(1):3-13. [doi: 10.1177/0146167211427309] [Medline: 22214884]
65. Green K, Broome H, Mirabella J. Postnatal depression among mothers in the United Arab Emirates: socio-cultural and physical factors. Psychol Health Med 2006 Nov;11(4):425-431. [doi: 10.1080/13548500600678164] [Medline: 17129919]
66. Walen HR, Lachman ME. Social support and strain from partner, family, and friends: costs and benefits for men and women in adulthood. J Soc Personal Relat 2016 Jun 30;17(1):5-30. [doi: 10.1177/0265407500171001]
67. Campbell-Voytal K, Fry McComish J, Visger JM, Rowland CA, Kelleher J. Postpartum doulas: motivations and perceptions of practice. Midwifery 2011 Dec;27(6):e214-e221 [FREE Full text] [doi: 10.1016/j.midw.2010.09.006] [Medline: 21055853]
68. Thomson S, Michelson D, Day C. From parent to 'peer facilitator': a qualitative study of a peer-led parenting programme. Child Care Health Dev 2015 Jan;41(1):76-83. [doi: 10.1111/cch.12132] [Medline: 24673488]
69. Banbury A, Nancarrow S, Dart J, Gray L, Parkinson L. Telehealth interventions delivering home-based support group videoconferencing: systematic review. J Med Internet Res 2018 Feb 02;20(2):e25 [FREE Full text] [doi: 10.2196/jmir.8090] [Medline: 29396387]
70. Singh R, Tay YY, Sankaran K. Causal role of trust in interpersonal attraction from attitude similarity. J Soc Personal Relat 2016 Jul 17;34(5):717-731. [doi: 10.1177/0265407516656826]
71. Barnes J, Senior R, MacPherson K. The utility of volunteer home-visiting support to prevent maternal depression in the first year of life. Child Care Health Dev 2009 Nov;35(6):807-816. [doi: 10.1111/j.1365-2214.2009.01007.x] [Medline: 19719770]
72. Kung WW. Cultural and practical barriers to seeking mental health treatment for Chinese Americans. J Community Psychol 2003 Jan;32(1):27-43. [doi: 10.1002/jcop.10077]
73. Mojaverian T, Hashimoto T, Kim HS. Cultural differences in professional help seeking: a comparison of Japan and the US. Front Psychol 2012;3:615 [FREE Full text] [doi: 10.3389/fpsyg.2012.00615] [Medline: 23426857]
74. Han M, Pong H. Mental health help-seeking behaviors among Asian American community college students: the effect of stigma, cultural barriers, and acculturation. J College Stud Devel 2015;56(1):1-14. [doi: 10.1353/csd.2015.0001]

XSL•FO
RenderX

Abbreviations

PIP: peer support intervention program
PND: postnatal depression
RCT: randomized controlled trial

Edited by G Eysenbach; submitted 27.11.18; peer-reviewed by P Stolte, Y Lau, T Rashid Soron, S Dauber; comments to author 22.03.19; revised version received 16.04.19; accepted 29.04.19; published 29.08.19.

Please cite as:
Shorey S, Ng ED
Evaluation of a Technology-Based Peer-Support Intervention Program for Preventing Postnatal Depression (Part 2): Qualitative Study
J Med Internet Res 2019;21(8):e12915
URL: http://www.jmir.org/2019/8/e12915/
doi:10.2196/12915
PMID:31469080

©Shefaly Shorey, Esperanza Debby Ng. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 29.08.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Associations Between Engagement With an Online Health Community and Changes in Patient Activation and Health Care Utilization: Longitudinal Web-Based Survey

Ruth E Costello[1], MSc; Amrutha Anand[2], MPH; Matt Jameson Evans[2], MRCS (Eng); William G Dixon[1,3], PhD

[1]Centre for Epidemiology Versus Arthritis, Manchester Academic Health Science Centre, The University of Manchester, Manchester, United Kingdom
[2]HealthUnlocked (Everything Unlocked Ltd), London, United Kingdom
[3]NIHR Manchester Biomedical Research Centre, Manchester University NHS Foundation Trust, Manchester, United Kingdom

Corresponding Author:
William G Dixon, PhD
Centre for Epidemiology Versus Arthritis
Manchester Academic Health Science Centre
The University of Manchester
Stopford building
Oxford Road
Manchester, M13 9PT
United Kingdom
Phone: 44 1612751642
Email: will.dixon@manchester.ac.uk

Abstract

Background: Participation in online health communities (OHCs) is a popular trend in the United Kingdom. However, so far, no evidence exists to indicate an association between participation in OHCs and improved health outcomes.

Objective: This study aimed to (1) determine changes in patient activation over 3 months in new users of an OHC, (2) describe patterns of engagement with an OHC, (3) examine whether patients' characteristics at baseline were associated with subsequent patterns of engagement, and (4) determine if patterns of engagement during the 3 months were associated with changes in patient activation, health care utilization, and health status.

Methods: Active new OHC users on HealthUnlocked (HU) were surveyed to measure demographics, levels of patient activation (describing a person's confidence in managing their own health; scale 0-100 with 4 categories), health care utilization, and health status using a Web-based survey at baseline and 3 months. Patient activation at baseline and 3 months was compared (aim 1). Alongside, for a sample of HU users and survey responders, daily OHC website usage data were automatically captured. This was used to identify clusters of engagement with HU (aim 2). For survey responders, baseline characteristics, patient activation, health care utilization, and health status were compared at baseline and 3 months, overall, and between engagement clusters using t tests and chi-square tests (aims 3 and 4).

Results: In 329 people who completed both surveys, baseline activation was most frequently level 3, described as taking action but still lacking confidence. At follow-up, a change of 2.6 points was seen, with the greatest change seen in those at lowest baseline activation levels. In addition, 4 clusters of engagement were identified: low, medium, high, and very high, who were active on HU for a mean of 4, 12, 29, and 59 days, respectively. Survey responders were more commonly high or very high engagers. Baseline activation was highest in low and very high engagers. Overall activation increased over time in all engagement groups. Very high engagers had the greatest improvement in activation (5 points), although the average change was not above what is considered clinically meaningful for any group. Fewer accident and emergency visits were seen at follow-up in those with higher engagement, although this trend was not seen for other health care utilization measures. There was no change in health status at 3 months.

Conclusions: This observational study provides some insight into how patterns of engagement with OHCs are associated with changes in patient activation, health care utilization, and health status. Over 3 months, overall, the change in activation was not clinically significant, and there were some indications that OHCs may be of benefit to particular groups. However, the study limitations prevent firm conclusions about causal relationships.

XSL•FO
RenderX

(J Med Internet Res 2019;21(8):e13477) doi:10.2196/13477

KEYWORDS

self-management; chronic disease; health information exchanges; digital health; peer support; peer-to-peer support; online support groups; internet

Introduction

The internet is used by more than 90% of the population in the United Kingdom, according to a national survey in 2017. Of those who had used the internet in the last 3 months, 53% of people had looked for health-related information [1]. In addition to accessing websites from trusted organizations, such as the National Health Service, and patient organizations or charities, for example, Cancer Research UK, people are increasingly using peer-to-peer online health communities (OHCs). OHCs are forums where people with specific conditions can share information and experiences with other people with the same condition through peer support and discussion. A survey of US adults found that 1 in 4 people read or watched commentary of another person's experiences of health or medical issues on the Web [2]. Learning about other people's experiences online may have a number of impacts, ranging from finding information and feeling supported to affecting behavior [3].

One way that OHCs may be particularly useful is in helping with self-management of chronic diseases. Studies analyzing OHC content have found that conversations may support self-efficacy [4,5]. A cross-sectional survey of users of OHCs for 3 different conditions (breast cancer, fibromyalgia, and arthritis) found that 74% felt they now had *the right knowledge to manage their illness* after participation in an OHC [6]. One might therefore argue that accessing online information and peer support would lead to empowerment, improved coping, reduced uncertainty, and potentially even reduced health care utilization [7]. Studies have found OHC use to be associated with increased empowerment [6,8,9] with themes of feeling better informed and social support [6-8,10,11]. Conversely, there are arguments that information in OHCs may not be reliable or accurate, and peer support, rather than information from health care professionals, may amplify anxiety [12,13].

Patient activation is defined as a person's knowledge, skills, and confidence in managing their health or chronic condition [14]. As a concept, patient activation differs slightly from empowerment; it builds on self-efficacy and aims to capture the development of a patient's engagement with managing their own health, from believing they have an active role to having the confidence to self-manage their own health when under stress [15]. The measure has been shown to be reliable and valid [15] in a variety of countries and populations [16-21]. Higher activation has been associated with positive health behaviors such as attending screenings and eating 5 or more fruits and vegetables per day [22]. Studies have shown that those with lower activation scores are more likely to be hospitalized and more likely to visit an accident and emergency department [23]. At present, it remains uncertain as to whether interventions to improve disease knowledge can improve activation and in turn lead to improved behaviors and health outcomes. Studies investigating similar concepts such as empowerment and self-efficacy are frequently cross-sectional [5,6]. To date, only 1 study has examined the relationship between OHCs and activation specifically. In this study, experienced users of an OHC had higher activation scores than new users, and activation scores in both groups increased after 3 months' OHC use, with higher scores in those who self-reported using the site more frequently [24]. This study did not investigate subsequent health care utilization.

Studying the relationship between engagement with OHCs and patient activation and health outcomes is complex because of (1) confounding by indication (where people accessing the platform are inherently different from those who do not, which in turn affects their probability of the outcomes of interest) [25], (2) natural changes in health care utilization at different stages of disease (eg, general practitioner [GP] visits may naturally be higher around the time of diagnosis than in subsequent periods), and (3) challenges in quantifying the *exposure* of OHC engagement [26]. However, because website visits leave digital traces, it is possible to measure how commonly people interact with a site, thereby allowing the relationship between different engagement patterns and outcomes to be studied.

This study aimed to (1) determine change in activation over 3 months in new users of an OHC, (2) describe patterns of engagement with an OHC, (3) examine whether patient characteristics at baseline were associated with subsequent patterns of engagement, and (4) determine if patterns of engagement during the 3 months were associated with changes in patient activation, health care utilization, and health status.

Methods

Setting

HealthUnlocked (HU) is a host to multiple OHCs with more than 4.5 million visitors each month. HU has more than 700 communities for a variety of health conditions as well as hosting communities to support aspects of well-being, for example, weight loss and healthy eating [27]. These OHCs are built in collaboration with patient organizations, who moderate the communities, to ensure safety of its users and verify credibility of content shared. Often, expert users with no formal association with a patient organization volunteer to moderate communities focused on health and well-being.

Once registered on the platform, a user can choose to follow communities relevant to their health interest and post questions, updates, or any information that they wish to share, or reply to previous posts from other users. In addition to text, users can post images on these OHCs. Users can also like other posts or follow other users to build a network around them. An exemplar post is included in Multimedia Appendix 1.

XSL•FO
RenderX

Data

This study used data from a survey (Multimedia Appendix 2) designed and run by HU in conjunction with a research team at King's Health Economics to understand health and economic outcomes in new users of HU. The anonymized survey data were sent to the University of Manchester after the survey ended, and this analysis was designed to meet our study aims.

Population

HU is freely available to the public, and people sign up with their email address and password. Those who signed up, followed at least one of the communities included in this study, and were active on the website between 48 and 72 hours after signing up were eligible for the study and were emailed the survey. People who completed the baseline survey were sent a follow-up survey 3 months later. Reminder emails for both surveys were sent 2 days after the original email. The survey started in September 2016 and continued until the sample size reached at least 300. This was based on the study having 90% power to identify a mean difference of 3 between baseline and follow-up in patient activation score.

Survey

The survey asked about demographics (eg, age, gender, occupation, education, and ethnicity); information about health: main diagnosis (collected as free text and verified against community group followed); disease duration in response to the question "How long since you were diagnosed with the condition?" with the options less than a year ago, 1 to 3 years, 4 to 6 years, 7 to 9 years, and 10 years or more; and comorbidities in response to the question "Do you currently have other long-term concerns in addition to your diagnosis?" with the options No, Yes I have 1 more, Yes I have 2 more, and Yes I have more than 3. Patient activation was measured using the Patient Activation Measure (PAM). The measure contains 13 statements where respondents indicate whether they strongly agree (4 points), agree (3 points), disagree (2 points), or strongly disagree (1 point) with each of the statements [15]. Using a standardized table, these scores were converted to a score out of 100 where a higher score indicates a person showing greater activation. People with a score of 100 at baseline or follow-up were removed from the cohort during analysis as a score of 100 is considered implausible (Personal communication, C Delaney, 2018). PAM score was then converted to 4 levels of activation, as defined by the authors: level 1 (PAM score <47): overwhelmed and passive in managing their own health; level 2 (PAM score 47.1-55.1): lack of knowledge and confidence; level 3 (PAM score 55.2-72.4): taking action but still lacking confidence; and level 4 (PAM score 72.5-100): have adopted good health behaviors but may have problems when under stress [15]. Health status was measured using the EuroQol-5D (EQ-5D), which contained 5 questions about mobility, self-care, usual activities, pain, and depression or anxiety. Each question was scored between 1 and 5, and the score for each of these questions was weighted using a value set for the UK population and a single score created with the anchors 0 and 1 [28]. The health care utilization questionnaire asked, "In the last 3 months, roughly how many times have you visited the following healthcare services?": GP, outpatient clinic, primary care nurse, accident & emergency (A&E) with the options never, 1 to 3 times, 4 to 6 times, and more than 6 times, and "In the last 3 months, roughly how many days have you spent admitted in a hospital?": a free-text box allowed respondents to indicate the number of days. For the first question, categories 4 to 6 times and more than 6 times were combined because of small numbers. Number of days in hospital was categorized into none, 1 to 5 days, 6 to 10 days, and more than 10 days based on the spread of the data.

HealthUnlocked Engagement

Engagement with HU was determined through the following measures, which are automatically captured daily: pages viewed, number of clicks anywhere on the website, number of community groups followed, number of users followed (subscribing to or following a community or user means posts from these communities or users will appear in the subscribers newsfeed), posts liked, written comments, and primary posts (starting a post) for each user. A daily count of each engagement measure was provided by HU for all people who completed both baseline and follow-up surveys and a random sample of 336 people who completed only the baseline survey and a random sample of 337 who completed neither survey.

Analysis

Mean PAM scores at baseline and follow-up were compared using a *t* test. The proportion of people at each PAM level at baseline and follow-up were compared. Users were then grouped into clusters based on their daily HU engagement data. First, for each day, a person was flagged as having engaged with HU on that day if any of the HU engagement measure counts were not zero. As this was time series data, a first-order Markov Mixture model with an expected maximization algorithm was used to identify clusters [29]. First, the model identified the states of engagement each day, with 3 latent states assumed: high engagement, low engagement, or disengaged. Everybody started at high engagement, and disengagement was assumed to be an *absorbing state* after which there would be no further engagement. People were then clustered based on transitional probabilities of changing engagement state. The optimum number of cluster groups was identified using the elbow method [30]. Baseline characteristics were compared among cluster groups, and the mean number of days of engagement for each cluster group was reported.

Patient activation score, EQ-5D score, and health care utilization at baseline and follow-up were compared between cluster groups. Health care utilization measures were reduced to binary measures of whether participants had any visits in the last 3 months because of low numbers of people with a high number of visits. Box and whisker plots were used to show the distribution of PAM at baseline and follow-up. Kruskal-Wallis tests checked if there was a statistically significant difference in PAM scores between cluster groups. The proportion with a change of PAM score by more than 5 points (a suggested clinically meaningful difference [14]) was reported by engagement cluster. PAM score was then categorized into levels and compared at baseline and follow-up, with the proportions where PAM had increased, remained stable, and reduced reported by cluster group. EQ-5D scores were compared at

XSL•FO
RenderX

baseline and follow-up, and a *t* test was used to see if there was a statistically significant difference. The percentage of people with each type of health care visit at baseline and follow-up for each cluster group was compared using chi-square tests.

Ethical Approval

As this survey was service evaluation conducted by HU and King's Health Economics, NHS ethical approval was not required. Ethical approval for the analysis was confirmed as not required by the University of Manchester's ethics committee, as the data were already collected and were anonymized.

Results

Survey Response

The survey was sent to 9469 people; 990 people completed the baseline survey, of whom 329 completed the follow-up survey and had HU usage data available. Of those who completed the follow-up survey, 78.5% (258/329) were aged 50 years or older, 76.6% (252/329) were female, and 93.0% (305/328) were white. 87 (26.4%) were from musculoskeletal community groups (fibromyalgia, lupus, rheumatoid arthritis, polymyalgia rheumatica and giant cell arteritis, pain), and 67 (20.4%) were from endocrine (diabetes and thyroid) community groups (Table 1).

Patient Activation

There were 15 people with a PAM score of 100 at baseline, follow-up, or both; therefore, change in PAM score is reported for only 314 people. For those who completed both surveys, the mean PAM scores at baseline and follow-up were 60.2 and 62.8, respectively, a statistically significant difference of 2.6 points (standard deviation: 8.4 points, $P<.001$). When stratified by baseline PAM level (1 and 2 vs 3 and 4), those at levels 1 and 2 had a statistically significant 5.5-point increase (95% CI 4.1-6.8; $P<.001$). Those at levels 3 and 4 had a nonsignificant 1.1-point increase (95% CI 2.3 to −0.05). When categorized into PAM levels, nearly half (49.4%, 155/314) were at level 3 (taking action but still lacking confidence), and overall, PAM level increased at follow-up (Multimedia Appendix 3).

Engagement With HealthUnlocked

HU activity data were available across the 3-month period for all 329 participants who completed both surveys, random samples of 336 people who did not complete either survey, and 337 people who completed baseline only (total 1002).

Those who completed both surveys engaged with HU more frequently (median: 47 days) than those who only completed baseline (median: 24 days) or did not complete either survey (median: 9 days), although there was a wide spread in the number of people engaged with HU in each response group (Multimedia Appendix 4). In terms of people's activities at visits to HU, 50.90% (510/1002) of participants posted at least once with a median 1 post per person (interquartile range [IQR]: 0-2) and a maximum of 84 posts over 3 months. A total of 15,431 comments were made by 63.07% (632/1002) of participants, with a median of 2 comments (IQR: 0-11) and a maximum of 1549 comments over 3 months. Those who did not complete either survey had fewer written posts, comments, and posts liked per visit to HU than those who completed both surveys, whereas those who completed the baseline survey only had similar numbers of written posts, comments, and likes per visit compared with those who completed both surveys.

The hidden Markov model identified 4 clusters: (1) low engagers: (142/1002, 14.17%) who were active on HU for a mean of 4.4 (SD 2.1) days before not visiting the platform further; (2) medium engagers: (216/1002, 21.55%) who were active on HU for a mean of 11.9 (SD 6.3) days; (3) high engagers: (338/1002, 33.72%) who were active for a mean of 29.1 (SD 13.0) days; and (4) very high engagers: (306/1002, 30.54%) who were active for a mean of 59.2 (SD 22.2) days (Figure 1).

The majority of those completing the follow-up survey were high or very high engagers (114/329, 34.7%; and 163/329, 49.5%, respectively). The mean number of active days was slightly higher in those who completed the survey (Table 2). All further results refer to those who completed both surveys.

XSL•FO
RenderX

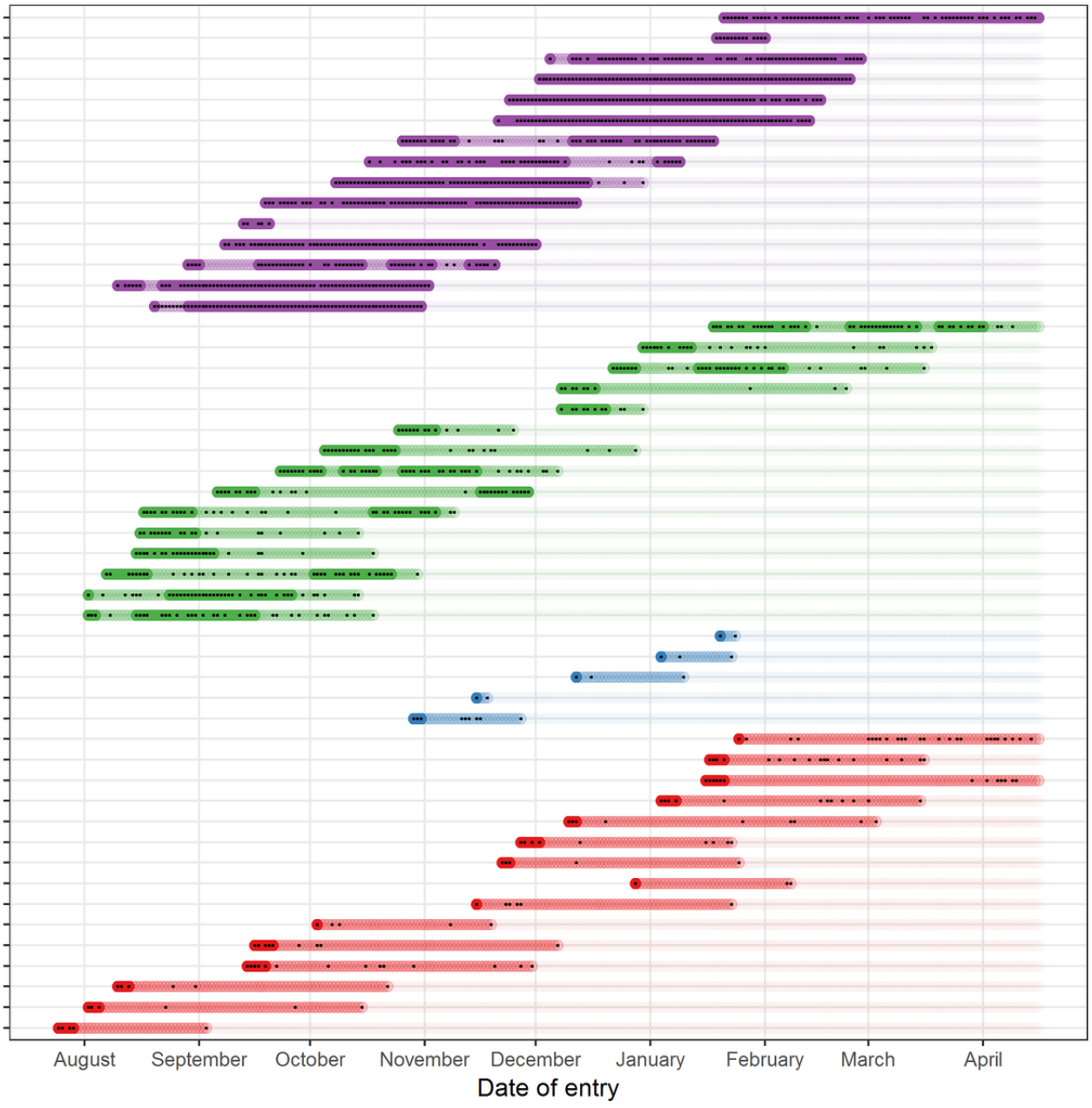
Table 1. Baseline characteristics of respondents (N=329).

Characteristics	Values, n (%)
Age (years)	
<40	26 (7.9)
40-49	45 (13.7)
50-59	94 (28.6)
60-69	118 (35.9)
>70	46 (14.0)
Gender	
Male	77 (23.4)
Female	252 (76.6)
Ethnicity	
White	305 (93.0)
Asian	6 (1.8)
Black/African/Caribbean	7 (2.1)
Hispanic/Latino	1 (0.3)
Multiple ethnicities	9 (2.7)
Missing	1 (0.3)
Employment status	
Employed	97 (29.6)
On sick leave, unable to work	37 (11.3)
Retired	160 (48.8)
Student	3 (0.9)
Unemployed	31 (9.5)
Missing	1 (0.3)
Education	
Primary school	5 (1.5)
Secondary school	156 (47.6)
University degree	116 (35.4)
Postgraduate degree	51 (15.5)
Missing	1 (0.3)
Comorbidities	
None	137 (42.0)
1	103 (31.6)
≥2	86 (26.4)
Missing	3 (0.9)
Community	
Cardiovascular	41 (12.5)
Respiratory	12 (3.6)
Cancers	34 (10.3)
Mental health	7 (2.1)
Digestive system	44 (13.4)
Endocrine	67 (20.4)
Genitourinary	16 (4.9)

XSL•FO
RenderX

Characteristics	Values, n (%)
Musculoskeletal	87 (26.4)
Nervous system	4 (1.2)
Blood disorders	16 (4.9)
Reproductive	1 (0.3)
Time since diagnosis (years)	
<1	113 (35.1)
1-6	105 (32.6)
>7	104 (32.3)
Missing	7 (2.1)

Figure 1. The number of days of engagement with HU and states of engagement by engagement cluster for a sample of all users (N=50). Each line represents a respondent, each dot represents a day the respondent engaged with the HU platform, the colors represent the engagement clusters, where blue indicates low engagers, red medium engagers, green high engagers, and purple very high engagers. The shading of the color represents the state of engagement where dark color indicates high engagement, and light color indicates low engagement. HU: HealthUnlocked.



XSL•FO
RenderX

Table 2. Number of days of engagement with HealthUnlocked by cluster.

Engagement cluster	Both surveys completed (N=329)		Baseline survey completed (N=337)		No surveys completed (N=336)	
	n (%)	Active days, mean (SD)	n (%)	Active days, mean (SD)	n (%)	Active days, mean (SD)
Low	18 (5.4)	6 (2.4)	31 (9.1)	5 (3)	93 (27.6)	3 (3)
Medium	34 (10.3)	14.3 (6.5)	69 (20.4)	13 (10)	113 (33.6)	9 (7)
High	114 (34.6)	34.4 (12.9)	144 (42.7)	26.5 (17.5)	80 (23.8)	22 (17)
Very high	163 (49.5)	66.4 (15.8)	93 (27.5)	60 (28)	50 (14.8)	55 (60)

Characteristics

The characteristics of each cluster of survey responders are shown in Table 3. High and very high engagers had a higher proportion of females (95/114, 83.3% and 124/163, 76.1%, respectively) compared with medium and low engagers (22/35, 64% and 11/18, 61%, respectively). Higher engagers had less comorbidity: 22.2% (n=36/162) of very high engagers had 2 or more comorbidities compared with 55.6% (n=10/18) of low engagers, and a shorter time since diagnosis: 39.0% (n=62/159) of very high engagers were diagnosed less than a year ago compared with 19% (n=3/16) of low engagers.

Patient Activation by Engagement Cluster

Figure 2 shows the distribution of PAM scores for each engagement cluster at baseline and follow-up. Median baseline PAM scores differed little across the 4 engagement groups with a difference of only 6 points from highest to lowest. PAM scores increased at follow-up in all engagement groups. Medium engagers had the lowest change of 0.5 points, and very high engagers had the highest change of 5.1 points.

When scores were categorized into PAM levels, 81% (13/16) of low engagers were at level 3 or 4 at baseline, increasing 94% (15/16) at follow-up. All other engagement clusters had around half at level 3 at baseline. Medium engagers had the highest proportion at level 2, and very high engagers had the highest proportion at level 4. At follow-up, the proportion at level 3 increased for all engagement clusters, and the only engagement cluster with an increased proportion at level 4 was the high engagement cluster (13.6% [15/110] at baseline vs 17.3% [19/110] at follow-up). Moreover, 22.2% (36/163) and 20.5% (33/163) of very high engagers were at level 4 at both baseline and follow-up, respectively (Table 4).

A PAM score increase of at least 5 points was seen in 35.0% (110/314) of respondents, with the highest proportion in the very high engagers (63/156, 40.4%) and lowest proportion in medium engagers (6/32, 18%). A PAM score decrease of at least 5 points was seen in 15.0% (47/314), with proportions similar across engagement clusters.

Health Status

Respondents had a mean EQ-5D score of 0.69 at baseline and 0.70 at follow-up, where 1 indicates perfect health, and less than zero indicates a state worse than death. There was little difference in average health status between baseline and follow-up within the 4 engagement clusters, with a maximum mean change within groups of 0.02 units.

Health Care Utilization

At baseline, 88.4% (283/320) of people visited their GP at least once and 37.5% (102/272) visited a primary care nurse at least once in the previous 3 months. 64.2% (190/296) people visited outpatients in the previous 3 months. 21.7% (62/286) visited A&E and 21.6% (71/328) were hospitalized in the previous 3 months. At follow-up, the proportion of people visiting a GP was slightly lower at 83.8% (268/320), and the proportion visiting a primary care nurse was similar to baseline. The proportion of people visiting outpatients or A&E or being hospitalized reduced, the biggest reduction being in A&E visits where only 12.6% (36/285) visited A&E at follow-up (Table 5). When stratified by engagement cluster, all engagement clusters had fewer people visiting a GP, outpatients, and being hospitalized at least once at follow-up, except medium engagers where the proportion of people visiting a GP remained the same. The proportion of people visiting a primary care nurse varied across engagement clusters with no clear pattern. The only statistically significant difference between engagement clusters was for A&E visits at follow-up, where there was a trend toward those with greater engagement with HU having a smaller proportion of people visiting A&E at follow-up. Low engagers had a 12% more people visiting A&E, and very high engagers had a 12.6% fewer people visiting A&E at follow-up (Table 5).

XSL•FO
RenderX

Table 3. Baseline characteristics of baseline and follow-up survey respondents by engagement cluster (N=329).

Baseline characteristic	Engagement cluster, n (%)			
	Low (n=18)	Medium (n=34)	High (n=114)	Very high (n=163)
Age (years)				
<40	1 (5)	3 (8)	14 (12.3)	8 (4.9)
40-49	3 (16)	7 (20)	17 (14.9)	18 (11.0)
50-59	6 (33)	6 (17)	36 (31.6)	46 (28.2)
60-69	6 (33)	15 (44)	33 (28.9)	64 (39.3)
>70	2 (11)	3 (8)	14 (12.3)	27 (16.6)
Gender				
Male	7 (38)	12 (35)	19 (16.7)	39 (23.9)
Female	11 (61)	22 (64)	95 (83.3)	124 (76.1)
Ethnicity				
White	17 (94)	32 (94)	101 (88.6)	155 (95.7)
Asian	1 (5)	1 (2)	2 (1.8)	2 (1.2)
Black/African/Caribbean	0 (0)	1 (2)	6 (0.0)	0 (0.6)
Latino	0 (0)	0 (0)	0 (5.3)	1 (0.0)
Multiple ethnicities	0 (0)	0 (0)	5 (4.4)	4 (2.5)
Missing	0	0	0	1 (0.6)
Employment status				
Employed	3 (16)	8 (23)	46 (40.4)	40 (24.7)
On sick leave, unable to work	2 (11)	4 (11)	13 (11.4)	18 (11.1)
Retired	9 (50)	17 (50)	44 3(8.6)	90 (55.6)
Student	1 (5)	1 (2)	1 (0.9)	0 (0.0)
Unemployed	3 (16)	4 (11)	10 (8.8)	14 (8.6)
Missing	0	0	0	1 (0.6)
Education				
Primary school	1 (5)	1 (2)	3 (2.6)	0 (0.0)
Secondary school	10 (55)	19 (55)	54 (47.4)	73 (45.1)
University degree	5 (27)	12 (35)	41 (36.0)	58 (35.8)
Postgraduate degree	2 (11)	2 (5)	16 (14.0)	31 (19.1)
Missing	0	0	0	1 (0.6)
Comorbidities				
None	4 (22)	18 (52)	46 (41.1)	69 (42.6)
1	4 (22)	5 (14)	37 (33.0)	57 (35.2)
≥2	10 (55)	11 (32)	29 (25.9)	36 (22.2)
Missing	0	0	2 (1.8)	1 (0.6)
Community				
Cardiovascular	2 (11)	3 (8)	14 (12.3)	22 (13.5)
Respiratory	1 (5)	1 (2)	4 (3.5)	6 (3.7)
Cancers	2 (11)	4 (11)	8 (7.0)	22 (13.5)
Mental health	5 (27)	1 (2)	2 (1.8)	2 (1.2)
Digestive system	4 (22)	5 (14)	12 (10.5)	22 (13.5)
Endocrine	4 (22)	7 (20)	21 (18.4)	35 (21.5)

XSL•FO
RenderX

Baseline characteristic	Engagement cluster, n (%)			
	Low (n=18)	Medium (n=34)	High (n=114)	Very high (n=163)
Genitourinary	5 (27)	2 (5)	11 (9.6)	3 (1.8)
Musculoskeletal	20 (30)	10 (29)	34 (29.8)	39 (23.9)
Nervous system	1 (1)	1 (2)	1 (0.9)	2 (1.2)
Blood disorders	1 (1)	1 (1)	6 (5.3)	10 (6.1)
Reproductive	0 (0)	1 (0)	1 (0.9)	0 (0.0)
Time since diagnosis (years)				
<1	3 (18)	8 (23)	40 (35.4)	62 (39.0)
1-6	5 (31)	12 (35)	32 (28.3)	56 (35.2)
≥7	8 (50)	14 (41)	41 (36.3)	41 (25.8)
Missing	2 (11.1)	0	1 (0.9)	4 (2.5)

Figure 2. Box and whisker plot of PAM scores by engagement cluster (N=314). Box plots represent the median (central line), interquartile range (box), range, excluding outliers (whiskers), and outliers (dots) of the percentage of patients within each engagement cluster. PAM: Patient Activation Measure.

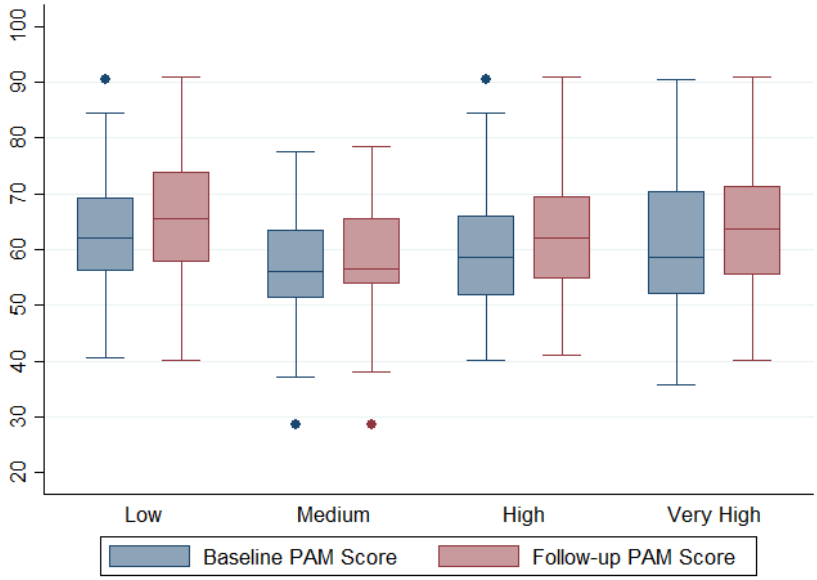


Table 4. Baseline and follow-up Patient Activation Measure level by engagement cluster.

PAM[a] level	Low, n (%)	Medium, n (%)	High, n (%)	Very high, n (%)
Baseline PAM level				
1	1 (6.3)	2 (6.3)	11 (10.0)	8 (5.1)
2	2 (12.5)	12 (37.5)	28 (25.5)	42 (26.9)
3	10 (62.5)	16 (50.0)	56 (50.9)	73 (46.8)
4	3 (18.7)	2 (6.3)	15 (13.6)	33 (21.2)
Follow-up PAM level				
1	1 (6.3)	3 (9.4)	6 (5.5)	3 (1.9)
2	0 (0.0)	7 (21.9)	22 (20.0)	31 (19.9)
3	11 (68.8)	20 (62.5)	63 (57.3)	90 (57.7)
4	4 (25.0)	2 (6.3)	19 (17.3)	32 (20.5)

[a]PAM: Patient Activation Measure.

XSL•FO
RenderX

Table 5. Health care utilization at baseline and follow-up by engagement cluster.

Healthcare utilization type	Low, n (%)	Medium, n (%)	High, n (%)	Very high, n (%)	Total, n (%)
At least one GP[a] visit in the last 3 months					
Baseline	15 (88.2)	28 (84.8)	103 (92.8)	137 (86.2)	283 (88.4)
Follow-up	14 (82.4)	28 (84.8)	97 (87.4)	129 (81.1)	268 (83.8)
At least one nurse visit in the last 3 months					
Baseline	7 (43.8)	6 (21.4)	34 (36.6)	55 (40.7)	102 (37.5)
Follow-up	5 (31.3)	9 (32.1)	27 (29)	60 (44.4)	101 (37.1)
At least one Outpatient visit in the last 3 months					
Baseline	11 (64.7)	16 (51.6)	65 (63.1)	98 (67.6)	190 (64.2)
Follow-up	9 (52.9)	14 (45.2)	62 (60.2)	85 (58.6)	170 (57.4)
At least one A&E[b] visit in the last 3 months					
Baseline	2 (11.8)	5 (17.2)	27 (27.8)	28 (19.6)	62 (21.7)
Follow-up	4 (23.5)	4 (13.8)	18 (18.6)	10 (7)	36 (12.6)
At least one hospitalization in the last 3 months					
Baseline	4 (22.2)	7 (21.2)	28 (24.6)	32 (19.6)	71 (21.6)
Follow-up	1 (5.6)	4 (12.1)	22 (19.3)	24 (14.7)	51 (15.5)

[a]GP: general practitioner.

[b]A&E: accident & emergency.

Discussion

Principal Findings

This study found that a group of HU users, who had completed a baseline and follow-up survey, had, on average, a moderate activation score at level 3 *taking action but lacking in confidence*. The improvement in activation over 3 months was, on average, only a modest 2.6 points overall. Overall, 1 in 3 respondents had a lower baseline PAM at levels 1 or 2: this group had the highest change in PAM, with an average increase of 5.8 points, a change thought to be clinically meaningful [14]. There were 4 different levels of engagement with the HU platform (low, medium, high, and very high engagers). Very high engagers used the platform on average 60 days over 3 months, were more frequently female, had no comorbidities, and a diagnosis within the previous year. Their activation increased the most over 3 months. Perhaps, indicating that those most recently diagnosed and with few comorbidities gain the most benefit from high engagement with HU. In terms of health care utilization, overall health care utilization reduced over follow-up. Those who engaged most with HU had fewer visits to A&E at follow-up, although this trend was not seen in other health utilization measures. If these findings represent a causal relationship (see below), it would have important implications for how OHCs can improve outcomes in patients with long-term conditions. Information provision from health care professionals and emerging initiatives such as social prescribing [31] could include directing patients to OHCs. Robust evidence on the effectiveness of OHCs as well as their cost-effectiveness would allow clearer positioning within the armamentarium of treatments for people living with health conditions.

Representativeness

The study reports on a population of people who completed a baseline survey and a follow-up survey 3 months later. We were able to compare this group with people who completed only the baseline survey and a random sample of other HU users. We noted that survey responders had higher levels of online engagement than those who did not complete the survey. This is perhaps unsurprising as those more motivated to engage with the platform may be more likely to complete the surveys. Nonetheless, this does not detract from comparisons among engagement groups in our study. Thinking further about representativeness, the population for this study was predominately female, older than 50 years, and White. A study of health-related social media users in the United States found similar proportions of health forum users were female [32]; therefore, this population may be a true representation of OHC users in terms of gender. It was expected that users of OHCs would be a younger population, as seen in other studies [6,24,32]; therefore, our older cohort may reflect some selection bias related to willingness to complete surveys. It has been shown that there are still digital disparities in terms of ethnicity, which may reflect why our sample is predominantly White [33]. It has been shown that African Americans had lower PAM levels; these were shown to be mediated through education and health literacy [34,35]. Unfortunately, the numbers are too small to investigate whether PAM was lower in those of non-Caucasian ethnicity in this study.

Previous Studies

Activation, as measured by the PAM, was similar to a UK sample whose mean PAM score was 59.4. Interestingly, when categorized into levels, only 17% were at level 4 at baseline in

XSL•FO
RenderX

this study compared with 21% of a random sample of the UK population, many with chronic conditions [36]. One might have expected patients accessing an OHC to have had higher levels of activation. In a study that examined activation in a Hebrew online social network, the authors found that people who were experienced users of the online social network, classified as those who had used the site for 6 months or more, had significantly higher PAM scores than new users, with a mean PAM 69.3 points for experienced users compared with 62.8 points for new users [24]. Both new and experienced users had higher PAM scores than scores in this study where responders had a mean score of 60.2 and 62.8 points at baseline and follow-up, respectively, although in this study *experienced users* had only used the OHC for 3 months.

Strengths and Limitations

This was a prospective study of changes to patient activation, health care utilization, and health status with the ability to associate changes with different patterns of engagement with an OHC, with a reasonably large sample size. Meaningful changes in activation were seen in some groups; however, we need to be careful in our interpretation of these findings and must consider some important limitations. First, there was no control group; therefore, we do not know how activation changes in people who did not use an OHC. This makes it difficult to make causal inferences: the small increases in PAM observed in all groups may well be an expected change from the point at which someone signs up to an OHC. It is indeed reasonable to hypothesize that people will sign up at times of greater clinical need. Over the course of the subsequent 3 months, their activation and health care utilization might change for the better regardless of OHC use. Although the clustering allowed identification of those who used HU very little and enabled comparisons across levels of engagement, the numbers were very small, with only 18 and 34 participants in the low and medium engagement groups who completed both baseline and follow-up surveys, making the comparisons across engagement groups less robust. Second, as already mentioned, this was a self-selecting population with only 329 people completing both baseline and follow-up surveys of over 9000 people contacted. This means the sample may not be representative of users of the OHC as a whole, with a skew toward those who engage more with the site. Third, there have been very few longitudinal studies where PAM has been repeatedly measured, which makes it difficult to interpret change in PAM over time. It has been suggested that a change of 5 points is a meaningful difference [14], and we found that 1 in 3 had an increase in PAM score of 5 points or more, although 1 in 10 had a decrease in PAM of 5 or more points. The appropriateness of this threshold for clinical importance is somewhat questionable having been derived from cross-sectional data, where 5 points was identified as the common difference in mean PAM score in people with healthy versus unhealthy behaviors [14]. Fourth, the study's follow-up was 3 months, which is not very long in terms of disease course and may not be long enough to identify a clear change in activation, health status, or health care utilization. We found a mean increase in PAM of 2.6 points, which is a small change—a longer follow-up may have allowed us to identify a larger change were one to transpire. We were unable to show, despite the small increase in activation, any significant change in health status. This may be because of the length of follow-up. Any interpretation of health status is again hampered by the lack of a control group. In the absence of any engagement with an OHC, it could be argued that health status would either improve (in response to a recent diagnosis and treatment) or worsen (because of progression of disease). We therefore do not know how the OHC engagement has influenced health status. Given these limitations, further investigation is warranted to see how activation changes seen compare with a control population and if certain groups of patients may benefit from OHC use, such as newly diagnosed patients. Understanding what functions within OHCs would deliver better outcomes would also be worthy of future investigation.

Conclusions

The main findings of this study are that HU users have varied levels of activation when they start using the platform. Patient activation seems to increase over time, although the extent of change did not seem to differ markedly between different levels of platform engagement. Activation increased the most in those with very high engagement with the HU platform and in those with low activation at baseline; however, it is unknown whether these improvements would have been seen irrespective of the use of the platform. Understanding the impact of participation in an OHC on health outcomes will require studies designed specifically to examine this putative causal association.

Acknowledgments

The authors would like to thank Professor Paul McCrone and Huajie Jin for their work in developing and refining the methodology for this study. They are grateful to Dr John McBeth for comments on a draft manuscript. They are grateful to David Selby for his help with hidden Markov modeling. This work was supported by the Arthritis Research UK Centre for Epidemiology: grant number 20380.

Authors' Contributions

MJE and AA contributed to survey design and were responsible for acquisition of the data, REC and WGD designed the analyses and interpreted the data, REC conducted the analyses and drafted the manuscript, and all authors critically revised the manuscript and approved the final version.

XSL•FO
RenderX

Conflicts of Interest

WGD has received consultancy fees from Google and Beyer. MJE is the Chief Medical Officer and cofounder of HealthUnlocked.com. AA is an employee of HU. REC has no conflicting interests to declare.

Multimedia Appendix 1
Exemplar post from HealthUnlocked.
[PDF File (Adobe PDF File), 342KB - jmir_v21i8e13477_app1.pdf]

Multimedia Appendix 2
Study survey.
[PDF File (Adobe PDF File), 375KB - jmir_v21i8e13477_app2.pdf]

Multimedia Appendix 3
Baseline and follow-up patient activation levels.
[PDF File (Adobe PDF File), 99KB - jmir_v21i8e13477_app3.pdf]

Multimedia Appendix 4
3. Box and whisker plot of the number of days of engagement by the number of surveys completed (N=1002).
[PDF File (Adobe PDF File), 89KB - jmir_v21i8e13477_app4.pdf]

References

1. Office for National Statistics. 2017. Internet Access – Households and Individuals, Great Britain: 2017 URL: https://www.ons.gov.uk/peoplepopulationandcommunity/householdcharacteristics/homeinternetandsocialmediausage/bulletins/internetaccesshouseholdsandindividuals/2017 [accessed 2019-08-09]
2. Pew Research Center. 2013. The Diagnosis Difference URL: https://www.pewresearch.org/science/2013/11/26/the-diagnosis-difference/ [accessed 2019-08-09]
3. Ziebland S, Wyke S. Health and illness in a connected world: how might sharing experiences on the internet affect people's health? Milbank Q 2012 Jun;90(2):219-249 [FREE Full text] [doi: 10.1111/j.1468-0009.2012.00662.x] [Medline: 22709387]
4. Wang Y, Willis E. Supporting self-efficacy through interactive discussion in online communities of weight loss. J Health Psychol 2018 Sep;23(10):1309-1320. [doi: 10.1177/1359105316653264] [Medline: 27338630]
5. Willis E. Patients' self-efficacy within online health communities: facilitating chronic disease self-management behaviors through peer education. Health Commun 2016;31(3):299-307. [doi: 10.1080/10410236.2014.950019] [Medline: 26325224]
6. van Uden-Kraan CF, Drossaert CH, Taal E, Seydel ER, van de Laar MA. Participation in online patient support groups endorses patients' empowerment. Patient Educ Couns 2009 Jan;74(1):61-69. [doi: 10.1016/j.pec.2008.07.044] [Medline: 18778909]
7. Allam A, Kostova Z, Nakamoto K, Schulz PJ. The effect of social support features and gamification on a web-based intervention for rheumatoid arthritis patients: randomized controlled trial. J Med Internet Res 2015 Jan 9;17(1):e14 [FREE Full text] [doi: 10.2196/jmir.3510] [Medline: 25574939]
8. Audrain-Pontevia AF, Menvielle L. Do online health communities enhance patient-physician relationship? An assessment of the impact of social support and patient empowerment. Health Serv Manage Res 2018 Aug;31(3):154-162. [doi: 10.1177/0951484817748462] [Medline: 29280679]
9. Aardoom JJ, Dingemans AE, Boogaard LH, van Furth EF. Internet and patient empowerment in individuals with symptoms of an eating disorder: a cross-sectional investigation of a pro-recovery focused e-community. Eat Behav 2014 Aug;15(3):350-356. [doi: 10.1016/j.eatbeh.2014.04.003] [Medline: 25064280]
10. Mo PK, Coulson NS. Are online support groups always beneficial? A qualitative exploration of the empowering and disempowering processes of participation within HIV/AIDS-related online support groups. Int J Nurs Stud 2014 Jul;51(7):983-993. [doi: 10.1016/j.ijnurstu.2013.11.006] [Medline: 24342449]
11. Bartlett YK, Coulson NS. An investigation into the empowerment effects of using online support groups and how this affects health professional/patient communication. Patient Educ Couns 2011 Apr;83(1):113-119. [doi: 10.1016/j.pec.2010.05.029] [Medline: 20599338]
12. van Uden-Kraan CF, Drossaert CH, Taal E, Shaw BR, Seydel ER, van de Laar MA. Empowering processes and outcomes of participation in online support groups for patients with breast cancer, arthritis, or fibromyalgia. Qual Health Res 2008 Mar;18(3):405-417. [doi: 10.1177/1049732307313429] [Medline: 18235163]
13. Moorhead SA, Hazlett DE, Harrison L, Carroll JK, Irwin A, Hoving C. A new dimension of health care: systematic review of the uses, benefits, and limitations of social media for health communication. J Med Internet Res 2013 Apr 23;15(4):e85 [FREE Full text] [doi: 10.2196/jmir.1933] [Medline: 23615206]

XSL•FO
RenderX

14. Fowles JB, Terry P, Xi M, Hibbard J, Bloom CT, Harvey L. Measuring self-management of patients' and employees' health: further validation of the patient activation measure (PAM) based on its relation to employee characteristics. Patient Educ Couns 2009 Oct;77(1):116-122. [doi: 10.1016/j.pec.2009.02.018] [Medline: 19356881]
15. 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]
16. Moljord IE, Lara-Cabrera ML, Perestelo-Pérez L, Rivero-Santana A, Eriksen L, Linaker OM. Psychometric properties of the patient activation measure-13 among out-patients waiting for mental health treatment: a validation study in Norway. Patient Educ Couns 2015 Nov;98(11):1410-1417 [FREE Full text] [doi: 10.1016/j.pec.2015.06.009] [Medline: 26146239]
17. Rademakers J, Nijman J, van der Hoek L, Heijmans M, Rijken M. Measuring patient activation in the Netherlands: translation and validation of the American short form patient activation measure (PAM13). BMC Public Health 2012 Jul 31;12:577 [FREE Full text] [doi: 10.1186/1471-2458-12-577] [Medline: 22849664]
18. Brenk-Franz K, Hibbard JH, Herrmann WJ, Freund T, Szecsenyi J, Djalali S, et al. Validation of the German version of the patient activation measure 13 (PAM13-D) in an international multicentre study of primary care patients. PLoS One 2013;8(9):e74786 [FREE Full text] [doi: 10.1371/journal.pone.0074786] [Medline: 24098669]
19. Ngooi BX, Packer TL, Kephart G, Warner G, Koh KW, Wong RC, et al. Validation of the patient activation measure (PAM-13) among adults with cardiac conditions in Singapore. Qual Life Res 2017 Apr;26(4):1071-1080. [doi: 10.1007/s11136-016-1412-5] [Medline: 27645458]
20. Packer TL, Kephart G, Ghahari S, Audulv A, Versnel J, Warner G. The patient activation measure: a validation study in a neurological population. Qual Life Res 2015 Jul;24(7):1587-1596. [doi: 10.1007/s11136-014-0908-0] [Medline: 25557496]
21. Maindal HT, Sokolowski I, Vedsted P. Translation, adaptation and validation of the American short form patient activation measure (PAM13) in a Danish version. BMC Public Health 2009 Jun 29;9:209 [FREE Full text] [doi: 10.1186/1471-2458-9-209] [Medline: 19563630]
22. Hibbard JH, Mahoney ER, Stock R, Tusler M. Do increases in patient activation result in improved self-management behaviors? Health Serv Res 2007 Aug;42(4):1443-1463 [FREE Full text] [doi: 10.1111/j.1475-6773.2006.00669.x] [Medline: 17610432]
23. Kinney RL, Lemon SC, Person SD, Pagoto SL, Saczynski JS. The association between patient activation and medication adherence, hospitalization, and emergency room utilization in patients with chronic illnesses: a systematic review. Patient Educ Couns 2015 May;98(5):545-552. [doi: 10.1016/j.pec.2015.02.005] [Medline: 25744281]
24. Grosberg D, Grinvald H, Reuveni H, Magnezi R. Frequent surfing on social health networks is associated with increased knowledge and patient health activation. J Med Internet Res 2016 Aug 10;18(8):e212 [FREE Full text] [doi: 10.2196/jmir.5832] [Medline: 27511272]
25. Kyriacou DN, Lewis RJ. Confounding by indication in clinical research. J Am Med Assoc 2016 Nov 1;316(17):1818-1819. [doi: 10.1001/jama.2016.16435] [Medline: 27802529]
26. Baltierra NB, Muessig KE, Pike EC, LeGrand S, Bull SS, Hightow-Weidman LB. More than just tracking time: complex measures of user engagement with an internet-based health promotion intervention. J Biomed Inform 2016 Feb;59:299-307 [FREE Full text] [doi: 10.1016/j.jbi.2015.12.015] [Medline: 26732997]
27. Agca R, Heslinga SC, Rollefstad S, Heslinga M, McInnes IB, Peters MJ, et al. EULAR recommendations for cardiovascular disease risk management in patients with rheumatoid arthritis and other forms of inflammatory joint disorders: 2015/2016 update. Ann Rheum Dis 2017 Jan;76(1):17-28. [doi: 10.1136/annrheumdis-2016-209775] [Medline: 27697765]
28. Devlin N, Shah KK, Feng Y, Mulhern B, van Hout B. Valuing health-related quality of life: an EQ-5D-5L value set for England. Health Econ 2018 Jan;27(1):7-22. [doi: 10.1002/hec.3564] [Medline: 28833869]
29. Druce KL, McBeth J, van der Veer SN, Selby DA, Vidgen B, Georgatzis K, et al. Recruitment and ongoing engagement in a UK smartphone study examining the association between weather and pain: cohort study. JMIR Mhealth Uhealth 2017 Nov 1;5(11):e168 [FREE Full text] [doi: 10.2196/mhealth.8162] [Medline: 29092810]
30. Alpaydin E. Introduction to Machine Learning. Cambridge, Massachusetts: MIT Press; 2010.
31. Drinkwater C, Wildman J, Moffatt S. Social prescribing. Br Med J 2019 Mar 28;364:1285. [doi: 10.1136/bmj.l1285] [Medline: 30923039]
32. Sadah SA, Shahbazi M, Wiley MT, Hristidis V. A study of the demographics of web-based health-related social media users. J Med Internet Res 2015 Aug 6;17(8):e194 [FREE Full text] [doi: 10.2196/jmir.4308] [Medline: 26250986]
33. 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]
34. Gwynn KB, Winter MR, Cabral HJ, Wolf MS, Hanchate AD, Henault L, et al. Racial disparities in patient activation: evaluating the mediating role of health literacy with path analyses. Patient Educ Couns 2016 Jun;99(6):1033-1037 [FREE Full text] [doi: 10.1016/j.pec.2015.12.020] [Medline: 26809936]
35. Eneanya ND, Winter M, Cabral H, Waite K, Henault L, Bickmore T, et al. Health literacy and education as mediators of racial disparities in patient activation within an elderly patient cohort. J Health Care Poor Underserved 2016;27(3):1427-1440 [FREE Full text] [doi: 10.1353/hpu.2016.0133] [Medline: 27524777]

XSL•FO
RenderX

36. Ellins J. www.picker.org. 2005 Nov 01. How engaged are people in their health care? URL: https://www.picker.org/wp-content/uploads/2014/10/How-engaged-are-people-in-their-health-care-....pdf

Abbreviations

A&E: accident & emergency
EQ-5D: EuroQol-5D
GP: general practitioner
HU: HealthUnlocked
IQR: interquartile range
OHCs: online health communities
PAM: Patient Activation Measure

Edited by G Eysenbach; submitted 04.02.19; peer-reviewed by PCI Pang, D Frohlich, P Schulz, M Peeples; comments to author 27.04.19; revised version received 20.06.19; accepted 05.07.19; published 29.08.19.

Please cite as:
Costello RE, Anand A, Jameson Evans M, Dixon WG
Associations Between Engagement With an Online Health Community and Changes in Patient Activation and Health Care Utilization: Longitudinal Web-Based Survey
J Med Internet Res 2019;21(8):e13477
URL: http://www.jmir.org/2019/8/e13477/
doi:10.2196/13477
PMID:31469082

©Ruth E Costello, Amrutha Anand, Matt Jameson Evans, William G Dixon. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 29.08.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Using Facebook Advertising to Recruit Representative Samples: Feasibility Assessment of a Cross-Sectional Survey

Lance Garrett Shaver[1], MPH; Ahmed Khawer[1], MPH; Yanqing Yi[1], PhD; Kris Aubrey-Bassler[2], MSc, MD; Holly Etchegary[3], PhD; Barbara Roebothan[1], RD, PhD; Shabnam Asghari[1], MD, PhD; Peizhong Peter Wang[1,4,5], MPH, MD, PhD

[1]Division of Community Health and Humanities, Faculty of Medicine, Memorial University of Newfoundland, St John's, NL, Canada

[2]Primary Healthcare Research Unit, Discipline of Family Medicine, Memorial University of Newfoundland, St John's, NL, Canada

[3]Clinical Epidemiology, Faculty of Medicine, Memorial University of Newfoundland, St John's, BC, Canada

[4]Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada

[5]School of Public Health and Management, Weifang Medical University, Weifang, China

Corresponding Author:
Peizhong Peter Wang, MPH, MD, PhD
Division of Community Health and Humanities
Faculty of Medicine
Memorial University of Newfoundland
Room 4M205
St John's, NL, A1B 3V6
Canada
Phone: 1 709 777 8571
Fax: 1 709 777 7382
Email: pwang@mun.ca

Abstract

Background: Facebook has shown promise as an economical means of recruiting participants for health research. However, few studies have evaluated this recruitment method in Canada, fewer still targeting older adults, and, to our knowledge, none specifically in Newfoundland and Labrador (NL).

Objective: This study aimed to assess Facebook advertising as an economical means of recruiting a representative sample of adults aged 35 to 74 years in NL for a cross-sectional health survey.

Methods: Facebook advertising was used to recruit for a Web-based survey on cancer awareness and prevention during April and May 2018; during recruitment, additional advertisements were targeted to increase representation of demographics that we identified as being underrepresented in our sample. Sociodemographic and health characteristics of the study sample were compared with distributions of the underlying population to determine representativeness. Cramer V indicates the magnitude of the difference between the sample and population distributions, interpreted as small (Cramer V=0.10), medium (0.30), and large (0.50). Sample characteristics were considered representative if there was no statistically significant difference in distributions (chi-square $P>.01$) or if the difference was small ($V\leq0.10$), and practically representative if $0.10<V\leq0.20$. The cost per recruit of Facebook advertising was compared with a quote for a random digit dialing (RDD)–recruited postal survey to determine if this method was economical.

Results: Facebook advertising is feasible and economical to conduct survey research, reaching 34,012 people, of which 2067 clicked on the ad, for a final sample size of 1048 people at Can $2.18 per recruit versus the quoted Can $23,316.05 for 400 recruits (Can $35.52 per recruit) via RDD. The sample was representative of rural and urban geography ($P=.02$; V=0.073), practically representative of age ($P=.003$; V=0.145) and income ($P<.001$; V=0.188), and over-representative of women ($P<.001$; V=0.507) and higher levels of education ($P<.001$; V=0.488). The sample was representative of the proportion of people with a regular health care provider ($P=.94$; V=0.025), diabetes prevalence ($P=.002$; V=0.096), and having had a colonoscopy or sigmoidoscopy ($P=.27$; V=0.034), and it was practically representative of smoking status ($P<.001$; V=0.14), and body mass index ($P<.001$; V=0.135). The sample was not representative of arthritis prevalence ($P<.001$; V=0.573), perceived health ($P<.001$; V=0.384), or time since last seasonal flu shot ($P<.001$; V=0.449).

XSL•FO
RenderX

Conclusions: Facebook advertising offers an easy, rapid, and economical means to recruit a partially representative (representative or practically representative of 8 of the 13 characteristics studied) sample of middle-aged and older adults for health survey research. As Facebook uses a nonrandom targeting algorithm, caution is warranted in its applications for certain types of research.

(J Med Internet Res 2019;21(8):e14021) doi:10.2196/14021

KEYWORDS

Facebook; health surveys; Canada; research subject recruitment; social media; internet; online recruitment

Introduction

Population-based survey research relies on the recruitment of participants and aims for collected samples to be representative of the underlying target population. However, traditional survey research can be limited by high recruitment costs, low response rates, and considerable time and personnel demands [1]. With reductions in landline telephone use, even traditional methods, such as random digit dialing (RDD), are challenged in recruiting representative samples of the population [2]. Only 72% of the Canadian households have landlines [3,4], limiting the ability of telephone recruitment in achieving representative samples [3].

Facebook has increasingly gained attention as a tool to recruit participants for research [5]. We propose that many of the aforementioned limitations associated with traditional methods of recruitment can be resolved through Web-based recruitment, specifically Facebook advertising. Facebook currently stands as one of the most popular social media platforms, with over 2 billion users across the globe [6]. Among internet users aged 18 years and older in Canada, the Atlantic provinces (which include Newfoundland and Labrador, ie, NL) have the highest Facebook usage, with 94% using Facebook and 73% using it daily [7]. Even in the age group that uses Facebook the least (individuals aged 55 years and older), 78% of those who use the internet said they used Facebook, with 52% using it daily [8]. With a broad geographic spread and unique health challenges, NL offers a unique population to study Facebook recruitment. These features make traditional recruitment with either probabilistic or nonprobabilistic sampling methods challenging, particularly because of the large geographic spread. Facebook also offers additional benefits with its ability to target advertisements to preferentially reach people based on demographics, location, interests, and behaviors [9].

Limited research has examined the use of Facebook for survey recruitment in Canada and, to our knowledge, none in NL. Several studies, including systematic reviews, have suggested that Facebook is economical and can yield representative survey samples [1,5]. Several reviews have found only limited evidence for Facebook recruitment of older adults, and thus it remains unclear whether a representative sample of adults in this age group can be recruited by this method [5,10]. The objective of our study is to investigate whether Facebook advertising can be a feasible (timely, economical, and with minimal human resource commitment) means of recruiting a representative sample of adults aged 35 to 74 years for a health research survey in NL. Our study will further seek to provide clarity in the matter of whether older populations can be effectively reached and recruited through Facebook. This investigation is a supplement to our primary project, a cross-sectional study examining cancer awareness, beliefs, and prevention-related health behaviors in the NL population.

Methods

Study Design

To participate in this survey, participants must have (1) been aged 35 to 74 years, (2) been living in NL for 2 or more years, and (3) provided consent to participate in the survey.

To ensure transparency in the study design and recruitment process, we followed the reporting guidelines for the checklist for reporting results of internet surveys, reporting these details throughout our methods [11]. A Facebook page was created for the study as a medium for posting advertisements, responding to inquiries, and for knowledge translation once the research was completed. No incentives were offered for participation. The central theme of our survey was cancer awareness and prevention, focusing largely on health behaviors. Google Forms was used to host the open survey and allow for automatic capture of responses into a spreadsheet. We did not use item randomization or adaptive questioning. There were 134 items and 13 screens, 9 for the survey, 1 for the introduction and consent, 2 for eligibility, and 1 to thank participants and offer the opportunity to contact us or provide feedback. No completeness check was used, but participants could return to previous pages and change responses.

Recruitment

The Facebook platform captures sociodemographic information that is self-reported by users (Facebook requires self-reporting for age and gender but not for education) and geographic location is captured based on data provided by the user, data from their device, and their Facebook activity [12]. Initially, we had just 1 generic advertisement (see Figure 1) targeted to all individuals aged 35 to 65 years and older (the maximum age that can be specified for Facebook ad targeting is *65 years and older*) living in NL. While recruitment was underway, we regularly evaluated the distribution of respondent demographics to assess if any were under-represented in our sample. We noticed that (1) men, (2) individuals from rural NL, (3) individuals aged 35 to 44 years, and (4) individuals with lower levels of education were under-represented in our sample. We created 4 different advertisements to target these specific groups as a means of purposive sampling to obtain a more representative sample. Details on these advertisements and the specific targeting criteria can be found in Multimedia Appendix 1. We then created multiple versions of the same survey, each with its own unique link, to track the respondents recruited by each advertisement.

XSL•FO
RenderX

Figure 1. Generic Facebook advertisement.

Daily advertisement spending was increased or decreased for each ad to increase or decrease the number and frequency of advertisements shown and, consequently, the rate at which individuals from specific demographics were being recruited. Advertisements ran for a total of 40 days, during April and May 2018, until the cutoff date for recruitment. Owing to the length and format of the survey, advertisements were targeted to desktop users but not mobile device users, as we believed the survey would be easiest to complete on a full-screen device.

For geographic targeting, we created an ad (rural ad) that targeted people aged 35 to 65 years and older in rural postal codes in NL. For the purpose of this study, we used Canada Post's classification system to determine whether an area was rural or urban, with rural areas being indicated by having a 0 for the second character in the postal code [13]. Facebook allows advertisements to be targeted geographically by forward sortation area (the first 3 characters of the postal code), and so we targeted the rural ad to all forward sortation areas in NL that contained a 0.

To target by age, we created an ad (age ad) that was shown to individuals aged 35 to 44 years. To target by gender, we created an ad (gender ad) that was only shown to men.

To target by education (education ad), we applied exclusion criteria rather than inclusion criteria for ad targeting. We excluded individuals with higher levels of education so that the ad was only shown to people who had graduated high school, had completed some high school, or had not specified their level of education. We targeted by exclusion because educational attainment is optionally self-reported and so we believed it was less likely to be reported accurately by Facebook users. Applying inclusion criteria would have been more specific but it would have excluded an estimated 36,000 Facebook users in our study population who did not specify their level of education, leaving a target audience of only 18,000 Facebook users who had specified they were high school grads or had some high school (estimate obtained using Facebook ad manager).

Exclusion Criteria

Respondents (those who submitted the survey) were excluded from the sample if they did not meet all of the inclusion criteria or if they had more than 10 missing variables on the remainder of the survey (not including questions on screening history or any text-based responses, as these might not have been relevant for all respondents). The age and postal code of each respondent was checked to confirm that they met the inclusion criteria for the study. If they did not, or this information was not provided, the respondents were excluded.

Statistical Analysis

Data analyses comprised the following 3 major aspects: (1) data cleaning and checking, (2) descriptive analysis of Facebook ad metrics and costs to determine cost per recruit, and (3) descriptive analyses to provide an overview of our sample and univariate analysis to examine whether the distributions of the sample's sociodemographic and selected health characteristics

XSL•FO
RenderX

were consistent with the target population. Census data were obtained from Statistics Canada for the 2016 census of the NL population for people aged 35 to 74 years [14]. The NL Centre for Health Information (NLCHI) provided data from the 2016 cycle of the Canadian Community Health Survey (CCHS) on NL residents aged 35 to 74 years. All statistical analyses were conducted using Microsoft Excel Version 16.17.

Campaign and Recruitment Measures

We investigated and reported advertisement campaign parameters, including the number of impressions, paid reach (unique individuals who saw the advertisement), unique link clicks, link clicks, ad spend, cost of administering the advertisements, cost per click, ad spend per recruit (only cost of advertisement included, which was used to compare results from different targeted advertisements), and cost per recruit (including ad spend and administrative costs). All costs are reported in Canadian dollars.

With respect to response rates, the number of unique site visitors could not be determined with Google Forms and so we used the numbers of unique link clicks provided by Facebook's ad manager. Because the advertisement link directed people to the first page of the survey, the number of unique link clicks on the ad over the number of unique individuals who saw the ad (Facebook's *paid reach* metric) will be considered the landing page view rate. Completion rate was defined as the number of recruits in our final sample over the unique link clicks. Google Forms does not track number of surveys started—only the number of surveys submitted—and there is no way to prevent, or identify when there are, multiple entries from the same individual, unless they were precise duplicates.

Feasibility

Feasibility was assessed both subjectively (the ease of conducting research and time commitment) and objectively (costs per recruit). Recruitment costs included both advertising costs and administrative costs during the entire recruitment period. We then compared this cost per recruit with an estimate provided by a local university-based research support unit for what they would have charged to conduct a postal survey using RDD sampling and telephone recruitment.

Representativeness

Representativeness of the Facebook sample was assessed by comparing sociodemographic characteristics of participants with the underlying population, obtained from the 2016 Census of the NL population between the ages of 35 and 74 years. These characteristics were as follows: age, gender, rural and urban geography, education, and household income. To further assess the representativeness of our sample, selected health characteristics were compared with the underlying population, using data from the 2016 CCHS of the NL population aged between 35 and 74 years. These characteristics were as follows: prevalence of arthritis, prevalence of diabetes, perceived health, having a regular health care provider, time since last seasonal flu shot, smoking status, body mass index (BMI), and whether they have ever had a colonoscopy or sigmoidoscopy.

Goodness-of-fit chi-square tests were conducted for each indicator to compare frequency counts from our sample with the expected relative frequency of the population to determine if the distribution of the Facebook sample was statistically consistent with the population. For the purpose of this analysis, $P<.01$ was considered to indicate rejection of the null hypothesis at significance level .01 (H_0: distribution consistent with population; H_a: distribution not consistent with the population). Likewise, $P>.01$ indicates failure to reject the null hypothesis at significance level .01, meaning that we can assume the sample distribution is consistent with the census distribution (ie, representative).

We then conducted post hoc tests for characteristics that had more than 2 categories using the methods described by Beasley and Schumacker [15] to find adjusted residuals (Z) and identify which categories were and were not consistent with the population. If the adjusted residual was $-2.58 \leq Z \leq 2.58$, the observed frequency for that category was considered similar to that expected under the null hypothesis at significance level .01 (ie, that the category was representative).

To determine the magnitude of difference between the sample distribution and the population, Cramer V posttest for effect size was calculated and interpreted as per Cohen [16], where values of 0.10, 0.30, and 0.50 corresponded to small, medium, and large effect sizes. We considered a characteristic to be *representative* of the population if (1) the chi-square test showed no statistically significant difference or, in case that there was a statistically significant difference, if (2) the Cramer V (df*=1) is less than 0.10. If Cramer V (df*=1), is between 0.10 and 0.20, we considered this small to medium effect size as *practically representative* of the population.

Targeting

The effectiveness of targeted advertisements at increasing representation of targeted demographics was assessed by comparing the distribution of the sample including the respondents from the targeted ad (observed) with the distribution of the sample excluding the respondents of the targeted ad (expected), using a chi-square statistical test at significance level of .01. If the targeted advertisement made a statistically significant change in the distribution that was closer to the target distribution of the underlying population, we considered the targeted advertising to be effective. We then compared advertisements based on dollars spent per recruit on advertising (ad spend per recruit).

Ethical Considerations

Ethical approval was obtained from the NL Health Research Ethics Authority. Consent was obtained before individuals could begin the survey and if it was not provided, participants were redirected away from the survey. The survey itself was anonymous and all efforts were made to maintain confidentiality of individuals who participated. Facebook was used only to advertise and direct interested individuals to the survey link, which was hosted with Google Forms. Data were securely kept on password-protected computers and cloud storage accounts.

XSL•FO
RenderX

Results

Statistical Analysis

Campaign and Recruitment Measures

There are numerous metrics for measuring Facebook ad performance. The primary metrics and results are in Table 1. The cost of Facebook advertising was Can $1750 and the cost of a graduate student to administer this survey over the 40 days was approximately Can $539 (24.5 hours work). Figure 2 details the recruitment and selection process. The final sample had 1048 recruits, at an average cost per recruit of Can $2.18 (Can $1.67 ad spend per recruit). Administrative duties included creating the advertisements, responding to comments and questions, answering emails, and reviewing results and optimizing advertisements and targeting.

It was costlier to recruit men than women, with an average Can $3.63 ad spend per man recruited versus Can $1.69 ad spend per woman recruited. Considering all advertisements, women were more likely to be shown the ad than men (reach=21,416 for women vs 11,188 for men), were more likely to click on the link—the landing page view rate was 7.10% (1522/21,416) for women versus 4.16% (465/11,188) for men, and cost less per link click (Can $0.64 ad spend for women vs Can $1.16 ad spend for men). Women (799/21,416, 3.77%) who saw that ad were more likely than men (242/11,188, 2.16%) to complete it (number of recruits over paid reach). However, the actual completion rate—the number of recruits per unique clicks on the ad link—was similar for women and men (799/1522, 52.50% and 242/465, 52.0%, respectively).

To provide a brief understanding of the sampling frame and size of the NL population based on the 2016 census, there were 295,300 people aged between 35 and 74 years in NL [14]. The estimated reach provided to us by Facebook's ad management tool was 110,000 active users aged 35 to 65 years and older. As Facebook's ad manager tool does not provide the option to set the age range to between 35 and 74 years, we cannot provide an accurate estimate of the percentage of the population that is represented in our sampling frame.

Table 1. Facebook advertising campaign and recruitment metrics.

Metric	Women	Men	Unknown	Total
Delivery[a], n (%)				
Impressions[b]	85,693 (64.91)	41,370 (31.34)	4958 (3.76)	132,021 (100)
Paid reach[c] (No. of unique Facebook users)	21,416 (62.97)	11,188 (32.89)	1409 (4.14)	34,012[d] (100)
Engagement[a], n (%)				
Unique link clicks[e]	1522 (73.63)	465 (22.50)	80 (3.87)	2067 (100)
Overall link clicks	1697 (73.27)	527 (22.75)	92 (3.97)	2316 (100)
Costs, Can $				
Ad spend	1084.63	609.60	55.77	1750
Administrative costs of recruitment[f,g]	269.50	269.50	—[h]	539
Performance				
Recruits, n (%)	799 (76.24)	242 (23.09)	7 (0.67)	1048 (100)
Landing page view rate (unique link clicks/paid reach)[g], %	7.10	4.16	5.68	6.08
Completion rate (recruits/unique link clicks)[g], %	52.5	52.0	—	50.7
Ad spend per recruit[g], Can $	1.36	2.52	—	1.67
Cost per recruit[g], Can $	1.69	3.63	—	2.18

[a]Metrics are estimated by Facebook.

[b]Impressions: the number of times any part of an ad appears on the user's screen.

[c]Paid Reach: the number of unique people who have seen the advertisement.

[d]The total of paid reach is listed as 34,012 in the Facebook Ad Manager data tables, even though the sum of the 3 categories (women, men, and unknown) is 34,013. This has been left unaltered for transparency.

[e]Link Clicks: the number of clicks on the ad's destination link to the survey.

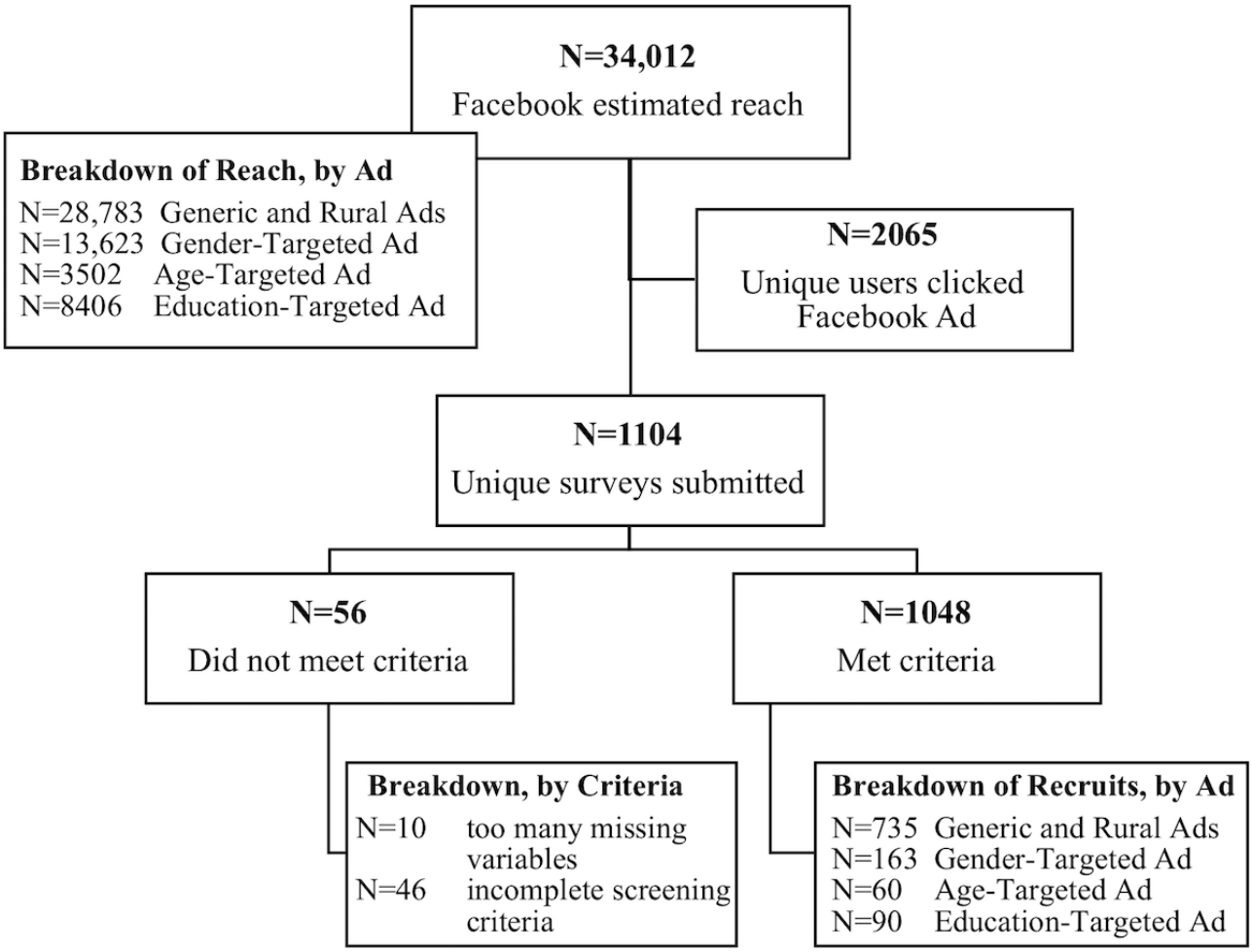
[f]Administrative costs (Graduate student's time: 24.5 hours or approximately 4.3 hours per week) include creating the advertisements, responding to comments/questions, answering emails, reviewing metrics and response to optimize advertisements.

[g]These metrics are not provided by Facebook but were calculated by the authors.

[h]Not applicable.

XSL•FO
RenderX

Figure 2. Details of the Facebook advertising and recruitment process. Note that the total estimated reach is lower than the sum of the number reached by individual advertisements, as some people were reached by multiple advertisements.



Feasibility

Economic considerations suggest Facebook advertising is a feasible means to recruit survey participants. The research was very easy to conduct, manageable by 1 graduate student as a research assistant committing 24.25 hours over 40 days. This method is also feasible for rapid recruitment of a large sample, with n=1048 participants being recruited in only 40 days and at a total cost of Can $2289 or Can $2.18 per recruit (excluding administrative costs, Can $1.67 ad spend per recruit). In comparison, we were quoted Can $21,316.05 as an estimate to contact 600 people with the anticipation of a final sample size of n=400 (Can $35.52 cost per recruit) using RDD. This considerably lower cost provides preliminary evidence to suggest that Facebook may be an economical recruitment method.

Representativeness

The sociodemographic characteristics of the sample are reported in Table 2. There was a statistically significant difference (*P*=.003) between the sample and population distributions for age. However, posttests showed a small effect size (V=0.145) between the 2 distributions. Post hoc analysis of adjusted residuals shows that the only age group that was not consistent with the population was the group aged 60 to 64 years, which was over-represented in our sample (*Z*=2.72: sample proportion is 180/1048, 17.17% vs population proportion is 14.23%). Considering the variation was small and was largely because of 1 age group, we concluded the age distribution of this sample is *practically representative*. Looking at the distribution of rural and urban recruits, there was no statistically significant difference between the sample and target population distributions at significance level .01 (*P*=.02), thus suggesting the distribution of rural and urban recruits is *representative* of the population distribution.

The distribution of annual household income was not consistent with the population (*P*<.001). Posttests showed a small to medium effect size (V=0.188). Post hoc analysis of adjusted residuals indicates that the variation is largely because of under-representation of individuals with household incomes under Can $30,000 (Z=−3.48: sample proportion is 141/942 [14.8%] vs population proportion is 19.84 %) and over-representation of individuals with household incomes between Can $50,000 and Can $59,999 (*Z*=3.43: sample proportion is 87/942, [10.2%] vs population proportion is 7.20%). All other income groups were consistent with the population. This, therefore, suggests that income is *practically representative* of the population.

XSL•FO
RenderX

Table 2. Comparison of selected sociodemographic characteristics of the Facebook sample with the underlying population of Newfoundland and Labrador from the 2016 census.

Characteristic	Facebook, n[a] (%)	2016 census, %	Goodness-of-fit, χ^2 (df)[b]	*P* value[b]	Cramer V (effect size)[c]	Adjusted residual, *Z*[d]	Interpretation[e], representative?
Gender			267.1 (1)	<.001	0.507 (large)	__[f]	No
Men	242 (23.25)	48.56	—	—	—	—	—
Women	799 (76.75)	51.44	—	—	—	—	—
Age (years)			21.9 (7)	.003	*0.145 (small-medium)*	—	*Practically*
35-39	96 (9.15)	10.41	—	—	—	*−1.32*	—
40-44	98 (9.34)	11.68	—	—	—	*−2.35*	—
45-49	125 (11.92)	13.09	—	—	—	*−1.12*	—
50-54	159 (15.16)	14.43	—	—	—	*0.68*	—
55-59	170 (16.21)	14.59	—	—	—	*1.49*	—
60-64	180 (17.16)	14.23	—	—	—	2.72	—
65-69	147 (14.01)	12.70	—	—	—	*1.29*	—
70-74	73 (6.96)	8.86	—	—	—	*−2.16*	—
Geography			*5.4 (1)*	*.02*	*0.073 (small)*	—	*Yes*
Rural	470 (44.85)	48.42	—	—	—	—	—
Urban	578 (55.15)	51.58	—	—	—	—	—
Education			248.8 (1)	<.001	0.488 (large)	—	No
No postsecondary	237 (22.72)	47.10	—	—	—	—	—
Postsecondary	806 (77.28)	52.90	—	—	—	—	—
Household income, Can $			33.6 (5)	<.001	*0.188 (small-medium)*	—	*Practically*
Less than 30,000	141 (14.8)	19.84	—	—	—	−3.48	—
30,000-49,999	188 (19.8)	18.01	—	—	—	*1.26*	—
50,000-59,999	87 (10.2)	7.20	—	—	—	3.43	—
60,000-79,999	140 (14.7)	12.82	—	—	—	*1.62*	—
80,000-99,999	114 (12.0)	10.36	—	—	—	*1.55*	—
More than 100,000	272 (28.6)	31.76	—	—	—	*−1.75*	—

[a]Totals may not match because of missing responses.

[b]Italicized chi-square and *P* values indicate failure to reject null (H_0: distribution consistent with population) at significance level .01, that is, consistent with census distribution.

[c]Italics indicate if Cramer V suggests the distribution is representative (Cramer V≤0.10) or practically representative (0.10<Cramer V≤0.20). Cramer V effect size indicates the size of the difference between the sample and population, with smaller being more representative.

[d]Italicized *Z* indicates post hoc adjusted residual Z is −2.58>Z<2.58, meaning observed number of cases is statistically similar to what would be expected if null hypothesis was true, at significance level of .01, that is, consistent with census proportion for category.

[e]Representativeness decision based on authors' interpretations of χ^2, post hoc adjusted residuals (*Z*), Cramer V effect size, and practical significance.

[f]Data not applicable.

XSL•FO
RenderX

Table 3. Comparison of selected health characteristics of the Facebook sample and the 2016 Canadian Community Health Survey for Newfoundland and Labrador (Canadian Community Health Survey data provided by the Newfoundland and Labrador Centre for Health Information).

Characteristic	Facebook, n[a] (%)	CCHS[b] 2016, %	Goodness-of-fit, χ^2 (df)[c]	*P* value[c]	Cramer V, (effect size)[d]	Adjusted residual, *Z*[e]	Interpretation[f], representative?
Has arthritis			344.4 (1)	<.001	0.573 (large)	__g	No
Yes	89 (8.48)	35.98	—	—	—	—	—
No	960 (91.52)	64.00	—	—	—	—	—
Has diabetes			9.7 (1)	.002	*0.096 (small)*	—	Yes
Yes	95 (9.06)	12.20	—	—	—	—	—
No	954 (90.94)	87.80	—	—	—	—	—
Self-perceived health			154.7 (4)	<.001	0.384 (medium-large)	—	No
Excellent	93 (8.87)	17.92	—	—	—	−6.92	—
Very Good	421 (40.17)	42.24	—	—	—	*−1.03*	—
Good	373 (35.59)	22.78	—	—	—	8.69	—
Fair	144 (13.74)	11.76	—	—	—	*1.86*	—
Poor	17 (1.62)	5.30	—	—	—	−5.17	—
Last flu shot[h]			207.0 (3)	<.001	0.449 (medium-large)	—	No
<1 year ago	461 (44.89)	29.57	—	—	—	9.03	—
1 year to <2 years ago	62 (6.04)	7.35	—	—	—	*− 1.55*	—
2 years ago or more	215 (20.93)	14.21	—	—	—	5.72	—
Never	289 (28.14)	48.87	—	—	—	−9.50	—
Has regular health care provider			*0.7 (1)*	*.94*	*0.025 (small)*	—	Yes
Yes	951 (90.83)	90.90	—	—	—	—	—
No	96 (9.17)	9.10	—	—	—	—	—
Smoking status			20.3 (1)	<.001	*0.141 (small-medium)*	—	Practically
Daily or Occasional	182 (17.86)	23.88	—	—	—	—	—
Do not smoke	837 (82.14)	76.12	—	—	—	—	—
Body mass index			18.7 (2)	<.001	*0.135 (small-medium)*	—	Practically
Normal to under-weight	246 (23.93)	27.06	—	—	—	—	—
Overweight	353 (34.34)	37.64	—	—	—	—	—
Obese	429 (41.73)	35.30	—	—	—	—	—
Has ever had a colonoscopy or sigmoidoscopy			1.2 (1)	*.27*	*0.034 (small)*	—	Yes
Yes	460 (43.89)	42.20	—	—	—	—	—
No (or no answer)	588 (56.11)	57.80	—	—	—	—	—

[a]Totals may not match because of missing responses.

[b]CCHS: Canadian Community Health Survey.

[c]Italicized chi-square and *P* value indicate failure to reject null (H_0: distribution consistent with population) at significance level .01, that is, consistent with census distribution.

[d]Italics is used to indicate if Cramer V suggests the distribution is representative (Cramer V≤0.10) or practically representative (0.10<Cramer V≤0.20). Cramer V effect size indicates the size of the difference between the sample and population, with smaller being more representative.

[e]Italicized *Z* indicates post hoc adjusted residual *Z* is −2.58>Z<2.58, meaning observed number of cases is statistically similar to what would be expected if null hypothesis was true, at significance level of .01, that is, consistent with census proportion for category.

XSL•FO
RenderX

[f]Representativeness decision based on authors' interpretations of χ^2, post hoc adjusted residuals (*Z*), Cramer V effect size, and practical significance.

[g]Data not applicable.

[h]Question asked slightly differently between CCHS and our Facebook Study, but results are unlikely to vary.

Gender was not representative of the population (*P*<.001; V=0.507), with women being considerably over-represented (see Table 2). Likewise, the distributions of high and low education in our sample were not consistent with population distributions (*P*<.001; V=0.488).

Table 3 shows the distribution of health characteristics of the sample in comparison with the underlying population. Our sample was representative of the proportion of people with diabetes in the population (*P*=.002; V=0.096), of the proportion of people with regular health care providers (*P*=.94), and of ever having a sigmoidoscopy or colonoscopy (*P*=.27; V=0.034). The sample was *practically representative* of smoking status (*P*<.001; V=0.14) and BMI (*P*<.001; V=0.135). The sample was not representative of the proportion of people with arthritis (*P*<.001; V=0.573), of perceived health status (*P*<.001; V=0.384), and of flu shot frequency (*P*<.001; V=0.449).

Caution should be used in considering the results for BMI mentioned in Table 3, as there is a potential data quality error for which we adjusted through imputation. When asked about body weight, participants were asked to *please specify pounds or kilograms*, as we presumed this choice would make it easier for the user to specify their weight. Unfortunately, a considerable number (n=225) provided a number but no units. To correct for this, any values that were equal or greater than 91 were assigned as pounds (n=221) and any values less than 91 were assigned as kilograms (n=4). This decision was made because pounds are most commonly used to specify weight in NL (and 95% of our sample that specified units did so in pounds). The cutoff of 91 was chosen because this was the sample's minimum value for weight in pounds. We then determined if the distribution of BMI class for the participants whose BMI was calculated based on our correction (Group 1) was similar to the distribution of the group of participants who needed no corrections to calculate BMI (Group 2). Using an independent samples chi-square test, we found no statistically significant differences between Groups 1 and 2 (χ^2_2=1.1; *P*=.57). On the basis of this test, we deemed our method of imputing units onto values appropriate because it did not change the distribution of BMI, and for this reason we decided to retain the participants with corrected weights in the sample.

Targeting

The results of the targeted advertisements are shown in Table 4, presenting sociodemographic distributions of recruits, by the advertisement used to recruit them.

XSL•FO
RenderX

Table 4. Sociodemographic characteristics of samples recruited with each Facebook advertisement and the costs associated with each advertisement.

Characteristic	Advertisement source			
	Generic and rural[a]	Gender (men)	Age (years; 35-44)	Education (low)
Gender, n[b] (%)				
Men	86 (11.7)	150 (92.0)	2 (3)	4 (4)
Women	644 (87.6)	13 (8.0)	57 (95)	85 (94)
Other/not specified	5 (0.7)	—[c]	1 (2)	1 (1)
Age (years), n[b] (%)				
35-39	51 (6.9)	10 (6.1)	31 (52)	4 (4.4)
40-44	61 (8.3)	5 (3.1)	27 (45)	5 (5.6)
45-49	94 (12.8)	18 (11.0)	2 (3)	11 (12.2)
50-54	127 (17.3)	20 (12.3)	—	12 (13.3)
55-59	123 (16.7)	29 (17.8)	—	18 (20.0)
60-64	131 (17.8)	38 (23.3)	—	11 (12.2)
65-69	100 (13.6)	29 (17.8)	—	18 (20.0)
70-74	48 (6.5)	14 (8.6)	—	11 (12.2)
Geography, n[b] (%)				
Rural	364 (49.5)	67 (41.1)	8 (13)	31 (34)
Urban	371 (50.5)	96 (58.9)	52 (87)	59 (66)
Education, n[b] (%)				
No postsecondary	164 (22.5)	37 (22.7)	6 (10)	30 (33)
Postsecondary	566 (77.5)	126 (77.3)	54 (90)	60 (67)
Associated costs, Can $				
Ad spend	486.62 (rural ad); 426.58 (generic ad)	495.72	83.6	257.48
Administrative cost[d]	378.02	83.83	30.86	46.29
Total cost[e]	1291.22	579.55	114.46	303.77
Ad spend per recruit	1.24	3.04	1.39	2.86
Total cost per recruit	1.76	3.56	1.91	3.38

[a]The first sample source (ad—generic and rural) included recruits from both the generic ad and the rural ad. Owing to an error in the survey links used in the rural and generic advertisements, we were not able to distinguish which ad the participants had been recruited with, and so we had to present them together.

[b]Totals may not match because of missing responses.

[c]No participants in this category were recruited by this ad.

[d]Administrative costs for each ad was calculated by multiplying the total administrative cost for all advertisements (Can $539) by the fraction of total recruits recruited by that ad.

[e]Total cost was calculated by adding ad spend together with administrative costs.

Targeting by Geography

Respondents to our survey were representative of both rural and urban geographies in NL. Owing to an error in the survey link, we were not able to differentiate which respondents were recruited from the generic ad and which were recruited from the rural ad, as both advertisements directed people to the same survey instead of 2 different surveys as intended. The respondents to the generic and rural advertisements were evenly distributed between rural and urban regions (364/735, 49.5% and 371/735, 50.5%, respectively), whereas the other advertisements, combined, favored urban regions (207/313, 66%) over rural (106/313, 34%). This suggests, but does not confirm, that targeting by geography is effective, and the cost of these advertisements (Can $1.24 ad spend per recruit) was relatively low.

Targeting by Gender

The gender-targeted ad was effective for targeting men, although costlier. Without targeting, the sample would have been 10.5%

XSL•FO
RenderX

men (92/878) and 89.5% women (786/878). With the gender-targeted ad, the final sample consisted of 23.25% men (242/1041) and 76.75% women (799/1041). Although the final sample did not achieve the same distribution as the census population, this showed that targeting by gender could increase the representation of the targeted gender in the sample in a statistically significant way ($P<.001$). However, the advertising costs were much greater than other targeted advertisements at Can $3.04 ad spend per recruit.

Targeting by Age

Without age targeting, the sample would have been 6.5% (65/988) of people aged 35 to 39 years and 7.2% (71/988) of people aged 40 to 44 years. With the age-targeted ad, the final sample was 9.16% (96/1048) of people aged 35 to 39 years and 9.35% (98/1048) of people aged 40 to 44 years. The difference was statistically significant ($P=.005$). The cost of age targeting (Can $1.39 ad spend per recruit) was not considerably more than the generic and rural targeted advertisements (Can $1.24 ad spend per recruit), and so this additional recruitment was effective and economical.

Targeting by Education

Without the education-targeted ad, 21.7% (207/952) of the sample had no postsecondary education and 78.4% (746/952) did. With the education-targeted ad, 22.72% (237/1043) had no postsecondary education and 77.28% (806/1043) did, but this difference was not statistically significant ($P=.43$). The cost to recruit these participants was considerably higher at Can $2.86 ad spend per recruit. Therefore, targeting by education was ineffective at increasing the proportion of participants who had lower levels of educational attainment.

Discussion

Principal Findings

This study is a novel investigation assessing whether or not Facebook advertising can be used to feasibly recruit a representative sample of middle-aged and older adults in the province of NL to complete a health survey. Representativeness was assessed by comparing numerous sociodemographic and health characteristics of our sample with the underlying population. Feasibility was assessed based on an assessment of costs, ease of use, and recruitment time. Moreover, this is the first Canadian study we know of to investigate Facebook advertising for the recruitment of adults aged up to 74 years; the vast majority of the national and international literature on social media recruitment focuses on the youth and young adults, with some focusing on middle-aged adult populations [5,10]. These results should be of considerable interest to academic researchers, community organizations, and governments, because they suggest this method can dramatically reduce barriers to conducting research with minimal sacrifices to representativeness.

Feasibility

The research was very easy to conduct, manageable by 1 graduate student as a research assistant committing 4.3 hours per week. This method is also feasible for rapid recruitment of a large sample, with 1048 participants being recruited in only 40 days. A major advantage over other nonprobability-based sampling methods is that it was easy to use targeted advertisements to improve representativeness of a sample and to recruit hard-to-reach populations. There was also essentially no footwork involved, such as putting up recruitment posters, and we believe there was less selection bias than using email listservs or snowball sampling.

We also found that Facebook advertising is an economical means of recruiting participants for survey research. As anticipated, compared with the quoted cost of RDD recruitment, Facebook advertising was considerably less expensive. Granted, this is comparing our nonprobability-based sample with a probability-based sample. Though perhaps not an entirely fair comparison to make, it nonetheless demonstrates the considerable degree of reduction in cost. In comparison with other Facebook-recruited samples in the literature, our research appears more economical. A total of 2 systematic reviews exploring Facebook recruitment for health, medical, or psychosocial research found median recruitment costs of US $17.48 and US $14.41 per completing participant across various study designs, and the cost per recruit in cross-sectional surveys similar to our own was US $11.46 [5,10]. Although costs are not directly comparable because of differences in geography, topic of research, and targeted demographics, it appears that our Facebook recruitment was considerably more economical than in these previous studies. This could be because of the higher interest of the population in cancer as an issue that affects everyone. Many other studies focused on less prevalent health issues—such as vaccine-hesitancy, Human Immunodeficiency Virus–related knowledge, vertigo, endometriosis, or risky sexual behavior [5]—that might not elicit the same response as our research on cancer awareness and prevention did. This would increase the costs associated with advertising to reach a larger number of people to obtain the same sample size.

Representativeness and Targeting

The Facebook-recruited sample was *representative* of the population with respect to age, geography, diabetes prevalence, having ever had a colonoscopy or sigmoidoscopy, and prevalence of having a regular health care provider, whereas it was *practically representative* with respect to income, smoking status, and BMI (see Tables 2 and 3). In their review, Thornton et al [5] found 16 Facebook-recruitment studies included a formal test of representativeness, with 9 testing the sample against the population of interest and 5 testing against traditional recruitment methods, such as phone and postal surveys, and 2 studies testing against both the population and traditional methods. Only 36% reported their samples were representative of their populations, but 86% were representative of samples recruited using traditional recruitment methods [5]. With 8 of 13 sociodemographic and health status variables representative or practically representative of the target population, our research further supports the conclusion that Facebook can yield a sample that is partially representative of the population. Targeted advertising was useful, but if we had begun targeted advertising earlier in the recruitment period and allocated more advertising dollars toward it, we believe we could have achieved a greater degree of representation from these groups.

XSL•FO
RenderX

Sociodemographic Characteristics

Targeting by age and gender was effective in increasing representation. Targeting men, although effective, was costlier. We suspect that targeting by geography helped increase the representation of rural participants in our study. However, because of the aforementioned error in the survey link for the rural ad, this is only a speculation. There was over-representation of people with higher levels of education, but targeting advertisements proved both costly and ineffective at improving representation of participants with lower levels of educational `attainment. We hypothesize this might be because targeting for education relies on the data that Facebook users optionally provide about their educational background. As such, it is possible that what is reported on a user's Facebook profile won't match their actual educational background. In contrast, age and gender targeting may have been relatively more effective because Facebook requires new users to report age and gender data when creating their Facebook account.

Lastly, even after specifically targeting men, our final sample had markedly more women than men. This is unsurprising because women are more likely to participate in health research [17]. The over-representation of women in this sample could be further explained by higher rates of Facebook use among women (60% women vs 40% men) in our targeted audience (information obtained on March 1, 2019 from Facebook's audience insights tool) and because 63% of the people who saw the ad were women. Gender-representation in our sample might have been improved by increased spending on the ad targeted toward men, but it was 2.14 times costlier to recruit men than women (Can $3.63 vs Can $1.69 ad spend per recruit).

Overall, our difficulties recruiting certain demographics and obtaining a fully representative sample are common in other sampling methods [17-20]. Being under-representative of people in the lowest income bracket is typical across both random and nonrandom sampling methods [17,18,20]. This may be because of socioeconomic barriers, such as not being able to afford internet or landline phones, which could reduce their chances of selection. Furthermore, people with lower socioeconomic status may be less interested in participating in research [17,21].

Health Characteristics

The self-reported health of participants in our sample was not representative of the population, with more participants self-reporting their perceived health status as *good*, and considerably less self-reporting it as *excellent* or *poor*. However, our recruited sample appeared to be more health conscious than the general population, according to various health characteristics compared in our study. For instance, compared with the population, a higher proportion of our sample had a flu shot within the last year. Influenza vaccination uptake is known to be associated with positive beliefs toward prevention, trust in the health care system, socioeconomic status, and a number of other personal, intermediate, and structural factors [22]. Overall, we believe the general health of the sample was at least partially representative of the population, given that 5 out of the 8 assessed health characteristics were representative of the population (see Table 3).

Strengths

Our study contributes to the evidence based on the use of social media for health research. We compared more sample characteristics with the population and used more formal tests of representativeness than many other studies have [1,5,10]. This method of Web-based recruitment works effectively in the context of NL. Furthermore, we have shown that Facebook can be used to recruit older adults to research surveys, where the vast majority of previous research on this subject has only considered younger populations. In addition, we went beyond simple tests of statistical significance by using Cramer V and post hoc adjusted residuals to add a greater depth to our analysis of how the sample distributions compare with the target population.

Limitations

It should be noted that large sample size constrains the value of statistical tests because even small differences can appear *significant* when dealing with large sample sizes [23]. The feasibility assessment we employed was not rigorous, based mostly on subjective interpretations and on a cost comparison between our Web-based survey recruited through nonprobability-based sampling and a postal survey recruited through random sampling. As our analysis compares representativeness with the population (census or CCHS), it is not possible to directly compare representativeness of samples recruited with Facebook with different methods of recruitment, such as RDD. We encourage future research to consider conducting studies that allow for direct comparisons with probabilistic and nonprobabilistic sampling methods.

With respect to targeting advertisements by level of educational attainment, our use of exclusion criteria for targeting (excluding any Facebook users who have any higher level of education reported), rather than inclusion criteria (targeting only Facebook users who self-report education as high school or less), might have been too inclusive. The choice to use exclusion criteria was because there were approximately 36,000 people (Facebook-estimated potential reach) whose education was *unspecified* that would not have been shown the ad if we targeted for inclusion of *high school grad* or *some high school*. Although using inclusion criteria would have resulted in a smaller potential reach, it would have likely resulted in a higher proportion of recruits with lower levels of educational attainment. Future research employing Facebook advertising for recruitment might consider more specific targeting by using inclusion criteria for educational attainment.

There were 2 potential limitations with our data itself. First, the aforementioned issue with the rural-targeted ad. This limits our ability to conclude whether geographic targeting was effective, but we strongly believe that it is worth using geographic targeting for advertisements, which will be of benefit to researchers who want a geographically representative sample. The second data limitation is related to BMI calculations, already discussed above. The correction we employed (by imputing units for weight where they were not specified) did not affect the distribution in a statistically significant way, and so we believe this limitation has minimal impact on our findings.

XSL•FO
RenderX

Considerations for Future Research

As our study examines effectiveness in a province with high Facebook usage, it may be more challenging to recruit representative samples in areas where Facebook use is lower. However, as NL has higher traditional landline usage than any other province (at 70% as of 2013), it is likely that traditional recruitment methods in other provinces would also yield less representative results [3]. For example, traditional landline usage rates are 43% in Quebec, 57% in British Columbia, and 61% in Ontario [3].

We believe there are many potential applications of Facebook advertising for researchers, governments, and community-based organizations who wish to learn about the populations they serve. Facebook advertising could also be used in qualitative research, sociological, and psychological research, where nonprobabilistic sampling is more common. Beyond research recruitment, we see Facebook advertising as having potential value in program planning and evaluation. For example, organizations planning a public health campaign may take a participatory approach by using Facebook advertising to reach individuals from their target population who they can ask to provide feedback on the campaign to optimize messaging and imagery.

Although the unique nature of Facebook advertising could warrant its own distinct sampling method classification, it is otherwise best described as nonprobabilistic purposive sampling. Although this technique is more common in qualitative research, it can be used to serve many purposes for both quantitative and qualitative studies [24]. A limitation of Facebook is that it is not possible to assess nonresponse bias, as there is little information on the nonrespondents, other than the age and gender of who has seen the advertisement. Another limitation of Facebook's advertising tool is that it does not allow for specification of age targets above 65 years, thus limiting us to choosing the range of *35 to 65+years*.

Facebook's health care research team recently published a scholarly paper suggesting the potential in using Facebook to advance the social determinants of health [25]. If Facebook has an interest in promoting health and health care research, we have several relatively simple recommendations on how this can be done. We strongly encourage Facebook to provide an option for random sampling of target audiences, as this will have considerable benefit in all areas of research—academic and otherwise. We further encourage Facebook to remove the *65+years* limitation for age targeting and allow users to specify any age range. This will be especially important as the number of older adults using Facebook continues to grow.

The greatest limitation of Facebook for research recruitment is, ironically, what makes it so useful as a marketing tool: Facebook's targeting algorithm, which learns from people's interactions with the ad, and then preferentially shows the ad to people with similar profiles. This produces an inherent sampling bias toward people who are more likely to respond to the advertisement. Therefore, because of the nonprobabilistic sampling nature of Facebook advertising, we caution against making certain inferences about the population—even if the sample is apparently representative of population demographic characteristics. Furthermore, given how the internet is constantly changing as technology grows, it is possible that this recruitment method may become more or less representative over time. Future changes to Facebook's targeting algorithm may also impact its reliability in recruiting samples.

Researchers should carefully weigh the importance of having a probabilistic sampling method in achieving their research objectives with the importance of achieving a robust sample size within resource and time constraints. If researchers believe the latter is of higher importance, then Facebook advertising should be considered. It could be especially effective in comparison with other nonprobabilistic sampling methods; other research has noted that Facebook may be able to obtain more representative samples than other types of nonprobabilistic methods [19]. Moreover, probabilistic sampling does not inherently result in representative inferences, so although some may criticize nonprobabilistic sampling, it can yield valid inferences with proper weighting and adjustment [26].

Finally, the literature addresses many concerns and provides guidance and tools for researchers to deal with issues, such as data ownership, privacy, bias, and communication between participants and potential participants [19,27-29]. The pertinent issue that arose in our study was that of communication with and between participants. Anyone who sees the advertisement may comment on it and so researchers should carefully consider how they will interact with participants, particularly in response to questions and comments. For example, after noticing a few inappropriate comments on our own advertisements, we deemed it necessary to create a comments policy to guide how we handled responses including profanity, spam, irrelevant remarks, unauthorized medical advice, stigmatizing messages, and false statements. We published this policy as a *note* on our Facebook page and referred to it as necessary. In summary, there are numerous ethical concerns to social media recruitment and internet research which should be considered by future investigators using social media for recruitment.

Conclusions

Overall, we achieved a partially representative sample, that is, it was representative or practically representative of 8 out of 13 sociodemographic and health characteristics assessed. Considering that this is a nonprobabilistic sampling method and the sociodemographic groups under-represented in our sample are commonly under-represented in probabilistic sampling methods, we believe these results are promising. These findings also suggest that purposively targeting subpopulations can improve representativeness. These findings suggest that Facebook advertising is a highly feasible and economical means of recruiting middle-aged and older adults for survey research. We, therefore, believe that Facebook advertising is useful for recruiting practically representative samples for many types of research and strongly encourage it in place of traditional nonprobabilistic methods where applicable. That said, there are inherent limitations to Facebook's current targeting algorithm that limit its usefulness when probabilistic sampling is required for inferences. We urge researchers who wish to employ Facebook advertising for recruitment to be detailed and highly transparent when describing their methods.

XSL•FO
RenderX

Acknowledgments

The authors wish to thank the NLCHI for providing the CCHS data. The authors would also like to thank Natalie Fedorak for her time proofreading the manuscript. This study was supported by the NL Healthy Aging Research Program through a research grant awarded to PPW (primary investigator).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Advertisement Campaign Details.
[PDF File (Adobe PDF File), 521KB - jmir_v21i8e14021_app1.pdf]

References

1. Batterham PJ. Recruitment of mental health survey participants using internet advertising: content, characteristics and cost effectiveness. Int J Methods Psychiatr Res 2014 Jun;23(2):184-191. [doi: 10.1002/mpr.1421] [Medline: 24615785]
2. Blumberg SJ, Ganesh N, Luke JV, Gonzales G. Wireless substitution: state-level estimates from the national health interview survey, 2012. Natl Health Stat Report 2013 Dec 18(70):1-16 [FREE Full text] [Medline: 24467831]
3. Canadian Radio-Television and Telecommunications Commission. 2017. Communications Monitoring Report URL: https://crtc.gc.ca/eng/publications/reports/PolicyMonitoring/2017/cmr2017.pdf [accessed 2019-03-01] [WebCite Cache ID 76XV3gXmM]
4. Technical Specifications - Statistics Canada. 2014. Residential Telephone Service Survey, 2013 URL: https://www150.statcan.gc.ca/n1/daily-quotidien/140623/dq140623a-eng.pdf [accessed 2019-03-01] [WebCite Cache ID 76YP5VdbE]
5. Thornton L, Batterham PJ, Fassnacht DB, Kay-Lambkin F, Calear AL, Hunt S. Recruiting for health, medical or psychosocial research using Facebook: systematic review. Internet Interv 2016 May;4:72-81 [FREE Full text] [doi: 10.1016/j.invent.2016.02.001] [Medline: 30135792]
6. Statista. 2018. Most Famous Social Network Sites Worldwide as of April 2019, Ranked by Number of Active Users (in Millions) URL: https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/ [accessed 2018-12-12] [WebCite Cache ID 76XUphZeU]
7. Statista. 2018. Frequency of Use of Facebook Products and Services in Canada as of March 2018, by Region URL: https://www.statista.com/statistics/822150/canada-facebook-products-services-use-by-region/ [accessed 2018-06-09] [WebCite Cache ID 76XTJUc3t]
8. Statista. 2018. Frequency of Use of Facebook Products and Services in Canada as of March 2018, by Age Group URL: https://www.statista.com/statistics/822184/canada-facebook-products-services-use-by-age/ [accessed 2018-06-09] [WebCite Cache ID 76XTnhBT7]
9. Facebook Business. How Does Facebook Detailed Targeting Work? URL: https://web.archive.org/web/20190301041538/https://www.facebook.com/business/help/182371508761821 [accessed 2018-12-12]
10. Whitaker C, Stevelink S, Fear N. The use of Facebook in recruiting participants for health research purposes: a systematic review. J Med Internet Res 2017 Dec 28;19(8):e290 [FREE Full text] [doi: 10.2196/jmir.7071] [Medline: 28851679]
11. 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]
12. Facebook Business. About Location Targeting URL: https://web.archive.org/web/20190301041700/https://www.facebook.com/business/help/202297959811696?helpref=related [accessed 2018-06-09]
13. Innovation, Science and Economic Development: Canada. 2015. Forward Sortation Area - Definition URL: https://www.ic.gc.ca/eic/site/bsf-osb.nsf/eng/br03396.html [accessed 2019-03-01] [WebCite Cache ID 76XVsO7Wg]
14. Statistics Canada. 2016. Data Tables, 2016 Census URL: https://www12.statcan.gc.ca/census-recensement/2016/dp-pd/dt-td/Index-eng.cfm [WebCite Cache ID 76XXQ4xKc]
15. Beasley TM, Schumacker RE. Multiple regression approach to analyzing contingency tables: post hoc and planned comparison procedures. J Exp Educ 1995 Oct;64(1):79-93. [doi: 10.1080/00220973.1995.9943797]
16. Gravetter FJ, Wallnau LB. Statistics For The Behavioral Sciences. Ninth Edition. Belmont, CA: Cengage Learning; 2013.
17. Korkeila K, Suominen S, Ahvenainen J, Ojanlatva A, Rautava P, Helenius H, et al. Non-response and related factors in a nation-wide health survey. Eur J Epidemiol 2001;17(11):991-999. [doi: 10.1023/A:1020016922473] [Medline: 12380710]
18. Etter JF, Perneger TV. Analysis of non-response bias in a mailed health survey. J Clin Epidemiol 1997 Oct;50(10):1123-1128. [doi: 10.1016/S0895-4356(97)00166-2] [Medline: 9368520]
19. Kosinski M, Matz SC, Gosling SD, Popov V, Stillwell D. Facebook as a research tool for the social sciences: opportunities, challenges, ethical considerations, and practical guidelines. Am Psychol 2015 Sep;70(6):543-556. [doi: 10.1037/a0039210] [Medline: 26348336]

XSL•FO
RenderX

20. Wang PP, Dicks E, Gong X, Buehler S, Zhao J, Squires J, et al. Validity of random-digit-dialing in recruiting controls in a case-control study. Am J Health Behav 2009;33(5):513-520. [doi: 10.5993/AJHB.33.5.4] [Medline: 19296741]
21. Richiardi L, Boffetta P, Merletti F. Analysis of nonresponse bias in a population-based case-control study on lung cancer. J Clin Epidemiol 2002 Oct;55(10):1033-1040. [doi: 10.1016/S0895-4356(02)00455-9] [Medline: 12464380]
22. Nagata JM, Hernández-Ramos I, Kurup AS, Albrecht D, Vivas-Torrealba C, Franco-Paredes C. Social determinants of health and seasonal influenza vaccination in adults ≥65 years: a systematic review of qualitative and quantitative data. BMC Public Health 2013 Apr 25;13:388 [FREE Full text] [doi: 10.1186/1471-2458-13-388] [Medline: 23617788]
23. Sullivan GM, Feinn R. Using effect size-or why the P value is not enough. J Grad Med Educ 2012 Sep;4(3):279-282 [FREE Full text] [doi: 10.4300/JGME-D-12-00156.1] [Medline: 23997866]
24. Patton MQ. Qualitative research. In: Everitt BS, Howell D, editors. Encyclopedia of Statistics in Behavioral Science. Volume 3. Chichester, England: John Wiley; 2005:1633-1636.
25. Abnousi F, Rumsfeld JS, Krumholz HM. Social determinants of health in the digital age: determining the source code for nurture. J Am Med Assoc 2019 Jan 22;321(3):247-248. [doi: 10.1001/jama.2018.19763] [Medline: 30570663]
26. Baker R, Brick JM, Bates NA, Battaglia M, Couper MP, Dever JA, et al. American Association for Public Opinion Research. 2013. Non-Probability Sampling URL: https://www.aapor.org/Education-Resources/Reports/Non-Probability-Sampling.aspx [accessed 2019-03-01] [WebCite Cache ID 76XSQwCM8]
27. King DB, O'Rourke N, DeLongis A. Social media recruitment and online data collection: a beginner's guide and best practices for accessing low-prevalence and hard-to-reach populations. Can Psychol 2014;55(4):240-249. [doi: 10.1037/a0038087]
28. Torous J, Nebeker C. Navigating ethics in the digital age: introducing connected and open research ethics (CORE), a tool for researchers and institutional review boards. J Med Internet Res 2017 Dec 8;19(2):e38 [FREE Full text] [doi: 10.2196/jmir.6793] [Medline: 28179216]
29. Gelinas L, Pierce R, Winkler S, Cohen IG, Lynch HF, Bierer BE. Using social media as a research recruitment tool: ethical issues and recommendations. Am J Bioeth 2017 Mar;17(3):3-14 [FREE Full text] [doi: 10.1080/15265161.2016.1276644] [Medline: 28207365]

Abbreviations

BMI: body mass index
CCHS: Canadian Community Health Survey
NL: Newfoundland and Labrador
NLCHI: Newfoundland and Labrador Centre for Health Information
RDD: random digit dialing

Edited by G Eysenbach; submitted 22.03.19; peer-reviewed by K Reuter, E Buchanan; comments to author 14.06.19; revised version received 21.06.19; accepted 27.06.19; published 19.08.19.

Please cite as:
Shaver LG, Khawer A, Yi Y, Aubrey-Bassler K, Etchegary H, Roebothan B, Asghari S, Wang PP
Using Facebook Advertising to Recruit Representative Samples: Feasibility Assessment of a Cross-Sectional Survey
J Med Internet Res 2019;21(8):e14021
URL: http://www.jmir.org/2019/8/e14021/
doi:10.2196/14021
PMID:31429409

©Lance Garrett Shaver, Ahmed Khawer, Yanqing Yi, Kris Aubrey-Bassler, Holly Etchegary, Barbara Roebothan, Shabnam Asghari, Peizhong Peter Wang. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 19.08.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Why Do Data Users Say Health Care Data Are Difficult to Use? A Cross-Sectional Survey Study

Ho Heon Kim[1*], BSN; Bora Kim[2*], BS; Segyeong Joo[3,4*], PhD; Soo-Yong Shin[5,6], PhD; Hyo Soung Cha[7], PhD; Yu Rang Park[1], PhD

[1]Department of Biomedical Systems Informatics, Yonsei University College of Medicine, Seoul, Republic of Korea

[2]Emphasis Information Technology Inc, Seoul, Republic of Korea

[3]Department of Biomedical Engineering, University of Ulsan College of Medicine, Seoul, Republic of Korea

[4]Biomedical Engineering Research Center, Asan Institute for Life Sciences, Asan Medical Center, Seoul, Republic of Korea

[5]Department of Digital Health, Samsung Advanced Institute for Health Sciences & Technology, SungKyunKwan University, Seoul, Republic of Korea

[6]Big Data Research Center, Samsung Medical Center, Seoul, Republic of Korea

[7]Cancer Big Data Center, National Cancer Center, Gyeonggi-do, Republic of Korea

[*]these authors contributed equally

Corresponding Author:
Yu Rang Park, PhD
Department of Biomedical Systems Informatics
Yonsei University College of Medicine
50-1 Yonsei-ro
Seodaemun-gu
Seoul, 03722
Republic of Korea
Phone: 82 10 5240 3434
Fax: 82 2 2228 2493
Email: yurangpark@yuhs.ac

Abstract

Background: There has been significant effort in attempting to use health care data. However, laws that protect patients' privacy have restricted data use because health care data contain sensitive information. Thus, discussions on privacy laws now focus on the active use of health care data beyond protection. However, current literature does not clarify the obstacles that make data usage and deidentification processes difficult or elaborate on users' needs for data linking from practical perspectives.

Objective: The objective of this study is to investigate (1) the current status of data use in each medical area, (2) institutional efforts and difficulties in deidentification processes, and (3) users' data linking needs.

Methods: We conducted a cross-sectional online survey. To recruit people who have used health care data, we publicized the promotion campaign and sent official documents to an academic society encouraging participation in the online survey.

Results: In total, 128 participants responded to the online survey; 10 participants were excluded for either inconsistent responses or lack of demand for health care data. Finally, 118 participants' responses were analyzed. The majority of participants worked in general hospitals or universities (62/118, 52.5% and 51/118, 43.2%, respectively, multiple-choice answers). More than half of participants responded that they have a need for clinical data (82/118, 69.5%) and public data (76/118, 64.4%). Furthermore, 85.6% (101/118) of respondents conducted deidentification measures when using data, and they considered rigid social culture as an obstacle for deidentification (28/101, 27.7%). In addition, they required data linking (98/118, 83.1%), and they noted deregulation and data standardization to allow access to health care data linking (33/98, 33.7% and 38/98, 38.8%, respectively). There were no significant differences in the proportion of responded data needs and linking in groups that used health care data for either public purposes or commercial purposes.

Conclusions: This study provides a cross-sectional view from a practical, user-oriented perspective on the kinds of data users want to utilize, efforts and difficulties in deidentification processes, and the needs for data linking. Most users want to use clinical and public data, and most participants conduct deidentification processes and express a desire to conduct data linking. Our study confirmed that they noted regulation as a primary obstacle whether their purpose is commercial or public. A legal system based on both data utilization and data protection needs is required.

XSL•FO
RenderX

(J Med Internet Res 2019;21(8):e14126) doi:10.2196/14126

KEYWORDS

data anonymization; privacy act; data sharing; data protection; data linking; health care data demand

Introduction

There has been considerable effort to use health care data [1,2], and many countries have implemented regulations to protect the privacy of patients and research subjects [3-5]. Owing to the sensitivity of health care data, privacy protection laws have limited its use [6]. Regulations that focus only on protecting privacy are emerging as a major challenge in using health care data [7-9].

Health care institutions and governments both generate a large amount of heterogeneous data [10]. To use these decentralized data, there have been dramatic increases in linking data from diverse sources [11]. By using big data analytic approaches, which leverage data drawn from multiple sources [12], data-driven research has the potential for widespread positive impact and global implications [13-16]. Efforts have been made to use health care data for the following purposes: ensuring a high level of evidence by using a large number of samples [17], identifying risk factors [18], and improving diagnosis and treatment standards [19].

However, in Korea, this use conflicts with the current regulations because data linking requires the data be identified and shared [20,21]. The privacy law of Korea is known as the strongest principle in Asia [22-24]. Although most discussions about privacy laws have centered on data protection, discussions about the privacy law are now about the need to facilitate the development of industries that utilize data beyond protection [8]. However, there has been no mention of what makes data usage and deidentification processes difficult or users' needs for data linking from a practical perspective.

The objective of this study is to investigate (1) the status of big data utilization in different medical areas (general hospitals, universities, industry, and academic society); (2) institutional obstacles and efforts in deidentification processes, which is an alternative approach for using health care data; and (3) users' data linking needs.

Methods

Study Design and Data Collection

This study is designed to investigate the demand for health care data, identify the difficulties in using health care data, and develop improvements for using health care data from the practical users' perspective. For this, we conducted a cross-sectional online survey. To recruit participants who use health care data, we (1) publicized the survey promotion campaign through social media (Facebook) and (2) sent official documents to academic societies encouraging participation in the online survey. Through the provided documents, anyone who used health care data was able to participate in the questionnaire (online open survey; see details in Multimedia Appendix 1).

The online questionnaire was developed and distributed using Office forms (Naver, Korea). This questionnaire could be accessed from mobile phones and personal computers. To ensure important questions were answered, seven mandatory items were designated among the 17 questions. This function was used to prevent participants from submitting responses without checking the answers on mandatory items before submission. However, the questionnaire did not verify data consistency. For example, respondents who replied that they did not have a demand for health care data could also select "clinical data" as a response to the question asking about required data. To ensure the validity of the questionnaire, the items on the questionnaire were developed through 15 revisions in consultation with eight experts over a period of approximately one month. The final questionnaire consisted of 16 items within five parts. Each screen contained one to eight questions; there were a total of eight screens in the survey (on mobile and PC screens).

Ethical clearance was obtained from the Public Institutional Review Board designated by The Korean Ministry of Health and Welfare (number: 2018-2199-001) before data collection.

Participant Recruitment

We selected five academic societies (Korean Society of Medical Informatics, Korean Society for Preventive Medicine, Korean Cancer Association, Korea Society of Artificial Intelligence in Medicine, and Korean Society of Epidemiology) that exhibit a high demand for health care data or were recommended by experts. Then, we encouraged participation in the survey by sending an official letter requesting cooperation for online surveys to the secretariat of each academic society.

A web link to access the survey was provided to interested respondents. Respondents were required to provide consent through this link. To receive consent from respondents, the first screen of the online questionnaire included the background, purpose, and duration of the research, as well as a description of the disadvantages or limitations. After respondents approved this introduction, the link led to the anonymous online questionnaire. As an incentive for participation, they were offered coffee gift vouchers by submitting their cell phone numbers. To transfer the coffee vouchers and exclude duplicate responses, informed consent to collect cell phone numbers was received separately. The cross-sectional online survey was conducted between October 5 and 19, 2018.

By the end of the survey period, 128 participants responded to the online survey. Responses that were contradictory (n=2) or did not exhibit a demand for health care data (n=8) were excluded; therefore, a total of 118 participants were included in the analysis. The overall eligible population of subjects was unknown because the online survey was sent to the five academic societies and was advertised through a social media promotion.

XSL•FO
RenderX

Among the responses (N=118), quality improvement of welfare services and research promotion were considered to be public purposes (81/115, 70.4%) and industrial development and profit generation were classified as commercial purposes (34/115, 29.6%); this classification excludes other minor purposes (n=3).

Questionnaire Items

The survey items were categorized into five parts. The first part included items that investigated the work experience and basic information of participants. The second part inquired about the type of data participants wanted. The third part related to obstacles and improvement suggestions for data use. The fourth part investigated the identification process, and the last part investigated data linking (details in Multimedia Appendix 2).

Statistical Analysis

Analyses were conducted using R (version 3.5.1) and Microsoft Excel (version 2016). Descriptive statistics for proportions of respondents, work profiles (eg, age, work experience, expertise area, working institution), and responses regarding data demand, data linking, and deidentification were explored.

For categorical variables, such as data needs, obstacles, and improvement suggestions, chi-square tests were performed to show these reponses were different between participants using data for public purposes and those using data for commercial purposes. We conducted chi-square tests with one section as the response to specific questions, such as obstacles to using health care data. Chi-square tests could not be performed for responses to questions that allowed participants to choose more than one answer (multiple response questions), such as data needs, because the responses were not independent. For questions that could have multiple responses, post hoc chi-square tests were performed (Multimedia Appendices 3 and 4). Post hoc pairwise chi-square tests involved testing each value of the nominal variable versus the sum of all others. After applying the same principle of chi-square to get the *P* value for each comparison, we then used Bonferroni correction to counteract the problem of type I error that occurs when multiple comparisons are made.

Results

Overall Population

The majority of online survey participants worked in a general hospital (62/118, 52.5%; multiple response question) or university (51/118, 43.2%; multiple response). Most participants were in the field of research (84/118, 71.2%), in their thirties (56/118, 47.5%), and had work experience between 1 and 5 years (56/118, 47.5%; Table 1).

Table 1. Profile of online survey respondents (N=118).

Characteristics	Respondents, n (%)
Age (years)	
20-29	21 (17.8)
30-39	56 (47.5)
40-49	34 (28.8)
50-59	5 (4.2)
Other	2 (1.7)
Institution (multiple response question)	
General hospital	62 (52.5)
University	51 (43.2)
Industry	15 (12.7)
Academic society	6 (5.1)
Other	3 (2.6)
Expertise	
Research	84 (71.2)
Data analysis	18 (15.3)
Planning	11 (9.3)
Device development	5 (4.2)
Expertise experience	
≥10 years	9 (7.6)
5 years to <10 years	32 (27.1)
1 year to <5 years	56 (47.5)
<1 year	21 (17.8)

XSL•FO
RenderX

Data Demand, Obstacles, and Improvement Suggestions

More than half of participants replied that they had a need for clinical data (82/118, 69.5%) and public data (76/118, 64.4%; Table 2). Only the general hospital group selected clinical data in a high proportion (56/62, 90.3%).

Participants reported that the most significant obstacles in trying to use health care data were conflicts with the law (53/118, 44.9%) and data standardization (50/118, 42.4%). However, the obstacles most frequently selected by each group were different. Overall, the four groups of respondents by institution (general hospital, university, industry, and academic society) reported data standardization problems and legal conflicts as the main challenges in using data.

Similarly, most participants indicated that legislation improvement was required to overcome these data utilization limitations (54/118, 45.8%), followed by the need for technical measures for data standards (47/118, 39.8%). Overall, participants suggested that law revision was the first priority of improvement (Table 2).

There was no statistically significant difference in the percentage of obstacles in groups that used health care data for either commercial or public purposes (P=.38). However, both groups indicated that data standardization and current laws function as constraints of health care data use (Table 3).

Table 2. Data needs, obstacles, and developmental proposals for data utilization.

Characteristics	Respondents, n (%)					Total (N=118), n (%)
	General hospital (n=62)	University (n=51)	Industry (n=15)	Academic society (n=6)	Other (n=3)	
Data needs (multiple response question)						
Clinical data (collected during care process in hospital)	56 (90.3)	29 (56.9)	10 (66.7)	5 (83.3)	0 (0.0)	82 (69.5)
Public data (managed by nation)	40 (64.5)	32 (62.7)	10 (66.7)	5 (83.3)	3 (100.0)	76 (64.4)
Research data (clinical research or trial data)	38 (61.3)	33 (64.7)	3 (2.0)	5 (83.3)	0 (0.0)	61 (51.7)
Life log data (patient generated health data)	17 (27.4)	16 (31.4)	8 (53.3)	4 (66.7)	1 (33.3)	36 (30.5)
Genetic data	13 (21.0)	18 (35.3)	1 (6.7)	2 (33.3)	1 (33.3)	28 (23.7)
Obstacle						
Conflict of laws	30 (48.4)	19 (37.3)	8 (53.3)	2 (33.3)	0 (0.0)	53 (44.9)
Data standardization	24 (38.7)	23 (45.1)	5 (33.3)	3 (50.5)	3 (100.0)	50 (42.4)
Strict social recognition	5 (8.1)	4 (7.8)	2 (13.3)	1 (16.7)	0 (0.0)	9 (7.6)
Other	1 (1.6)	1 (4.0)	0 (0.0)	0 (0.0)	0 (0.0)	2 (1.7)
None	2 (3.2)	3 (5.9)	0 (0.0)	0 (0.0)	0 (0.0)	4 (3.4)
Improvement suggestion (multiple response question)						
Law revision	29 (46.8)	22 (43.1)	7 (46.7)	4 (66.7)	2 (66.7)	54 (45.8)
Technical measures	22 (35.5)	23 (43.1)	6 (40.0)	3 (50.0)	1 (33.3)	47 (39.8)
Utilization support	12 (19.4)	7 (13.7)	0 (0.0)	0 (0.0)	0 (0.0)	16 (13.6)
Public consensus	6 (9.7)	11 (21.6)	2 (13.3)	1 (16.7)	1 (33.3)	22 (18.6)

XSL•FO
RenderX

Table 3. Comparison of data demand, obstacles, and improvement suggestions between health care data use for commercial and public purposes (N=115). Sample size excludes the n=3 for other minor purposes.

Measures	Respondents, n (%)		Chi-square (df)	*P* value
	Public purpose (n=81)	Commercial purpose (n=34)		
Data needs (multiple response question)			Not applicable	Not applicable
Clinical data	53 (65.4)	28 (82.4)		
Public data	53 (65.4)	23 (67.6)		
Research data	44 (54.3)	16 (47.1)		
Life log data	18 (22.2)	18 (52.9)		
Genetic data	17 (21.0)	9 (26.5)		
Obstacles			2.9 (4)	.38
Conflict with laws	38 (46.9)	13 (38.2)		
Data standardization	35 (43.2)	15 (44.1)		
Strict social recognition	6 (7.4)	3 (8.8)		
None	1 (1.2)	2 (5.9)		
Other	1 (1.2)	1 (2.9)		
Improvement (multiple response question)			Not applicable	Not applicable
Law revision	38 (46.9)	15 (44.1)		
Technical method	33 (40.7)	14 (41.2)		
Data utilization support	15 (18.5)	1 (2.9)		
Public consensus	14 (17.3)	6 (17.6)		

Deidentification

When using health care data, 101 participants responded that they conduct deidentification measures (101/118, 85.6%). The majority of participants reported that multiple deidentification methods are used (64/101, 63.4%). The most common method was pseudonymization (72/101, 71.3%), followed by masking (57/101, 56.4%). Most respondents who conducted deidentification considered privacy issues induced by rigid social culture as the biggest problem for deidentification (28/101, 27.7%), followed by the absence of clear criteria for deidentification measures (24/101, 23.8%; Table 4).

XSL•FO
RenderX

Table 4. Responses about the current state of data deidentification (N=118).

Measures	Respondents, n (%)
Deidentify when using health care data (n=118)	
Yes	101 (85.6)
No	17 (14.4)
Number of applied deidentification methods (n=101)	
1 method	37 (31.4)
2 methods	33 (28.0)
3 methods	18 (15.3)
4 methods	4 (3.4)
5 methods	9 (7.6)
Applied methods (n=101; multiple response question)	
Pseudonymization	72 (71.3)
Masking	57 (56.4)
Data reduction	37 (36.6)
Data suppression	30 (29.7)
Aggregation	22 (21.8)
Difficulties when deidentifying data (n=101)	
Strict social culture	28 (27.7)
Absence of clear deidentification guideline	24 (23.8)
Usefulness of deidentified data	15 (14.9)
Lack of understanding of deidentification policy and technology	14 (13.9)
Lack of relevant institution support	11 (10.9)
Lack of deidentification measure for unstructured data	9 (8.9)

Data Linkage

The majority of participants answered that they require data linking (98/118, 83.1%). The difference in the proportion of respondents who wanted to use data linkage for public or commercial purposes was not statistically significant (P=.64). The 98 respondents who said that data linking was necessary indicated that the purpose of linking data was to obtain longitudinal data (62/98, 63.3%). In addition, deregulation and data standardization comprised a large proportion of data linking improvement suggestions (33/98, 33.7% and 38/98, 38.8%, respectively). In the two items that investigated the reason for data linkage and suggestions to facilitate data linking, the proportion of responses in both the public purpose and commercial purpose groups did not significantly differ (P=.16 and P=.47, respectively).

The groups that used data for public purposes responded that health care data are to be used to develop health care policy (41.8%, 28/81). On the other hand, the group that used data for commercial purposes primarily responded that data was to be used for the development of diagnostic technology (n=12; Table 5).

XSL•FO
RenderX

Table 5. Demand for health care data linking.

Measures	Participants, n (%)				Chi-square (df)	*P* value[a]
	Public purpose (n=81)	Commercial purpose (n=34)	Other (n=3)	Total N=118)		
Data linking					0.2 (1)	.64
Required	67 (82.7)	30 (88.2)	1 (33.3)	98 (83.1)		
Not required	14 (17.3)	4 (11.8)	2 (66.7)	20 (16.9)		
Reason for data linking (n=98)					3.6 (2)	.16
Obtain longitudinal data	39 (58.2)	23 (76.7)	0 (0.0)	62 (63.3)		
Obtain larger number of subjects	15 (22.4)	5 (16.7)	0 (0.0)	20 (20.4)		
Develop policy predicated on data	13 (19.4)	2 (6.7)	1 (100.0)	16 (16.3)		
Suggestions for facilitating health care data linking (n=98)					2.5 (3)	.47
Deregulation	22 (32.8)	11 (36.7)	0 (0.0)	33 (33.7)		
Data standardization	28 (41.8)	10 (33.3)	0 (0.0)	38 (38.8)		
Effective guidelines including procedure, responsibility, and technology	11 (16.4)	8 (26.7)	1 (100.0)	20 (20.4)		
Improvement of social recognition	6 (9.0)	1 (3.3)	0 (0.0)	7 (7.1)		
Usage details (n=98)					18.8 (6)	.003
Development of health care policy	28 (41.8)	2 (6.7)	1 (100.0)	31 (31.6)		
Development of diagnostic technology	15 (22.4)	12 (40.0)	0 (0.0)	27 (27.6)		
Development of treatment modality	12 (17.9)	4 (13.3)	0 (0.0)	16 (16.3)		
General research	8 (11.9)	4 (13.3)	0 (0.0)	12 (12.2)		
Development of medical device	2 (3.0)	6 (20.0)	0 (0.0)	8 (8.2)		
Development of new drug	1 (1.5)	1 (3.3)	0 (0.0)	2 (2.0)		
Other	1 (1.5)	1 (3.3)	0 (0.0)	2 (2.0)		
Subtotal	67 (100.0)	30 (100.0)	1 (100.0)	98 (100.0)		

[a]Public versus commercial.

Discussion

Principal Findings

The primary finding of this study was the clarification of each health care area's need for data. Most wanted to use clinical data and public data, except for university respondents. Considering the amount of stored data depending on the health care field [25], it is understandable that clinical data are in high demand.

Secondly, most participants who use health care data conduct deidentification measures before data use. The majority of deidentification measures are implemented using more than one method. This survey was not able to distinguish between cases in which deidentification was not conducted when required and cases in which it was not conducted because the data was not identifiable (whether due to the exclusion of personal information or the lack of legal deidentification requirements).

Although it is not clear whether these respondents voluntarily implemented deidentification measures or were obligated to do so, it appears they consider health care data to be sensitive information. Their use of multiple deidentification measures may be considered proof of action to mitigate concerns about privacy infringement. However, they pointed out that rigid social culture acts as a primary obstacle in data deidentification. Therefore, if we prove that privacy is guaranteed, we can achieve social consensus and relieve sociocultural rigidity.

Lastly, the proportion of respondents who need to link data was significantly larger than the proportion of respondents who do not; these respondents indicated that deregulation and standardization are necessary to facilitate data linkage. This suggests that many of these respondents face difficulties due to intensive regulation. Data users may experience legal conflicts when they want to link data from external data sources. When linking with external data, an identifier is required, which is often personal information. If consent has been obtained for other research purposes previously, this identifier can be used; however, in big data analysis, there are limitations on obtained consent [21]. Furthermore, for personal information to be provided to third parties, they must obtain the consent of the subjects (article 17, Personal Information Protection Act [PIPA]). Practically, it has been burdensome for controllers to

XSL•FO
RenderX

recontact individual subjects and obtain consent; thus, they may be obstructed by law [26]. Data sharing and linkage are limited by the PIPA [27].

In recent years, some countries have attempted to revise their information protection legislation to prepare for the development of a new information industry [28,29]. The United States has enacted the Final Rule, a revision of the Common Rule, to reduce the regulatory burden and create a new concept of broad consent to enhance both the use and protection of data [4,30]. The European Union, by enacting the General Data Protection Regulation, has strengthened data protection principles while including principles such as the right to data portability [31]. In Japan, the concept of anonymizing processed information is defined by law, and the use of personal information is being promoted [32].

Yet, privacy remains an issue in countries that are trying to implement centralized electronic health records (EHRs), such as Canada. Centralized EHRs could have interoperability in terms of data structure because the same data schema enables data linking and communication. This would reduce the obstacle for health care data use. However, in terms of comprehensive use and communication of data, the privacy issue must be handled for secondary use. Therefore, as long as privacy remains an issue, there will also be a need for data linking. For example, a study on a Canada-wide EHR system stated that privacy systems should address the issues of deidentifying health care data and privacy concerns [33,34]; skeptics have warned against adopting a Canada-wide EHR system until then [35]. Furthermore, even with centralized EHRs, the privacy issue will remain in situations of linking with privately collected data, such as mobile data and data collected by wearable devices. In a survey conducted in the United States to identify digital health adoption and sentiments of consumers, results showed people are rarely willing to provide personal health data to pharmaceutical companies, research institutes, or information technology companies [36].

Considering this global trend, the regulation of personal information in Korea does not reflect these changes [37]. There have been many studies on methods by which regulation can be improved to reflect changes in secondary data usage; however, to provide a basis for these legislative improvements, there was a need for evidence to show that actual users experienced these difficulties and needs for data linking.

The results of this study confirm that the use of health care data conflicts with the law, which leads to the implication that legislation should be revised to facilitate data utilization. However, it should not simply be deregulated, but balanced between protection and utilization, as is the case of major countries. To improve this legal system, a survey of opinions on the use of health care data also should be conducted on the data supplier and beneficiary side (the general population). In the United States, these surveys about digital health consumer's sentiments have been conducted, and most respondents remain wary about sharing their health data with technology companies [36]. Likewise, surveys on how the opinions of hospital's data managers differ from those of the users in our study should be conducted to achieve a better social consensus and reconcile the two areas of data utilization and protection.

Limitations

The respondents in this study were primarily involved in general hospitals and universities, whereas the respondents in academia and industry were few. Respondents affiliated with universities are considered to hold concurrent positions in general hospitals. In addition, we did not obtain significant information on the characteristics of the entire population in the survey because survey promotion was conducted through social network services and the transmission of official documents. However, considering the number of medical institutions in Korea (tertiary hospitals or secondary hospitals) and the number of universities, it is natural that many respondents belong to medical institutions and universities. Although this may not directly represent the opinion of the entire population in need of health care data, in the absence of previous studies that directly investigate the opinions of data users, this study has the advantage of illuminating the present status of Korean data users' perspectives in a cross-sectional way.

In addition, to represent the overall opinion of the population, it is necessary to select the population for each institution and extract a sample using a stratified sampling method. We examined the current circumstances of health care data use from data users' perspectives, but data managers and beneficiaries should also be surveyed for policy development to ensure that all parties are considered in bridging the gap between data privacy and utilization.

Conclusion

This study provides a cross-sectional view from a practical user-oriented perspective on the types of data users find valuable, the efforts and obstacles that characterize deidentification processes, and users' needs for data linking. Most respondents seek to use clinical and public data. Moreover, most implement deidentification measures. We confirmed that they want to link data but are limited by regulations regardless of whether their purpose is commercial or public. A legal system that is founded on both the utilization and protection of data is necessary.

Acknowledgments

This study was supported by a new faculty research seed money grant of Yonsei University College of Medicine for 2018 (32-0044), the Basic Science Research Program through the National Research Foundation of Korea (NRF), funded by the Ministry of Education (NRF-2017R1D1A1B03035762) and the grants (2017-544) from the Asan Institute for Life Sciences (Seoul, South Korea).

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1

Notice for survey.

[PNG File, 78KB - jmir_v21i8e14126_app1.png]

Multimedia Appendix 2

Online questionnaire on health care data utilization.

[DOCX File, 23KB - jmir_v21i8e14126_app2.docx]

Multimedia Appendix 3

Post hoc pairwise chi-square test for comparison of data needs, obstacles, and improvement.

[DOCX File, 17KB - jmir_v21i8e14126_app3.docx]

Multimedia Appendix 4

Post hoc chi-square test results: demand for health care data linking.

[DOCX File, 19KB - jmir_v21i8e14126_app4.docx]

References

1. Gu D, Li J, Li X, Liang C. Visualizing the knowledge structure and evolution of big data research in healthcare informatics. Int J Med Inform 2017 Dec;98:22-32. [doi: 10.1016/j.ijmedinf.2016.11.006] [Medline: 28034409]
2. Dhindsa K, Bhandari M, Sonnadara R. What's holding up the big data revolution in healthcare? BMJ 2018 Dec 28;363:K3557. [doi: 10.1136/bmj.k5357] [Medline: 30593447]
3. Suominen H, Müller H, Ohno-Machado L, Salanterä S, Schreier G, Hanlen L. Prerequisites for international exchanges of health information for record research: comparison of Australian, Austrian, Finnish, Swiss, and US policies. Stud Health Technol Inform 2017;245:1312. [Medline: 29295395]
4. Federal Register. 2018. Federal Policy for the Protection of Human Subjects: Six Month Delay of the General Compliance Date of Revisions While Allowing the Use of Three Burden-Reducing Provisions During the Delay Period URL: https://www.federalregister.gov/documents/2018/06/19/2018-13187/federal-policy-for-the-protection-of-human-subjects-six-month-delay-of-the-general-compliance-date
5. Voss W. European Union data privacy law reform: General Data Protection Regulation, privacy shield, and the right to delisting. Business Lawyer 2017 Jan 5;72(1):221-233 [FREE Full text]
6. Dyke S, Dove E, Knoppers BM. Sharing health-related data: a privacy test? NPJ Genom Med 2016 Aug 17;1(1):160241-160246 [FREE Full text] [doi: 10.1038/npjgenmed.2016.24] [Medline: 27990299]
7. Roski J, Bo-Linn G, Andrews TA. Creating value in health care through big data: opportunities and policy implications. Health Aff (Millwood) 2014 Jul;33(7):1115-1122. [doi: 10.1377/hlthaff.2014.0147] [Medline: 25006136]
8. Park DW, Lee SM, Jung SH. Korea Health Industry Development Institution. 2017 Dec. [Investigation of latest health medical Big data law trend] URL: https://www.khidi.or.kr/board/view?pageNum=1&rowCnt=10&linkId=48719246&menuId=MENU00085 [accessed 2019-02-14]
9. van Panhuis WG, Paul P, Emerson C, Grefenstette J, Wilder R, Herbst A, et al. A systematic review of barriers to data sharing in public health. BMC Public Health 2014 Nov 05;14(1):1144 [FREE Full text] [doi: 10.1186/1471-2458-14-1144] [Medline: 25377061]
10. Archenaa J, Anita EM. A survey of big data analytics in healthcare and government. Procedia Comput Sci 2015;50:408-413. [doi: 10.1016/j.procs.2015.04.021]
11. Lowrance W. Learning from experience: privacy and the secondary use of data in health research. J Health Serv Res Policy 2003 Jul;8 Suppl 1:S1:2-S1:7. [doi: 10.1258/135581903766468800] [Medline: 12869330]
12. Jagadish H, Gehrke J, Labrinidis A, Papakonstantinou Y, Patel J, Ramakrishnan R, et al. Big data and its technical challenges. Commun ACM 2014 Jul 01;57(7):86-94. [doi: 10.1145/2611567]
13. Obermeyer Z, Emanuel EJ. Predicting the future-big data, machine learning, and clinical medicine. N Engl J Med 2016 Sep 29;375(13):1216-1219 [FREE Full text] [doi: 10.1056/NEJMp1606181] [Medline: 27682033]
14. Kostkova P, Brewer H, de Lusignan S, Fottrell E, Goldacre B, Hart G, et al. Who owns the data? Open data for healthcare. Front Public Health 2016 Feb;4:7 [FREE Full text] [doi: 10.3389/fpubh.2016.00007] [Medline: 26925395]

XSL•FO
RenderX

15. Wang Y, Kung L, Byrd T. Big data analytics: understanding its capabilities and potential benefits for healthcare organizations. Technol Forecast Soc 2018 Jan;126:3-13. [doi: 10.1016/j.techfore.2015.12.019]
16. Raghupathi W, Raghupathi V. Big data analytics in healthcare: promise and potential. Health Inf Sci Syst 2014;2(1):3 [FREE Full text] [doi: 10.1186/2047-2501-2-3] [Medline: 25825667]
17. Collins F, Hudson K, Briggs J, Lauer M. PCORnet: turning a dream into reality. J Am Med Inform Assoc 2014;21(4):576-577 [FREE Full text] [doi: 10.1136/amiajnl-2014-002864] [Medline: 24821744]
18. Seong S, Kim Y, Park S, Khang Y, Kim H, Park J, et al. Cohort profile: the National Health Insurance Service-National Health Screening Cohort (NHIS-HEALS) in Korea. BMJ Open 2017 Sep 24;7(9):e016640 [FREE Full text] [doi: 10.1136/bmjopen-2017-016640] [Medline: 28947447]
19. Tomczak K, Czerwińska P, Wiznerowicz M. The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. Contemp Oncol (Pozn) 2015;19(1A):A68-A77 [FREE Full text] [doi: 10.5114/wo.2014.47136] [Medline: 25691825]
20. Kim D, Kim M, Kim K, Youm Y. Improved privacy statement for big data industry activation. In: Proceedings of Symposium of the Korean Institute of communications and Information Sciences. 2017 Jun Presented at: Korean Institute of Communication Sciences Summer General Conference; June 21-23, 2017; Jeju, Korea p. 257-258 URL: http://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE07218090
21. Park M. A study on legal framework for secondary use of public data in scientific research on health: focusing on secondary use of bio·medical information and genomics data. Kor J Med Law 2016 Dec 31;24(2):69-101. [doi: 10.17215/kaml.2016.12.24.2.69]
22. Greenleaf G, Park W. South Korea's innovations in data privacy principles: Asian comparisons. Comput Law Secur Rev 2014 Oct;30(5):492-505. [doi: 10.1016/j.clsr.2014.07.011]
23. Chesterman S. Data Protection Law in Singapore: Privacy and Sovereignty in an Interconnected World. Singapore: Academy Publishing; 2014.
24. Geller T. In privacy law, it's the US vs. the world. Commun ACM 2016 Jan 25;59(2):21-23. [doi: 10.1145/2852233]
25. Schlick C, Castle J, Bentrem DJ. Utilizing big data in cancer care. Surg Oncol Clin N Am 2018 Dec;27(4):641-652. [doi: 10.1016/j.soc.2018.05.005] [Medline: 30213409]
26. Kim SH. The principal issues and prospects of biobank management. Kor Assoc Health Med Sociol 2014;35:263-283 [FREE Full text]
27. Kim H, Kim S, Joly Y. South Korea: in the midst of a privacy reform centered on data sharing. Hum Genet 2018 Aug;137(8):627-635 [FREE Full text] [doi: 10.1007/s00439-018-1920-1] [Medline: 30121900]
28. Mostert M, Bredenoord A, Biesaart M, van Delden JJ. Big Data in medical research and EU data protection law: challenges to the consent or anonymise approach. Eur J Hum Genet 2016 Dec;24(7):956-960 [FREE Full text] [doi: 10.1038/ejhg.2015.239] [Medline: 26554881]
29. Gostin L, Halabi S, Wilson K. Health data and privacy in the digital era. JAMA 2018 Jul 17;320(3):233-234. [doi: 10.1001/jama.2018.8374] [Medline: 29926092]
30. US Department of Health & Human Services. HHS.gov. 2017 Jan 18. Final rule enhances protections for research participants, modernizes oversight system URL: http://wayback.archive-it.org/3926/20170127095200/https://www.hhs.gov/about/news/2017/01/18/final-rule-enhances-protections-research-participants-modernizes-oversight-system.html [accessed 2019-02-14]
31. Politou E, Alepis E, Patsakis C. Forgetting personal data and revoking consent under the GDPR: challenges and proposed solutions. J Cybersecur 2018;4(1):1. [doi: 10.1093/cybsec/tyy001]
32. Kim W, Kim H, No M. Weekly ICT Trends. 2017. [Comparison of personal information de-identification law for use of big data information by major countries] URL: http://www.itfind.or.kr/WZIN/jugidong/1792/file19146369765231286 97-1792.pdf [accessed 2019-02-14]
33. Francois M, Obisike E. Accelerating the national implementation of electronic health records in Canada. Eur Sci J 2016 May 30;12(15):65. [doi: 10.19044/esj.2016.v12n15p65]
34. Office of the Auditor General of Ontario. 2016 Annual Report Volume 1. Toronto, ON: Queen's Printer for Ontario; 2016. Electronic health records' implementation status URL: http://www.auditor.on.ca/en/content/annualreports/arreports/en16/2016AR_v1_en_web.pdf [accessed 2019-06-21]
35. Atherley G. The Risks of Electronic Health Records. Fraser Forum 2009:29-32 [FREE Full text]
36. Sean D, Megan Z. Rock Health. 2019. Beyond wellness for the healthy: digital health consumer adoption 2018 URL: https://rockhealth.com/reports/beyond-wellness-for-the-healthy-digital-health-consumer-adoption-2018/ [accessed 2019-02-14]
37. Seoul Asan Medical Center. Ministry of the Interior and Safety, Personal Information Protection Commission. 2016. [A study on improvement of the legal system for processing of personal information in research purpose: focusing on the field of healthcare research] URL: http://www.prism.go.kr/homepage/researchCommon/downloadResearchAttachFile.do?work_key=001&file_type=CPR&seq_no=001&pdf_conv_yn=Y&research_id=1079930-201600006 [accessed 2019-02-14]

Abbreviations

EHR: electronic health record
PIPA: Personal Information Protection Act

XSL•FO
RenderX

Edited by G Eysenbach; submitted 28.03.19; peer-reviewed by K Dhindsa, K Kim; comments to author 15.06.19; revised version received 27.06.19; accepted 29.06.19; published 06.08.19.

Please cite as:
Kim HH, Kim B, Joo S, Shin SY, Cha HS, Park YR
Why Do Data Users Say Health Care Data Are Difficult to Use? A Cross-Sectional Survey Study
J Med Internet Res 2019;21(8):e14126
URL: https://www.jmir.org/2019/8/e14126/
doi:10.2196/14126
PMID:31389335

©Ho Heon Kim, Bora Kim, Segyeong Joo, Soo-Yong Shin, Hyo Soung Cha, Yu Rang Park. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 06.08.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Public Views on Models for Accessing Genomic and Health Data for Research: Mixed Methods Study

Kerina H Jones[1], BSc, PhD; Helen Daniels[1], BSc, PhD; Emma Squires[1], BSc; David V Ford[1], MBA

Population Data Science, Swansea University Medical School, Swansea University, Swansea, United Kingdom

Corresponding Author:
Kerina H Jones, BSc, PhD
Population Data Science
Swansea University Medical School
Swansea University
Singleton Park
Swansea, SA28PP
United Kingdom
Phone: 44 1792602764
Email: k.h.jones@swansea.ac.uk

Abstract

Background: The literature abounds with increasing numbers of research studies using genomic data in combination with health data (eg, health records and phenotypic and lifestyle data), with great potential for large-scale research and precision medicine. However, concerns have been raised about social acceptability and risks posed for individuals and their kin. Although there has been public engagement on various aspects of this topic, there is a lack of information about public views on data access models.

Objective: This study aimed to address the lack of information on the social acceptability of access models for reusing genomic data collected for research in conjunction with health data. Models considered were open web-based access, released externally to researchers, and access within a data safe haven.

Methods: Views were ascertained using a series of 8 public workshops (N=116). The workshops included an explanation of benefits and risks in using genomic data with health data, a facilitated discussion, and an exit questionnaire. The resulting quantitative data were analyzed using descriptive and inferential statistics, and the qualitative data were analyzed for emerging themes.

Results: Respondents placed a high value on the reuse of genomic data but raised concerns including data misuse, information governance, and discrimination. They showed a preference for giving consent and use of data within a safe haven over external release or open access. Perceived risks with open access included data being used by unscrupulous parties, with external release included data security, and with safe havens included the need for robust safeguards.

Conclusions: This is the first known study exploring public views of access models for reusing anonymized genomic and health data in research. It indicated that people are generally amenable but prefer data safe havens because of perceived sensitivities. We recommend that public views be incorporated into guidance on models for the reuse of genomic and health data.

(J Med Internet Res 2019;21(8):e14384) doi:10.2196/14384

KEYWORDS

human genome; genetic databases; public opinion; data storage; data linkage

Introduction

Background

We are witnessing a rapid expansion in the availability and use of genomic data to inform research and clinical care. The trajectory is for this trend to increase with the falling cost of whole-genome sequencing, rapid technical advances, and rise of data management facilities [1]. Work with genomic data spans the whole spectrum, from large-scale research using many thousands of records to precision medicine at the individual level. To date, the majority of genomic data have been used in large-scale research such as genome-wide association studies (GWAS) and exome-wide association studies (EWAS), among others. GWAS and EWAS are essential observational studies comparing DNA sequences and exploring variants that may be associated with a phenotypic trait. Precision medicine, where treatments and medication regimes are informed by individual

XSL•FO
RenderX

genetic status, seeks to use the findings of observational studies to inform tailored clinical care. There are strong drivers for precision medicine, such as reducing the use of poorly effective interventions, thereby leading to better patient outcomes and saving costs. The UK Chief Medical Officer has expressed her dream for genomic medicine to become commonplace: "I want the NHS across the whole breadth to be offering genomic medicine - that means diagnosis of our genes - to patients where they can possibly benefit" [2]. The Department of Health program seeking to sequence 100,000 genomes has achieved this target, thereby creating a valuable data resource toward this goal [3]. The US National Institute of Health has initiated a program called All of Us, aiming to gather data from at least 1 million citizens to accelerate the introduction of precision medicine into all areas of health and health care on a large scale [4]. The Canadian government has made a major investment to advance cutting-edge developments in genomics research [5]. Successes in precision medicine are growing, but there is much further work to be done to gain the advantages and for it to become mainstream.

Between the ends of the spectrum, large scale studies using genomic data and precision medicine using individual genomic information, there is a vast range of work and many permutations of research to enable meaningful findings to be taken forward. With the exception of certain single-gene conditions such as cystic fibrosis and sickle cell anemia, the relationship between genotype and phenotype is highly complex, involving multiple genes and expression profiles. Research using genomic data in conjunction with health data (defined here as health records, phenotypic data, and lifestyle data) in condition-specific cohorts or population-level studies plays a unique role. In this type of research, it is often genomic derivatives (eg, variants and risk scores) rather than sequence data that are used with health data. Such studies allow the relationships between genomic markers, lifestyle factors, and phenotypes of interest to be explored. However, for this research to take place, genomic data collected for research studies need to be available for reuse. To be most effective, the data need to be linkable at the individual level so that genomic and wider factors can be taken into account.

There are some prevalent concerns about social acceptability and the risks posed for individuals and their kin in the use of genomic data for research. These include possible discrimination in relation to insurance coverage and employment opportunities for people with genetic conditions or high perceived risks of developing a disease [6]. However, public views on the use of genomic data have been found to be variable. The Global Alliance for Genomics and Health (GA4GH) has been running a public survey extending across numerous countries seeking views on the use of genomic data [7]. Among over 10,000 respondents, 52% felt information from DNA was different from other medical data, with 48% unsure or making no distinction. In relation to web-based data, respondents considered their banking data as needing most protection, above medical and genomic information [8]. However, although there is public engagement on various aspects of this topic, there is a lack of information about public views on models of data access. From a review on published studies using genomic data (to be published separately), we categorized 3 main data access models: open access where data are publicly available on a website, curated data released to specified researchers, and data accessed by specified researchers within a safe haven. We define a data safe haven as a secure virtual environment within which data are managed and analyzed and from which anonymized results can be released [9].

Study Aim

To date, the majority of extant genomic data have been collected for research studies; the reuse of these datasets in deidentified form formed the focus of our study. The main aim of this study was to address the lack of information on the social acceptability of access models for reusing genomic data collected for research in conjunction with health data. This included public views on the use of the data with informed consent, without informed consent, and without consent but with notification for each of the 3 models. The reuse of any genomic data collected for clinical care and incorporated into the electronic health record (EHR) is out of scope for this study. It is assumed that the sharing of EHR data is subject to health provider mechanisms and agreements with recipient parties in line with relevant jurisdictional legislation.

Methods

Study Design

The study used a mixed methods approach, collecting and integrating quantitative and qualitative data. Public views were ascertained using a series of 8 workshops held between February and November 2018. Groups were selected purposively on the basis of having an interest in health care, health research, biological science, or data linkage research. The reasoning for this was 2-fold: we wanted to gain the views of groups with an interest in at least one of these areas as the study was breaking new ground; and it meant that we could tailor and reduce the explanation of concepts to meet the needs of the audience in the time available for the workshops. The groups comprised pupils at sixth-form college (in Neath Port Talbot); students at a further education college (in Pembrokeshire); university staff and students; a business professionals group; a general public consumer panel; science festival attendees, a grand round of health professionals; and University of Third Age members. All workshops were held in Wales, and where the location is not mentioned, the meetings took place on Swansea University premises. As the workshops took place at preexisting meetings with no registration process, it was not possible to control the numbers attending or influence participant selection.

Ethical approval for research with public participants was obtained from the Swansea University Medical School Research Ethics and Governance Committee. We note that in working with the public, we generally referred to genetic data rather than genomic, as genetics was a more familiar term to the participants. We did, however, explain the terms and the difference between them.

The workshops were led by KHJ and HD and ran for 1 hour. Notes of the discussions were taken (by HD and KHJ) and compared for consensus. The presentation of the study and

XSL•FO
RenderX

research examples was given by KHJ, who also initiated the discussions. The information was presented in a deliberately neutral way not to preempt or influence viewpoints. The information was presented consistently across all workshops with each following the same format.

Initial Discussions on Public Knowledge

This involved asking the audience about the latest news item they had heard about genetics, their awareness of genetic data being used in research, and how they believe the data are used.

Introduction to the Study

This included describing the study purpose; the focus on the reuse in deidentified form of genetic data collected for research, as distinct from the process of data collection for clinical purposes; types of genetic data and wider health data; and how data can be accessed in terms of the 3 main models (open access, released externally, or within a data safe haven). We included a brief explanation of genetic data sequence and derivatives such as traits, variants, and risk scores.

Examples of organizations operating differing access were genetic and health data made openly available in the Personal Genome Project [10], UK Biobank collates and releases anonymized linked genetic and health data to approved researchers for specific studies [11], and the Sax Institute provides access to anonymized genetic and health data to approved researchers for specific studies within a data safe haven [12]. It was explained that, although data may be used in anonymous form, identifiable data are needed to process the primary research data for reuse and enable linkage to health data. Through these examples, we provided practical descriptions of each operating model, and how the ethical and other regulatory permissions needed for researchers to access the data can vary.

Participants were provided with examples of research studies that have used genetic data, particularly with health data, and arising from the commercial and noncommercial sectors. Studies included large-scale work to identify variants of interest, considering lifestyle factors in relation to the BRCA mutations, and medication monitoring based on genetics. We also included an introduction to direct-to-consumer genetic testing companies, such as Patients Like Me [13] and 23andMe [14], which provide sequencing services and personal genetic information to individuals, then retain and use the resulting data for research.

Discussions on Public Views

At this stage in the workshop, an open discussion was encouraged by asking the audience how they felt about these kinds of research taking place and what could be done to address any concerns they have.

Exit Questionnaire

Participants were asked to complete a questionnaire at the end of the workshop (Multimedia Appendix 1). This asked about knowledge and views on the use of genetic data and specifically asked about the relative acceptability of the 3 models of access. The questionnaire data were collected in anonymized form, and it gave the participants the opportunity to provide their views individually and privately.

Quantitative responses to the questionnaire were analyzed in IBM SPSS (version 22). Descriptive statistics were used to characterize the respondents, and for frequencies, the chi-square test was used to assess independence between categories, and the two-proportion z-test was used to compare proportions with Bonferroni correction where appropriate [15]. Free-text qualitative responses from the questionnaire were analyzed thematically by manual assessment and comparison between members of the research team (HD and KHJ) for consensus on theme identification and data convergence. A similar thematic analysis was conducted on the topics arising in the open discussions. Most analyses were based on the questionnaire responses, with the information from the open discussions being used more generally for context setting.

Results

Overview

The initial discussions (step 1) on public knowledge raised topics from news stories such as *3-parent* babies and designer babies, heritable genetic conditions such as Huntingdon disease, the potential for cancer research, gene editing, and invasive testing of embryos for genetic problems. Participants perceived the great value and opportunities becoming available with the increasing use of genetic data. These benefits were reiterated in the open discussion (step 3), but participants also drew attention to various concerns about privacy and confidentiality depending on data use and parties concerned. These included statements such as, “I would be worried about being discriminated against (insurance, work etc.),” “I am fearful of data in the hands of commercial companies,” and “I’m concerned about legislative and attitudinal ‘creep’—what we enforce now will, no doubt creep over time.” In terms of what could be done to address their concerns, participants highlighted the need for robust governance, data anonymization, data security, and greater transparency in data use.

Characterizing the Respondents

Information about the attendees was collected in part A of the questionnaire shown in Multimedia Appendix 1. There were 116 respondents in total: 54 men and 62 women. The denominator in all percentage values is 116 unless otherwise stated. The age profile was as follows: 16 to 25 years: 18.9% (22/116); 26 to 35 years: 31.8% (37/116); 36 to 45 years: 15.5% (18/116); 46 to 55 years: 9.4% (11/116); 56 to 65 years: 6.8% (8/116); and older than 65 years: 17.2% (20/116). As the data were collected in age bands, mean age and standard deviation are not known. This profile was compared with the 2011 UK census figures [16] to gauge representativeness of the sample. The age bands are slightly different in the census but are close enough to provide an indicative measure. From this, we observed that our sample was overrepresented in the 26 to 35 year age band. Among the respondents, 50 (43.1%, 50/116) had children, and the remainder did not. Just more than half (n=58, 52%) held a degree with almost one-third (n=32, 33%) holding a degree in a biological subject. This higher rate than in the general population was to be expected because of the nature of the groups. By comparison, approximately 40% of the UK population are graduates [17].

XSL•FO
RenderX

Participants were asked to gauge their own background knowledge of genetics in 5 categories from none to very good. This was a matter of personal perception of one's own ability, and as a general summary, the frequencies were no knowledge: 6.8% (8/116); a little: 34.4% (40/116); middling: 36.2% (42/116); good: 14.6% (17/116); and very good: 6.8% (8/116). This was examined further by taking into account the highest education level of the respondents in a biological subject and separately in any subject. All the qualifications are UK based, apart from degrees and professional qualifications, which are more generic categories: General Certificate of Secondary Education (GCSE) examinations are taken at age 16 years, Advanced Level (A level) examinations are taken at 18 years, Higher National Certificate/Diploma qualifications are often taken in post-16 Further Education colleges, and National Vocational Qualifications are work-based examinations. When all the categories of education level in Q5a or Q5b were included, there was no association between attainment level and self-reported knowledge. This may have been because of the variability between the categories of qualification. However, when the categories were restricted to those known to be hierarchical (GCSE, A level, and degree), attainment in a biological subject was found to be associated with higher self-reported knowledge (chi-squared *P*=.02). The relationship between self-reported knowledge and attainment across all subjects remained insignificant.

Respondents' Views on Data Use

The majority (78/116, 67.2%) of questionnaire respondents placed a high or very high value on using genetic data for research, with only 8 people (7/116, 6.8%) considering the value to be low or very low. No specific definition of value was presented to allow participants to make their own interpretations. Placing higher value on the use of genetic data was associated with higher attainment in a hierarchical biological subject (chi-square *P*=.005) and all subjects (chi-square *P*=.02).

There was no association between levels of value of using genetic data and levels of concern across all respondents. The most frequent level of concern about the use of genetic data for research was moderate, with 53/116 people (48%) indicating this response, 38/116 people (34%) showing low or very low concern, and 20/116 people (19%) having high or very high concern. The level of concern was not associated with educational attainment in either a biological subject or other subjects. It was also not associated with age or workshop attended. It was, however, associated with gender, with women showing higher levels of concern than men (chi-squared *P*=.01). A variety of themes emerged from the free-text responses in relation to causes for concern (Table 1), with the most frequent being misuse of data tied with concerns about information governance. Respondents noted that their concerns would be allayed if they could be assured of appropriate data use, data security, and proper information governance to avoid the data falling into the hands of agents who might discriminate against them.

Table 1. Concerns about the use of genetic data for research (N=116).

Type of concern	Respondents, n (%)	Percentage of total concerns
Data misuse	30 (25.9)	23
Information governance	30 (25.9)	23
Purpose of use	23 (19.8)	18
Discrimination	22 (19.0)	17
Security	16 (13.8)	13
Disclosure risk	7 (6.0)	6

Respondents' concerns about the use of genetic data for research are given in themed areas. Each participant was asked to provide up to 3 concerns, which were then grouped into themes.

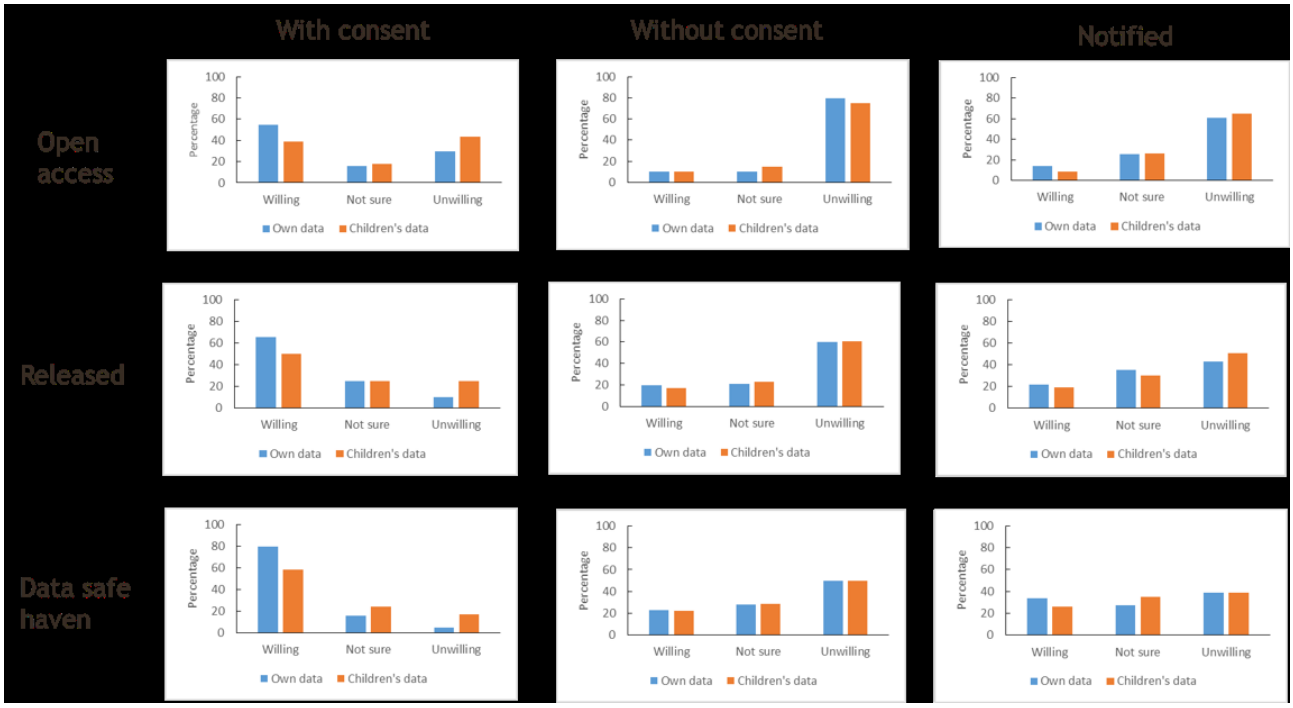
The questionnaire asked participants to provide their views on access to their genetic data in each of the 3 main access models: openly accessible on a website, released to researchers, and access within a data safe haven. For each option, they were asked to indicate whether they were willing, not sure, or unwilling in relation to whether they were consented, not consented, or not consented but notified. The results of analyzing these categorical data are shown in Figure 1. Participants showed a preference for informed consent over the other options and a preference for data use within a safe haven compared with other access models. A similar, but more cautious, pattern was observed for the use of data pertaining to participants' children. Interestingly, less than 5% of respondents stated they would be unwilling for their genetic data to be reused in a data safe haven if they were asked for consent to do so.

For each model, attendees indicated their willingness for their data to be used with consent, without being consented, and notified.

The significance of the apparent differences was tested by comparing the proportions of respondents willing for their data to be reused across consent options within each data access model (ie, intramodel seen horizontally in Figure 1). Similarly, we compared the proportions willing between models of data access (ie, intermodel seen vertically in Figure 1). We repeated these 2 analyses for willingness for children's data reuse. We also compared proportions willing for the use of their own and their children's data by access model for a given consent status.

XSL•FO
RenderX

Figure 1. Participant views on access to their and their children's genetic data in each of the three main access models.



Intramodel Comparison of Data Access Preferences

We found that there was a significant difference between willingness for the reuse of respondents' genetic data with versus without consent and with consent versus notified within all 3 access models. This was the case with and without the Bonferroni correction. There was, however, no significant difference in the proportion of respondents willing for their data to be reused between the options of without consent and notified. The same pattern was observed in their willingness for their children's data to be reused (Table 2). This indicates that the respondents' preferences are to be consented to the reuse of genomic data collected for research. It was not considered appropriate for reuse of genomic research data to proceed without consent, nor sufficient to be merely notified.

Table 2. Intramodel comparison: with consent, without consent, and notified.

Consent status—*access model*	With consent: without consent		With consent: notified		Without consent: notified	
	P value[a]	CI	*P* value[a]	CI	*P* value[a]	CI
Open: own	≤.001	0.33 to 0.56	≤.001	0.29 to 0.52	.40	−0.13 to 0.05
Open: children	≤.001	0.17 to 0.41	≤.001	0.18 to 0.42	.81	−0.10 to 0.08
Release: own	≤.001	0.34 to 0.58	≤.001	0.31 to 0.56	.72	−0.09 to 0.13
Release: children	≤.001	0.27 to 0.54	≤.001	0.25 to 0.52	.70	−0.10 to 0.14
DSH[b]: own	≤.001	0.46 to 0.68	≤.001	0.34 to 0.58	.09	−0.02 to 0.24
DSH: children	≤.001	0.22 to 0.51	≤.001	0.18 to 0.47	.57	−0.01 to 0.17

[a]*P* values above .05 were not considered significant.

[b]DSH: data safe haven.

This table displays comparisons of respondent preferences between consent options within each of the 3 models of data access: openly accessible, released externally to researchers, and accesses within a safe haven. Preferences in relation to their own and children's data are shown. The significance level is 95%. The CIs are the interval lower and upper limits for the difference in proportions using the two-proportion z-test. Frequency denominators varied between questions with none less than 100 (of 116 participants) and were used accordingly. A Bonferroni correction was applied (for 3 tests in each set, thus requiring a *P* value <.017 to remain significant at the 95% level), but in this case, it did not affect the results.

Intermodel Comparison of Data Access Preferences

We found no significant differences in respondents' willingness for their genomic data to be reused when the open access model was compared with the release model. The results were similarly not significant between the release model and access within a data safe haven, with the Bonferroni correction. The same pattern was seen for the reuse of children's data. There were, however, some significant differences between the open access and data safe haven models, corresponding with the greater degree of respondent concern about data being openly accessible compared with accessed within a safe haven (Table 3).

XSL•FO
RenderX

Table 3. Intermodel comparison: open access, release, and within data safe haven.

Access model—consent status	Open: release		Open: DSH[a]		Release: DSH	
	P value[b]	CI	*P* value[b]	CI	*P* value[b]	CI
With consent: own	.11	−0.24 to 0.02	<.001	0.13 to 0.37	.02[c]	0.02 to 0.26
With consent: children	.14	−0.04 to 0.26	.01[c]	0.05 to 0.34	.27	−0.07 to 0.24
Without consent: own	.07	−0.01 to 0.19	.02[c]	0.02 to 0.23	.60	−0.08 to 0.15
Without consent: children	.21	−0.04 to 0.17	.04[c]	0.01 to 0.24	.39	−0.07 to 0.18
Notified: own	.17	−0.03 to 0.18	.001	0.08 to 0.31	.06	0.00 to 0.24
Notified: children	.06	0.00 to 0.21	.004	0.06 to 0.29	.30	−0.06 to 0.20

[a]DSH: data safe haven.

[b]*P* values above .05 were not considered significant.

[c]Not significant when Bonferroni correction applied.

This table displays comparisons of respondent preferences between access models for each of the 3 consent options: with consent, without consent, and without consent but notified. Preferences in relation to their own and children's data are shown. The conditions for the two-proportion z-test and Bonferroni correction were as for Table 2. Some results were no longer significant when adjusting for multiple testing.

Own Versus Childrens' Data Reuse Comparison

The comparisons showed no significant differences in the reuse of respondents' own against children's data for any model when the options were without consent or without consent but notified. As can be seen from Figure 1, the proportions of respondents willing for their, or their children's, data to be reused under these options were low. There were some differences in willingness for the reuse of respondents' own against children's data when the option was with consent. However, the only 1 remaining significant after Bonferroni correction was reuse within a data safe haven. This was the model favored by the respondents with a greater degree of caution seen in relation to the reuse of children's data compared with their own (Table 4).

Table 4. Comparison between the use of own and children's data by access model for a given consent status.

Consent status—*access model*	With consent—own: children		Without consent—own: children		Notified—own: children	
	P value[a]	CI	*P* value[a]	CI	*P* value[a]	CI
Open	.027[b]	0.02 to 0.30	.94	−0.09 to 0.09	.27	−0.04 to 0.15
Release	.033[b]	0.01 to 0.29	.62	−0.09 to 0.14	.66	−0.09 to 0.15
DSH[c]	.002	0.08 to 0.34	.92	−0.12 to 0.13	.27	−0.06 to 0.21

[a]*P* values above .05 were not considered significant.

[b]Not significant when Bonferroni correction applied.

[c]DSH: data safe haven.

Respondent preferences between the use of their own and children's data by access models and consent option are shown. The conditions for the two-proportion z-test and Bonferroni correction were as for Table 2. Most results were not significant, and this was further reduced when the Bonferroni correction was applied.

Respondents' Reasons for Viewpoints

Participants elaborated on reasons for their views in relation to each model. Their free-text viewpoints provided context to the numerical data. Their main concerns included data security, protection of identity, the right to make informed choices, control over data use, who might access the data under the various models, and concern about unknown future developments. Example responses for each of the access models (openly accessible, data released externally to researchers, and data accessed within a data safe haven) from a variety of participants are given below. Information about each respondent has been included for context.

In relation to data being openly accessible, viewpoints included the following:

> *Potential for re-identification by unscrupulous people/organisations.* [female, aged 16-25 years, no children, A level in a biological subject, professional qualification in another subject, data use: high value and moderate concern]
>
> *[I am] sufficiently unhappy with the idea that it would be an active deterrent from having children at all/going abroad to have them.* [male, aged 26-35 years, no children, degree in a biological subject, data use: moderate value and very high concern]

XSL•FO
RenderX

> *It's not my place to give info of my children to strangers for any reason.* [male, aged 26-35 years, has children, degree in a biological subject, data use: very high value and low concern]
>
> *...want to help with research...my DNA can't be used to re-identify me yet.* [female, aged 16-25 years, no children, GCSE in a biological subject, data use: high value, low concern]

In relation to data released externally, viewpoints included the following:

> *I don't know enough about it [but] I'd like to help researchers find cures and things.* [female, aged 16-25 years, no children, GCSE in a biological subject, data use: moderate value, moderate concern]
>
> *My primary concern is the security of data if it is sent out to researchers.* [male, aged >65 years, has children, degree in a biological subject, data use: very high value, low concern]
>
> *If going to researchers, then it is less likely to be abused by others online etc. who are not researchers.* [male, aged 46-55 years, has children, professional qualification in a biological subject, data use: moderate value and low concern]
>
> *Not sure if people could be identified* [female, aged 26-35 years, no children, degree in a biological subject, data use: very high value and moderate concern]

In relation to data accessed within a safe haven, viewpoints included the following:

> *[Data used] to what end is my main concern. But overall happier in a safer environment.* [male, aged 26-35 years, no children, A level in a biological subject, degree in another subject, data use: very high value and low concern]
>
> *I fully expect this to be used—missing a trick otherwise.* [female, aged 36-45 years, has children, degree in a biological subject, data use: high value and moderate concern]
>
> *I would want to be reassured of safeguards.* [female, aged 26-35 years, has children, A level in a biological subject, degree in another subject, data use: very high value and moderate concern]
>
> *Happy for my data to be used as I'm and adult and I'm told. Not for my children. They need to be able to make that decision.* [male, aged 26-35 years, has children, GCSE in a biological subject, degree in another subject, data use: very high value and low concern]

The viewpoints reflected the workshop discussions being premised on the use of genetic data with wider health data in line with the focus of our study, but we included a question (in the questionnaire) to ask specifically about views on genetic data use linked with health record data to clarify any additional views. Many of the participants' viewpoints expressed were the same or similar to their previous responses; however, some expressed stronger concerns, and none were less concerned. Where additional views were given, they further polarized the preference for consent and data reuse in a safe haven.

In relation to linked genetic and wider health data being openly accessible in anonymized form, participant views included the following:

> *Scholars need these websites to check their daily work*
>
> *Fine as long as there is consent*
>
> *I would need to see what the data looks like. There are too many concerns for me to agree to this*
>
> *I feel it's not safe and I don't want the discrimination*

For the release of linked anonymized genetic and wider health data, responses included the following:

> *I do not mind as long as it is anonymous*
>
> *Acceptable as long as there is consent and data is held securely*
>
> *Dependent on the research question and access limitations*
>
> *Unsure—would depend on safeguards imposed*

For accessing linked anonymized genetic and wider health data in a safe haven, viewpoints included the following:

> *Most comfortable with this*
>
> *Safe, secure, governance, auditable—okay*
>
> *Data being used by verified researchers for public benefit is a good thing. Having the data kept safe and controlled is a must*
>
> *I am happy for this to happen provided it is safe, not sold etc*

Discussion

Principal Findings

This study has begun to address the lack of knowledge on the social acceptability of access models for reusing genomic data collected for research in conjunction with wider health data. This is the first known study to address this topic. As noted earlier, it does not relate to the reuse of genomic data collected for clinical care, incorporated into the health record. Our findings indicate that most public participants place a high to very high value on the reuse of genomic data. This viewpoint was associated with higher educational attainment in a biological subject but was also present across all educational subjects. Their areas of concern included data misuse, how data are governed, disclosure risk, possible discrimination, and the purpose of data use and were in accord with the large-scale survey conducted by the GA4GH [7,8]. Levels of concern were not, however, associated with educational attainment but were associated with gender with women showing greater concern than men. This might be due, in part, to the fact that women show greater levels of anxiety than men in the general population [18].

A comparison of response frequencies indicated that participants preferred to give informed consent for the reuse of their own or children's genomic data that had been collected for research, over being notified or not consented. Tests of statistical

XSL•FO
RenderX

significance between without consent and without consent but notified suggest the respondents saw little difference between these 2 options and bolstered the preference for consent. Participants favored data use within a safe haven compared with the release and open access models in terms of response frequencies. However, this difference was only significant between the open access and safe haven models of data access. Free-text responses provided information on reasons for preferences. These included the desire to support research but strong views against open access, which go some way to explaining the observed results.

Although there was little statistical significance, participants expressed more caution in relation to children's data than their own in terms of the response frequencies on access model and consent options and their elaborations on reasons for their choices. Respondents felt it was important that children are able to make their own choices in relation to the reuse of their genomic data. In practice, this depends on the age of valid consent but supports the reconsenting of young people once that age is attained. However, we also acknowledge that this might not always be practicable and propose that participants coming of age should always be taken into account at the outset so that it can be accommodated in the research plans.

We have limited this study to public views on the reuse of genomic data collected for research because this is the main source of extant genomic data and because the sharing of data contained in the health record is subject to health provider data governance, including repurposing in line with jurisdictional legislation. There are numerous long-established enterprises across the world that reuse population health data in anonymized form for research, using datasets drawn from existing health records rather than engaging in primary data collection [19]. We recognize this as distinct from data collected for research and are not making any recommendations in relation to the work of these enterprises other than to state that an insistence on additional consent for the reuse of health record data would be impractical and highly detrimental to such research [20].

The use of genomic data through web-based open access is current practice in large-scale GWAS and EWAS because of the compute capacity required in processing the data files and the need to access data across multiple sources [21]. GWAS and EWAS often use genomic data without associated wider health data because they are concerned with profiling variants of possible significance [22]. We acknowledge the indispensable value of such studies, such as for sequence alignment and variant calling. The GA4GH has proposed a system of registered access for web-based use of genomic data with health data to facilitate the reuse of data within the bounds of consent restrictions and other ethical obligations [23]. We welcome this as an improvement, as our findings with little favor for sharing genomic data through open web-based access suggest that open access would not be socially acceptable as an extensible model for research linking genomic to wider health data. Release to specified researchers or especially access within a data safe haven was preferred and also reflects the general trend in working with linked health data [9].

Limitations

We recognize the limitations of our study. The sample size was relatively small but included a range of ages, interest areas, and backgrounds. In the interests of privacy, we did not collect the full demographic details of our participants, and having grouped age into 10-year bands, we could not drill down further on this variable. The sample size also meant that we were limited in our options for stratified analyses. The workshops took place across South Wales, and although we do not know if public opinion on the reuse of genomic data with other health data would be significantly different in other parts of the United Kingdom or wider world, we acknowledge location as a possible limitation. As noted, we also acknowledge that a greater number of our participants were educated to degree level than the general population. However, further work could be undertaken to expand the study and address this possible source of bias in the opinions expressed. We have not included a consideration of legal and ethical requirements and how they vary, as we are preparing a further publication to do these issues justice and keep the focus of this paper on public views.

Recommendations for Future Work

The reuse of genomic data with health data is an expanding area and one that needs much further work with the public and other stakeholders. Although our participants expressed a preference for consent, there are questions around the purpose and nature of the consent. If the data are to be reused in anonymized form, then strictly speaking, consent is not required for that reuse. But identifiable data are necessary to create anonymized data, thus calling for consent for data processing into anonymized form. This would need to be made clear on the participant information sheet and in the consent process.

We recommend that consent for the reuse of research data be incorporated into the consent form at data collection to avoid subsequent difficulties with reuse. The lead author has proposed this to the UK Health Research Authority for all primary research using personal data, not limited to genomic data. It is being taken forward as advice to be given to researchers by ethics committees and institutional review boards as part of the UK integrated research application system. Example wording for the participant information sheet and consent form is given in Multimedia Appendix 2. This is being piloted and has begun to see acceptance by research ethics committees [24]. This simple idea has the potential to revolutionize data reuse by avoiding the lack of appropriate consent. However, we also acknowledge that consent is not the ultimate panacea [25], particularly with the degree of unknownness in genomic data. Further work on consent for reuse needs to ensure the public properly understand this characteristic of genomic data.

Recommendation 1: The inclusion of consent to use personal data for deidentification so data can be reused for research should be incorporated into consent forms and participant information sheets for studies collecting primary data.

Biobanks commonly use a broad consent model where participants agree to a framework for future research of certain types. In the past, this involved the use of biological samples, but in this genomic era, it also involves the use of data generated

XSL•FO
RenderX

from the samples, thus raising greater privacy concerns for individuals and their kin [6]. Dynamic consent usually involves recontacting individuals to ask consent for particular uses [26]. Although there seem to be pros and cons with each, it would be difficult to enact a meaningful dynamic consent model for the reuse of data in anonymized form because the whole point of reusing data in this form is to protect (and not know) identity. But this too needs further exploration with the public and other stakeholders.

Recommendation 2: Public engagement should be conducted on a range of consent models to gain viewpoints on the acceptability of models for the reuse of genomic and health data, taking into account ethical and legal issues, and practicalities such as research utility and computing constraints.

The GA4GH proposal for registered access to web-based genomic and health data is a novel development and one requiring an assessment of public views to gauge its relative acceptability. This could follow a similar model as we have used here and could take into account factors such as types of research, parties accessing the data, requirements and constraints on data processing, compute capacity, and the pros and cons of other access models.

Recommendation 3: Public engagement should be conducted on the acceptability of registered access to web-based genomic and health data, in comparison with other access options and conditions.

As the use of genomic data is still relatively new, the level of public awareness needs to be enhanced to enable people to make informed choices. This should include a balanced explanation of the known perils and promise of genetic research and precision medicine and should be conducted transparently in a 2-way dialog, acknowledging that there are unknowns. This is also true for many health professionals, hence the rise in Genomic Medicine education. We also propose that education on this subject begins early by being properly incorporated into curricula for secondary school pupils (aged 11-18 years) to enable current and future generations to make informed choices.

Recommendation 4: Greater efforts are needed to raise awareness, engage in public dialog, and improve education about the perils and promise of genetic data research and precision medicine.

Genomic data are not singled out from other health data in data protection legislation (at least in the European Union) [27], and it is important not to bias public opinion and stifle research by exceptionalizing the risks in reusing genomic data [28]. Nonetheless, it is imperative that robust safeguards are in place so that genetic privacy and confidentiality (including in relation to kin) are secured. We propose that there is a need for a risk-based model incorporating public views into a flexible suite of controls to protect identities and maximize data utility.

Recommendation 5: Public views should be incorporated into the development of a risk-based, flexible suite of controls for accessing genomic and health data for research.

Conclusions

There is undoubtedly great potential in the use of genomic and health data for large-scale research and precision medicine. However, concerns about social acceptability and the risks posed for individuals and their kin need to be addressed. Although there is much public engagement on health data sharing in general and on various aspects of genomic data reuse, there has been a lack of information about public views on models for accessing combined genomic and health data. To date, most extant genomic data have been collected for research studies, and these datasets formed the focus of our study. This is the first known study to explore public views of access models for reusing anonymized genomic and health data in research. It indicated that people are generally amenable but prefer data safe havens over external release to specified researchers and over open access because of perceived sensitivities. We recommend further public engagement, and that public views be incorporated into guidance on models for the reuse of genomic and health data.

Acknowledgments

This study was funded by the UK Medical Research Council: MC_PC_16035. The authors acknowledge the advice of the Wales Genomic Medicine Centre and the support of the Wales Gene Park in sourcing some of the public groups. They acknowledge that they are associated with the SAIL Databank, which is a data safe haven. Every effort was made to present the access models objectively and not to influence participants' opinions.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Copy of questionnaire for Genetic Data Integration (GeDI) workshop participants.
[PDF File (Adobe PDF File), 72KB - jmir_v21i8e14384_app1.pdf]

Multimedia Appendix 2
Example wording for inclusion in the Participant Information Sheet and Consent Form for the reuse of data collected for research.
[PDF File (Adobe PDF File), 32KB - jmir_v21i8e14384_app2.pdf]

XSL•FO
RenderX

References

1. National Human Genome Research Institute. 2018. The Cost of Sequencing a Human Genome URL: https://www.genome.gov/sequencingcosts/
2. GOV.UK. 2016. Chief Medical Officer Annual Report 2016: Generation Genome URL: https://www.gov.uk/government/publications/chief-medical-officer-annual-report-2016-generation-genome
3. Genomics England. 2019. The 100,000 Genomes Project URL: https://www.genomicsengland.co.uk/about-genomics-england/the-100000-genomes-project/
4. National Institutes of Health (NIH) — All of Us. 2019. URL: https://allofus.nih.gov/
5. Genome Canada. 2019. URL: https://www.genomecanada.ca/en
6. Naveed M, Ayday E, Clayton EW, Fellay J, Gunter CA, Hubaux J, et al. Privacy in the genomic era. ACM Comput Surv 2015 Sep;48(1):pii: 6 [FREE Full text] [doi: 10.1145/2767007] [Medline: 26640318]
7. Your DNA, Your Say. 2019. URL: https://surveys.genomethics.org/survey/yourdnayoursay/en
8. Your DNA Your Say. 2019. Feedback URL: https://societyandethicsresearch.wellcomegenomecampus.org/sites/default/files/media/item/your-dna-your-say-at-the-global-alliance-for-genomics-and-health-plenary/files/20171017-orlando-florida-usa-ga4gh-plenary-video-of-slides.mp4
9. Lea NC, Nicholls J, Dobbs C, Sethi N, Cunningham J, Ainsworth J, et al. Data safe havens and trust: toward a common understanding of trusted research platforms for governing secure and ethical health research. JMIR Med Inform 2016 Jun 21;4(2):e22 [FREE Full text] [doi: 10.2196/medinform.5571] [Medline: 27329087]
10. The Personal Genome Project. 2019. URL: https://www.personalgenomes.org/gb
11. Medical Research Council. 2019. UK Biobank URL: https://mrc.ukri.org/research/facilities-and-resources-for-researchers/biobank/
12. Sax Institute. 2019. 45 and Up Study URL: https://www.saxinstitute.org.au/our-work/45-up-study/
13. Patients Like Me. 2019. URL: https://www.patientslikeme.com/
14. 23 and Me: DNA Genetic Testing & Analysis. 2019. URL: https://www.23andme.com/en-gb/
15. Henry R. Etymologia: Bonferroni correction. Emerg Infect Dis 2015 Feb;21(2):289 [FREE Full text] [doi: 10.3201/eid2102.et2102] [Medline: 25786274]
16. Office for National Statistics. 2011. 2011 Census: Population and household estimates for the United Kingdom URL: http://www.ons.gov.uk/ons/rel/census/2011-census/population-and-household-estimates-for-the-united-kingdom/rft-table-3-census-2011.xls
17. Office for National Statistics. 2017. Graduates in the UK Labour Market URL: https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/employmentandemployeetypes/articles/graduatesintheuklabourmarket/2017
18. Crawford JR, Henry JD, Crombie C, Taylor EP. Normative data for the HADS from a large non-clinical sample. Br J Clin Psychol 2001 Nov;40(Pt 4):429-434. [doi: 10.1348/014466501163904] [Medline: 11760618]
19. Jones KH, Ford DV. Population data science: advancing the safe use of population data for public benefit. Epidemiol Health 2018;40:e2018061 [FREE Full text] [doi: 10.4178/epih.e2018061] [Medline: 30703857]
20. Stanley F. Privacy or public good? Why not obtaining consent may be best practice. Significance 2010 May 18;7(2):72-75. [doi: 10.1111/j.1740-9713.2010.00423.x]
21. Front Line Genomics. 2018. Genomic Data 101: An Introduction To Genomic Data URL: http://www.frontlinegenomics.com/wp-content/uploads/3090_FLG_101_Feb_2017_GenomicData_FullMagazine_Mar17-1.pdf
22. GeneCards. 2019. URL: https://www.genecards.org/
23. Dyke SO, Linden M, Lappalainen I, de Argila JR, Carey K, Lloyd D, et al. Registered access: authorizing data access. Eur J Hum Genet 2018 Dec;26(12):1721-1731 [FREE Full text] [doi: 10.1038/s41431-018-0219-y] [Medline: 30069064]
24. Wales Gene Park. 2019. Contact Us URL: http://www.walesgenepark.cardiff.ac.uk/contact-us/
25. Laurie G, Ainsworth J, Cunningham J, Dobbs C, Jones KH, Kalra D, et al. On moving targets and magic bullets: can the UK lead the way with responsible data linkage for health research? Int J Med Inform 2015 Nov;84(11):933-940 [FREE Full text] [doi: 10.1016/j.ijmedinf.2015.08.011] [Medline: 26342668]
26. Steinsbekk KS, Myskja BK, Solberg B. Broad consent versus dynamic consent in biobank research: is passive participation an ethical problem? Eur J Hum Genet 2013 Sep;21(9):897-902 [FREE Full text] [doi: 10.1038/ejhg.2012.282] [Medline: 23299918]
27. General Data Protection Regulation (GDPR). 2016. URL: https://gdpr-info.eu/
28. Chin JJ, Campbell AV. What - if anything - is special about 'genetic privacy'? In: Kaan TS, Ho CW, editors. Genetic Privacy: An Evaluation of the Ethical and Legal Landscape. London: Imperial College Press; 2013:223-252.

Abbreviations

A level: Advanced Level
EHR: electronic health record
EWAS: exome-wide association studies
GA4GH: Global Alliance for Genomics and Health

XSL•FO
RenderX

GCSE: General Certificate of Secondary Education
GWAS: genome-wide association studies

Edited by G Eysenbach; submitted 14.04.19; peer-reviewed by A Jimenez, J Oldenburg; comments to author 25.06.19; revised version received 05.07.19; accepted 07.07.19; published 21.08.19.

Please cite as:
Jones KH, Daniels H, Squires E, Ford DV
Public Views on Models for Accessing Genomic and Health Data for Research: Mixed Methods Study
J Med Internet Res 2019;21(8):e14384
URL: http://www.jmir.org/2019/8/e14384/
doi:10.2196/14384
PMID:31436163

©Kerina H Jones, Helen Daniels, Emma Squires, David V Ford. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 21.08.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Understanding Public Attitudes Toward Researchers Using Social Media for Detecting and Monitoring Adverse Events Data: Multi Methods Study

Su Golder[1], PhD; Arabella Scantlebury[1], PhD; Helen Christmas[2], MPH

[1]Department of Health Sciences, University of York, York, United Kingdom

[2]Leeds Teaching Hospitals NHS Trust, Leeds, United Kingdom

Corresponding Author:
Su Golder, PhD
Department of Health Sciences
University of York
Heslington
York, YO10 5DD
United Kingdom
Phone: 44 1904321904
Email: su.golder@york.ac.uk

Abstract

Background: Adverse events are underreported in research studies, particularly randomized controlled trials and pharmacovigilance studies. A method that researchers could use to identify more complete safety profiles for medications is to use social media analytics. However, patient's perspectives on the ethical issues associated with using patient reports of adverse drug events on social media are unclear.

Objective: The objective of this study was to explore the ethics of using social media for detecting and monitoring adverse events for research purposes using a multi methods approach.

Methods: A multi methods design comprising qualitative semistructured interviews (n=24), a focus group (n=3), and 3 Web-based discussions (n=20) with members of the public was adopted. Findings from a recent systematic review on the use of social media for monitoring adverse events provided a theoretical framework to interpret the study's findings.

Results: Views were ascertained regarding the potential benefits and harms of the research, privacy expectations, informed consent, and social media platform. Although the majority of participants were supportive of social media content being used for research on adverse events, a small number of participants strongly opposed the idea. The potential benefit of the research was cited as the most influential factor to whether participants would give their consent to their data being used for research. There were also some caveats to people's support for the use of their social media data for research purposes: the type of social media platform and consideration of the vulnerability of the social media user. Informed consent was regarded as difficult to obtain and this divided the opinion on whether it should be sought.

Conclusions: Social media users were generally positive about their social media data being used for research purposes; particularly for research on adverse events. However, approval was dependent on the potential benefit of the research and that individuals are protected from harm. Further study is required to establish when consent is required for an individual's social media data to be used.

(J Med Internet Res 2019;21(8):e7081) doi:10.2196/jmir.7081

KEYWORDS

adverse effects; social media; ethics; research; qualitative research; digital health; infodemiology; infoveillance; pharmacovigilance; surveillance

XSL•FO
RenderX

Introduction

Background

Adverse Event Mentions on Social Media

Patient reports of adverse drug events from social media have great potential to improve the detection and monitoring of adverse effects of medications. Increasingly, people are posting information on their experiences of adverse events on social media, including discussion boards and Twitter. An overall prevalence of adverse events reports on social media has been estimated at 0.2% on generic social media platforms such as Twitter to 8% of all posts in patient forums [1]. The order of magnitude of data and the speed at which the data are made available (approaching real time) make social media a tool with the potential to revolutionize drug surveillance. This has led to a massive surge in the development of techniques for social media analytics of adverse event posts [2] and the assessment of data on adverse events from social media [1].

Ethical Challenges of Social Media Research

However, these new research avenues are not without ethical challenges [3-7]. Potentially difficult considerations surround the purpose and value of the research, benefits and harms to participants as well as privacy, informed consent, and confidentiality [8]. The ethical issues of social media research have been much debated [3,8-11]. However, there has been little formal investigation of patients' views on the use of social media for research purposes. We previously undertook a systematic review with 17 included studies, which aimed to ascertain attitudes toward the ethics of research using social media [8]. However, only 5 studies were specifically concerned with health-related research [8]. No studies explored the views of social media users on the use of social media for pharmacovigilance or identifying adverse events of nondrug interventions [8].

Objectives

We aim to address this gap by exploring the attitudes of the public toward their social media posts being used to monitor adverse events for research purposes. More specific research questions include the following:

1. What are the views and experiences of social media users on reporting adverse events on social media?
2. What are social media users' attitudes toward social media being used as a source for research data?
3. What are social media users' attitudes and ethical concerns toward social media being used as a source for adverse event information?
4. What do social media users perceive to be the ethical barriers to the use of social media in research and pharmacovigilance?

Methods

Overall Approach

This qualitative exploratory study used interviews, virtual discussions, and a focus group to explore social media users' views and attitudes toward the use of social media to monitor adverse events. A multi methods approach was selected to ensure that we obtained the perspectives from a varied sample, which was particularly important given the exploratory nature of the study and target population—members of the public [12].

Interviews

Social media users are universally diverse and so to capture this diversity, we used a convenience sampling frame and 5 different methods to recruit participants to interviews.

Posters and Flyers

Posters and flyers advertising the study were displayed across the University of York campus and at community centers in York.

Local Facebook Groups

Recruitment advertising was undertaken on York-based Facebook groups.

University of York Staff Networks

Study information were posted in staff newsletters.

Local Networks

Local parents, breastfeeding networks, and exercise groups were contacted to identify potential participants. Previous research has shown these groups to have a higher response rate than patient forums [13].

Snowballing Technique

Participants were encouraged to invite friends and family to contact the researcher if they were interested in taking part.

Interviews were semistructured and followed a topic guide, which was based on previous literature identified in a systematic review and expert opinion [8]. To refine the topic guide, we conducted 3 pilot interviews with University of York staff who would have been eligible for the study. The topic guide included a *core* set of questions to give some consistency but was used flexibly to allow for new and unanticipated responses to be introduced (Multimedia Appendix 1: Interview topic guide). Interviews permitted conversations to flow as naturally as possible while ensuring that the following topics were covered: *social media and health information on the Web*, *reporting of side effects of drugs or other treatments*, *attitudes toward research and different types of research or researchers*, *privacy expectations*, and *research conduct*.

Interview participants were sampled until no new themes emerged, as we aimed for theoretical saturation [14]. A total of 24 face-to-face interviews were conducted by either SG or HC (Table 1). Each interview lasted between 15 and 45 min and was undertaken in York between April and September 2018.

XSL•FO
RenderX

Table 1. Participant characteristics.

Participant type	Participant total, N	Gender, n			Age (years), n			
		Male	Female	Unknown	18-30	31-45	46-65	≥65
Interviewees	24	5	19	0	13	3	5	3
Virtual discussants	20	13	4	3	NR[a]	NR	NR	NR
Focus group participants	3	0	3	0	0	1	2	0
Total	47	18	26	3	13+[b]	4+	7+	3+

[a]Not reported.

[b]Refers to the fact that the number might have been higher if the age of the virtual discussants had been available.

Virtual Discussions

We contacted 7 patient forums (from general health to specific forums such as cancer or mental health forums) to obtain permission for virtual discussion. The individual forums are not listed here as we do not have permission to do so. Virtual discussions were stimulated by creating threads on selected social media sites and monitoring the Web-based posts responding to the threads (Multimedia Appendix 2: Posts to create Web-based discussion). Permission to begin discussion threads was granted by 3 of the 7 patient forum moderators, and these 3 threads resulted in 31 replies from 20 posters (Table 1). This allowed us to tap into different populations, as participants were not limited to the local geographical area, and provided an anonymous method of communication for individuals who might post on social media but were unwilling to participate in a focus group or interview. It was anticipated that those posting on health information sites might be more likely to have had posts used in research and, therefore, might hold a different perspective to those recruited through interviews and focus groups. There were no limits on the number of participants and discussions were open for 3 weeks in August 2018.

Focus Group

A focus group was conducted to encourage spontaneous generation of ideas through group dialog and interchange via a series of 8 statements and 6 scenarios (Multimedia Appendix 3: Focus group topic guide). The 8 statements followed the same themes used in the interviews—*social media and health information on the Web*, *reporting of side effects of drugs or other treatments*, *attitudes toward research and different types of research or researchers*, *privacy expectations*, and *research conduct*. The scenarios, on the other hand, suggested research on side effects being undertaken by different types of researchers or institutions, with different intentions, types of data collected (eg, numerical vs textual quotes), and social media platforms.

The focus group took place in December 2018 and lasted 60 min. A total of 3 mothers, recruited from a baby group at a Sure Start Community Centre, who met approximately 6 times a year took part in the focus group (Table 1).

Approval Process

Ethical approval was obtained from Department of Health Sciences' Research Governance Committee at the University of York. Written informed consent was obtained from all participants at the start of the focus group and interviews. At the beginning of each interview and focus group, verbal consent was obtained, assurances regarding participants' anonymity and confidentiality were made, and participants were informed that they could withdraw from the study at any time. All participants were provided with a participant information sheet before providing consent.

Analysis

The interviews and focus group were audio recorded and transcribed verbatim. All participants were assigned pseudonyms for data reporting and analysis. Analysis was facilitated by use of the qualitative data management software package Nvivo (version 11, QSR International). Information from the interviews, virtual discussion, and focus groups is presented together throughout in the reporting of our results.

Initially, we analyzed the data from each data collection method separately, using the recommended stages of thematic analysis as described by Braun and Clarke, 2006 [15]: familiarization, generating initial codes, searching for themes, reviewing themes, defining and naming themes, and data reporting. Throughout the analysis, theme and subtheme development was largely deductive, based on the topic guide and in the second stage of the analysis, the underpinning framework. However, themes were also allowed to emerge.

We then undertook a second level of analysis, where our initial themes and subthemes from all data sources were mapped onto a framework that emerged from a recent systematic review of the ethics of social media research [8]. Adopting a 2-stage approach to our analysis enabled our findings to be translated from an initial list of descriptive themes to broader overarching ones and in doing so, placed the study findings in a broader context and in line with the current evidence. For transparency, Table 2 demonstrates how our initial themes map onto the framework. This framework was selected as it was the most comprehensive framework available, drawing on the results of 17 included studies, and adapted as necessary. The framework proposes 4 main influencing factors when considering whether individuals are willing for researchers to use their social media data to monitor adverse events: the research, the social media users, consent, and responsibilities (Table 2).

XSL•FO
RenderX

Table 2. Theoretical framework.

Themes and subthemes adapted from Golder 2017 [8]	Description
Researchers	
Patients' and/or members of the public's views on researchers using social media	General reactions to the concept of using social media content for research on adverse effects
What is the purpose of the research?	Overall outcome or intention of the research
Who is conducting the research?	The affiliation of the researcher (such as university or commercial company)—this is associated with the perceived benefit of the research
Quality of research using social media data	High or low quality of the methodology used, including risk of bias
Potential harm to researchers	Any risks of harm that the researcher is exposed to
Social media users	
Perceived risk of harm to social media users	Any risks of harm that the social media users are exposed to, either individually or as a group
Perceived risk of harm to vulnerable groups	Any risks of harm for particular groups—groups could be determined as vulnerable by either their individual characteristics or the topic discussed
The original intended purpose of the posts	The intent of the poster at the time the message was placed
Privacy on social media	The public versus private nature of social media and the need for anonymity or confidentiality
Consent	
Informed consent	Permission for posts to be used in a study—this includes issues around the terms of service (also known as *terms of use* or *terms and conditions*). These are the rules agreed to for using social media sites
Research disclosure	Researchers being transparent and honest about their intent—this can be either up-front or at a later stage
Responsibilities	
Social media user	The issue of self-regulation is whereby individuals control content through personal censorship
Social media platform	The type of social media platform, for example, closed or open, personal or professional, or social norms connected with the platform
Site administrators	Site administrators, list administrators, or list moderators are often in charge of maintaining a discussion or mailing list
Governments	Refers to legal issues, regulation, or government oversight and includes issues of copyright

We adopted a reflexive approach to data collection and analysis. A total of 2 researchers undertook data collection (SG, HC) and analysis (SG, HC, and AS), with regular discussions held between the research team throughout data collection and code and theme development. The research team are considered to be in a neutral position in relation to any past expectations that may have influenced our ability to collect or analyze the data. SG is a researcher with a background in systematic reviews and social media research. HC is a Public Health Registrar. AS is a mixed-methods researcher with a background in qualitative evaluations in electronic health and systematic reviews.

Results

Presentation of Data

First, we present data relating to participants' experience of using social media and reporting adverse side effects, which provide important context about the participants represented in our sample. The remainder of our study's findings are then presented according to the 3 overarching themes outlined in the framework: the 2 actors (*researchers* and *social media users*), consent, and context (in terms of responsibilities, such as regulations or code of conduct), which provide insight into social media users' attitudes toward social media data being used for research and the various challenges associated with this.

Participants' Experience of Using Social Media and Reporting Side Effects

Experience of Using Social Media

All participants used social media. However, younger interviewees (<30 years) tended to use more social media platforms and more frequently. For instance, younger participants discussed frequency of daily usage, whereas older participants discussed accessing social media on a times-per-week basis. The most commonly used social media platform was Facebook, followed by Twitter, Instagram, LinkedIn, and Snapchat. Other platforms included the following: Mumsnet, patient forums, Tumblr, Reddit, and YouTube. Participants reported using Facebook, largely for personal

XSL•FO
RenderX

reasons, such as keeping in touch with friends and family, whereas platforms such as Twitter were largely for professional use.

Only 3 participants reported using social media to explore possible medication side effects. This was largely because the majority of participants had either not considered the possibility of doing so or had concerns about the trustworthiness of information on social media. Those who did use social media to access health information viewed it as a unique and valuable information source that provided access to information that was not accessible elsewhere. For example, information posted on social media was perceived to give a greater insight into people's lived experiences than a *list of possible side effects*:

> *There's a strong mental health community on Tumblr, and so I did search for a few tags about the withdrawal symptoms associated with this, because I was looking for more personal experiences rather than a list of possible side effects.* [Rebecca, Female 18-30 Interviewee]

Experience of Reporting Side Effects

None of the participants from the interviews or focus group had reported an adverse event to a regulatory agency (eg, the Yellow Card system in the United Kingdom). Although supportive of such official reporting channels, few had heard of them and many felt that they would have to experience a very serious adverse event to warrant using them.

Less than half of interview and focus group participants had mentioned a side effect on social media, with some of these individuals no longer willing to do so. Many considered side effects to be too personal or preferred to talk to a health professional, family member, or friend. This was largely because of confidentiality concerns, wanting an expert opinion, or viewing social media as unhelpful and in some cases scaremongering. Those that did discuss side effects on social media mostly did so to discuss the practical implications of experiencing side effects to converse with people with similar experiences or help others:

> *I've used sort of social media, mainly Facebook, to say, for example, I wanted to come off the medication and try another one, will that affect my driving?* [Jamie, Male 31-45 Interviewee]

> *I search the hashtag epilepsy a lot, if somebody's put like "oh, I'm on Keppra..." which is a medication that I used to be on "...and it's making me feel like this, does anybody else get this?" I will put "I was taken off that because of the same reason" or something like that.* [Carol, Female 18-30 Interviewee]

However, a proportion of those who were willing to discuss side effects on social media admitted that they would only discuss minor side effects that they did not feel warranted discussions with a doctor, or they were only willing to do so on a closed group:

> *Say you've got, something like breast cancer and you were wanting to deal with the side effects of that and you were part of a closed breast cancer support group. You could imagine in that case you might go, "oh my God. I'm finding chemo hard to deal with. What have other people done?" So I can imagine posting in a closed social media forum but not in an open forum.* [Arabella, Female 31-46 Focus Group Participant]

Researchers

Patients' and/or Members of the Public's Views on Researchers Using Social Media

Although the majority of participants were positive about the idea of their social media data being used for research purposes, 1 interviewee and 3 virtual discussants were *completely against it* and felt social media data should never be used for research in any circumstances. Even of those that were supportive of the use of their social media data for research purposes, a sizable proportion felt that their approval would depend on certain factors, such as the purpose and quality of the research and who is conducting it.

The purpose of the research was considered one of the main factors influencing whether participants would approve the use of their social media data for research purposes. Participants were particularly supportive of their data being used for research where it was clear that there would be a societal benefit and liked the idea of helping others. Health care research, and in particular pharmacovigilance, was seen as a *good thing* over and above other types of research, with potential benefit to others. Participants felt that the ubiquitous use of social media, and consequently, large amount of unique, valuable, and accessible data, meant that it could be used to improve *patient safety*, *save lives*, and help *pick up side effects quicker*. For example, participants discussed how social media made side effect information more readily available, and it was felt that some people might be more willing to post about their side effects on social media than report them to a health professional or through traditional channels:

> *It could be very beneficial because you could be finding out side effects that aren't so readily reported to health professionals.* [Imogen, Female 18-30 Interviewee]

> *I guess it's useful in a way that people would post on social media to report what's happening in their daily lives...maybe it's easier to post on social media rather than talk to the doctor, cos they don't have to make a trip to the doctor's.* [Carla, Female 20 Interviewee]

This extended to data on medication adherence and rare adverse events for which social media might be able to uncover information more easily:

> *There must be loads of people that get side effects and rather than going and tell anyone about it, they just stop taking the medication. So, nobody would ever know. And they are more likely to respond and say, "oh yeah—that happened to me and then I stopped taking it." So, there could be millions of people that are getting the same side effect, give up*

XSL•FO
RenderX

> *on the tablet and then just don't say anything. So, I think, it probably could be a really good way to find those people that are hidden that haven't reported it officially.* [Kate, Female 46-65 Interviewee]

Who is Conducting the Research?

Participants were divided as to whether they thought that the affiliation or type of researcher (such as an academic, a government official, or an industry representative) made a difference to the ethical implications of research using social media data. Although some felt it did not matter, as researchers were *all working toward similar goals* and thus would *put them in the same pot*, others categorized researchers according to the underlying motives of their research. The biggest issue was whether the motives were for profit, and participants often stated that they would be less supportive of commercial organizations because of their *vested interests*, whereas *altruistic* motives were met with greatest approval:

> *The only one I would not like would be commercial. If you're trying to make life better for people, then by all means, if you're trying to generate profit, then no I think that's not acceptable.* [Jamie, Male 31-45 Interviewee]

Participants also raised concerns about the motives of research undertaken by the following: Government, charities, and students, who were described as having their *own agenda*, influencing public opinion, or obtaining a degree or qualification. Participants tended to be more comfortable with academic institutions, which were viewed as more likely to adhere to good ethical conduct and not have ulterior motives:

> *Anybody who is making any kind of income I'd be very unhappy about. So, strangely enough, I would actually feel more comfortable if it was an academic researcher, because I would hope that people had some awareness of research ethics. And weren't particularly getting any direct benefits from it, because I assume for a lot of academic researchers whether 100 people said, this was rubbish, and, 100 people said, "this was really good", it doesn't actually matter to them. So, they haven't got a particular axe to grind.* [Megan, Female 46-65 Interviewee]

> *I would have much more faith in an academic institution or institutions than I would in a corporate company. I think I would trust academic institutions more than I would the government to be setting up anything. I just feel like there is much more pressure on the government to behave in ways that, God, I sound like my parents! I do think there is that pressure to benefit the corporations and the big conglomerates more than potentially there is to protect the individual privacy or just kind of adhere to basic ethics.* [Gill, Female 31-45 Interviewee]

Quality of Research Using Social Media Data

The quality of social media content was a particular concern. Some participants used the internet for health information but preferred trustworthy sites, such as National Health Service (NHS)–managed sites rather than social media sites. The validity of any research conducted using social media data was bought into question, and this was considered to be a particular concern when the research related to important issues, such as health or side effects. It was acknowledged that social media might not reflect the truth as it was considered easier to make false claims or exaggerate the truth on social media than to a health professional, and people's motives for posting might reflect the content. For example, people may post to try to provoke a reaction or particular response. In addition, participants were wary of side effects being wrongly attributed by the public to a particular drug and cited the measles, mumps, and rubella vaccine and autism as an example:

> *I don't trust the common layperson...I think people are much too willing to see a causal link where there isn't one and that especially with something like this. This would really worry me.* [Gill, Female 31-45 Interviewee]

> *I find that quite worrying that they base research on what people post because I don't always think what people post is necessarily the truth and may exaggerate their possible drug use or it might be that they exaggerate their symptoms of an illness. So, I don't think a forum like Facebook or something like that would be very reliable.* [Mandy, Female 31-45 Interviewee]

> *There is loads of issues about going just to social media—are all these people telling the truth? How representative are they? Why should I believe them? Are they real? Do these people really exist? Is what they're saying true?* [Megan, Female 46-65 Interviewee]

Another concern with the validity of social media research related to the representativeness of social media users to the general population. It was thought that a lot of people did not use social media or, if they used it, did not post on health issues, which might mean that data were *skewed*. Participants felt that individuals who posted on social media might be more likely to have had either extremely positive or negative experiences or represented the more extreme people in society. Social media users were also perceived to be younger than the general population, and this was seen as a particular problem when studying health conditions that were more prevalent in elderly populations. People with some types of conditions were also thought to be more active social media users and more vocal than people with other conditions:

> *If I was reading a study that was like "we took information from social media", I'd be cautious of reading it because with social media you're only getting a percentage of the population. Are you getting like a broad enough snapshot of the population to get a proper sort of outcome? It's not like going door to door, right, where everyone has a chance to have their say.* [Helen, Female 18-30 Interviewee]

A potential solution to the low validity of the research data when using social media was to use these data as an adjunct to other

XSL•FO
RenderX

sources of information to provide a *rough idea* for background information:

> *I suppose it gives them a rough idea. But, it's the fact that people exaggerate. It's not controlled, is it? You know, I guess, it's good for them to use maybe for background information to start work. But I wouldn't see it as a reliable source to produce some findings from. Because of that fact that people exaggerate...So, I guess, I feel like it's a really good way to get an idea of trends in things. But, shouldn't be the only sort of thing for research. There should be other sources to back it up.* [Kate, Female 46-65 Interviewee]

Social Media Users

Themes discussed by participants relating to *social media users* were concerned with the potential risks or harm to social media users and, in particular, vulnerable groups.

Perceived Risk of Harm to Social Media Users

Although some participants emphasized how the public nature of social media meant that there was little or no harm in the posts being used in research, others felt that the risks to social media users should be considered even if the data were in the public domain. Any risks of harm to social media users were perceived to be dependent on the type of information posted, with personal or sensitive information, such as side effects, warranting the need for more stringent ethical safeguards than opinions on trivial issues, such as television programs:

> *Even if it's completely open, these people who are doing the research still need to think through what the implications are and think very carefully about how they're using that information even when there's no deceit required to get that information. I still think they need to think about the possible implications.* [Arabella, Female 30-45 Focus Group Participant]

Perceived Risk of Harm to Vulnerable Groups

Particular groups of people were perceived to be at greater risk of harm from research using social media content. Vulnerability was recognized in terms of age. The younger participants tended to perceive the older generation as more vulnerable because of a lack of understanding of how social media worked, as they had not grown up with it. On the other hand, the older generation perceived the younger people as more likely to share personal details without considering the implications of what they are posting and not listen to warnings from older generations. It was also acknowledged that irrespective of age, people could be *naïve*, and there was a danger of *getting carried away in a conversation* and forgetting how public social media posts were:

> *I've got a lot of friends who have been in the Retreat hospital, like the mental hospital, and I follow all of them on Instagram and Facebook and they very often have discussions how they're feeling on the medication and so on. And then when they're discussing it with each other sometimes I'm looking at it and I'm thinking, I don't think everybody should be knowing that. That's quite personal. And I think once they start a conversation they forget that the world can watch and they continue with it and you get caught up in it and you forget that actually everybody can read that now.* [Kate, Female 46-65 Interviewee]

Vulnerability was also linked to people's illnesses or conditions, with some conditions or topics, such as side effects or mental health, considered more sensitive than others:

> *If someone has disclosed something about a mental health issue or something which is more potentially stigmatized, particularly something like HIV or something like that, I think you would have to be more careful about using that data because I think, they're going to put themselves, potentially, in a more vulnerable position than if someone has had a bout of chicken pox or the flu.* [Claire, Female 31-45 Interviewee]

Users of patient forums were viewed as vulnerable by interviewees and focus group participants. This was reinforced by the virtual discussants, who considered themselves to be vulnerable and felt that they had enough to contend with already, stating "it's hard enough to manage what we have—I don't think we require any voyeur over that period" (Male, virtual discussant, paraphrased). Indeed, knowing researchers were looking through their posts may be very harmful to some virtual discussants who even stated that this would have put them off posting:

> *This discussion forum was my only method of connecting with individuals who were experiencing the same illness as me...Given I was debilitating paranoid, in the event that I had known researchers were potentially trawling posts for information, I doubt very much I would have posted.* [Unknown gender, virtual discussant, paraphrased]

The Original Intended Purpose of the Posts

The original purpose of the posts was mentioned by a few participants who were keen to emphasize that their reasons for using social media included the following: for support, to encourage open discussions on important issues, to help benefit others, or to find or communicate with followers, friends, and family. Ethical issues were, therefore, seen to be associated with their social media data being used for research purposes, particularly when considering that this was not one of the motives for posting and the fact that many people might not even be aware that researchers could use the data. Associated with this were concerns that their data might be taken out of context or misinterpreted:

> *If you put a public profile up...you weren't thinking about researchers using your statuses, you were thinking oh maybe I'll find more friends that way. Like maybe people will be able to find me better. So, I feel like it's not valid to say that because it's public there are no ethical issues because people weren't intending for you to use their data in that way.* [Sophie, Female 18-30 Interviewee]
>
> *Maybe it is public but then they're putting it in a public domain for a certain reason and then as a*

XSL•FO
RenderX

> *researcher you're changing that use. So, that's like me going into an interview and telling people that this is for research and then going off and actually using it for my own private company, you know, that's not ethical so, it's sort of a similar thing. I know it's not the same but it's got similar, kind of, parallels, in a way because you're changing what that person did with that information...a researcher comes along and they take that information and use it in a completely different light to what it was originally intended. So, I think that, for me, is the main ethical issue with it. That intent and reformulation of that information. But, not to say, it shouldn't be done.* [Imogen, Female 18-30 Interviewee]

Privacy on Social Media

Many participants were fully aware that they did not know how much of the information they posted was public and who could see or use their data. Participants were also aware that privacy settings existed but did not always keep these up to date or know what they were currently set as. Many also stated that they would like to know more about how privacy worked on social media but understood that this was a complicated issue:

> *I think it's quite complicated to understand. You really have to make an effort to know what's private and what's not, because it's just not as easy as that, because on Facebook you can send private messages to people, which are obviously only for that person, or you can send it in a big group, or can you make parts of your profile or the messages you post private and other parts not. So, I think you have to really make an effort yourself to know what you're sharing with people and what you're not sharing, and who with, and what parts of social media are accessible to different people.* [Joanne, Female 18-30 Interviewee]

> *I've only recently sort of found the privacy settings. So, yeah—I've had a Facebook page for a few years and it was only last year that I realized that it was going public. And then I changed the settings.* [Mandy, Female 31-45 Interviewee]

The difference between an in-person conversation and a conversation on social media was also highlighted, as face-to-face conversations even in public spaces were seen as more private than social media. Furthermore, although some participants considered all social media data to be public in one way or another, others considered there to be different ethical considerations depending on the social media platform. The more public the platform, for example, the more likely the participants were to agree that research was ethical on these sites. Concern was also raised that researchers might be able to access private accounts or areas of social media that individuals considered to be private:

> *All of them are very public so it doesn't matter and on Facebook you've fifty or a thousand friends, it's still public. I mean, I know people who say whatever you post on Facebook it's yours, it's private, but it really isn't.* [Evie, Female 18-30 Interviewee]

Consent

The issue of *consent* was discussed with emphasis on when and how informed consent was required and whether research disclosure was necessary.

Informed Consent

Whether informed consent is necessary for researchers to use an individual's social media data was met with uncertainty and polarized opinions. A small number of participants felt that informed consent was a requirement for any research and that using *any information or observations from individuals without their knowledge or permission is unethical*. For the majority of participants, whether informed consent is required depended on a number of issues. The most frequently discussed factor was whether anonymity could be retained. This was seen as crucial to protecting the individual poster. Linked to anonymity and identity was the use of quotes. Although participants were accepting of quantifiable data (such as 63 people reported suffering from insomnia after taking drug X), they were less accepting of researchers using direct quotes that could lead to people being traceable, for example, by a simple Google search. It was, therefore, recommended that informed consent be obtained for direct quotes or that quotes are paraphrased.

Another important issue impacting on whether informed consent was considered necessary was whether the post was public or not and the type of platform it was posted on. It was generally accepted that informed consent was less of an issue for a social media platform such as Twitter than Facebook. Others also talked about the sensitivity of the data and how this might impact on whether informed consent was deemed necessary. For instance, it was not thought necessary to obtain consent for trivial information but data discussing health issues were considered more likely to require informed consent:

> *That's a really interesting question, because you're putting that information out there for the whole world to see...Depends what kind of...oh, I don't know, what kind of research they were doing maybe?* [Carol, Female 18-30 Interviewee]

The logistical problems of obtaining informed consent were acknowledged and described as an *utter nightmare*. This was both in terms of the number of people who would need to be contacted and beliefs that direct messaging was intrusive and unlikely to get a response.

There was overwhelming agreement that terms and conditions of the social media site would not be an acceptable or effective way to seek informed consent. This was mainly attributed to the fact that no one ever reads them because of the size of the print and their length, people are already *bombarded* with too much information, and they are constantly updated and revised. It was also pointed out that it would be very difficult for consent to be covered by the terms and conditions because of the variety of research that could be conducted. For instance, a social media user may be happy for their data to be used in some research but not in other research, and it would be impossible for all scenarios to be covered:

> *Conditions saying like your information can be used for research—what research? Who's it going to, etc,*

XSL•FO
RenderX

> *there are so many different topics that you can research on social media that I might be fine with one but not the other, so I think to that extent, it's a totally an unfair condition clause in my eyes.* [Helen, Female 18-30 Interviewee]

Participants proposed an *opt in* or *opt out* option as a potential alternative to relying on terms and conditions. For some individuals, concerns about the ethics and practicalities of obtaining consent led to suggestions for researchers to use social media to recruit participants or conduct a Web-based survey instead of using social media data:

> *I think there's a number of reasons like a direct personal message wouldn't work, cos like if someone's going to reply to a message that was like I'm doing research on social media, I'd be like oh, this person's bugging me, you know what I mean, I wouldn't even reply to it. So, I think you might not get responses there, I think it's just tricky with social media, but then again, how else are you going to ask for their consent.* [Helen, Female 18-30 Interviewee]

> *I suppose it depends on how many people you want to quote. For me, I'd be kind of wondering, potentially, can you contact the individual. But then, if you're looking at 100/200 perhaps more people that is not going to be viable.* [Gill, Female 31-45 Interviewee]

Research Disclosure

The majority of participants felt that researchers should disclose their identity and the purpose of the research and believed that not to do so would be *sneaky*. However, there were occasions where disclosure was considered unnecessary, such as in situations where only numerical data are required, or where doing so would distort the research. It was thought that people might have changed what they said if they knew they were being studied, and thus different results might be obtained. It was also argued that other people on social media did not have to declare who they were. Indeed, part of the attraction of social media was perceived to be that it provided a space where people could pretend to be someone else or have a different persona to real life:

> *I guess nobody who is on there has to disclose who they are, do they? I mean they could be patients or they can be people who are just interested, or family, so I guess you don't have to say that you're a researcher...it's nice to be more transparent, but then you might get different results. So yes, I think it's not necessary, but it is a bit—a bit tricky.* [Joanne, Female 18-30 Interviewee]

Situations where research disclosure was seen as important included when it involved direct interaction with an individual as opposed to being observational:

> *If they're just sitting and effectively using it to surf stuff that's posted openly, then I'm not sure that if you had a Twitter account, I don't think you need to say "social researcher @Twitter" kind of thing. But if you're, if you're effectively prompting people to get information. If you're starting to pose questions saying, "has anybody got any experience of this?" At that point, to do that without being open about your background, that seems very deceitful.* [Arabella, Female 31-46 Focus Group Participant]

Responsibilities

The issue of whose *responsibility* it is to oversee research practice and protect users was also discussed by the participants. Responsibilities tended to fall directly to the social media users, site (including the platform, such as Twitter or Facebook, and site administrators), or regulations.

Social Media User

Self-regulation was mentioned by many participants as the answer to protecting individual privacy and thus moving responsibility to the social media users. Some participants felt strongly that it was the responsibility of social media users to self-regulate, and this was largely attributed to the *voluntary* nature of posting:

> *If people are stupid enough to put really personal things out there and people see it then it's your own fault. It's fair game.* [Harry, Male 18-30 Interviewee]

> *If it's out there, it's out there...If you put it on social media, it's there for everybody. That's it.* [Hilda, Female over 65 Interviewee]

Some participants felt that posting about health issues was a strange or *somewhat bizarre* thing to do and could not understand why others might do so:

> *I'm astonished that people do [post on their health]. I find that incredible that something so personal people are quite happy to post questions online about it.* [Megan, Female 46-65 Interviewee]

Many participants also stated that over time they had changed their Web-based behavior. Although in the past they might have been more open and trusting of social media, over time they had become more cautious resulting in increased self-regulation:

> *I have in the past when I was younger, I think, I posted more but that was, I think, before I learned how often things go wrong.* [Gill, Female 31-45 Interviewee]

Participants had also become more cautious over time after hearing negative stories about social media from friends and family and/or in the media and made particular reference to media coverage of Cambridge Analytica:

> *In the past I've used it for work as well but as sort of, you know, you hear news stories about somebody has said something on social media and it's got them into trouble at work, even though it's nothing to do with work so, as I've got older and become a little bit more wary of that sort of thing.* [Jamie, Male 31-45 Interviewee]

Some participants also stated that they were likely to self-regulate more after taking part in this research, considering the issues around researchers using social media data:

> *I am slightly concerned now how open my Facebook page might be and how it's being used—it would*

XSL•FO
RenderX

definitely make me be cautious of what I post. [Mandy, Female 31-45 Interviewee]

Social Media Platform

Although terms and conditions of the social media platforms were rejected as a means to protect individuals, careful selection of social media platforms was considered important. The majority of participants differentiated between different platforms in terms of the ethical considerations of researchers using social media. Participants spoke about the unwritten rules or purpose of different types of social media. For example, Twitter was thought of as a broadcasting platform to publicize information and described as a *way of deliberately grabbing attention* or *posting out to the world* and was, therefore, perceived to be open for use by researchers. In contrast, Facebook was thought of as a place to share information with family and friends, with even public information on Facebook not necessarily considered open for use by researchers:

> *I feel like it's different on different social medias because Twitter is less personal and more kind of flaunty, you post to everybody on Twitter...I think it's different on Twitter than it is on Facebook because people are not talking to their family and their friends on that, they're just writing things that they want others to see. So, I don't see such an ethical problem of using people's data from Twitter, as I do people using their stuff from Facebook.* [Sophie, Female 18-30 Interviewee]

> *I think something like Twitter you're posting out to the world...if you're posting on a public forum, for example, the BBC News Have Your Say, your comment is going on a forum that anyone in the world can see. If you're doing it on Facebook, where you've explicitly said I want this group, these people to see my data, see my opinion or see my, you know, about side effects, then that's strictly controlled, then I think that, yes, that's the difference to me...Facebook's different but Twitter, yes, Twitter's the one where you're posting out to the world.* [Jamie, Male 31-45 Interviewee]

Site Administrators

Site administrators or moderators were rarely mentioned. However, a couple of participants felt that researchers should disclose who they were to the administrators or moderators and ask permission from them to use the site in research.

Governments

Laws on social media research were generally seen as a positive by moving responsibility to governments. However, there was understandable confusion about what laws or regulations already existed and whether they could be applied to research using social media data. Ownership and copyright were also mentioned, and it was recognized that once posted, the social media platform might have ownership rights over information, with the user unable to reclaim what they posted. Although this was considered a cause for concern, participants did not see a way around this.

Participants were unsure whether there should be strict rules and regulations or something which was less formal and instead provided guidance on how to conduct ethical social media research. A *code of conduct* or *best practice* was supported by most participants to give reassurance to social media users and be helpful to researchers:

> *Some more national guidance on the ethics of social media research would be helpful. Because, it does seem unfortunate if everybody's having to reinvent the wheel and different universities have different standards. You'd think that there would be some sort of national expectations on, you know maybe not hard and fast rules but at least guidance of best practice to actually make it easier for everybody. To make it easier for the researchers but also to make it easier for people who are using social media so that they know what people might get up.* [Arabella, Female 31-46 Focus Group Participant]

The logistics of regulation was also seen as a hindrance. It was thought that implementing any laws would be difficult, if not impossible to do. Questions were asked on how this could be policed or monitored and how, from an international perspective, this would work, as different countries would have different laws and social media data are available worldwide:

> *There should be [laws], but nothing's ever going to change...I can't imagine a time when...anybody would be prosecuted for it, because who is going to keep looking at it?...If it can be policed properly or...I don't think it can be. I don't...who...you know, it's national regulations, international, European, how do you get through it all, how do you get that? I don't know. Yeah. It should be, but...I'm a...realist.* [Joan, Female over 65 Interviewee]

Discussion

Summary of Results

Our study provides information on members of the public's views on the use of social media data for research, with a focus on adverse effects research. Opinion was divided, with some supportive of social media data being used freely as it was in the public domain, whereas others felt concerned about vulnerable groups, sensitive topics, and issues with people's awareness of privacy regulations and how their data could be used. Our study found that participants were generally more positive about the study of adverse effects using social media data than the general use of social media in research. However, this support was caveated and dependent on a number of conditions being met. The most important condition was that social media data were used for research where there was clear societal benefit and users were protected from any potential harm that might arise from their data being used. The most powerful benefit to social media research for side effects was the potential to save lives.

The most alarming potential harm expressed was that by a virtual discussant who stated that they would not have used social media had they known researchers were trawling their

XSL•FO
RenderX

posts. This is of great concern as social media have become a great source of support for many people, and this may be particularly the case for mental illness—where people are already very vulnerable [16].

There was divided opinion as to whether consent was required to use social media data for research purposes. This uncertainty could partly be attributed to the fact that many participants had neither realized that researchers used social media posts nor thought about any ethical issues, the *tricky* nature of social media research, and the different types of social media and research.

Comparison With Existing Literature

We used a framework derived from a systematic review of the literature [8] to aid how we analyzed and interpreted the study's findings. This framework evolved mostly from non–health-related research with an international coverage, which often included researcher's views and those from social media users.

Many of the themes within this chosen framework are interrelated. For example, *self-regulation* is related to *harm*, as this leads to a decline in freedom and puts restraints on the supportive nature of social media. In addition, *self-regulation* is related to *privacy*, as people choose to keep certain information private, and *responsibilities*, as it puts the onus firmly in the hands of the social media user.

Most issues in this study were in line with previous study findings, regarding research using social media data in other areas, and this was demonstrated by the fact that no new themes emerged that were not already in the framework. But there were different issues raised within some of the themes—in particular, the responsibilities of social media moderators and governments were not emphasized as an issue in this study.

For instance, more emphasis was placed in this study on the beneficial aspects of social media research, validity of the social media data, and power of self-regulation than in previous studies. This is likely to be because of the potential societal benefit of adverse effects research, the perceived importance of high-quality research for adverse effects, and how increasingly savvy users are beginning to become in respect to lack of privacy on social media.

Unlike previous studies, in this study, no consideration was given to the potential harm to the researchers. This may be unsurprising given that the participants were more focused on the potential harms to social media users, as they were social media users themselves. The harms perceived for social media users in other types of research included bullying, abuse, and persecution, whereas here, the concerns in this study were more in line with preventing users from gaining support through heaving self-regulation or leaving sites. This is likely to be because of the international perspective of other studies and some of the topics (such as homosexuality and sexual abuse) covered. For instance, in some countries outside the United Kingdom, homosexuality is illegal or at least taboo.

Implications

It is clear that social media users are in favor of some sort of overarching guidance for all institutions to follow. Our findings will not only help direct future research but will also provide social media websites, universities, ethics boards, pharma companies, and policymakers with evidence to inform policy and guidance on the use of social media data for research.

This research shows the variety of responses received when asking social media users about the ethics of using social media for a specific subject area. However, the complexity of ethical considerations is largely understood by social media users, as are the different considerations for different types of social media data collection, even within 1 research area. Although overall, there was support because of the large potential benefits, care must still be exercised when conducting social media research into adverse effects.

By harnessing technology, social media research can help inform research on adverse effects in a relatively easy and effective way. Although there may be no 1 rule that fits all regarding the ethical considerations, this research demonstrates the need for careful consideration of the ethics and increasing awareness of how social media are used in research. Participants expressed surprise at current practices in social media research or were unsure about what is currently being done with their posts or data and were eager to know more about privacy settings and how their data could be used. Education and information provision may be one of the most appropriate ways forward to help protect individuals. The use of patient and public involvement to develop consent and ethics processes for this research area may also help the approval processes for social media research. Thus, public and researchers could work together on an on-going basis, via means, such as committees or review panels.

The participants indicated that their opinions on the use of social media in research changed over time. Social media are constantly evolving communication means, with changing popularity between different platforms for different demographics. The use of people's data is also increasingly becoming more apparent, particularly with recent news events. It is, therefore, imperative that the users of social media are consulted over time, as users reflect on new developments. Although this research demonstrates the value of ascertaining the views of users through interviews, virtual discussions, and focus groups, there is value in other types of research, such as Web-based surveys.

Strengths and Limitations

The study adds to a limited qualitative evidence base on the use of social media data for research. The multi methods approach and range of recruitment strategies adopted ensured that a wide range of views were captured. Despite this, interview and focus group participants were predominately female and represented a limited geographical area. Our study also highlights the potential for using *virtual discussants*, a previously underutilized source of qualitative data collection, which can provide a viable source of qualitative data. However, when using this method in the future, researchers should be aware that obtaining

XSL•FO
RenderX

permissions from social media platforms can limit access to these data, as was the case in this study, and because of the nature of the data collection and anonymity of these forums, gaining insight into the demographics of the participants can be difficult. Interviews and focus groups were also conducted by an academic researcher and a public health researcher. Although this may have influenced how participants responded to the idea of researchers using their social media data, criticisms and concerns were still provided.

Conclusions

There appears to be a wide disparity in attitudes toward research using social media data from those who believe ethical approval (such as approval from an institutional review board, a research ethics board, or research governance body) is not necessary to those who support the idea that ethics should prevent such research taking place. Adverse effects are viewed as personal and, therefore, more likely to attract ethical consideration. However, adverse effects are also seen as an important area of research, which may have enormous benefit to society. All future adverse effects research should consider the ethical implications with an aim to minimize harm and maximize benefit. This research indicates the value that the public place on these aspects to aid researchers in the development of their research methods.

Acknowledgments

This study is an independent research arising from a Postdoctoral Research Fellowship, Su Golder PDF-2014-07-041 supported by the National Institute for Health Research (NIHR). The views expressed in this paper are those of the authors and not necessarily those of the NHS, NIHR, or Department of Health.

Authors' Contributions

SG developed the study protocol, conducted the focus group and interviews, coded the transcriptions, analyzed the results, and drafted the paper. AS commented on the study protocol, assisted with the analysis of the results, and commented on drafts of the paper. HC commented on the study protocol, conducted interviews, coded the transcriptions, and commented on drafts of the paper.

Conflicts of Interest

SG had support from the NIHR for the submitted paper; SG, AS, and HC had no financial relationships with any organizations that might have had an interest in the study; SG, AS, and HC had no other relationships or activities that could appear to have influenced the study.

Multimedia Appendix 1
Interview topic guide.
[PDF File (Adobe PDF File), 96KB - jmir_v21i8e7081_app1.pdf]

Multimedia Appendix 2
Posts to create online discussion.
[PDF File (Adobe PDF File), 54KB - jmir_v21i8e7081_app2.pdf]

Multimedia Appendix 3
Focus group topic guide.
[PDF File (Adobe PDF File), 85KB - jmir_v21i8e7081_app3.pdf]

References

1. Golder S, Norman G, Loke YK. Systematic review on the prevalence, frequency and comparative value of adverse events data in social media. Br J Clin Pharmacol 2015 Oct;80(4):878-888 [FREE Full text] [doi: 10.1111/bcp.12746] [Medline: 26271492]
2. Sarker A, Ginn R, Nikfarjam A, O'Connor K, Smith K, Jayaraman S, et al. Utilizing social media data for pharmacovigilance: a review. J Biomed Inform 2015 Apr;54:202-212 [FREE Full text] [doi: 10.1016/j.jbi.2015.02.004] [Medline: 25720841]
3. Eysenbach G, Till JE. Ethical issues in qualitative research on internet communities. Br Med J 2001 Nov 10;323(7321):1103-1105 [FREE Full text] [doi: 10.1136/bmj.323.7321.1103] [Medline: 11701577]
4. Elgesem D. What is special about the ethical issues in online research? Ethics Inf Technol 2002;4:195-203. [doi: 10.1023/A:1021320510186] [Medline: 15977360]
5. Flicker S, Haans D, Skinner H. Ethical dilemmas in research on internet communities. Qual Health Res 2004 Jan;14(1):124-134. [doi: 10.1177/1049732303259842] [Medline: 14725180]
6. Heilferty CM. Ethical considerations in the study of online illness narratives: a qualitative review. J Adv Nurs 2011 May;67(5):945-953. [doi: 10.1111/j.1365-2648.2010.05563.x] [Medline: 21231955]

XSL•FO
RenderX

7. Swirsky ES, Hoop JG, Labott S. Using social media in research: new ethics for a new meme? Am J Bioeth 2014;14(10):60-61 [FREE Full text] [doi: 10.1080/15265161.2014.948302] [Medline: 25229594]
8. Golder S, Ahmed S, Norman G, Booth A. Attitudes toward the ethics of research using social media: a systematic review. J Med Internet Res 2017 Jun 6;19(6):e195 [FREE Full text] [doi: 10.2196/jmir.7082] [Medline: 28588006]
9. Capurro R, Pingel C. Ethical issues of online communication research. Ethics Inf Technol 2002;4:189-194. [doi: 10.1023/A:1021372527024.pdf]
10. Whitehead LC. Methodological and ethical issues in internet-mediated research in the field of health: an integrated review of the literature. Soc Sci Med 2007 Aug;65(4):782-791. [doi: 10.1016/j.socscimed.2007.03.005] [Medline: 17512105]
11. Herron M, Sinclair M, Kernohan WG, Stockdale DJ. Ethical issues in undertaking internet research of user-generated content: a review of the literature. Evid Based Midwifery 2011;9(1):9-15 [FREE Full text]
12. Pope C, Mays N, editors. Qualitative Research in Health Care. Third Edition. Oxford, UK: Blackwell/ BMJ; 2006.
13. Mikal J, Hurst S, Conway M. Ethical issues in using Twitter for population-level depression monitoring: a qualitative study. BMC Med Ethics 2016 Apr 14;17:22 [FREE Full text] [doi: 10.1186/s12910-016-0105-5] [Medline: 27080238]
14. Baker SE, Edwards R, Doidge M. National Center for Research Methods. 2012. How Many Qualitative Interviews Is Enough? Expert Voices and Early Career Reflections on Sampling And Cases in Qualitative Research URL: http://eprints.ncrm.ac.uk/2273/4/how_many_interviews.pdf [accessed 2019-08-08]
15. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
16. Shepherd A, Sanders C, Doyle M, Shaw J. Using social media for support and feedback by mental health service users: thematic analysis of a Twitter conversation. BMC Psychiatry 2015 Feb 19;15:29 [FREE Full text] [doi: 10.1186/s12888-015-0408-y] [Medline: 25881089]

Abbreviations

NHS: National Health Service
NIHR: National Institute for Health Research

Edited by G Eysenbach; submitted 06.06.19; peer-reviewed by M Ameen, K Smith, G Goh; comments to author 03.07.19; revised version received 05.07.19; accepted 23.07.19; published 29.08.19.

Please cite as:
Golder S, Scantlebury A, Christmas H
Understanding Public Attitudes Toward Researchers Using Social Media for Detecting and Monitoring Adverse Events Data: Multi Methods Study
J Med Internet Res 2019;21(8):e7081
URL: http://www.jmir.org/2019/8/e7081/
doi:10.2196/jmir.7081
PMID:31469079

©Su Golder, Arabella Scantlebury, Helen Christmas. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 29.08.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Implementation and Evaluation of a Web-Based Distribution System For Anesthesia Department Guidelines and Standard Operating Procedures: Qualitative Study and Content Analysis

Kaspar F Bachmann[1], MD; Christian Vetter[1], MD; Lars Wenzel[1], MAS; Christoph Konrad[2], MD; Andreas P Vogt[1], MD

[1]Department of Anaesthesiology & Pain Medicine, Inselspital, Bern University Hospital, University of Bern, Bern, Switzerland

[2]Department of Anaesthesiology & Pain Medicine, Kantonsspital Lucerne, Lucerne, Switzerland

Corresponding Author:
Kaspar F Bachmann, MD
Department of Anaesthesiology & Pain Medicine
Inselspital, Bern University Hospital
University of Bern
Inselspital Freiburgstrasse 10
Bern, 3010
Switzerland
Phone: 41 316322111
Email: kaspar.bachmann@insel.ch

Abstract

Background: Digitization is spreading exponentially in medical care, with improved availability of electronic devices. Guidelines and standard operating procedures (SOPs) form an important part of daily clinical routine, and adherence is associated with improved outcomes.

Objective: This study aimed to evaluate a digital solution for the maintenance and distribution of SOPs and guidelines in 2 different anesthesiology departments in Switzerland.

Methods: A content management system (CMS), WordPress, was set up in 2 tertiary-level hospitals within 1 year: the Department of Anesthesiology and Pain Medicine at the Kantonsspital Lucerne in Lucerne, Switzerland, as an open-access system, followed by a similar system for internal usage in the Department of Anaesthesiology and Pain Medicine of the Inselspital, Bern University Hospital, in Bern, Switzerland. We analyzed the requirements and implementation processes needed to successfully set up these systems, and we evaluated the systems' impact by analyzing content and usage.

Results: The systems' generated exportable metadata, such as traffic and content. Analysis of the exported metadata showed that the Lucerne website had 269 pages managed by 44 users, with 88,124 visits per month (worldwide access possible), and the Bern website had 341 pages managed by 35 users, with 1765 visits per month (access only possible from within the institution). Creation of an open-access system resulted in third-party interest in the published guidelines and SOPs. The implementation process can be performed over the course of 1 year and setup and maintenance costs are low.

Conclusions: A CMS, such as WordPress, is a suitable solution for distributing and managing guidelines and SOPs. Content is easily accessible and is accessed frequently. Metadata from the system allow live monitoring of usage and suggest that the system be accepted and appreciated by the users. In the future, Web-based solutions could be an important tool to handle guidelines and SOPs, but further studies are needed to assess the effect of these systems.

(J Med Internet Res 2019;21(8):e14482) doi:10.2196/14482

KEYWORDS

standards; computer communication networks; anesthesiology; decision making, computer-assisted

XSL•FO
RenderX

Introduction

Guidelines and Standard Operating Procedures

Generally, guidelines and standard operating procedures (SOPs) are an integral part of perioperative medicine, and these are particularly an integral part of anesthesiology. They have found their way into daily clinical routine, and they form the basis for patient safety algorithms [1-3]. Adherence to guidelines has been associated with improved outcomes in the fields of anesthesiology and intensive care, and it has an impact on patient safety, employee training, and overall quality [4-7]. Guidelines have also been shown to motivate a team, especially if the employees were involved in creating the content [8]. However, the creation, maintenance, and distribution of these guidelines within an institution can be challenging, and the potential benefits and drawbacks remain unclear [9], especially as scientific data on the benefits of guidelines are scarce. Furthermore, measuring quality and safety in anesthesia remains a challenge [10], especially as there are only a few validated indicators, and evidence of their scientific validity is low [11].

Digitization in Medical Care

Before the age of computers, many institutions distributed their guidelines and SOPs in paper form. Over time, the increased use and availability of computers has led to the digitization of medical care [12,13]. During this transformation, many printed guidelines were transformed into digital files, often using PDF, and these files were commonly stored on local servers. However, this approach can lead to outdated files, and availability to users (eg, the anesthesia providers) is limited, with no search function or linking of content. Managing and reviewing content is challenging and laborious. In the era of digitization, with increased access to computers in the operating room, a fully computerized approach to this problem seems practical, and improved adherence to guidelines can be expected because of improved availability [14]. Digitization has enabled solutions that provide fast navigation, a broad overview, and new formats for content, such as movies. Digitized learning material and mobile learning, in general, can be effective [15]. With digital solutions, content is easily accessible and can be managed in a centralized database, and updates are easy and time saving. The choice of a content management system (CMS) depends on various factors, such as the publication process, accessibility, open-platform support, the implementation process, costs, and security [16]. Scientific evidence to support the implementation of digital distribution and CMSs remains scarce, and issues, such as insufficient security or unsatisfactory publication processes, have been raised with certain workflows [17].

Aim

This paper presents a digital solution for the development, distribution, and management of guidelines and SOPs. We have successfully implemented our Web-based solution in 2 different anesthesiology departments in Switzerland. In this paper, we describe the necessary requirements, the implementation process, and the metadata generated by the users (ie, employees of the department). Furthermore, we lay out the process for content management and development.

Methods

Content Management Systems

The CMS is a software used to create, update, and organize content produced by a defined group. This is usually done with a Web-based solution. The CMS allows creation and structuring of websites, without advanced knowledge or training in coding. The most common Web-based CMSs are WordPress, Joomla, Drupal, and TYPO3 [18]. Through the front end of the CMS (the interface for the regular user), users can access the content either directly or through a search function and receive information about upcoming updates and news within the system. Content managers can log in to the system and access the back end, through which content is created, updated, and distributed to the users. Administrators are in charge of system updates, as well as troubleshooting. Multiple users can access the CMS simultaneously, and role-based access control maps users to roles and different levels of permission. The goal of the CMS is to make permission management convenient by grouping users into different roles and enabling them to work within the CMS.

WordPress as a Content Management System

Various CMSs exist, and they can be used to distribute guidelines and SOPs within an institution. We chose WordPress (WordPress Foundation), as it can be implemented independently, is free, can be adapted to individual needs, and is intuitive to use. WordPress is a Web-based app, running on an Apache HTTP Server (Apache Software Foundation) in a Hypertext Preprocessor (PHP) environment (The PHP Group, php.net). It requires database access, such as MySQL (Oracle Corporation) or MariaDB (MariaDB Corporation). In November 2018, WordPress was the most frequently used CMS worldwide (used by 60.7% of all websites whose CMS is public knowledge) [18]. This results in extensive resources and a huge support community.

Implementation Process

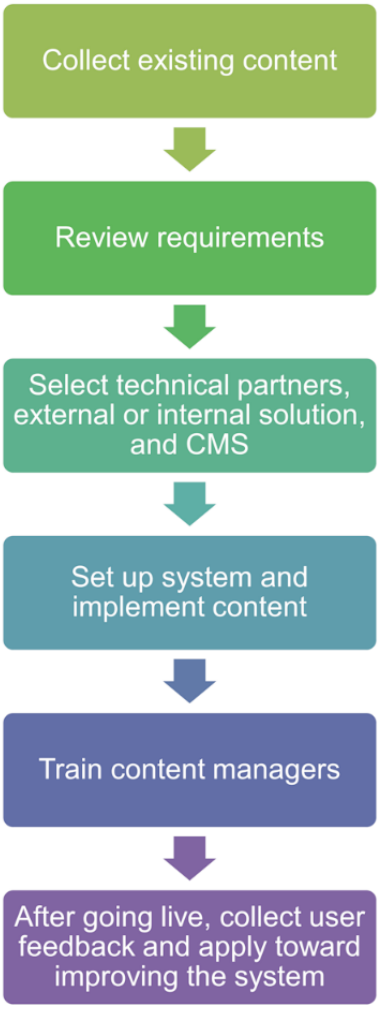
The system's main purpose is the distribution of general and local SOPs and guidelines. The system will never store any kind of patient-related data. An overview of the implementation process is shown in Figure 1.

Before implementation of the system, the existing guidelines and SOPs were collected to gain an overview of the existing content. After that, we performed a review of requirements (Textbox 1).

Depending on the technical knowledge available within a department, a technical partner for the system implementation may be needed. The system can either be run internally within the institution or externally on a public server that is accessible worldwide. This ultimately depends on the sensitivity of the information published and the technical possibilities within an institution. Both solutions offer advantages and disadvantages (Table 1).

XSL•FO
RenderX

Figure 1. Process of implementing the content management system in an institution. CMS: content management system.



Textbox 1. System requirements.

Requirements for the back end.

- The system must have a centralized database, which can be accessed through different devices
- The system can be managed and developed simultaneously by different users. Personnel shortages must not impede the performance
- Drafting, commenting, and revising content must be possible before publishing
- All changes and revisions must be trackable
- Data must be backed up on a regular basis
- Updates and maintenance must be performed regularly

Requirements for the front end.

- The user must be able to see which contents were updated and when
- Users must be able to find the desired content using simple navigation or a fast-performing search function
- There should be no limitations as far as devices or operating systems are concerned

XSL•FO
RenderX

Table 1. Advantages and disadvantages of external and internal systems

System type	Advantages	Disadvantages
External system	• Accessible on all devices at any time • Accessible from home and while on the go • Publicity for the institution	• External costs might apply • Data might not be stored internally • Security risks might arise, especially if updates are not performed regularly
Internal system	• Security and privacy • Internal support systems are available	• Depending on the technical settings, an internal setup might be complex • Content can only be accessed through internal devices. Usually, access from home or on the go is not possible

After we had chosen a hosting platform and decided to develop an internal system (Bern) and an external system (Lucerne), the systems could be set up. Given the open-source character of the software used, documentation was easily available. An operational concept with specified roles and responsibilities was drafted, and this concept was approved by the head of the department. Requirements pertaining to availability and security were defined.

Content managers (ie, users in charge of different sections, such as attending specialists in charge of certain anesthesia divisions) needed training. One of the big advantages of such a system is that content management can be delegated to a number of employees within the department. This precludes a bottleneck that could develop if only a single user or a few users are in charge of content management. Finally, an in-depth analysis of metadata and usage was performed.

Statistical Analysis and Metadata

Metadata generated by user access were recorded with either WP Statistics (Verona Labs) or Visitor Statistics Pro. Data were imported and analyzed using Microsoft Excel (Version 2016) or Sigmaplot 13.0 (Systat Software). Continuous variables were expressed as means (SD); categorical variables were presented as frequencies and percentages.

Results

Implementation and Content

The system was implemented in 2 tertiary-level hospitals in Switzerland: first, it was implemented in the Department of Anesthesiology and Pain Medicine at the Kantonsspital Lucerne in Lucerne, Switzerland, and second, it was implemented in the Department of Anaesthesiology and Pain Medicine of the Inselspital, Bern University Hospital, in Bern, Switzerland. The content was divided into various sections (eg, clinical anesthesia, regional anesthesia, and airway management), and each section was overseen by a senior specialist. This included creating new content, as well as updating existing pages. Content was reviewed at least once a year. Users could report directly to the senior specialist in charge if they noticed a need for changes or for implementation of new content. Content could be navigated using a menu bar or accessed directly through a search function. The most frequently accessed content was guidelines on regional anesthesia, followed by various SOPs used with clinical anesthesia.

Department of Anesthesiology and Pain Medicine, Kantonsspital Lucerne

WordPress was set up on an external server, providing worldwide access to the department's SOPs and guidelines (Figure 2). The system went live in May 2014. The external server is hosted by a Swiss hosting company, and it provides the necessary infrastructure, such as the latest versions of PHP and MySQL, and it includes a preinstalled version of WordPress. There is no need for manual app setup or updates, apart from WordPress itself. Implementation took place over a period of roughly 1 year. This involved a requirements analysis, the collection of already existing guidelines and SOPs, the setting up of a test website, and the gradual transfer of the content to the CMS. There are 269 pages of content within the CMS. These pages are divided into clinical SOPs (eg, SOPs for neuroanesthesia, cardiac anesthesia; 196 pages), regional anesthesia guidelines (13 pages), emergency guidelines (7 pages), airway guidelines (6 pages), patient management guidelines (eg, patients with diabetes, kidney disease; 18 pages), guidelines on drugs (10 pages), SOPs for monitoring (6 pages), and checklists (13 pages). In the 365 days ending on November 1, 2018, there were 155,379 visitors to the website, corresponding to 1,057,492 website requests. As the content is in German, the website is primarily accessed by people in German-speaking countries (73,136 visitors from Germany, 41,262 visitors from Switzerland, and 9852 visitors from Austria). Access to the website was primarily gained using iPhone (113,692 visitors, 35.80%) or Windows (113,213 visitors, 35.28%; Table 2). There are 44 registered users involved in managing the content. The cost of this setup is minimal. Other than buying a domain name (CHF 70) and paying for a hosting service (CHF 100 per year), there were no financial investments. This did not include the time invested by the department's employees. Roughly 300 to 400 hours were needed for system setup and 50 hours per year for system maintenance. This resulted in an overall cost of less than CHF 1000 for the entire project. The hosting company updates all server applications on a regular basis. The administrators perform WordPress core updates and updates of all installed plug-ins multiple times per year. This ensures that security flaws are promptly fixed. Furthermore, we ran All In One WP Security (Tips and Trick HQ), which protects the website from unwarranted access, with an additional firewall function. No patient-relevant data were stored on the website. Content was backed up on a weekly basis, and backups were kept for half a year.

XSL•FO
RenderX

Figure 2. Screenshot of the content management system running in Lucerne.

Table 2. Operating systems most commonly used to access the Lucerne Hospital's hosting system.

Operating system	Overall use, %
Windows	35.28
iPhone	34.80
Android	7.11
iPad	6.81
Macintosh	11.01
Linux	0.99
Unknown	3.99

Department of Anaesthesiology and Pain Medicine, Inselspital, Bern University Hospital

The system was set up on an internal departmental server and went live in March 2018 (Figure 3). External access is impossible, as the system operates in a separate domain behind the hospital firewall. The infrastructure runs on a Windows Server (Version 2012 R2); Apache HTTP Server, PHP, MariaDB, and WordPress were installed manually. The entire setup process was documented. As installations are maintained manually and the systems need constant development, the setup comprises a test server, as well as a live server. All updates and major changes to the system are first established within the test environment before going live. The project started in March 2017, and realization was possible within 12 months. Content is presented on 341 pages managed by 35 active users. Content is divided into the following sections: in-hospital SOPs (eg, local phone numbers, operating room schedule, and hygiene; 16 pages), clinical anesthesia (236 pages), patient management (18 pages), airway guidelines (6 pages), guidelines on monitoring (3 pages), regional anesthesia (8 pages), SOPs on pain therapy (41 pages), guidelines for drugs (6 pages), and SOPs for postanesthesia care (7 pages).

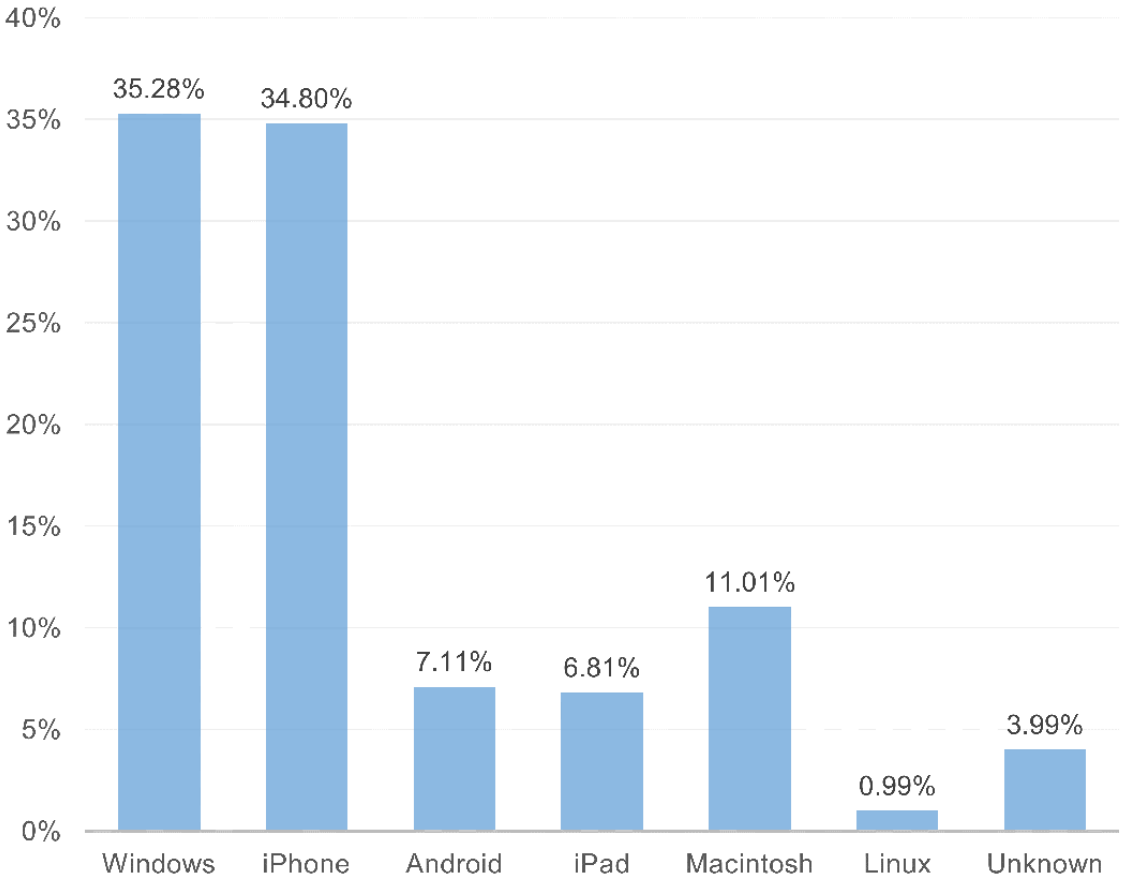
Since going live, the system has been visited 13,856 times, corresponding to 45,284 page views (3.27 page views per visit). Users access the website predominantly from outside the operating room or patient care, with only 25% of users using computers positioned directly at the anesthesia station and 75 % using computers outside of the operating room. As the system is locked behind the department's firewall, virtual private network was not provided for mobile phones, and no mobile phone access was possible. Hourly usage peaked twice daily, between 9 am and 10 am, with 8.4 (SD 5.9) visitors, and between 3 pm and 4 pm, with 8.7 (SD 6.1) visitors (Figure 4). As there are no expenses for external servers, the system has not produced any expenses, except for the time which system developers and users dedicated to the platform.

As the system runs on an internal platform with in-hospital access only, digital attacks from the World Wide Web are not possible, and the website is protected by the corporate firewall. However, all apps and plug-ins are updated regularly, which ensures that security holes are closed. The system does not

XSL•FO
RenderX

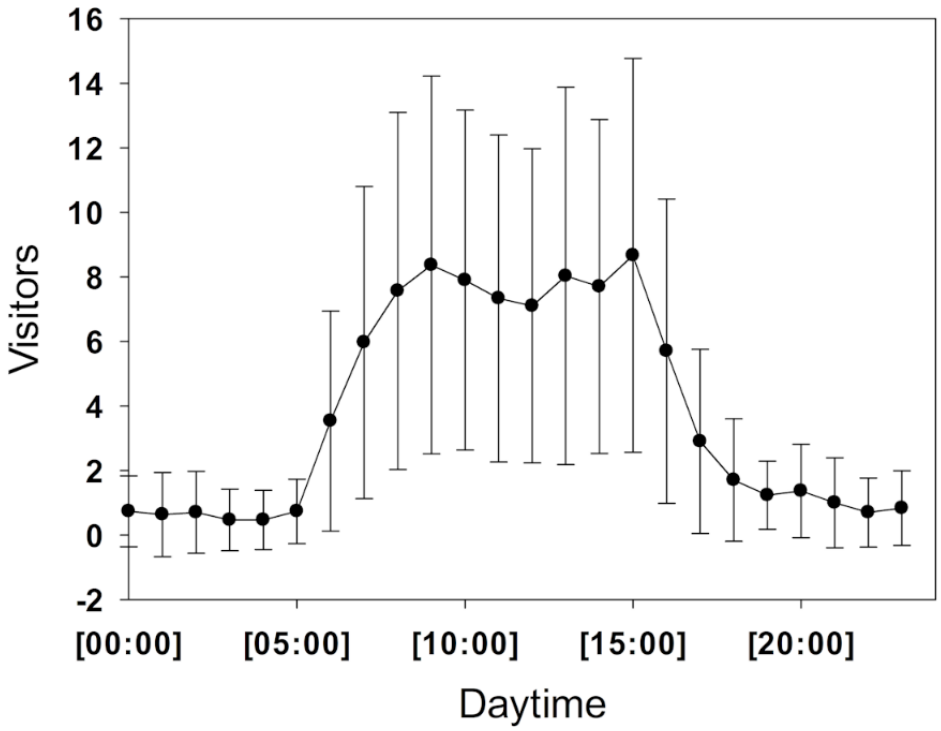
contain any patient-relevant information, and it is backed up daily. System security was discussed with the information technology department.

Figure 3. Screenshot of the internal content management system running in Bern.



XSL•FO
RenderX

Figure 4. Hourly usage of the internal content management system for October and November 2018 at the Bern University Hospital. Dots represent means and whiskers represent SD.



Discussion

Principal Findings

The implementation process described in this paper has been successfully performed twice in large anesthesia departments in Switzerland. Although the process requires some technological knowledge or at least an appropriate technical partner, setup itself is easy, and it can be done within a reasonable period of time. There are various CMS options available, and CMSs have been shown to be an ideal tool for efficient and consistent management of guidelines [16]. Our detailed requirement review allowed us to choose an appropriate CMS and showed that WordPress would be superior to regular wiki-based systems. Metadata generated by the systems show high usage. Use during the day reflects the active hours in our operating rooms. However, as the database is primarily accessed independent of direct patient care, it is safe to assume that its primary use is as a reference book. As a limitation to our study, direct comparison of the systems is not possible, as the Lucerne system is hosted externally, and its metadata are skewed by worldwide access. However, the statistics from Lucerne show that there is international interest in these guidelines and SOPs.

Content Management Systems in the Health Care Sector and Digital Maturity

WordPress has already been used as an electronic portfolio system [19], as a platform for the dissemination of evidence-based medicine [20], and as a centralized in-hospital database to share and distribute information, with low costs [21]. However, it has never been applied as a distribution system for guidelines and SOPs in anesthesia or emergency care. With WordPress, content can be easily managed by our senior staff, who are able to create and update pages. It is a major advantage that these content managers are anesthesiologists working in clinical practice on a daily basis and are thus in close contact with the users. This ensures a direct feedback loop and prompts implementation of new content or updates to existing content. The possibility of having multiple users manage content simultaneously prevents the system from being dependent on a single person or a small group, as might be the case with an individually coded app (eg, iOS or Android). WordPress is the most commonly used CMS worldwide, appreciated for its flexibility and features. However, this also makes it a target for security breaches and attacks. The best way to protect a system is by always running the latest stable release [22]. To avoid problems, we do not keep any sensitive information, especially patient-related data, in our databases. The modernization of the health care sector in Western Europe was reported to increase spending in technology and informatics from US $13.2 billion in 2013 to US $14.6 billion in 2018 [23]. In an assessment of Swiss hospitals, investment in hardware and software was found to be the most promising way to improve digital maturity of a health care organization [23]. The concept of digital maturity represents the ability to respond to changing needs and challenges in a computerized world. A strong link was found between usage intensity and digital maturity. The implementation of a Web-based CMS might be seen as an investment in software solutions, and therefore, it might be seen as an improvement of the digital maturity of the 2 corresponding health care organizations. It seems reasonable to record usage intensity over time to assess whether these apps might increase digital maturity through wider usage.

XSL•FO
RenderX

Mobile Technology Enhancing Accessibility and Usability

Our metadata concerning operating systems show that mobile phones are frequently used to access our guidelines and SOPs. Mobile learning has been shown to be as effective as traditional learning [15], suggesting that mobile phones apps are a viable tool to access learning and reference materials. The increasing availability of these mobile phones devices and the possibility of accessing the content on the go suggest that there could be a further peak in usage if the internal system in Bern is opened up for mobile phones access. The effect on patient care remains unclear, but usage patterns suggest that electronic access to guidelines is highly appreciated. Mobile phones apps are an emerging tool in health care [24-27], and they are leading to new possibilities, such as applications in patient management, resource distribution, and quality control [28]. Our Web-based system and most of the apps available fall in the category of patient management [28-30]. Some trials investigating the effect of mobile technology have shown that it significantly improves outcomes related to disease management [29].

Improving Outcome and Cost-Effectiveness

Studies concerned with the implementation of digital solutions have shown variable results with regard to guideline adherence [14,31-33], but the scientific data on this topic are limited. To further develop the CMS to meet the needs of our users, we intend to assess the effect of the implemented systems through frequent administration of questionnaires, and if feasible, an outcome-related study will be considered. Owing to its high usage, ease of use, and low cost, our Web-based repository for health care guidelines and operating procedures could potentially contribute to the digitization of other health care organizations with needs similar to ours [34]. Solutions that can be developed rapidly and implemented easily may be crucial for the survival of organizations in the health care landscape [34].

Conclusions

We have demonstrated that WordPress is a suitable solution for distributing and managing the internal SOPs and guidelines of 2 tertiary anesthesia departments. Although our study was performed solely in anesthesia departments, implementation in different areas of health care seems feasible. Metadata allow live monitoring and feedback. The systems are cost effective and can be handled from within the department, without depending on third-party support.

Acknowledgments

The authors wish to thank Jeannie Wurz, medical writer and editor, Department of Anaesthesiology and Pain Medicine, Inselspital, for the careful editing of the manuscript. This research did not receive funding from agencies in the public, commercial, or not-for-profit sectors.

Authors' Contributions

KFB drafted the manuscript, developed the project idea, and implemented the system. CV was responsible for conceptual design and content. LW is a technical partner, and LW helped implement the system. CK is head of the department in Lucerne, and sop.klifairs.ch was developed under his guidance. APV was the senior staff member supervising the project in Bern, and APV revised the manuscript. All authors have read and approved the final version of this manuscript.

Conflicts of Interest

None declared.

References

1. Schuster F, Johannsen S, Roewer N. [Helsinki declaration on patient safety in anaesthesiology - part 3: SOP for malignant hyperthermia]. Anasthesiol Intensivmed Notfallmed Schmerzther 2013 Mar;48(3):162-164. [doi: 10.1055/s-0033-1342900] [Medline: 23592246]
2. Wiesmann T, Bornträger A, Steinfeldt T, Wulf H. [Helsinki declaration on patient safety in anaesthesiology -- SOP for local anesthetics intoxication]. Anasthesiol Intensivmed Notfallmed Schmerzther 2013 Jan;48(1):32-35. [doi: 10.1055/s-0032-1333075] [Medline: 23364822]
3. Happel O, Roewer N, Kranke P. [Helsinki declaration on patient safety in anaesthesiology--part 8: SOP for checking equipment and drugs]. Anasthesiol Intensivmed Notfallmed Schmerzther 2013 Sep;48(9):538-542. [doi: 10.1055/s-0033-1355233] [Medline: 24048662]
4. Nachtigall I, Deja M, Tafelski S, Tamarkin A, Schaller K, Halle E, et al. Adherence to standard operating procedures is crucial for intensive care unit survival of elderly patients. J Int Med Res 2008;36(3):438-459. [doi: 10.1177/147323000803600308] [Medline: 18534125]
5. Nachtigall I, Tamarkin A, Tafelski S, Deja M, Halle E, Gastmeier P, et al. Impact of adherence to standard operating procedures for pneumonia on outcome of intensive care unit patients. Crit Care Med 2009 Jan;37(1):159-166. [doi: 10.1097/CCM.0b013e3181934f1b] [Medline: 19050600]

XSL•FO
RenderX

6. Gutsche JT, Kornfield ZN, Speck RM, Patel PA, Atluri P, Augoustides JG. Impact of guideline implementation on transfusion practices in a surgical intensive care unit. J Cardiothorac Vasc Anesth 2013 Dec;27(6):1189-1193. [doi: 10.1053/j.jvca.2013.05.040] [Medline: 24064204]
7. Rognås L, Hansen TM, Kirkegaard H, Tønnesen E. Standard operating procedure changed pre-hospital critical care anaesthesiologists' behaviour: a quality control study. Scand J Trauma Resusc Emerg Med 2013 Dec 5;21:84 [FREE Full text] [doi: 10.1186/1757-7241-21-84] [Medline: 24308781]
8. de Treville S, Antonakis JM, Edelson NM. Can standard operating procedures be motivating? Reconciling process variability issues and behavioural outcomes. Total Qual Manag Bus 2010 Dec 3;16(2):231-241 [FREE Full text] [doi: 10.1080/14783360500054236]
9. Woolf SH, Grol R, Hutchinson A, Eccles M, Grimshaw J. Clinical guidelines: potential benefits, limitations, and harms of clinical guidelines. Br Med J 1999 Feb 20;318(7182):527-530 [FREE Full text] [doi: 10.1136/bmj.318.7182.527] [Medline: 10024268]
10. Haller G, Stoelwinder J, Myles PS, McNeil J. Quality and safety indicators in anesthesia: a systematic review. Anesthesiology 2009 May;110(5):1158-1175. [doi: 10.1097/ALN.0b013e3181a1093b] [Medline: 19352148]
11. Chazapis M, Gilhooly D, Smith AF, Myles PS, Haller G, Grocott MP, et al. Perioperative structure and process quality and safety indicators: a systematic review. Br J Anaesth 2018 Jan;120(1):51-66 [FREE Full text] [doi: 10.1016/j.bja.2017.10.001] [Medline: 29397138]
12. Kvedar J, Coye MJ, Everett W. Connected health: a review of technologies and strategies to improve patient care with telemedicine and telehealth. Health Aff (Millwood) 2014 Feb;33(2):194-199. [doi: 10.1377/hlthaff.2013.0992] [Medline: 24493760]
13. Neubeck L, Lowres N, Benjamin EJ, Freedman SB, Coorey G, Redfern J. The mobile revolution--using smartphone apps to prevent cardiovascular disease. Nat Rev Cardiol 2015 Jun;12(6):350-360. [doi: 10.1038/nrcardio.2015.34] [Medline: 25801714]
14. McEvoy MD, Hand WR, Stoll WD, Furse CM, Nietert PJ. Adherence to guidelines for the management of local anesthetic systemic toxicity is improved by an electronic decision support tool and designated 'Reader'. Reg Anesth Pain Med 2014;39(4):299-305 [FREE Full text] [doi: 10.1097/AAP.0000000000000097] [Medline: 24956454]
15. Dunleavy G, Nikolaou CK, Nifakos S, Atun R, Law GC, Tudor CL. Mobile digital education for health professions: systematic review and meta-analysis by the digital health education collaboration. J Med Internet Res 2019 Feb 12;21(2):e12937 [FREE Full text] [doi: 10.2196/12937] [Medline: 30747711]
16. Reddy S, Herring S, Gray A. Identifying an appropriate content management system to develop clinical practice guidelines: a perspective. Health Informatics J 2017 Dec;23(1):14-34. [doi: 10.1177/1460458215616264] [Medline: 26635323]
17. Barwise A, Garcia-Arguello L, Dong Y, Hulyalkar M, Vukoja M, Schultz MJ, et al. Checklist for early recognition and treatment of acute illness (CERTAIN): evolution of a content management system for point-of-care clinical decision support. BMC Med Inform Decis Mak 2016 Dec 3;16(1):127 [FREE Full text] [doi: 10.1186/s12911-016-0367-3] [Medline: 27716243]
18. W3Techs - Extensive and Reliable Web Technology Surveys. Usage Statistics and Market Share of Wordpress URL: https://w3techs.com/technologies/details/cm-wordpress/all/all [accessed 2019-04-24] [WebCite Cache ID 77rz0LEJg5D]
19. Avila J, Sostmann K, Breckwoldt J, Peters H. Evaluation of the free, open source software WordPress as electronic portfolio system in undergraduate medical education. BMC Med Educ 2016 Jun 3;16:157 [FREE Full text] [doi: 10.1186/s12909-016-0678-1] [Medline: 27255920]
20. Dyson MP, Newton AS, Shave K, Featherstone RM, Thomson D, Wingert A, et al. Social media for the dissemination of cochrane child health evidence: evaluation study. J Med Internet Res 2017 Dec 1;19(9):e308 [FREE Full text] [doi: 10.2196/jmir.7819] [Medline: 28864427]
21. Kotoulas A, Stratis I, Goumenidis T, Lambrou G, Koutsouris DD. Short-term adoption rates for a web-based portal within the intranet of a hospital information system. BMJ Health Care Inform 2019 Apr;26(1):1-4 [FREE Full text] [doi: 10.1136/bmjhci-2019-100004] [Medline: 31039127]
22. Rosso S, Abrahamson B, Adams M, Cave J, HouSandí H, Hulse D, et al. WordPress. 2015. Security URL: https://wordpress.org/about/security/
23. Mettler T, Pinto R. Evolutionary paths and influencing factors towards digital maturity: an analysis of the status quo in Swiss hospitals. Technol Forecast Soc Change 2018 Aug;133:104-117 [FREE Full text] [doi: 10.1016/j.techfore.2018.03.009]
24. Gupta RK, McEvoy MD. Initial experience of the American Society of Regional Anesthesia and Pain Medicine COAGS regional smartphone application: a novel report of global distribution and clinical usage of an electronic decision support tool to enhance guideline use. Reg Anesth Pain Med 2016;41(3):334-338. [doi: 10.1097/AAP.0000000000000391] [Medline: 27076369]
25. Machado GC, Pinheiro MB, Lee H, Ahmed OH, Hendrick P, Williams C, et al. Smartphone apps for the self-management of low back pain: a systematic review. Best Pract Res Clin Rheumatol 2016 Dec;30(6):1098-1109. [doi: 10.1016/j.berh.2017.04.002] [Medline: 29103552]
26. 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]

XSL•FO
RenderX

27. Haskins BL, Lesperance D, Gibbons P, Boudreaux ED. A systematic review of smartphone applications for smoking cessation. Transl Behav Med 2017 Dec;7(2):292-299 [FREE Full text] [doi: 10.1007/s13142-017-0492-2] [Medline: 28527027]
28. Olivero E, Bert F, Thomas R, Scarmozzino A, Raciti IM, Gualano MR, et al. E-tools for hospital management: an overview of smartphone applications for health professionals. Int J Med Inform 2019 Dec;124:58-67. [doi: 10.1016/j.ijmedinf.2019.01.010] [Medline: 30784427]
29. Free C, Phillips G, Watson L, Galli L, Felix L, Edwards P, et al. The effectiveness of mobile-health technologies to improve health care service delivery processes: a systematic review and meta-analysis. PLoS Med 2013;10(1):e1001363 [FREE Full text] [doi: 10.1371/journal.pmed.1001363] [Medline: 23458994]
30. Thomairy NA, Mummaneni M, Alsalamah S, Moussa N, Coustasse A. Use of smartphones in hospitals. Health Care Manag (Frederick) 2015;34(4):297-307. [doi: 10.1097/HCM.0000000000000080] [Medline: 26506291]
31. Semler MW, Weavind L, Hooper MH, Rice TW, Gowda SS, Nadas A, et al. An electronic tool for the evaluation and treatment of sepsis in the ICU: a randomized controlled trial. Crit Care Med 2015 Aug;43(8):1595-1602 [FREE Full text] [doi: 10.1097/CCM.0000000000001020] [Medline: 25867906]
32. Kessler ME, Carter RE, Cook DA, Kor DJ, McKie PM, Pencille LJ, et al. Impact of electronic clinical decision support on adherence to guideline-recommended treatment for hyperlipidaemia, atrial fibrillation and heart failure: protocol for a cluster randomised trial. BMJ Open 2017 Dec 4;7(12):e019087 [FREE Full text] [doi: 10.1136/bmjopen-2017-019087] [Medline: 29208620]
33. Sim EY, Tan DJ, Abdullah HR. The use of computerized physician order entry with clinical decision support reduces practice variance in ordering preoperative investigations: a retrospective cohort study. Int J Med Inform 2017 Dec;108:29-35. [doi: 10.1016/j.ijmedinf.2017.09.015] [Medline: 29132628]
34. van Velthoven MH, Cordon C, Challagalla G. Digitization of healthcare organizations: the digital health landscape and information theory. Int J Med Inform 2019 Dec;124:49-57. [doi: 10.1016/j.ijmedinf.2019.01.007] [Medline: 30784426]

Abbreviations

CMS: content management system
SOP: standard operating procedure

Edited by G Eysenbach; submitted 24.04.19; peer-reviewed by X Garcia-Eroles, R Lee; comments to author 06.06.19; revised version received 28.06.19; accepted 29.06.19; published 15.08.19.

Please cite as:
Bachmann KF, Vetter C, Wenzel L, Konrad C, Vogt AP
Implementation and Evaluation of a Web-Based Distribution System For Anesthesia Department Guidelines and Standard Operating Procedures: Qualitative Study and Content Analysis
J Med Internet Res 2019;21(8):e14482
URL: https://www.jmir.org/2019/8/e14482/
doi:10.2196/14482
PMID:31418427

©Kaspar F Bachmann, Christian Vetter, Lars Wenzel, Christoph Konrad, Andreas P Vogt. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 15.08.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

Adherence Reporting in Randomized Controlled Trials Examining Manualized Multisession Online Interventions: Systematic Review of Practices and Proposal for Reporting Standards

Ina Beintner[1*], PhD; Bianka Vollert[1*], Dipl Psych; Anna-Carlotta Zarski[2], PhD; Felix Bolinski[3], MSc; Peter Musiat[4], PhD; Dennis Görlich[5], PhD; David Daniel Ebert[2], PhD; Corinna Jacobi[1], PhD

[1]Faculty of Psychology, School of Science, Technische Universität Dresden, Dresden, Germany

[2]Institute of Psychology, Faculty of Humanities, Social Sciences, and Theology, Friedrich-Alexander-University Erlangen-Nürnberg, Erlangen, Germany

[3]Department of Clinical, Neuro- and Developmental Psychology, Faculty of Behavioural and Movement Sciences, Vrije Universiteit Amsterdam, Amsterdam, Netherlands

[4]Department of Psychological Medicine, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, United Kingdom

[5]Institute of Biostatistics and Clinical Research, Faculty of Medicine, Westfälische Wilhelms-Universität Münster, Münster, Germany

[*]these authors contributed equally

Corresponding Author:
Ina Beintner, PhD
Faculty of Psychology
School of Science
Technische Universität Dresden
Chemnitzer Straße 46
Dresden, 01062
Germany
Phone: 49 351 463 ext 37469
Fax: 49 351 463 37208
Email: mail@ina-beintner.de

Abstract

Background: Adherence reflects the extent to which individuals experience or engage with the content of online interventions and poses a major challenge. Neglecting to examine and report adherence and its relation to outcomes can compromise the interpretation of research findings.

Objective: The aim of this systematic review is to analyze how adherence is accounted for in publications and to propose standards for measuring and reporting adherence to online interventions.

Methods: We performed a systematic review of randomized controlled trials on online interventions for the prevention and treatment of common mental disorders (depression, anxiety disorders, substance related disorders, and eating disorders) published between January 2006 and May 2018 and indexed in Medline and Web of Science. We included primary publications on manualized online treatments (more than 1 session and successive access to content) and examined how adherence was reported in these publications.

Results: We identified 216 publications that met our inclusion criteria. Adherence was addressed in 85% of full-text manuscripts, but only in 31% of abstracts. A median of three usage metrics were reported; the most frequently reported usage metric (61%) was intervention completion. Manuscripts published in specialized electronic health journals more frequently included information on the relation of adherence and outcomes.

Conclusions: We found substantial variety in the reporting of adherence and the usage metrics used to operationalize adherence. This limits the comparability of results and impedes the integration of findings from different studies. Based on our findings, we propose reporting standards for future publications on online interventions.

(J Med Internet Res 2019;21(8):e14181) doi:10.2196/14181

XSL•FO
RenderX

KEYWORDS

adherence; compliance; usage; attrition; ehealth; e-mental health; mental health; behavior change; reporting standards; CONSORT eHealth; review

Introduction

Online interventions have become popular in the prevention and treatment of mental disorders, and they have been shown to be effective in clinical trials for a wide range of common mental disorders [1-6]. These interventions typically include multiple interactive self-help lessons to improve mental health by using established psychotherapeutic techniques. These lessons can be delivered via a Web browser or mobile app [7].

However, the behavior changes that online interventions aim to induce are very unlikely to occur if participants expose themselves to the intervention only briefly or do not do so at all. Adherence can be conceptualized as the extent to which individuals experience or engage with the content of an online intervention [8]. Poor adherence is a major issue in almost all these interventions [9] and even more so if they are unguided [10].

Attention to adherence has increased over time. However, even 13 years after Eysenbach's landmark paper "The law of attrition" [11], referring to the finding that a significant proportion of participants in electronic health (eHealth) research do not fully use the studied technology, adherence is still not consistently and systematically examined and reported in studies of online interventions. Additionally, operationalization of adherence varies substantially across trials [12], limiting the comparability of results between trials. However, neglecting to examine adherence and its impact on outcomes in online intervention trials can compromise the interpretation of research findings, and, in turn, lead to inappropriate recommendations and decisions regarding the use and implementation of such interventions.

If an intervention is not effective even though the participants have been using it the way they were supposed to, it is very likely that the intervention itself is not working and that the core components of the intervention need to be changed or that there is a mismatch between user needs and intervention components. If, however, the intervention is not effective while people are not exposing themselves to a sufficient "dose" of it, implications for further research might be quite different. A mismatch between user needs and the intervention or its components is likely in these cases. Poor adherence may then lead to systematic underestimation of the potential intervention effects. Instead of changing core components of the intervention that teach skills and prompt change in behaviors related to mental health, the intervention may need to be augmented with components to improve adherence, or recruitment strategies may need to be changed to reach those who are open to actually using the interventions. Thus, it is vital to study both intervention effects of and adherence to online interventions and their interactions simultaneously. Furthermore, it is important to identify differential usage patterns in multicomponent interventions and user characteristics that are associated with these patterns [13-15]. In order to achieve this, multiple usage metrics are needed [15].

Although adherence has received increasing attention in the study of online interventions and been addressed in existing reporting guidelines [16], the field is still lacking common standards for addressing and reporting adherence. Various terms, definitions, and measures have been used to describe how users engage with online interventions. Some terms have been adopted from the field of pharmacotherapy [17] and others, from guidelines to describe participant flow in clinical research trials [16,18]. Although the term adherence is widely used, some authors also use the terms compliance, (session) attendance, engagement, user retention, persistence, exposure, intervention usage, or polar opposite terms—attrition or (treatment or study) dropout. Even when authors use the same term, they do not necessarily mean the same thing. For example, the term dropout can either refer to the premature cessation of treatment (treatment dropout) or the noncompletion of postintervention assessments (study dropout), although some investigators equate both [19]. In a similar way, the term attrition is used to refer to the loss of participants in the intervention of a (clinical or epidemiological) trial. For trials examining online interventions, it has been proposed that "nonusage attrition" (comparable to treatment dropout) and "dropout attrition" (comparable to study dropout) should be distinguished from each other, and it has been postulated that these forms of attrition are related to each other [11]. For this review, we chose the term adherence as an umbrella term for describing how participants use online interventions, because this term implies that they have to actively engage with an intervention [20]; the term can also be used outside of clinical trials.

In addition to a variety of terms that describe how users adhere to interventions, there are also numerous ways to measure adherence. Most online intervention platforms store log data (eg, time spent on the intervention page), which allow us to track if and how users interact with the intervention. Despite some shortcomings of this data collection method, such as not knowing with certainty whether the person who used the program was the same person who signed up, or whether a user actually engaged with the content or just opened the pages without further engagement, it provides us with objective and comparable usage metrics. However, there is much variety in the usage metrics reported for online interventions (eg, percentage of participants completing all modules/sessions, percentage of participants who visited/revisited websites, average duration of visits, average number of log-ins, and average number of pages visited) [12,21]. In addition, usage related to specific program components can only be reported for interventions that include the component (eg, a discussion board or diaries). In addition, some usage metrics only apply to guided interventions (eg, the number of messages sent to a coach). Thus, the number and type of appropriate usage metrics clearly depend on the design and delivery mode of an intervention [9,22]. In addition, the way usage is tracked and

XSL•FO
RenderX

stored has an impact on which usage data and types of adherence are available afterward. In a recent review on the concept of adherence to health technology, Sieverink et al [12] brought up an additional aspect: In order to define adherence, the intended use of an intervention would have to be both defined and justified by the developer (comparable to the optimum dose of a medication) beforehand. However, we still know very little about the necessary dose of online interventions to achieve optimal outcomes. The authors of the review point out that, all too often, developers of online interventions implicitly assume that their interventions work best if all users expose themselves to all parts of the content, and other patterns of use are rarely considered.

Michie et al [15] have pointed out that engagement with online behavior change interventions is complex, depends on the intervention context, and is not limited to technology usage (adherence; micro level of engagement) but extends to behavior change outside the intervention (macro level of engagement). It has been argued that usage metrics such as the completion of exercises, homework, or diaries (where an active input from the user is required) might be linked more closely to intervention outcome rather than measures reflecting passive consumption of content [14]. Macro level engagement is very specific to the behavior change intended with an intervention and likely more complex than engagement at the micro level. Thus, quantitative measures of macro level engagement will always be specific to the type of intervention. Quantitative measures for micro level engagement or adherence on the other hand reflect the structure of an intervention rather than its content and goals. They can therefore be harmonized across interventions. These measures can also be utilized to identify usage patterns.

Although higher adherence has been shown to be linked to larger intervention effects in numerous trials [23-28], other trials [29-33] found no impact of adherence on outcomes. Heterogeneous methods for measuring and reporting adherence as well as examining the dose-response relationship between adherence and outcome and neglecting to consider differential usage patterns may contribute to these conflicting findings.

The CONSORT-eHEALTH (Consolidated Standards of Reporting Trials of Electronic and Mobile HEalth Applications and onLine TeleHealth) guidelines [16] offer recommendations on how to report usage data for online interventions. It is worth noting that in these guidelines, splitting up reports of research findings from one trial into several publications (eg, main outcomes and adherence) is explicitly discouraged ("salami publication"), and the reporting of information on usage is expected in primary publications. Furthermore, CONSORT-eHEALTH highly recommends to report usage metrics both in the abstract and the results sections and to describe usage parameters in the methods section, including details on what recommendations and instructions were given to the users. Moreover, subgroup analyses including only participants who used the intervention are highly recommended. Discussing bias due to nonusage is considered essential. However, there are no precise recommendations on which usage metrics should be chosen for different types of online interventions (eg, guided vs unguided, single vs multisession), because there was no consensus at the time CONSORT-eHEALTH was first published.

The aim of our systematic review was to analyze how adherence has been addressed and which usage metrics have been reported in primary publications on randomized controlled trials (RCTs) evaluating manualized, multisession online interventions (including interventions that are delivered through mobile devices) for common mental disorders (ie, depression, anxiety, substance use disorders, and eating disorders). Specifically, we examined whether adherence was reported in the abstract, the results section, and the CONSORT flowchart of each publication; which usage metrics were reported; whether usage or adherence were addressed in the discussion section; and whether usage metrics were analyzed in relation to outcome.

Based on our findings, we propose common standards for addressing adherence, including specific recommendations for usage metrics that the existing CONSORT-eHEALTH guidelines are currently not specifying.

Methods

Inclusion and Exclusion Criteria

We included articles that met the following criteria in the review: (1) the article was published in a peer reviewed journal after the publication of Eysenbach's seminal "The Law of attrition" [11], between January 2006 and May 22, 2018; (2) the article reported research on an online intervention targeting a common mental disorder (depression [without bipolar disorder, postpartum depression, and complicated grief], anxiety disorders [without posttraumatic stress disorder and obsessive-compulsive disorder], substance use disorders, and eating disorders); (3) the article reported the main findings from an RCT; (4) the trial examined a manualized, multisession (two or more) online intervention (Web- or app-based); (5) participants received sequential access to the intervention content; (6) the intervention taught the participants skills; and (7) the article was written in English or German.

We excluded articles that met the following criteria from the review: (1) The article described research on an online intervention targeting common mental disorders in patients with a primary somatic disorder (eg, diabetes or cancer) or an online intervention targeted at carers or parents of patients; (2) the trial examined a highly individualized intervention without common core content; (3) the trial examined a blended intervention, and (4) the trial examined an intervention purely based on text messaging, email, online discussion boards, or online chat groups.

This review has not been preregistered, and the review protocol has not been published.

Search Strategy

We conducted a literature search using the Medline and Web of Science databases. We used the following search terms: "online," "internet," "webbased," "mobile"; "treatment," "psychotherapy," "therapy," "self-help," "prevention," "intervention"; and "depression," "depressive," "anxiety," "phobia," "phobic," "eating disorder," "disordered eating,"

XSL•FO
RenderX

"anorexia," "anorexic," "bulimia," "bulimic," "binge eating," "substance abuse," "substance related disorder," "alcohol," "nicotine," and "cannabis" (Multimedia Appendix 1).

Study Selection

Studies were selected in two steps. First, titles and abstracts were screened by author IB to exclude publications that were clearly out of the scope of the review; did not report studies on interventions targeting common mental disorders; described study protocols, reviews, and meta analyses; reported secondary analyses only; or had not been published in a peer review journal. Second, authors IB, BV, PM, and AZ assessed the remaining full-text articles for eligibility. Each publication was assessed by two authors. We coded the following variables (along the bibliographic data) for each publication: (1) Is this an RCT? (2) Is this an online intervention? (3) Is it manualized? (4) Does it have multiple sequential sessions/modules? (5) Does it teach skills? (6) Is this the main publication? (7) Study registration number.

Data Extraction

Data extraction was conducted in three steps. First, authors BV and IB coded the following variables for all included studies: (1) Is adherence addressed in the abstract? (2) Is adherence addressed in the results section? (3) Is adherence addressed in the CONSORT statement? (4) Is adherence examined in relation to outcome? (5) Is adherence addressed in the discussion section?

In the next step, authors BV and IB extracted the information on adherence reported in the results section and the CONSORT statement. For that purpose, a data extraction form was developed, which captured the following usage metrics: full intervention completion (eg, "XX%/N completed the full intervention" or "XX%/N failed to complete the full intervention"), completion of a set minimum of sessions/modules (eg, "XX%/N completed 6 out of 8 sessions" or "XX%/N completed less than 2 out of 5 sessions"), average number of completed sessions/modules, specified point of intervention dropout (last opened session or module, sometimes illustrated by a graph), intervention dropout (not specified, eg, "XX%/N did not complete the intervention"), number of participants who were allocated to the intervention but never logged on, number of times participants logged on, proportion of patients accessing the treatment site per week, total time spent on the program, time spent on the program site per week/per login, and number of participants who completed a survey that is part of the intervention (not just assessment for the clinical trial), number of entries in a diary, number of completed exercises, number of messages to a coach, number of participants who posted to a discussion board, number of participants who accessed a discussion board (without necessarily posting something), number of visits to the discussion board, number of participants who shared diary entries with other participants, average percentage of pages read, and average percentage of screens viewed. The use of this form was piloted and revised in a group meeting (IB, PM, and BV) on a subset of the included articles (N=150). The resulting data extraction form was then used to extract data from the remaining studies (IB, FB, PM, BV, and AZ). Disagreements regarding the coding were discussed between IB and BV until a consensus was reached.

Data Analysis

Data were entered into an Excel spreadsheet, which was then converted into an SPSS (Statistical Package for the Social Sciences) file. Each publication was treated as a separate case. Descriptive analyses were performed using SPSS [computer software] (Version 24.0. IBM Corp, Armonk, NY). Absolute and relative frequencies were used as the primary measures for the adherence reporting and defined usage metrics. We tested differences in adherence reporting and usage metrics between studies published in specialized eHealth journals versus nonspecialized journals. Two-sided Chi-squared tests were used, and P values<.05 were considered to indicate significant differences between the two groups.

Results

We identified 216 publications reporting on the primary outcomes of RCTs investigating online interventions for common mental disorders (Figure 1, Multimedia Appendix 1). Of these, 34 were published in specialized eHealth journals (eg, Journal of Medical Internet Research, Internet Interventions, and Computers in Human Behavior) and 182 were published in nonspecialized journals.

Interventions for depression (n=73) were evaluated most frequently, followed by interventions for anxiety disorders (n=65). Substance use disorders (n=27) and eating disorders (n=24) were targeted less frequently. Transdiagnostic interventions were investigated in 27 trials. The majority of trials evaluated interventions aiming to treat mental disorders (n=177); prevention (n=34) and aftercare (n=5) interventions were less frequently investigated.

Table 1 provides details on how adherence was reported in the overall sample of publications as well as how publications from specialized eHealth journals differed from publications in nonspecialized journals. Adherence was not reported at all in 28 publications (13%). The majority of publications (83.3%) included information on adherence in the results section, while less than half (41.2%) included information in the CONSORT statement. Adherence was addressed in the discussion in most publications (69.4%). Approximately one in three publications (30.1%) included information on adherence in the abstract.

Although roughly one in two publications in specialized eHealth journals included information on how adherence was related to outcomes, this was only true for one in four publications in nonspecialized journals (χ^2_1=4.6, P=.30). In 25 publications (11.6%), correlation or regression analyses were used to investigate the relationship between program usage and outcomes. In 23 publications (10.6%), results of a per protocol analysis that included only participants who completed a preset minimum of the intervention were reported. In eleven publications (5.1%), comparisons between treatment completers and noncompleters or high and low adherers were reported.

In the 188 publications that contained information on adherence, a median of three usage metrics was reported. Of the total of

XSL•FO
RenderX

216 publications, 23 (10.6%) included one metric, 46 (21.3%) included two metrics, 56 (25.9%) included three metrics, and 63 (29.2%) included four or more metrics. Most metrics were related to session/module progression, while the usage of additional intervention features (eg, diaries, discussion boards, and messaging) was rarely reported. The most frequently reported metric was "full intervention completion" (Table 2).

Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses flow diagram. CMHD: common mental health disorders; RCT: randomized controlled trial.

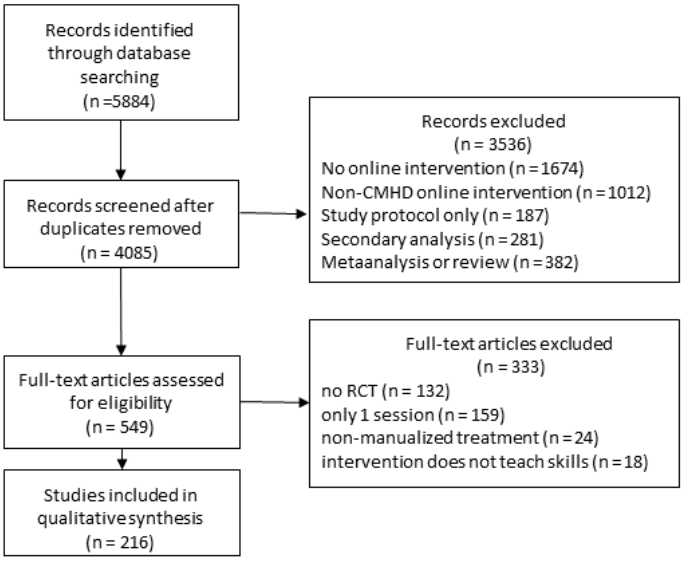


Table 1. Adherence reporting.

Adherence reporting	Overall (N=216), n (%)	Studies published in eHealth[a] journals (n=34), n (%)	Studies published in nonspecialized journals (n=182), n (%)	*P* value
Adherence not addressed at all	28 (13.0)	5 (14.7)	23 (12.6)	.74
Adherence addressed in the abstract	65 (30.1)	12 (35.3)	53 (29.1)	.47
Adherence addressed in the results section	180 (83.3)	29 (85.3)	151 (83.0)	.74
Under own heading	144 (66.7)	24 (70.6)	120 (65.9)	.60
Adherence related to outcome	62 (28.7)	15 (44.1)	47 (25.8)	.03
Adherence addressed in CONSORT[b] chart	89 (41.2)	11 (32.4)	78 (42.9)	.22
Adherence addressed in discussion	150 (69.4)	24 (70.6)	126 (69.2)	.88

[a]eHealth: electronic health.

[b]CONSORT: Consolidated Standards of Reporting Trials.

XSL•FO
RenderX

Table 2. Usage metrics (used in at least 10 publications).

Usage metrics	Overall (N=216), n (%)	Studies published in eHealth[a] journals (n=34), n (%)	Studies published in nonspecialized journals (n=182), n (%)	*P* value
Full intervention completion	127 (58.8)	19 (55.9)	108 (59.3)	.71
Completion of set minimum of sessions/modules	99 (45.8)	21 (61.8)	78 (42.9)	.04
Average number of completed sessions/modules	82 (38.0)	9 (26.5)	73 (40.1)	.13
Specified point of intervention dropout (last opened/completed session/module)	71 (32.9)	12 (35.3)	59 (32.4)	.74
Intervention dropout (point not specified)	28 (13.0)	3 (8.8)	5 (2.7)	.43
Number of participants who were allocated to the intervention, but never logged in	68 (31.5)	13 (38.2)	55 (30.2)	.36
Number of times participants were logged on	18 (8.3)	2 (5.8)	16 (8.8)	.57
Total time spent on program	21 (9.7)	0 (0)	21 (11.5)	.04
Number of entries in a diary	12 (5.6)	3 (8.8)	9 (4.9)	.37
Number of messages to a coach	17 (7.9)	0 (0)	17 (9.3)	.06

[a]eHealth: electronic health.

Discussion

Principal Findings

The aim of this systematic review was to analyze reporting of adherence to online interventions for common mental disorders and to propose recommendations for reporting adherence in future publications. The majority of publications included information on program usage, but in 13% of the publications, adherence was not referred to at all. Adherence was typically addressed in the results section (often under its own heading), less often in the discussion section, and upfront in the abstract in only in one-third of the publications.

In the majority of publications, multiple usage metrics were reported, which is in line with recommendations given in previous reviews on adherence [9]. Most authors reported usage metrics related to session/module progression, with full intervention completion being the usage metric reported most commonly, while use of intervention components (eg, diaries and discussion boards) was reported less often. Our results are similar to those found in a previous review on predictors of adherence by Donkin and colleagues [14], where module completion was, after the number of logins, the usage metric most commonly reported, whereas usage metrics related to specific intervention components were reported less often. In general, we found considerable variability in the type and number of reported usage metrics. Some of the metrics were only used by specific research groups; for example, the number of pages opened/viewed was only used in seven publications on different trials investigating the eating disorder prevention program “StudentBodies” [34].

Some usage metrics might be more useful or valid than others in terms of how they reflect actual usage behavior. For example, measuring time spent in the intervention may not be the most appropriate metric, as it is still not trivial to determine whether recordings are related to actual intervention use or some other activity in another tab of the same browser or even outside the browser. Completion of exercises throughout the online intervention, on the other hand, might be a user metric that can capture deeper content-focused engagement with the intervention.

In our review, only one in four publications addressed adherence in relation to outcome; publications from specialized eHealth journals did so significantly more often than publications from nonspecialized journals. A previous review [14] reported a slightly higher rate (48%) of studies investigating the impact of adherence to outcome. The authors also investigated the impact of usage metrics on the association between adherence and treatment outcome and suggested that the number of logins and completed modules were associated with effectiveness. In our review, we did not aim to evaluate the relationship between adherence and outcome. However, as findings on the adherence-outcome relationship are inconsistent, it is crucial to conduct such analyses in addition to the primary analysis in future studies on eHealth interventions.

Proposal for Reporting Standards

Based on our review and previous reviews, we propose reporting standards regarding adherence and usage metrics (Textbox 1). Most importantly, adherence should be addressed in every publication regarding online interventions (ie, in the main outcome paper). Interventions may include different intervention components, but many have similarities regarding their design, such as multiple (consecutive) sessions or modules. Therefore, while some usage metrics are specific to components (eg, use of a diary), others are universal (eg, average number of completed sessions/modules) and should be reported for every intervention. Utilization and reporting of the same usage metrics across interventions facilitates comparison between interventions and studies and allows pooling of data from multiple studies. Hence, it seems reasonable to include usage metrics that have been used most often in the past (ie, information on completed sessions/modules) and to complement them with additional metrics that are appropriate for the intervention based on its

XSL•FO
RenderX

design and goals. It is key to include detailed information on how adherence was operationalized and how usage metrics were obtained (eg, how "full intervention completion" was defined). Information on adherence should be included in the abstract, the results, and the CONSORT flow chart. Detailed information on user retention should be included in the results section (eg, in a line chart) to illustrate the rate of use of the intervention by participants according to the module/session. Dichotomization of usage metrics (eg, intervention completion vs noncompletion) should be avoided in favor of continuous measures.

The assessment and report of multiple usage metrics is encouraged, as it has several advantages. First, the use of individual components can be investigated; thus, multiple usage patterns can be identified and ultimately be linked to outcome [12,14]. Second, it offers the possibility to create a composite score consisting of multiple adherence measures that might reflect more facets of adherence, in general. However, it is essential to explain how such composite scores were built and what its single components are. Donkin et al [14] suggested a composite measure including time spent online, completion of activities, and other measures related to an active engagement with the program to be a suitable measure of adherence. Third, reporting of different usage metrics facilitates comparison between interventions and studies on multiple dimensions.

If an intervention includes components like diaries, discussion boards, or messaging tools that are considered an essential part of the intervention by the developers, information on the usage of these components should be provided to allow examination of the actual benefit of the component. Participants in interventions with multiple components may exhibit different usage patterns, and adherence measures should reflect this by including use metrics related to those different components. Analyses of multiple usage metrics can extend our knowledge on the most relevant measures (ie, those closest related to outcome) or parameters a composite score for "overall adherence" should contain.

If possible, adherence should be addressed from two perspectives: progress through the intervention and the level of active engagement with the intervention content. This may help distinguish between users who only "consume" the content (eg, read texts and watch videos) and comply with the protocol and those who actively engage with the intervention (eg, write messages, use diaries, and implement behavior chances). It is, however, a challenge to measure this active engagement, which should therefore be a priority for future research to investigate.

The possible impact of adherence on intervention outcomes should be addressed in the discussion. As appropriate, secondary analyses investigating this impact should be undertaken.

Textbox 1. Standards for reporting adherence.

- Address adherence in every publication regarding online interventions.
- Provide details on how adherence was operationalized and how usage metrics were obtained in the methods section.
- Include information on adherence in your abstract.
- Provide detailed information on adherence in the results section and the Consolidated Standards of Reporting Trials flow chart:
  - Include detailed information on user retention per session or module. Avoid only reporting dichotomized data (eg, treatment completers vs noncompleters, low vs high adherers).
  - Include at least the following metrics: average number of completed modules/session and number of participants who were allocated to the intervention but never logged in.
  - In interventions with multiple components, include metrics that reflect those different components.
  - Differentiate between passive components (eg, reading assignments and videos) and active components that require engagement (eg, diaries and discussion board). If possible, report adherence to exercises in daily life between intervention sessions.
  - If appropriate, analyze how adherence is related to outcome.
- Address the possible impact of adherence on intervention outcomes in the discussion section.

Strengths and Limitations

Strengths of this review are the systematic approach and application of independent coding by at least two of the authors, which was employed in every step of the review after initial title and abstract screening. This reduced the risk of selection bias. Our selection procedure led to the inclusion of studies on interventions that are comparable in core design characteristics (eg, multiple sequential sessions). In addition, a large number of publications could be included, reflecting the growth of the field in the past decade.

This review also has some limitations. Initial title and abstract screening was conducted by only one person. Although only publications that were clearly out of the scope of our review were excluded at this step, we cannot completely rule out the possibility that publications were excluded erroneously. Our review is further limited to studies evaluating interventions targeting common mental disorders (depression, anxiety, eating disorders, and substance-related disorders). Interventions targeting other disorders (eg, psychosis or bipolar disorders) were excluded because these conditions are not viewed as common mental disorders. Hence, our findings cannot be generalized across the whole field of electronic mental health research. Moreover, this review cannot draw any conclusions regarding the impact of adherence on outcome (eg, whether there are consistent findings in terms of strength or direction of associations), but examined whether a dose-response relationship was addressed at all in the individual publications. Secondary

XSL•FO
RenderX

publications on adherence were not considered, as the CONSORT-HEALTH guidelines (Eysenbach, 2011 [16]) explicitly suggest having information on adherence in main outcome papers.

Conclusions

In summary, most publications included information on adherence and addressed adherence in the discussion. The most frequently reported usage metric was full intervention completion. There was substantial variety in the usage metrics utilized to operationalize adherence, which impedes comparisons regarding adherence between studies and interventions. Only one in three publications reported on adherence in the abstract. Publications often are screened by abstract and sometimes even evaluated only by the abstract. Results presented in the abstract are more likely to be disseminated by journalists than results presented elsewhere in a manuscript [35]. Hence, to prevent misinterpretation of study results, the abstract should tell the full story, including information on how an intervention was used and how this may have impacted outcomes.

Acknowledgments

The ICare Consortium has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement number 634757.

Conflicts of Interest

None declared.

Multimedia Appendix 1
List of included studies.
[PDF File (Adobe PDF File), 451KB - jmir_v21i8e14181_app1.pdf]

References

1. 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]
2. 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]
3. Ebert DD, Cuijpers P, Muñoz RF, Baumeister H. Prevention of Mental Health Disorders Using Internet- and Mobile-Based Interventions: A Narrative Review and Recommendations for Future Research. Front Psychiatry 2017;8:116 [FREE Full text] [doi: 10.3389/fpsyt.2017.00116] [Medline: 28848454]
4. Josephine K, Josefine L, Philipp D, David E, Harald B. Internet- and mobile-based depression interventions for people with diagnosed depression: A systematic review and meta-analysis. J Affect Disord 2017 Dec 01;223:28-40. [doi: 10.1016/j.jad.2017.07.021] [Medline: 28715726]
5. 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 2016 Mar 12;3:CD011565. [doi: 10.1002/14651858.CD011565.pub2] [Medline: 26968204]
6. 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 Dec;20(1):1-8. [doi: 10.1002/erv.1130] [Medline: 21796737]
7. Ebert DD, Van Daele T, Nordgreen T, Karekla M, Compare A, Zarbo C, et al. Internet- and Mobile-Based Psychological Interventions: Applications, Efficacy, and Potential for Improving Mental Health. European Psychologist 2018 Apr;23(2):167-187. [doi: 10.1027/1016-9040/a000318]
8. 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]
9. Beatty L, Binnion C. A Systematic Review of Predictors of, and Reasons for, Adherence to Online Psychological Interventions. Int J Behav Med 2016 Dec;23(6):776-794. [doi: 10.1007/s12529-016-9556-9] [Medline: 26957109]
10. 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. [doi: 10.1016/j.invent.2014.08.003]
11. 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]
12. Sieverink F, Kelders SM, van Gemert-Pijnen JEWC. Clarifying the Concept of Adherence to eHealth Technology: Systematic Review on When Usage Becomes Adherence. J Med Internet Res 2017 Dec 06;19(12):e402 [FREE Full text] [doi: 10.2196/jmir.8578] [Medline: 29212630]

XSL•FO
RenderX

13. 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]
14. 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]
15. Michie S, Yardley L, West R, Patrick K, Greaves F. Developing and Evaluating Digital Interventions to Promote Behavior Change in Health and Health Care: Recommendations Resulting From an International Workshop. J Med Internet Res 2017 Jun 29;19(6):e232 [FREE Full text] [doi: 10.2196/jmir.7126] [Medline: 28663162]
16. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. J Med Internet Res 2011 Dec 31;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]
17. Horne B, Weinman J, Barber N, Elliott R, Morgan M. NETSCC. London: National Co-ordinating Centre for NHS Service Delivery and Organisation; 2005. Concordance, adherence and compliance in medicine taking URL: http://www.netscc.ac.uk/hsdr/files/project/SDO_FR_08-1412-076_V01.pdf [accessed 2019-08-02]
18. Moher D, Hopewell S, Schulz KF, Montori V, Gøtzsche PC, Devereaux PJ, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. BMJ 2010 Mar;340:c869 [FREE Full text] [Medline: 20332511]
19. Little RJ, D'Agostino R, Cohen ML, Dickersin K, Emerson SS, Farrar JT, et al. The prevention and treatment of missing data in clinical trials. N Engl J Med 2012 Oct 4;367(14):1355-1360 [FREE Full text] [doi: 10.1056/NEJMsr1203730] [Medline: 23034025]
20. Aronson JK. Compliance, concordance, adherence. Br J Clin Pharmacol 2007 Apr;63(4):383-384 [FREE Full text] [doi: 10.1111/j.1365-2125.2007.02893.x] [Medline: 17378797]
21. 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 Jan 06;13(1):e2 [FREE Full text] [doi: 10.2196/jmir.1639] [Medline: 21212045]
22. Morrison LG, Geraghty AWA, Lloyd S, Goodman N, Michaelides DT, Hargood C, et al. Comparing usage of a web and app stress management intervention: An observational study. Internet Interv 2018 Jun;12:74-82 [FREE Full text] [doi: 10.1016/j.invent.2018.03.006] [Medline: 30135771]
23. Andersson E, Hedman E, Wadström O, Boberg J, Andersson E, Axelsson E, et al. Internet-Based Extinction Therapy for Worry: A Randomized Controlled Trial. Behav Ther 2017 May;48(3):391-402. [doi: 10.1016/j.beth.2016.07.003] [Medline: 28390501]
24. Mira A, Bretón-López J, García-Palacios A, Quero S, Baños RM, Botella C. An Internet-based program for depressive symptoms using human and automated support: a randomized controlled trial. NDT 2017 Mar;Volume 13:987-1006. [doi: 10.2147/ndt.s130994]
25. Stolz T, Schulz A, Krieger T, Vincent A, Urech A, Moser C, et al. A mobile app for social anxiety disorder: A three-arm randomized controlled trial comparing mobile and PC-based guided self-help interventions. J Consult Clin Psychol 2018 Jun;86(6):493-504. [doi: 10.1037/ccp0000301] [Medline: 29781648]
26. Ip P, Chim D, Chan KL, Li TMH, Ho FKW, Van Voorhees BW, et al. Effectiveness of a culturally attuned Internet-based depression prevention program for Chinese adolescents: A randomized controlled trial. Depress Anxiety 2016 Dec;33(12):1123-1131. [doi: 10.1002/da.22554] [Medline: 27618799]
27. Oromendia P, Orrego J, Bonillo A, Molinuevo B. Internet-based self-help treatment for panic disorder: a randomized controlled trial comparing mandatory versus optional complementary psychological support. Cogn Behav Ther 2016 Jun;45(4):270-286. [doi: 10.1080/16506073.2016.1163615] [Medline: 27007256]
28. Schulz A, Stolz T, Vincent A, Krieger T, Andersson G, Berger T. A sorrow shared is a sorrow halved? A three-arm randomized controlled trial comparing internet-based clinician-guided individual versus group treatment for social anxiety disorder. Behav Res Ther 2016 Sep;84:14-26. [doi: 10.1016/j.brat.2016.07.001] [Medline: 27423374]
29. Berger T, Urech A, Krieger T, Stolz T, Schulz A, Vincent A, et al. Effects of a transdiagnostic unguided Internet intervention ('velibra') for anxiety disorders in primary care: results of a randomized controlled trial. Psychol Med 2017 Jan;47(1):67-80. [doi: 10.1017/S0033291716002270] [Medline: 27655039]
30. de Zwaan M, Herpertz S, Zipfel S, Svaldi J, Friederich H, Schmidt F, et al. Effect of Internet-Based Guided Self-help vs Individual Face-to-Face Treatment on Full or Subsyndromal Binge Eating Disorder in Overweight or Obese Patients: The INTERBED Randomized Clinical Trial. JAMA Psychiatry 2017 Oct 01;74(10):987-995 [FREE Full text] [doi: 10.1001/jamapsychiatry.2017.2150] [Medline: 28768334]
31. Nyström MBT, Stenling A, Sjöström E, Neely G, Lindner P, Hassmén P, et al. Behavioral activation versus physical activity via the internet: A randomized controlled trial. J Affect Disord 2017 Dec;215:85-93 [FREE Full text] [doi: 10.1016/j.jad.2017.03.018] [Medline: 28319696]

XSL•FO
RenderX

32. Ebert DD, Buntrock C, Lehr D, Smit F, Riper H, Baumeister H, et al. Effectiveness of Web- and Mobile-Based Treatment of Subthreshold Depression With Adherence-Focused Guidance: A Single-Blind Randomized Controlled Trial. Behav Ther 2018 Dec;49(1):71-83. [doi: 10.1016/j.beth.2017.05.004] [Medline: 29405923]
33. Calear AL, Batterham PJ, Poyser CT, Mackinnon AJ, Griffiths KM, Christensen H. Cluster randomised controlled trial of the e-couch Anxiety and Worry program in schools. J Affect Disord 2016 May 15;196:210-217. [doi: 10.1016/j.jad.2016.02.049] [Medline: 26926660]
34. 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]
35. Randerson J. The Guardian. 2012. Should science journalists read the papers on which their stories are based? A survey of science, health and environment journalists reveals the importance they attach to reading research papers The Guardian URL: https://www.theguardian.com/science/blog/2012/mar/28/science-jounalists-read-papers-stories [accessed 2019-08-01]

Abbreviations

eHealth: electronic health
CONSORT-EHEALTH: Consolidated Standards of Reporting Trials of Electronic and Mobile HEalth Applications and onLine TeleHealth
RCT: randomized controlled trial

Edited by G Eysenbach; submitted 29.03.19; peer-reviewed by J Apolinário-Hagen, C Young, A Beleigoli; comments to author 25.04.19; revised version received 07.06.19; accepted 27.06.19; published 15.08.19.

Please cite as:
Beintner I, Vollert B, Zarski AC, Bolinski F, Musiat P, Görlich D, Ebert DD, Jacobi C
Adherence Reporting in Randomized Controlled Trials Examining Manualized Multisession Online Interventions: Systematic Review of Practices and Proposal for Reporting Standards
J Med Internet Res 2019;21(8):e14181
URL: http://www.jmir.org/2033/8/e14181/
doi:10.2196/14181
PMID:31414664

©Ina Beintner, Bianka Vollert, Anna-Carlotta Zarski, Felix Bolinski, Peter Musiat, Dennis Görlich, David Daniel Ebert, Corinna Jacobi. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 15.08.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Virtual Reality for Screening of Cognitive Function in Older Persons: Comparative Study

Sean Ing Loon Chua[1], BEngSc (Hons), MSc; Ngiap Chuan Tan[1,2], MBBS, MMed, FCFPS; Wei Teen Wong[1,2], MBBS, MMed; John Carson Allen Jr[1], PhD; Joanne Hui Min Quah[1,2], MBBS, MMed; Rahul Malhotra[1], MBBS, MPH, MD; Truls Østbye[1,3], MPH, MD, PhD

[1]Duke-NUS Medical School, Singapore, Singapore

[2]SingHealth Polyclinics, Singapore, Singapore

[3]Duke University, Durham, NC, United States

Corresponding Author:
Ngiap Chuan Tan, MBBS, MMed, FCFPS
SingHealth Polyclinics
167 Jalan Bukit Merah
Connection One Tower 5 #15-10
Singapore, 150167
Singapore
Phone: 65 63777136
Email: tan.ngiap.chuan@singhealth.com.sg

Abstract

Background: The prevalence of dementia, which presents as cognitive decline in one or more cognitive domains affecting function, is increasing worldwide. Traditional cognitive screening tools for dementia have their limitations, with emphasis on memory and, to a lesser extent, on the cognitive domain of executive function. The use of virtual reality (VR) in screening for cognitive function in older persons is promising, but evidence for its use is sparse.

Objective: The primary aim was to examine the feasibility and acceptability of using VR to screen for cognitive impairment in older persons in a primary care setting. The secondary aim was to assess the module's ability to discriminate between cognitively intact and cognitively impaired participants.

Methods: A comparative study was conducted at a public primary care clinic in Singapore, where persons aged 65-85 years were recruited based on a cut-off score of 26 on the Montreal Cognitive Assessment (MoCA) scale. They participated in a VR module for assessment of their learning and memory, perceptual-motor function, and executive function. Each participant was evaluated by the total performance score (range: 0-700) upon completion of the study. A questionnaire was also administered to assess their perception of and attitude toward VR.

Results: A total of 37 participants in Group 1 (cognitively intact; MoCA score≥26) and 23 participants in Group 2 (cognitively impaired; MoCA score<26) were assessed. The mean time to completion of the study was 19.1 (SD 3.6) minutes in Group 1 and 20.4 (3.4) minutes in Group 2. Mean feedback scores ranged from 3.80 to 4.48 (max=5) in favor of VR. The total performance score in Group 1 (552.0, SD 57.2) was higher than that in Group 2 (476.1, SD 61.9; P<.001) and exhibited a moderate positive correlation with scores from other cognitive screening tools: Abbreviated Mental Test (0.312), Mini-Mental State Examination (0.373), and MoCA (0.427). A receiver operating characteristic curve analysis for the relationship between the total performance score and the presence of cognitive impairment showed an area under curve of 0.821 (95% CI 0.714-0.928).

Conclusions: We demonstrated the feasibility of using a VR-based screening tool for cognitive function in older persons in primary care, who were largely in favor of this tool.

(J Med Internet Res 2019;21(8):e14821) doi:10.2196/14821

KEYWORDS

virtual reality; feasibility studies; mental status and dementia tests; technology; video games; dementia; cognitive dysfunction

XSL•FO
RenderX

Introduction

Background and Rationale

Dementia is becoming more prevalent worldwide. About half of the dementia cases are in Asia, and the total number of cases worldwide is forecasted to increase to 63 million in 2030 [1]. In Singapore, one in ten people aged ≥60 years may have dementia [2]. Patients with dementia demonstrate significant cognitive decline in at least one or more cognitive domains, including complex attention, executive function, learning and memory, language, perceptual-motor function, and social cognition [3]. Mild cognitive impairment (MCI) represents a "middle ground" between normal ageing and dementia, and there has been growing interest in its timely diagnosis [4]. Although the Mini-Mental State Examination (MMSE) is the most widely applied test for dementia screening, the Montreal Cognitive Assessment (MoCA) is considered the best alternative for screening of MCI [5].

Besides the MMSE and MoCA, there are more than 40 other tests available for cognitive screening in health care settings [5]. A challenge with the application of commonly used paper-and-pencil or even digitalized screening tools is the limited cognitive domains that they assess. This is at the expense of other cognitive domains like executive function [6] and perceptual-motor function, which, when deficient, are associated with a high risk of progression to dementia [7]. Furthermore, the scoring in many of these screening tools is influenced by factors such as education level and cultural background [8]. Functional status scales used to assess the severity of dementia, such as the Barthel Index for Activities of Daily Living (ADL), also heavily depend on subjective observational measures. One plausible solution to overcome some of these limitations is the employment of virtual reality (VR) technology.

Virtual reality is a technology that provides interaction between a user and artificially generated environments. In recent years, with technological advancement, the use of VR has become more widespread. Beyond entertainment purposes, VR has also found purpose within certain fields of medicine [9] such as cognitive rehabilitation and training [10]. As a screening tool, VR has shown to have greater ecological validity [11], which reflects how well these tests predict real-world settings. This is in contrast with the contrived testing environment around the routine screening tests used today.

Several studies have attempted to investigate the use of VR to screen for cognitive impairment in older persons. Tong et al [12] used a tablet technology to screen for abnormal cognitive status in the emergency department, while Zygouris et al [13] developed a cognitive training app and compared its performance with established neuropsychological tests. Neither of these studies applied VR to cognitive screening in primary care, which is the most relevant setting for early detection of cognitive impairment [14].

Study Aims

The primary aim of this study was to examine the feasibility and acceptability of using VR to screen for cognitive impairment in older persons in a primary care setting. The VR platform used in this study is a new tool (referred to as the RE@CH assessment module or VR module) developed by the Institute of Technical Education (ITE), College West, Singapore. Feasibility is defined by the proportion of participants who successfully complete the RE@CH assessment module within a stipulated time. Acceptability is based on the feedback received at the end of the study with regard to acceptance, perception, and experience with VR.

The secondary aim was to assess the ability of the RE@ACH assessment module to discriminate between cognitively intact and cognitively impaired participants. Performance was assessed by first establishing a scoring algorithm on the module, followed by comparing the performance scores between the two groups. Validity was assessed by examining the correlation of the scores against other routine cognitive screening assessments including the Abbreviated Mental Test (AMT) [15], MMSE [16], and MoCA [17]. This was followed by receiver operating characteristic (ROC) curve analysis to determine a useful cut-off score.

Methods

Study Design

This study was conducted at SingHealth Polyclinics-Outram, a public primary care clinic located in the central region of Singapore and serving a multiethnic Asian population.

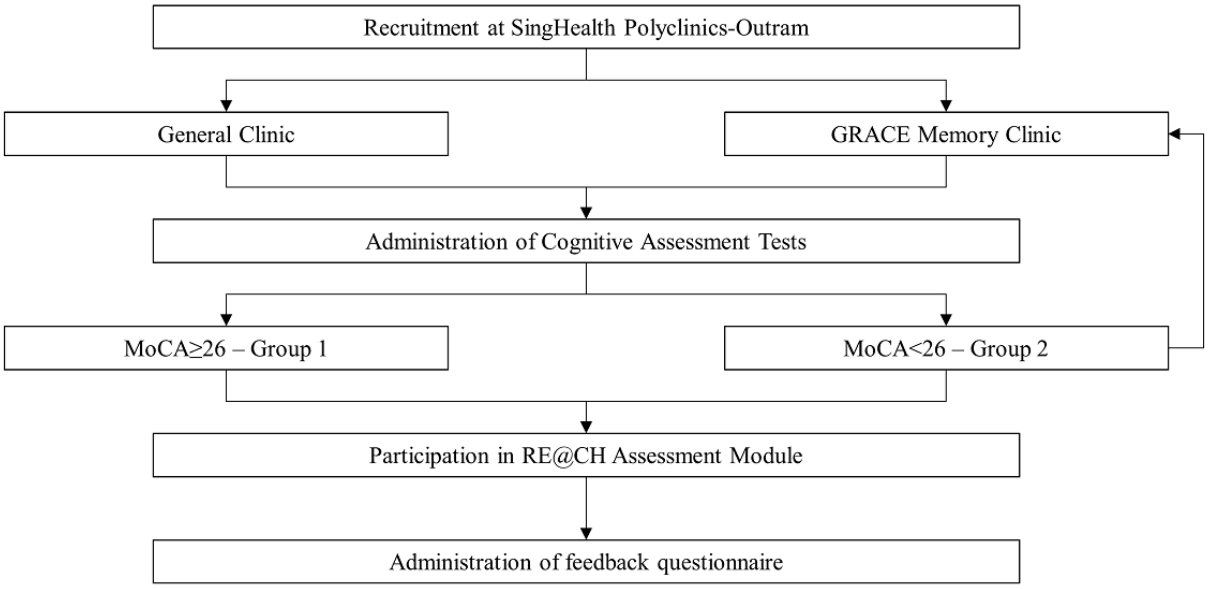
Most patients attend these polyclinics to visit their primary care physicians or family physicians for management of their noncommunicable diseases such as hypertension, dyslipidemia, and type 2 diabetes mellitus. Subspecialized clinics operate on selected days of the month at this study site, such as the GeRiAtric serviCE (GRACE) memory clinic, which screens patients for suspected cognitive impairment including MCI and dementia.

Participants and Eligibility

The participants were older persons registered at SingHealth Polyclinics-Outram, aged between 65 and 85 years. Sixty participants were targeted based on a small to medium standardized effect size [18] of the secondary aim (difference between mean performance scores and correlation coefficient), calculated using the upper confidence limit approach [19]. Of the 137 participants who were approached, 60 (43.8%) were recruited and enrolled after providing consent. All were recruited directly at the study site from March 2019 to April 2019 and were assessed by either the primary investigator (PI) or co-PI of the study project. Potential participants were patients waiting to see the doctor in the general clinic or GRACE memory clinic (Figure 1). They were either screened selectively at the waiting area or referred by physicians from those clinics.

XSL•FO
RenderX

Figure 1. Study flow. GRACE: GeRiAtric serviCE; MoCA: Montreal Cognitive Assessment.



All participants were required to understand the procedure for using the RE@CH assessment module, perform the movements involved in the study, and have the mental capacity to provide written informed consent. Those with poor vision, inability to follow verbal commands, aphasia, impairment of kinetic abilities that could inaccurately affect their performance on the assessment module, and unwillingness or inability to comply with the study protocol were excluded. Those with severe functional impairment based on the Lawton instrumental ADL (iADL) scale [20] were also excluded.

Enrolled participants provided written informed consent before entering the study, and this included consent to access their electronic medical records. All participants had an acceptable mental capacity, and consent was obtained in the presence of a witness. The MoCA cognitive assessment screen was performed and participants were divided into two groups (Figure 1), each meeting specific eligibility criteria:

- Group 1: Cognitively intact individuals, as defined by a MoCA score of ≥26. Those with pre-existing formal diagnosis of cognitive impairment of any degree or a history of cerebrovascular accident or neurological deficits were excluded.
- Group 2: Cognitively impaired individuals, as defined by a MoCA score of <26.

RE@CH Assessment Module

The RE@CH assessment module (Figure 2) uses VR and motion sensor (Leap motion) technology to replicate activities encountered in daily living as 3D games (Figure 3). The virtual environment was projected on a 55-inch 2D screen that recreates an immersive experience for the user. Through these activities, several cognitive domains were assessed. These included learning and memory, perceptual-motor function, and executive function, which had the greatest emphasis.

The study team designed a scoring algorithm (Table 1) to appraise the participant's performance on the RE@CH assessment module. Seven relevant tasks were selected and the scored depending on the participants' ability to complete the task correctly within the stipulated time, the number of attempts, and the proportion of tasks performed correctly.

Participants had to complete various tasks via the module using appropriate hand gestures. Before the formal VR assessment, they were guided through one orientation task to familiarize themselves with the mechanics of the system. Their performance on the next seven key activities was scored manually using the scoring algorithm (Table 1) in the following order:

1. Opening a door using the correct key and passcode number
2. Making a phone call by recalling a predefined 8-digit number
3. Identifying: (a) Famous people, (b) Advertisement of groceries, and (c) 4-digit number on a lottery slip on a newspaper
4. Sorting household objects according to category
5. Picking an outfit appropriate for a specified occasion
6. Withdrawing cash from an automated teller machine
7. Shopping for groceries in a provision shop

Prior to each task, the study team guided the participant on the requirement of the task through a short tutorial.

XSL•FO
RenderX

Figure 2. Setup of the RE@CH assessment module in the consultation room.

Figure 3. Main screen page of the RE@CH assessment module, showing two 3D games: Secret Door (opening a door using the right key) and Speed Dialling (making a phone call by recalling a predefined 8-digit number).

XSL•FO
RenderX

Table 1. Scoring algorithm on RE@CH assessment module.

Cognitive domain, task, content	Score					Remarks
	0	25	50	75	100	
Perceptual-Motor						
1—Opening door with correct key and passcode number	No attempt	Unable to complete Step 1	Complete Step 1	Complete Step 1 and 2	Complete Steps 1, 2, and 3	• Step 1: Pick the correct key • Step 2: Open the door • Step 3: Key in the correct number
Learning and Memory						
2—Make a phone call, recalling the 8-digit number in predefined sequence	No attempt	Unable to complete on 3 attempts	Complete on 3 attempts	Complete on 2 attempts	Complete on 1 attempt	• Every reattempt after a wrong digit is keyed in, is counted as 1 attempt
Executive Function						
3—Identify items in each category from the newspaper: (1) famous people, (2) advertisement of groceries, and (3) 4-digit number on a lottery slip	No attempt	Fail to identify any category correctly	Identify 1 category correctly	Identify 2 categories correctly	Identify all 3 categories correctly	• 1 item in from each category
4—Housekeeping: Sort things inside the room	No attempt	Sort 0 items correctly	Sort 1 item correctly	Sort 2 items correctly	Sort all 3 items correctly	• Sorting 1 round of 3 items according to their appropriate category
5—Dressing/grooming: Pick appropriate outfit for occasion	No attempt	Pick appropriate outfit for 0 of 3 tries	Pick appropriate outfit for 1 of 3 tries	Pick appropriate outfit for 2 of 3 tries	Pick appropriate outfit for all 3 tries	• Picking the appropriate outfit for 3 different occasions
6—Handling finances: Withdrawing cash from ATM[a]	No attempt	Unable to complete Step 1	Complete Step 1	Complete Step 1 and 2	Complete Steps 1, 2, and 3	• Step 1: Insert ATM card • Step 2: Enter correct PIN[b] • Step 3: Select and withdraw correct amount
7—Handling finances: Shopping at provision shop	No attempt	Unable to complete Objective 1 with 0 correct item	Complete Objective 1 with 1 correct item	Complete Objective 1 with 2 or more correct items	Complete Objective 1 and 2	• Objective 1: Pick 3 correct items • Objective 2: Pay the correct amount

[a]ATM: automated teller machine.

[b]PIN: personal identification number.

Data Collection

First, AMT, MMSE, and MoCA cognitive screening tests were carried out in order and scored as part of the cognitive function assessment to determine eligibility and classification into either of the two groups.

Next, baseline characteristics were collected, including demographics, functional status (Barthel Index and Lawton iADL), clinical parameters, and past medical history. All data were obtained directly from the participants, except for their past medical history, which was from their electronic medical records.

Subsequently, the VR component of the study began once the PI or co-PI of the study started the briefing of the participant on the module and ended after the completion of the seventh activity on the RE@CH assessment module. The start and end time were recorded and used to calculate the participant's time spent on the module, which included the time spent on the tutorial and explanation by the facilitator.

Finally, the participants completed an interviewer-administered questionnaire to assess their perception toward VR (Textbox 1). There were six questions, and the answers were rated on a Likert scale (1 - strongly disagree, 2 - disagree, 3 - neutral, 4 - agree, 5 - strongly agree). Total feedback scores were calculated by summing the scores from the six questions.

XSL•FO
RenderX

Textbox 1. Sample questionnaire assessing subjects' perception toward virtual reality experience, categorized by individual questions.

1. The VR software I have experienced is easy to use.
2. The amount of time I spent with the VR software is acceptable to me.
3. The use of VR software helps to make the experience in the clinic more interactive.
4. The use of VR technology to help diagnose medical condition appeals to me.
5. I would not mind seeing more new technologies being used by the doctor/medical staff during the consultation.
6. Overall, I enjoyed the VR experience in the clinic.

Outcome Measures

To achieve the primary aim of examining the feasibility and acceptability of the VR module, the following outcomes were measured: percentage of participants who successfully completed the RE@CH assessment module, time taken to complete the RE@CH assessment module, and scores from the feedback questionnaire.

To achieve the secondary aim of assessing the performance and validity of the RE@CH assessment module, the following outcomes were measured: performance scores on the RE@CH assessment module and scores on other cognitive assessment tests including AMT, MMSE, and MoCA. Additional statistical analysis was performed to derive the correlation between these outcome measures and to obtain the ROC curve.

All outcome measures were recorded in both study groups.

Statistical Analysis

Baseline characteristics were compared between the groups using a two-sample *t* test and the Fisher exact test for continuous and categorical variables, respectively. Summary statistics were calculated individually for the recruitment statistics and questionnaire-dependent feedback scores. Groups 1 and 2 were compared using the two-sample *t* test.

Performance scores were analyzed at each task level, given a total score, and compared between the two groups using the methods described above, as appropriate. Correlation between performance scores and other variables was assessed using Pearson correlation. A logistic regression was performed based on the total performance scores, followed by ROC analysis to assess its predictive capability to discriminate between cognitively intact and cognitively impaired individuals.

A *P* value<.05 was considered statistically significant (two-sided). Analyses were performed using SAS University Edition software (Version 9.4M6 of the SAS System for Windows; SAS Institute Inc, Cary, NC).

Results

Sample Description

Based on their MoCA scores, 37 participants with a score≥26 were placed in Group 1 (cognitively intact) and 23 participants with a score<26 were placed in Group 2 (cognitively impaired). Of those in Group 2, 10 were new cases pending referral to the GRACE memory clinic. Baseline characteristics of participants in both groups were compared (Table 2). There were no statistically significant differences in age and gender between the two groups. Mean cognitive assessment scores, based on AMT, MMSE and MoCA, were all higher in Group 1 than in Group 2.

XSL•FO
RenderX

Table 2. Baseline characteristics by cognitive status.

Characteristics	Cognitively intact (n=37)	Cognitively impaired (n=23)	*P* value[a]
Age (years), mean (SD)	70.7 (3.6)	73.2 (5.4)	.06
Gender, n (%)			.21
Female	26 (70.3)	15 (65.2)	
Male	11 (29.7)	8 (34.8)	
Years of education, n (%)			<.001
<6	4 (10.8)	12 (52.2)	
6-10	23 (62.2)	8 (34.8)	
>10	10 (27.0)	3 (13.0)	
Functional status scores, mean (SD)			
ADL[b]	98.5 (2.6)	98.3 (3.9)	.78
iADL[c]	22.8 (0.5)	21.7 (1.4)	.002
Cognitive assessment scores, mean (SD)			
AMT[d]	9.6 (0.6)	8.3 (1.3)	<.001
MMSE[e]	28.8 (1.0)	25.9 (3.4)	<.001
MoCA[f]	27.8 (1.2)	22.2 (3.3)	<.001

[a]In categories where the mean values can be directly compared, the *P* values are given individually. In categories where the distribution across subcategories are compared, only *P* values for the main categories are given.

[b]ADL: activities of daily living, maximum score of 100.

[c]iADL: instrumental activities of daily living, maximum score of 23.

[d]AMT: Abbreviated Mental Test, maximum score of 10.

[e]MMSE: Mini-Mental State Examination, maximum score of 30.

[f]MoCA: Montreal Cognitive Assessment, maximum score of 30.

Primary Aim: Feasibility and Acceptability

Recruitment Statistics and Time to Completion

All 60 (100%) enrolled participants successfully completed the study. The mean total time spent on the RE@CH assessment module was not significantly different between the two groups (Group 1: 19.1 [SD 3.6] minutes; Group 2: 20.4 [SD 3.4] minutes; *P*=.17; Table 3).

Assessment of Perception Toward Virtual Reality

Results from the questionnaire were favorable toward the VR experience. Mean feedback scores for each question (over a scale of 1-5) ranged from 3.80 to 4.48 (Table 3). The largest proportion of responses to all statements was “Agree” or “Strongly agree.” The difference in total feedback scores between the study groups was not statistically significant (*P*=.62; Table 3).

XSL•FO
RenderX

Table 3. Total time spent on the RE@CH assessment module and feedback scores between study groups.

Factor	Cognitively intact (n=37)		Cognitively impaired (n=23)		*P* value	Overall[a] (N=60), mean (SD)
	Mean (SD)	Median	Mean (SD)	Median		
Time spent on the virtual reality module						
Time (minutes)	19.1 (3.6)	17	20.4 (3.4)	18	.17	19.6 (3.6)
Feedback score from the questionnaire						
Mean feedback score by question[b]						
Question 1	3.76 (1.01)	3	3.87 (0.97)	3	.67	3.80 (0.99)
Question 2	4.11 (0.77)	4	4.13 (0.87)	4	.92	4.12 (0.80)
Question 3	4.08 (0.76)	4	4.17 (0.94)	4	.68	4.12 (0.83)
Question 4	3.92 (1.04)	3	4.26 (0.81)	4	.18	4.05 (0.96)
Question 5	4.30 (0.81)	4	4.26 (0.69)	4	.86	4.28 (0.76)
Question 6	4.51 (0.56)	4	4.43 (0.95)	4	.72	4.48 (0.72)
Total feedback score[c]	24.7 (3.6)	25	25.1 (3.9)	26	.62	24.8 (3.7)

[a]Overall mean values from both study groups.

[b]Description of the individual questions can be found in Textbox 1.

[c]Sum of the numerical scores (ranging from 1 to 5) given for each of the six questions.

Secondary Aim: Performance and Validity

Discriminating Performance Between Cognitively Intact and Cognitively Impaired Participants

In general, the mean scores for each task was higher in the cognitively intact compared to the cognitively impaired participants, except for task 1 (Table 4). These differences were statistically significant for tasks 2, 3 and 7, and the total performance score (sum of task 1 to task 7 scores) exhibited statistically significant differences. Results of the individual performance scores in the RE@CH assessment module are summarized in Table 4.

Table 4. Performance scores on the RE@CH assessment module by cognitive status for individual tasks and total scores.

Task	Cognitively intact (n=37), mean (SD)	Cognitively impaired (n=23), mean (SD)	*P* value
1—Opening door	96.6 (14.6)	96.7 (11.4)	.97
2—Making phone call	67.6 (33.3)	41.3 (23.4)	.002
3—Reading newspaper and identifying specific components	87.8 (20.9)	65.2 (26.9)	<.001
4—Sorting items	81.1 (19.0)	71.7 (29.5)	.18
5—Choosing the right clothes	66.9 (25.0)	57.6 (23.2)	.16
6—Withdrawing money at ATM[a]	53.4 (10.5)	50.0 (0.0)	.06
7—Shopping at supermarket	98.6 (5.7)	93.5 (11.2)	.049
Total performance score[b]	552.0 (57.2)	476.1 (61.9)	<.001

[a]ATM: automated teller machine.

[b]Sum of individual scores (ranging from 0 to 100) in each of the seven tasks.

Correlation of the RE@CH Assessment Module and Routine Cognitive Screening Assessment

The total performance score showed moderate positive correlation with scores from the other routine cognitive screening assessment tools, namely, AMT (0.312, *P*=.02), MMSE (0.373, *P*=.003), and MoCA (0.427, *P*<.001). In addition, the correlation of the total performance score with age was negative and poor (–0.291, *P*=.02). Table 5 presents the correlation matrix between the different variables.

XSL•FO
RenderX

Table 5. Correlation matrix demonstrating the degree of association among ADL score, iADL score, AMT score, MMSE score, MoCA score, age, education, total feedback, and total performance scores.

	ADL[a]	iADL[b]	AMT[c]	MMSE[d]	MoCA[e]	Age	Education	Feedback score	Performance score
ADL									
Pearson correlation coefficient	—[f]	–0.01	–0.04	0.30	0.12	–0.40	0.07	0.20	–0.03
P value	—	.94	.79	.02	.35	.002	.61	.13	.85
iADL									
Pearson correlation coefficient	–0.01	—	0.22	0.30	0.49	–0.23	0.003	–0.03	0.20
P value	.94	—	.09	.02	<.001	.07	.98	.79	.13
AMT									
Pearson correlation coefficient	–0.04	0.22	—	0.73	0.73	–0.22	0.37	0.05	0.31
P value	.79	.09	—	<.001	<.001	.10	.00	.68	.02
MMSE									
Pearson correlation coefficient	0.30	0.30	0.73	—	0.76	–0.31	0.57	0.07	0.37
P value	.02	.02	<.001	—	<.001	.02	<.001	.62	.00
MoCA									
Pearson correlation coefficient	0.12	0.49	0.73	0.76	—	–0.23	0.40	–0.07	0.43
P value	.35	<.001	<.001	<.001	—	.08	.002	.59	<.001
Age									
Pearson correlation coefficient	–0.40	–0.23	–0.22	–0.31	–0.23	—	–0.13	–0.07	–0.29
P value	.002	.07	.10	.02	.08	—	.33	.59	.02
Education									
Pearson correlation coefficient	0.07	0.003	0.37	0.57	0.40	–0.13	—	–0.09	0.24
P value	.61	.98	.003	<.001	.002	.33	—	.48	.06
Feedback score									
Pearson correlation coefficient	0.20	–0.03	0.05	0.07	–0.07	–0.07	–0.09	—	–0.02
P value	.13	.79	.68	.62	.59	.59	.48	—	.85
Performance score									
Pearson correlation coefficient	–0.03	0.20	0.31	0.37	0.43	–0.29	0.24	–0.02	—
P value	.85	.13	.02[g]	.003[g]	<.001[g]	.02[g]	.06	.85	—

[a]ADL: activities of daily living.

[b]iADL: instrumental activities of daily living.

[c]AMT: Abbreviated Mental Test.

[d]MMSE: Mini-Mental State Examination.

[e]MoCA: Montreal Cognitive Assessment.

[f]Not applicable.

[g]Significance at $P<.05$.

Receiver Operating Characteristic Curve Analysis

ROC curve analysis (Figure 4) was conducted over the continuous total performance score to assess its capability to discriminate between cognitively intact (MoCA≥26) and cognitively impaired individuals (MoCA<26). The area under the curve (AUC) was found to be 0.821 (95% CI 0.714-0.928). An optimal statistical cutoff is achieved at a cut-off score of 500 (78.2% sensitivity, 75.7% specificity, 66.7% positive predictive value, and 84.8% negative predictive value; Figure 5).

XSL•FO
RenderX

Figure 4. Receiver operating characteristic curve for the RE@CH assessment module's ability to discriminate between the two groups through the total performance scores.

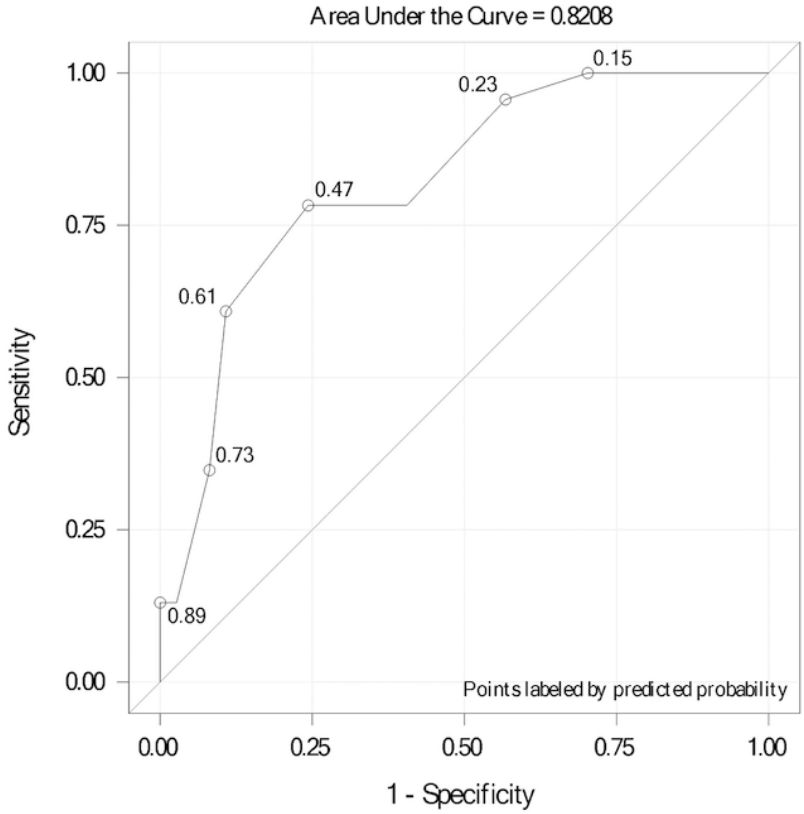
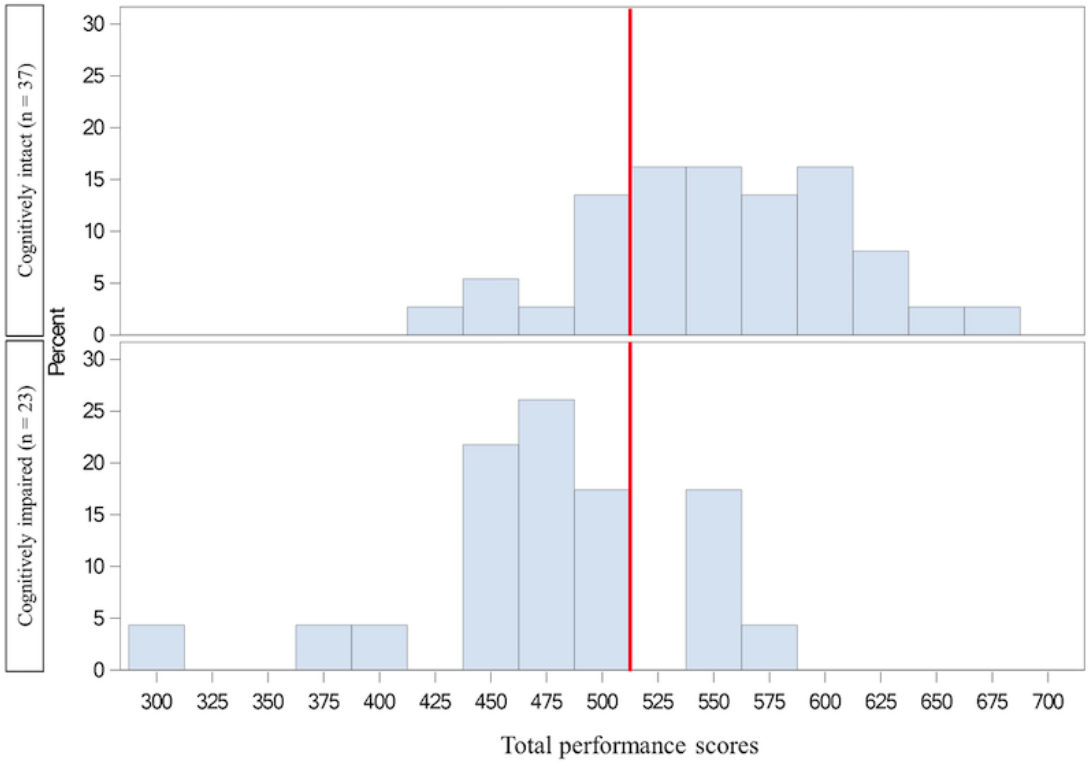


Figure 5. Distribution of the total performance scores of the RE@CH assessment module by group. The red line indicates the optimal cut-off score of 500.



XSL•FO
RenderX

Discussion

Primary Aim: Feasibility and Acceptability

Assessing Feasibility of the Virtual Reality Module in Primary Care

The primary aim of this study was to determine the feasibility and acceptability of VR to screen for cognitive impairment among older persons in primary care. We found the RE@CH assessment module feasible for this setting. All participants recruited successfully completed the study, regardless of their performance score. There were no withdrawals or any immediate adverse effects reported. The time to completion was comparable to the total time to completion of all three of the routine cognitive screening tools (AMT, MMSE, and MoCA) [5] used at SingHealth Polyclinics-Outram for memory-related complaints. Most of the participants agreed that the time spent on the VR was acceptable, as supported by a mean feedback score of 3.8 of 5 for Question 1 (Table 3).

Assessing Acceptance of Virtual Reality Among Participants

The RE@CH assessment module was well accepted by both participant groups, as evidenced by the positive feedback scores across all six questions. The response rate of 43.8% in this study suggests that almost one in two older persons are receptive to the use of novel technology for assessing their cognitive function. The scores for immersion in the VR experience (Table 3) were overall above average. This was supported by earlier studies that showed that older adults generally have a positive attitude toward the VR environment [21] and new technology [22]. A few felt that the VR software was not easy to use, but this could be because of unfamiliarity with VR technology.

Secondary Aim: Performance and Validity

Comparing Performance Between Cognitively Intact and Impaired Participants

The secondary aim of this study was to determine the performance and validity of the RE@CH assessment module in screening for cognitive impairment. We found that although the total performance scores were significantly higher in the cognitively intact group, it was attributed to a few individual tasks. Tasks 2, 3, and 7 were better at discriminating between the two groups. The remaining tasks could be affected by the technical limitations inherent in the VR module and the structure of the scoring algorithm. Tasks 4 and 5 were challenging to both groups, because the motion sensor was not precise enough to pick up all hand gestures. Tasks 1 and 6 were relatively easy, and all participants performed well. This might have led to underperformance or overperformance across the groups, but it is not likely to have differentiated between the two groups well.

Assessing Validity of the Virtual Reality Module Compared to Other Screening Assessments

The RE@CH assessment module generated valid total performance scores that had a moderate positive correlation with all the other three validated cognitive screening assessments. A high positive correlation was not expected, since the latter focused specifically on a few cognitive domains and the VR module could have identified cognitive deficits that other screening assessments were unable to detect. The ROC analysis indicated relatively strong prognostic classification capability [23] with an AUC of 0.821 when benchmarked against MoCA scores.

Clinical Implications

Our findings provide preliminary evidence that VR modalities, such as the RE@CH assessment module, can be used for cognitive screening among older person in primary care. We built upon the positive results from previous studies [11-13,24] and introduced another integral component of a VR setup—the Leap motion sensor—that detects hand gestures [25]. Instead of a single game, we introduced multiple short VR-based activities to assess the different cognitive domains.

Our findings show that participants could still perform most of the tasks within a time frame of about 20 minutes (mean 19.6 minutes). In addition, we showed that VR was not only feasible, but also relatively practical to execute within the premise of a primary care clinic. The screening can be executed prior to consultations with the primary care physician to optimize their clinic visits by at-risk patients. Implementation studies are needed in future research to assess its roll-out in clinical practice.

Limitations

This study had several limitations. First, there was a significant difference in the level of education between the two groups. More participants with MoCA scores≥26 had higher educational levels than those with lower MoCA scores, which could impact the outcomes of their performance. The disparity could be related to the recruitment process. Participants who were better educated and more likely to have been exposed to health technology and innovations were likely and willing to be enrolled into the study, even in the absence of cognitive symptoms. The sample size was small in this feasibility study. An adequately powered randomized controlled trial, stratified by education level, is planned to further assess the use of VR in cognitive assessment.

Second, the participants might not be representative of the entire target population at risk of cognitive impairment (older persons aged≥65 years) [26], since recruitment was carried out only at one location with multiple exclusion criteria. Furthermore, the study population was classified solely on the basis of the MoCA scores, which is a screening tool and not considered diagnostic of cognitive impairment. Participants with subjective cognitive impairment but normal MoCA scores could have been assigned erroneously to the cognitively intact group.

Lastly, the RE@CH assessment module was a prototype that was originally used for rehabilitation and then adapted for cognitive screening. It included tasks that might be challenging to complete because of the prototype content design, and not due to the participant's cognitive impairment. A specially designed VR program that will be developed by us to assess cognitive function has been recently awarded funding by a local public information technology agency, which will address the limitations of the current prototype. The next prototype will

XSL•FO
RenderX

eliminate tasks that do not differentiate between the two groups to both reduce redundancy and optimize the time spent to complete the cognitive assessment. The next study will include time motion measurements to evaluate accuracy, efficiency, and cost-effectiveness of our tool in comparison with the conventional screening tools such as MoCA.

Conclusions

The study successfully demonstrated the feasibility of a VR-based screening tool in primary care. The target population expressed a positive perception of and attitude toward VR and were open to the use of this type of technology for their cognitive assessment. The results of this feasibility study are invaluable in the design of a novel VR program and study protocol by validating its use as a comprehensive, multidomain cognitive function screening tool in the next phase of development.

Acknowledgments

We would like to extend our gratitude to our health care colleagues at SingHealth Polyclinics (Outram) who supported us in various ways, including screening and referring patients for the study. We also appreciate our collaborators at ITE College West, Mr Teh Tuan Ann and Dr Lim Soon Huat, for development of the RE@CH assessment module and their technical assistance. We are thankful to the SingHealth-Duke NUS Family Medicine Academic Clinical Programme (FM ACP) secretariat (Ms Winnie Seah, Winnie Ong, Eugene Quek) and SingHealth Polyclinics Department of Research Administration (Ms Patricia Kin, Caris Tan, and Usha Sankari) for their assistance in ethics board review and procurement of equipment. This study received funding from the FM ACP Seed Grant and the Duke-NUS MSF Grant.

Authors' Contributions

SILC, WTW, JHMQ, TO, and NCT designed the study protocol. SILC and WTW implemented the study and collected the data. JCA and SILC analyzed the data. SILC, JCA, NCT, RM, and TO reviewed and interpreted the results. SILC drafted the manuscript. All authors reviewed the draft and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Wimo A, Winblad B, Aguero-Torres H, von Strauss E. The magnitude of dementia occurrence in the world. Alzheimer Dis Assoc Disord 2003;17(2):63-67. [Medline: 12794381]
2. Subramaniam M, Chong SA, Vaingankar JA, Abdin E, Chua BY, Chua HC, et al. Prevalence of Dementia in People Aged 60 Years and Above: Results from the WiSE Study. J Alzheimers Dis 2015;45(4):1127-1138. [doi: 10.3233/JAD-142769] [Medline: 25672767]
3. Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition. Washington, DC: American Psychiatric Association Publishing; 2013.
4. Dubois B, Padovani A, Scheltens P, Rossi A, Dell'Agnello G. Timely Diagnosis for Alzheimer's Disease: A Literature Review on Benefits and Challenges. J Alzheimers Dis 2016;49(3):617-631 [FREE Full text] [doi: 10.3233/JAD-150692] [Medline: 26484931]
5. Tsoi KKF, Chan JYC, Hirai HW, Wong SYS, Kwok TCY. Cognitive Tests to Detect Dementia: A Systematic Review and Meta-analysis. JAMA Intern Med 2015 Sep;175(9):1450-1458. [doi: 10.1001/jamainternmed.2015.2152] [Medline: 26052687]
6. Cullen B, O'Neill B, Evans JJ, Coen RF, Lawlor BA. A review of screening tests for cognitive impairment. J Neurol Neurosurg Psychiatry 2007 Aug;78(8):790-799 [FREE Full text] [doi: 10.1136/jnnp.2006.095414] [Medline: 17178826]
7. Gauthier S, Reisberg B, Zaudig M, Petersen RC, Ritchie K, Broich K, International Psychogeriatric Association Expert Conference on mild cognitive impairment. Mild cognitive impairment. Lancet 2006 Apr 15;367(9518):1262-1270. [doi: 10.1016/S0140-6736(06)68542-5] [Medline: 16631882]
8. Yokomizo JE, Simon SS, Bottino CMDC. Cognitive screening for dementia in primary care: a systematic review. Int Psychogeriatr 2014 Nov;26(11):1783-1804. [doi: 10.1017/S1041610214001082] [Medline: 25023857]
9. Satava RM. Medical applications of virtual reality. J Med Syst 1995 Jun;19(3):275-280. [doi: 10.1007/BF02257178]
10. Pensieri C, Pennacchini M. Overview: Virtual Reality in Medicine. JVWR 2014 Jan 19;7(1):1-34. [doi: 10.4101/jvwr.v7i1.6364]
11. Tarnanas I, Schlee W, Tsolaki M, Müri R, Mosimann U, Nef T. Ecological validity of virtual reality daily living activities screening for early dementia: longitudinal study. JMIR Serious Games 2013 Aug 06;1(1):e1 [FREE Full text] [doi: 10.2196/games.2778] [Medline: 25658491]
12. Tong T, Chignell M, Tierney MC, Lee J. A Serious Game for Clinical Assessment of Cognitive Status: Validation Study. JMIR Serious Games 2016 May 27;4(1):e7 [FREE Full text] [doi: 10.2196/games.5006] [Medline: 27234145]

XSL•FO
RenderX

13. Zygouris S, Giakoumis D, Votis K, Doumpoulakis S, Ntovas K, Segkouli S, et al. Can a virtual reality cognitive training application fulfill a dual role? Using the virtual supermarket cognitive training application as a screening tool for mild cognitive impairment. J Alzheimers Dis 2015;44(4):1333-1347. [doi: 10.3233/JAD-141260] [Medline: 25428251]
14. Olazarán J, Torrero P, Cruz I, Aparicio E, Sanz A, Mula N, et al. Mild cognitive impairment and dementia in primary care: the value of medical history. Fam Pract 2011 Aug;28(4):385-392. [doi: 10.1093/fampra/cmr005] [Medline: 21402661]
15. Hodkinson HM. Evaluation of a mental test score for assessment of mental impairment in the elderly. Age Ageing 1972 Nov;1(4):233-238. [doi: 10.1093/ageing/1.4.233] [Medline: 4669880]
16. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975 Nov;12(3):189-198. [Medline: 1202204]
17. Nasreddine ZS, Phillips NA, Bédirian V, Charbonneau S, Whitehead V, Collin I, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. J Am Geriatr Soc 2005 Apr;53(4):695-699. [doi: 10.1111/j.1532-5415.2005.53221.x] [Medline: 15817019]
18. Coe R. It's the effect size, stupid: What effect size is and why it is important. 2002 Presented at: Annual Conference of the British Educational Research Association; September 12-14, 2002; University of Exeter, England.
19. Sim J, Lewis M. The size of a pilot study for a clinical trial should be calculated in relation to considerations of precision and efficiency. J Clin Epidemiol 2012 Mar;65(3):301-308. [doi: 10.1016/j.jclinepi.2011.07.011] [Medline: 22169081]
20. Graf C. The Lawton instrumental activities of daily living scale. Am J Nurs 2008 Apr;108(4):52-62. [doi: 10.1097/01.NAJ.0000314810.46029.74] [Medline: 18367931]
21. Torres ACS. Cognitive effects of video games on old people. International Journal on Disability and Human Development 2011 Mar 09;10(1):55-58. [doi: 10.1515/ijdhd.2011.003]
22. Lee B, Chen Y, Hewitt L. Age differences in constraints encountered by seniors in their use of computers and the internet. Computers in Human Behavior 2011 May;27(3):1231-1237. [doi: 10.1016/j.chb.2011.01.003]
23. Zhu W, Zeng N, Wang N. Sensitivity, specificity, accuracy, associated confidence interval and ROC analysis with practical SAS implementations. In: NESUG proceedings: Health Care and Life Sciences. 2010 Presented at: NESUG 2010; November 14-17, 2010; Baltimore, Maryland p. 6724.
24. Werner P, Rabinowitz S, Klinger E, Korczyn AD, Josman N. Use of the virtual action planning supermarket for the diagnosis of mild cognitive impairment: a preliminary study. Dement Geriatr Cogn Disord 2009;27(4):301-309. [doi: 10.1159/000204915] [Medline: 19252401]
25. Martín-Gutiérrez J. Virtual Technologies Trends in Education. Eurasia J Math Sci Tech Ed 2016 Dec 16;13(1):469-486. [doi: 10.12973/eurasia.2017.00626a]
26. Nagaendran K, Chen LHC, Chong MS, Chua EV, Goh CKS, Kua J, et al. Ministry of health clinical practice guidelines: dementia. Singapore Med J 2013 May;54(5):293-299. [doi: 10.11622/smedj.2013112] [Medline: 23716158]

Abbreviations

ADL: activities of daily living
AMT: Abbreviated Mental Test
AUC: area under curve
iADL: instrumental activities of daily living
GRACE: GeRiAtric serviCE
MCI: mild cognitive impairment
MMSE: Mini-Mental State Examination
MoCA: Montreal Cognitive Assessment
PI: primary investigator
ROC: receiver operating characteristic
VR: virtual reality

Edited by G Eysenbach; submitted 27.05.19; peer-reviewed by E Siegmund-Schultze, N Kumar; comments to author 21.06.19; revised version received 24.06.19; accepted 27.06.19; published 01.08.19.

Please cite as:
Chua SIL, Tan NC, Wong WT, Allen Jr JC, Quah JHM, Malhotra R, Østbye T
Virtual Reality for Screening of Cognitive Function in Older Persons: Comparative Study
J Med Internet Res 2019;21(8):e14821
URL: https://www.jmir.org/2019/8/e14821/
doi:10.2196/14821
PMID:31373274

XSL•FO
RenderX

©Sean Ing Loon Chua, Ngiap Chuan Tan, Wei Teen Wong, John Carson Allen Jr, Joanne Hui Min Quah, Rahul Malhotra, Truls Østbye. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 01.08.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

The Importance of Systematically Reporting and Reflecting on eHealth Development: Participatory Development Process of a Virtual Reality Application for Forensic Mental Health Care

Hanneke Kip[1,2], MSc; Saskia M Kelders[1,3], PhD; Yvonne H A Bouman[2], PhD; Lisette J E W C van Gemert-Pijnen[1,4], Prof Dr

[1]Centre for eHealth and Wellbeing Research, Department of Psychology, Health and Technology, University of Twente, Enschede, Netherlands

[2]Department of Research, Stichting Transfore, Deventer, Netherlands

[3]Optentia Research Focus Area, North-West University, Vanderbijlpark, South Africa

[4]Faculty of Medical Sciences, Universitair Medisch Centrum Groningen, Groningen, Netherlands

Corresponding Author:
Hanneke Kip, MSc
Centre for eHealth and Wellbeing Research
Department of Psychology, Health and Technology
University of Twente
Drienerlolaan 5
Enschede, 7522 NB
Netherlands
Phone: 31 534896536
Email: h.kip@utwente.nl

Abstract

Background: The use of electronic health (eHealth) technologies in practice often is lower than expected, mostly because there is no optimal fit among a technology, the characteristics of prospective users, and their context. To improve this fit, a thorough systematic development process is recommended. However, more knowledge about suitable development methods is necessary to create a tool kit that guides researchers in choosing development methods that are appropriate for their context and users. In addition, there is a need for reflection on the existing frameworks for eHealth development to be able to constantly improve them.

Objective: The two main objectives of this case study were to present and reflect on the (1) methods used in the development process of a virtual reality application for forensic mental health care and (2) development model that was used: the CeHRes Roadmap (the Centre for eHealth Research Roadmap).

Methods: In the development process, multiple methods were used to operationalize the first 2 phases of the CeHRes Roadmap: the contextual inquiry and value specification. To summarize the most relevant information for the goals of this study, the following information was extracted per method: (1) research goal, (2) explanation of the method used, (3) main results, (4) main conclusions, and (5) lessons learned about the method.

Results: Information on 10 methods used is presented in a structured manner. These 10 methods were stakeholder identification, project team composition, focus groups, literature study, semistructured interviews, idea generation with scenarios, Web-based questionnaire, value specification, idea generation with prototyping, and a second round of interviews. The lessons learned showed that although each method added new insights to the development process, not every method appeared to be the most appropriate for each research goal.

Conclusions: Reflection on the methods used pointed out that brief methods with concrete examples or scenarios fit the forensic psychiatric patients the best, among other things, because of difficulties with abstract reasoning and low motivation to invest much time in participating in research. Formulating clear research questions based on a model's underlying principles and composing a multidisciplinary project team with prospective end users appeared to be important in this study. The research questions supported the project team in keeping the complex development processes structured and prevented tunnel vision. With regard to the CeHRes Roadmap, continuous stakeholder involvement and formative evaluations were evaluated as strong points. A suggestion to further improve the Roadmap is to explicitly integrate the use of domain-specific theories and models. To create a tool kit with a broad range of methods for eHealth development and further improve development models, studies that report and reflect on development processes in a consistent and structured manner are needed.

XSL•FO
RenderX

(J Med Internet Res 2019;21(8):e12972) doi:10.2196/12972

KEYWORDS

eHealth; technology development; virtual reality; forensic psychiatry; community-based participatory research; human-centered design; case study

Introduction

Electronic health (eHealth)—a technology to support health, well-being, and health care—can offer many benefits, such as increased quality of care, easily accessible health care, and increased self-management [1]. However, these benefits are often not fully realized in practice [2]. A possible explanation for this is that technology does not optimally fit the needs, wishes, and characteristics of the involved end users and their context [3-5]. A way to improve this fit is thorough participatory eHealth development in which potential end users are structurally involved in the development process [1,6-8]. Consequently, many efforts have been made to create models, approaches, and guidelines for development of eHealth technologies. Examples are the CeHRes Roadmap (the Centre for eHealth Research Roadmap) [9], the person-based approach [10], the accelerated creation-to-sustainment model [11], intervention mapping [12], the persuasive system design model [13], and the agile science approach [14]. Most of these models and approaches do not offer concrete prescriptions for ready-to-use research methods that fit specific contexts and people. Instead, they present abstract guidelines for development to support researchers in shaping their development process. Although a step-by-step, detailed prescription of a specific development process does not seem feasible because of different characteristics of contexts, people, and technologies, there does seem to be a need for more knowledge and guidelines on how to apply these models in practice [14]. To support researchers in operationalizing development models, we propose that a general tool kit with a broad range of eHealth development methods might be developed. Such a tool kit can provide an overview of broad-range development methods and guidelines on when and how to apply them. In this way, it can support researchers in choosing appropriate methods for the context and end users with which they are working and different phases of their development process. Using a tool kit can prevent other researchers from having to reinvent the wheel and result in more efficient and better substantiated development processes.

To create a tool kit, more generalizable knowledge on eHealth development methods is necessary. To build this knowledge base, more case studies that explain and reflect on specific development methods used seem to be necessary [15]. On top of that, there also should be more critical reflection on eHealth development models [2,10], mostly to be able to constantly improve these models to keep them in line with the most recent insights. Although there are several studies that describe development processes of eHealth technologies [16-20], there seems to be no standardized way of reporting and reflecting on the methods used. Also, an in-depth critical reflection on the development model used is often lacking. To fill these gaps in the literature, this case study presents and reflects on the development process of a virtual reality (VR) application for forensic mental health care. This study had 2 main goals. First, it aimed to increase knowledge on suitable methods for participatory eHealth development. This contributes to creating the aforementioned tool kit. Second, it aimed to reflect on the development model used to guide the process: the CeHRes Roadmap. Combined with other studies that reflect on this model, this can result in further improvement of the Roadmap.

Methods

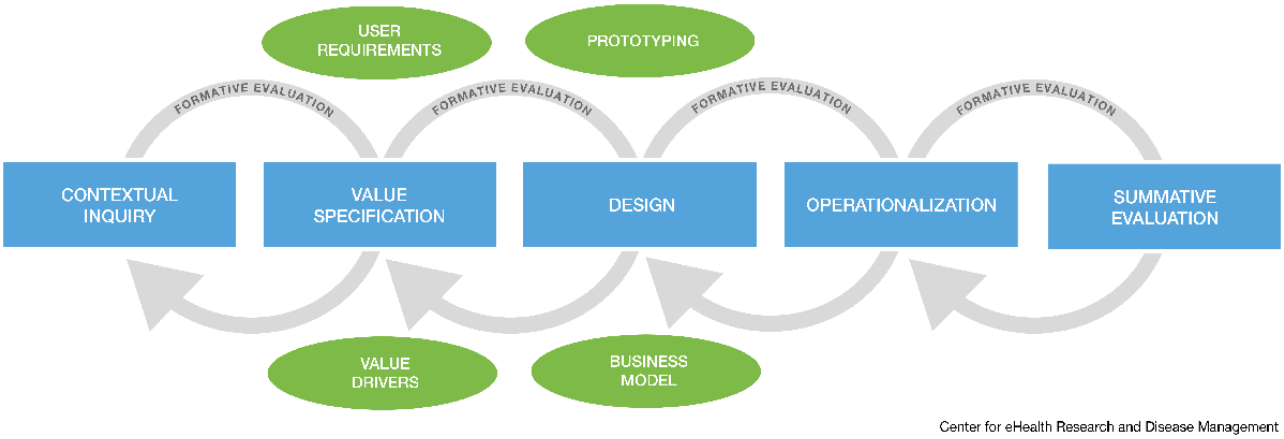
The CeHRes Roadmap

In this study, the aforementioned CeHRes Roadmap [9] was applied to shape the development process of the VR application. This development model specifically focuses on eHealth development, implementation, and evaluation with structural stakeholder involvement [1,6,9]. The Roadmap has been proven useful for eHealth development in multiple settings [16,18,21] and seems to be suitable for development in complex contexts [9], such as forensic mental health care. The Roadmap is based on 5 principles that are also acknowledged by other studies on eHealth development:

- eHealth development should be a participatory process—structurally and actively involving stakeholders during development is important [7,10,12,21].
- eHealth should not be seen as a separate, stand-alone tool but has to be integrated in a health care context, which also implies changes in the way health care is delivered [5,22,23].
- eHealth development and implementation should be intertwined; implementation is a very complex activity that should be accounted for from the start of the development process [24,25].
- eHealth technologies should be based on theories from persuasive design, which can be used to support behavior and attitude change via technology [13].
- Continuous, formative evaluation in eHealth development is important to enable creating by evaluation [7,8,14,26].

XSL•FO
RenderX

Figure 1. The CeHRes Roadmap (the Centre for eHealth Research Roadmap) [9].



These principles are translated into a model with 5 phases with accompanying goals, which are presented in Figure 1 [9]. This model can be used by developers to shape their development approach [3,6,9]. As the aim of this paper is to describe the development of eHealth technology and not the implementation or evaluation, the focus lies on the first 2 phases of the Roadmap: the contextual inquiry and value specification. These phases aim to create a thorough foundation for a technology and account for the interrelationship among the context, the people involved, and the technology. In the contextual inquiry, relevant stakeholders are identified, their roles, tasks, and opinions are analyzed, and the current situation and its weak and strong points are described to determine if and in what way technology can contribute. In the value specification, the values of the key stakeholders have to be identified and prioritized to determine what the added value of a technology should be. These values have to be translated into specific requirements that state what the technology should be able to do and look like [6].

Case

Due to the involvement of 2 of the researchers in the development process, this research can be labeled as an action study. In this study, the development process of a VR application for the treatment of forensic psychiatric patients is presented. This project was initiated and mostly took place at Transfore, a forensic hospital in the east of the Netherlands, which offers forensic mental health care to both in- and outpatients. Forensic mental health care is a complex branch of mental health care, which is situated at the intersect between mental health care and the law because it deals with the combination of mental illness and delinquent behavior. In forensic mental health care, inpatients who reside in a closed setting and outpatients who are living at home are treated for sexual or aggressive criminal behavior [20,21]. A primary goal is to prevent criminal recidivism by means of treatment of offense-related factors, such as antisocial behavior or coping skills. Owing to their low motivation for treatment, low educational levels, and comorbid psychiatric disorders [22-24], forensic psychiatric patients can be characterized as a vulnerable patient population [25,26], which can be hard to include in research [27].

Multiple studies have pointed out the potential of VR for the assessment and treatment of forensic psychiatric patients [27-29]. VR offers the possibility to practice coping skills instead of talking about them, can be used to overcome practical issues for inpatients residing in clinics, and can enable therapists to observe patients' reactions to offense-related stimuli or situations, such as children, drugs, or aggressive persons [29-31]. In VR, users enter computer-generated worlds that substitute their real-world sensory experiences with virtual ones [32], resulting in a feeling of presence: a sense of actually being in a virtual place [33]. Although VR applications have been used in mental health care, especially in exposure therapy for phobias [34], not much is known about its application in the treatment of forensic psychiatric patients [27]. Furthermore, little attention has been paid to how VR interventions should be developed for mental health care in general [32]. In our recent systematic review, we found that there are hardly any studies that discuss the development of technologies for forensic mental health [28]. However, especially in such a complex context in which there is little experience with the application of VR, thorough development is important [10,27]. Consequently, a thorough contextual inquiry and value specification to provide a good foundation for the application were especially important.

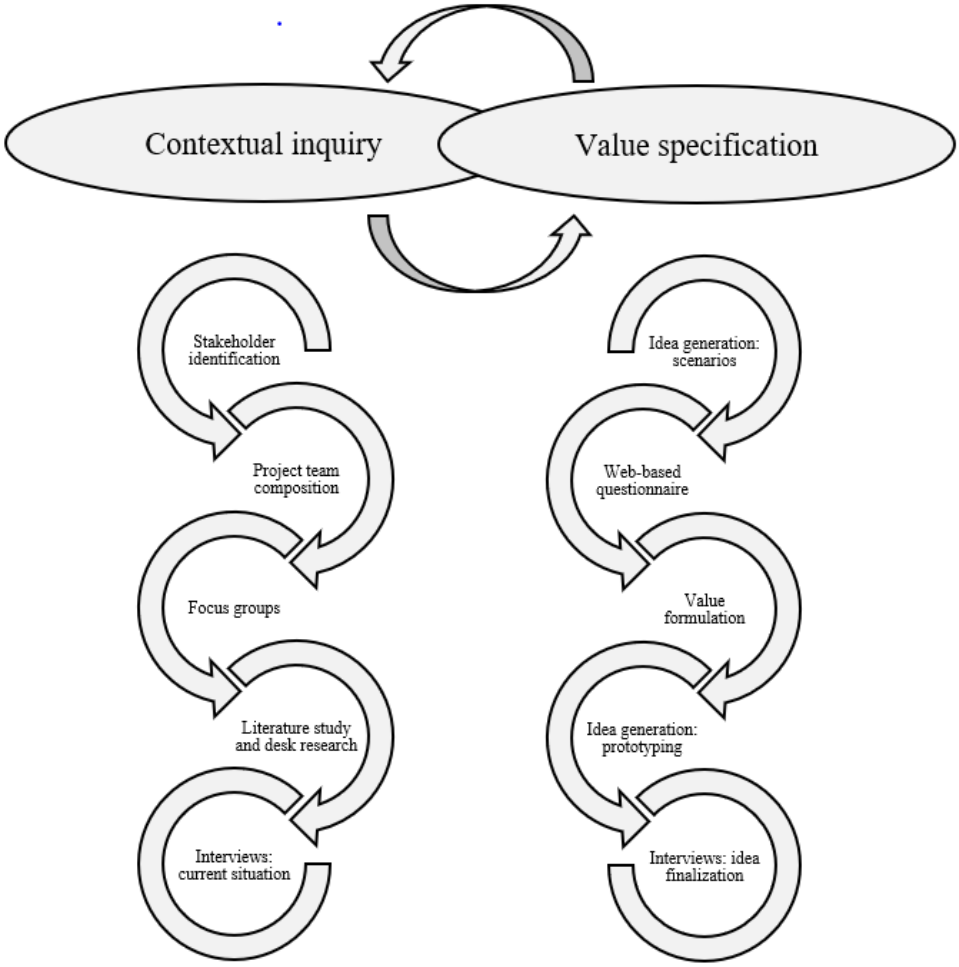
Materials and Procedures

In this study, multiple methods were used to operationalize the first 2 phases of the CeHRes Roadmap. The development process started with the contextual inquiry. In this phase, the stakeholders were identified, a literature review was conducted, and a multidisciplinary project team to coordinate the project was constituted. Also, focus groups and interviews with forensic patients and therapists were held. In the value specification phase, 6 scenarios with concepts for VR applications were generated by the multidisciplinary project team. These concepts were presented to the patients, therapists, and stakeholders in a Web-based questionnaire. Next, values were formulated and used to create a concept for a VR app. This concept was visualized in a low-fidelity prototype and presented to the patients and therapists in an interview to examine their opinions and preferences. These activities were not performed sequentially: several methods were conducted alongside each other or were updated throughout the process [18]. Figure 2 provides an overview of the methods used in the development process. The arrows represent the iterative nature of the process and show that the methods and results of the contextual inquiry

XSL•FO
RenderX

and value specification are not strictly separated but overlap. For more in-depth information about the results of the interviews and questionnaire, we refer readers to 2 other papers [29,30] that focus more on the content of the results and potential of VR for forensic mental health instead of a reflection on the methods and overall development process.

Figure 2. An overview of the used methods in the contextual inquiry and value specification phases of this study.



Analysis

To reflect on the suitability of the methods and overall development process, we provided the most relevant information about each research method in a comprehensive table. The aim of this table is to present the goals, methods, results, and experiences with each method as clearly and concisely as possible. For each development method, the following information is reported:

- *Research question:* The research question for the development activity.
- *Method:* The name of the method, including the most relevant methodological information.
- *Target group:* If applicable, a description of the target group of which the data were collected and characteristics of the participants.
- *Main results*: A summary of the most important results and, if necessary, a reference to a Multimedia Appendix with further information about these results.
- *Conclusions:* The main conclusions and recommendations for further steps of the development process, which were drawn based on the results.
- *Lessons learned*: A reflection on the suitability of the method for the specific development phase, target group, and research question.

Results

Contextual Inquiry

In the contextual inquiry, we generated an overview of relevant stakeholders and their roles and tasks. Furthermore, the current situation and its points of improvement were analyzed to determine if and in what way VR could contribute to treatment in forensic mental health [6]. We used multiple methods that are provided in Table 1 below.

XSL•FO
RenderX

Table 1. An overview of the methods and outcomes and reflection on these methods of the contextual inquiry.

Research goal	Method	Target group	Main results	Conclusions	Lessons learned
Creating an overview of people and organizations who had a stake in the development process	*Stakeholder identification:* Via desk research, expert recommendations, and snowball sampling [31], constantly updated throughout the project	Not applicable	Identification of a broad range of stakeholders, such as end users, financers of care, knowledge institutes, and other forensic care organizations—see Multimedia Appendix 1 (1.1) for a visualization of the identified stakeholders	Stakeholder identification was useful to identify potential financers, participants, or institutions for data collection and to look for potential development partners	• This method served well as a starting point for the project, but as in-depth information about (key) stakeholders was lacking, additional research into stakeholder perspectives was necessary, for example, via interviews. • The stakeholder identification was constantly revised over the course of the project to keep it up-to-date. • The identification proved to be important in preventing the relevant stakeholders from being overlooked in the development process and also in supporting the researchers in identifying participants for studies.
Constituting a multidisciplinary project team comprising patients, therapists, managers, and researchers to coordinate the project	*Project team composition:* In total, 5 potential end users (patients and therapists) were asked to join the team by the policy advisor of the organization in which the project took place (convenience sampling) to coordinate the project [6]. A total of 2 researchers were added for methodological and theoretical knowledge	Not applicable	The project team with 2 patients, 3 therapists, 2 researchers, and 1 policy advisor (n=8) was responsible for content-related and practical activities, such as structuring the development process, setting up studies, and accompanying research goals, interpreting results, decision making, and planning	The multidisciplinary project team was found to be essential for the coordination of the project, mostly because of the integration of different perspectives.	• Including potential end users in the project team was useful to ensure that decisions were aligned with their perspective. In hindsight, the team might have benefited from someone with more technical knowledge on VR[a], for example, a developer. • Practical issues can influence the project team composition, for example, sometimes therapists or patients did not have enough time. It was important to make agreements on what to do when this occurred. • Structure was needed to keep members involved: setting regular meetings, clear communication in between meetings, and keeping minutes of meetings. Coordination by a project manager was important to achieve this. • The project team members had individual, concrete, and specific tasks that helped in keeping everyone actively involved. • Patients indicated that participating in the project team gave them a sense of purpose and helped them with their treatment.

XSL•FO
RenderX

Research goal	Method	Target group	Main results	Conclusions	Lessons learned
Determining how far there is support and enthusiasm for VR[a] in forensic mental health care and identifying the ideas of therapists and patients about potential ways of using VR in treatment	*Focus groups:* Structure: presentations on VR by 2 companies, trying out VR by participants, individually coming up with ideas about VR in treatment, creating ideas in groups of 4, presenting the ideas to the entire group ; The duration of the focus groups was 2 hours and data were collected via researchers' notes and templates filled in by the participants (see Multimedia Appendix 1: 1.2).	Patients (n=14) and therapists (n=23)	Most participants were very positive about VR. There was a broad range of ideas about using VR, for example, to improve skills, enhance insight by therapists or loved ones, or treat specific disorders, such as psychosis or posttraumatic stress disorder. See Multimedia Appendix 1 (1.3) for a table with the main results of the focus groups.	There appeared to be many possibilities, but further specification and insight into why and how VR should be used was required	• Focus groups were a good and efficient way to start this broad, complex project with many possible outcomes, mostly to get an idea of attitudes and potential end users. • These focus groups aimed to generate idea, so provided little in-depth information about needs and goals. It was necessary to complement them with other methods, such as interviews. • The way this focus group was set up was seen as a strong point: there was a clear structure without much steering on content, which enabled all participants to brainstorm freely and individually. This resulted in a very broad range of ideas, which was relevant for this phase of the development process. • It was relatively easy to find participants for the focus groups. An important reason for this seemed to be the possibility to learn more about and try out VR.
Gaining an overview of all studies and current initiatives concerning VR in treatment of (forensic) psychiatric patients	*Literature study and desk research:* Scientific database, search string (virtual reality OR VR OR augmented reality OR AR) AND (treatment OR intervention OR therapy) AND (forensic OR offend* OR crim*) and searching the internet, talking to stakeholders, and visiting conferences	Not applicable	In July 2017, only 6 relevant studies were found, mostly focused on the assessment of sexual delinquents [32-35] or general literature studies on VR [36,37]. Multiple ongoing projects were identified via desk research but with no accompanying scientific publications or available products	Not much is known about VR in forensic mental health care in both practice and research, so there appeared to be a need for a bottom-up development process to identify why and in what way VR could be used	• Especially, desk research proved to be relevant for the project because there were no publications (yet) about many recent, ongoing initiatives/projects. • The strategy for desk research could have been more structured, for example, by creating an activity plan and planning recent updates of desk research. • It was important to look outside of the focus of the project (eg, studies on VR in general), either by conducting a literature study (which is time consuming) or by searching for published reviews or meta-analyses. • It might have been useful to systematically collect the literature on theories and models on delinquent behavior, as in this project, it was done in a more ad hoc manner.
Identifying points of improvement in the existing forensic mental health treatment of in- and outpatients and possible applications of VR, which could improve the current situation, according to therapists and patients	*Interviews: current situation:* The first part of the interview scheme focused on points of improvement of the current treatment (regardless of VR) [21], the second part focused on the possibilities of VR to improve the current treatment. The outcomes of the focus groups were used to structure the interview scheme	Therapists (n=8) and patients (n=3), working or treated at multiple locations of Transfore, the forensic hospital.		The interviews gave much information about why and how VR could be of added value. However, there were still too many possible directions to make a grounded decision about the goal and content of VR. Additional research into the needs and wishes of end users was required	

XSL•FO
RenderX

Research goal	Method	Target group	Main results	Conclusions	Lessons learned
			Via inductive coding [36], 2 types of codes were identified in line with the 2 research questions. Points of improvement were related to patients' return to society; specific patient characteristics, such as treatment motivation; and treatment characteristics, such as skills training. Possibilities of VR were skills training with interaction, observation of patients' reactions, and creating insight for others. The codes can be found in Multimedia Appendix 1 (1.4)		• The participants were asked to provide scenarios about their own experiences and ideas in an open, explorative manner to prevent too much steering by the researchers. To gain in-depth information, good interviewing skills and probing questions appeared to be important. • Eliciting scenarios in participants proved to be unsuitable for (most) patients, mostly because of the broad questions that required much abstract reasoning. The part with examples from the focus groups worked better but was still experienced as difficult. Also, the interview took 1 hour, which proved to be a threshold for participating and resulted in difficulties with inclusion. • The type of information collected via the interviews would have been hard to retrieve via questionnaires because of the need for probing questions. The research questions might have also been answered by means of (small) in-depth focus groups, which might have been less time-consuming.

[a]VR: virtual reality.

Value Specification

In the value specification phase, the outcomes of the contextual inquiry were used to further specify what the added value of a technology should be according to the key stakeholders. Again, multiple methods were used to identify the stakeholders' preferences and opinions on VR in forensic treatment and prototypes to specify these abstract values were created. These methods are provided in Table 2.

XSL•FO
RenderX

Table 2. An overview of the methods and outcomes and reflection on these methods of the value specification.

Research goal	Method	Target group	Main results	Conclusions	Lessons learned
Generating multiple ideas on the use of VR[a] in forensic mental health care, based on the outcomes of the contextual inquiry	*Idea generation—scenarios:* In 3 sessions, all project team members individually brainstormed about ideas for VR applications. The 6 most promising ideas were worked out in a template (see Multimedia Appendix 1: 1.5) by multiple project team members. On the basis of these templates, scripts were written and 6 short videos were filmed	Not applicable	A short video was created for each of the 6 ideas. All videos had the same underlying structure: the goal of VR, its use during treatment, an example, and the desired outcomes. The videos (with English subtitles) can be watched on YouTube [38]. An example of a scenario can be found in Multimedia Appendix 1 (1.6)	The videos made clear that there are a lot of promising possibilities for VR in forensic mental health, so it appeared to be necessary to make decisions about what to prioritize and why	• The structured approach in which multiple templates were used worked well in this project: it forced all different members of the project team to work and think in a similar way. • Each member of the project team had a clear role with individual responsibilities. This was experienced as helpful in motivating the team members and ensuring that all of their perspectives were present in the 6 ideas. • Creating scripts and videos was very time-consuming, so motivated members who are willing to invest time and effort and enough budget were necessary preconditions for making videos.
Identifying (1) the preferences of stakeholders of the 6 ideas and (2) the stakeholders' values regarding VR in forensic mental health care	*Web-based questionnaire:* After asking sociodemographic questions, the 6 videos were presented to the participants in random order. After each video, the PII[b] [39,40] a question about the participant's grade for the idea, and 3 open questions on positive points, points of improvement, and suggestions for the idea were provided	Patients (n=19); therapists (n=89); other stakeholders (n=38), such as parole officers or researchers from different Dutch forensic institutions	There were no significant differences between the grades and PII scores for ideas. A broad range of positive and negative aspects and remarks were identified via inductive coding. These can be found in Multimedia Appendix 1 (1.7)	The results of the questionnaire were mostly in line with the interviews but provided more detailed and specific information, for example, how VR should be personalized and which skills should be trained	• The answers of the patients fitted the research questions of the questionnaire better than the answers that were given by patients in the interviews. This indicated that the concrete, scenario-based videos were a better way to include the patient perspective than the broad, abstract interviews. • Although the goal was to make this method less time-consuming, filling in the questionnaire still took about 30 minutes, which might explain why a large share of the participants (55.4%) did not fully complete it. A shorter questionnaire might have led to more response but also would mean that less information would have been retrieved. • The quantitative measures indicated no major differences between opinions about ideas. Although it was not clear if this was an issue regarding validity or if there actually were no differences, it was still useful to ask for a grade for each idea. The PII was not of added value in this questionnaire. • Although this method proved to be useful to further specify previously found results, it would not have been suitable as an initial method to gather in-depth information, partly because no probing questions could be asked, and answers were relatively short.

XSL•FO
RenderX

Research goal	Method	Target group	Main results	Conclusions	Lessons learned
Formulating values that capture what the added value of the technology should be for people and context, according to the stakeholders	*Value formulation:* On the basis of all previous results, 2 researchers created attributes that summarized the needs or wishes of stakeholders [17]. On the basis of categories of related attributes, accompanying values that stated what VR should achieve, improve, or add according to the stakeholders [6,17] were formulated. The values were discussed by the project team and minor adjustments were made accordingly	Not applicable	A total of 43 attributes and 13 values were formulated. An example of how a value was created can be found in Multimedia Appendix 1 (1.8). The following values were formulated: fit with patient; improvement of skills; insights into behavior, thoughts, and feelings; bridge between treatment room and practice; generalization of skills to daily life; safety; treatment motivation; unique addition to current treatment; ease of use within treatment; cooperation between patient and therapist; wide applicability; affordability; and constant adaptation of the application	Formulating values proved to be a very good way to *get to the point* and summarize the essence of the results so far. It forced the project team to critically think about the overall added value and goals of the VR app and prevented them from getting lost in details or a tunnel vision	• Values might be difficult to understand for outsiders as they are abstract, concise summaries of the needs and wishes. Consequently, clear definitions of the values were provided to prevent misunderstandings. • Besides their importance for development, the project team determined that values could also be useful to determine what to evaluate: to what extent was the added value actually achieved in practice? This way of thinking about values allowed the project team to think ahead in terms of implementation and evaluation and facilitated a broader view on the VR application. • In hindsight, the process of formulating values was more complex than expected. The project team had to account for the results of all used research methods, combine them in an abstract way, and make decisions about conflicting values, such as the importance of visual realism. A clear guideline for formulating values would have been useful.
Generating a concept for a VR application based on the values and previously gathered results	*Idea generation—prototyping:* The project team discussed the values, attributes, and outcomes of all research activities and their implications for a VR application. Via multiple brainstorming sessions in which multiple low-fidelity prototypes were created, a first version of an idea was developed	Not applicable	The main goal of the VR application was to support therapists and patients in identifying *triggers* that can elicit undesired behavior and search for *helpers* that can support the patient in dealing with these triggers. Patient and therapist can together build personalized scenarios via a dashboard with several categories that contain elements that can be added to a scenario (see Multimedia Appendix 1: 1.9 for the prototype)	The developed concept was a combination of elements of all 6 videos that were created by the project team. Also, important concepts that already arose from the interviews were present in the idea, for example, personalization, skills training, and new insights	• To ensure the consistency of the development process, the idea generation process started with discussing the implications of all earlier conducted studies, even though it was more appealing for the project team to start creating the idea right away. • Visualization of ideas via low fidelity (lo-fi) prototypes appeared to work well during the idea generation process to make abstract concepts more concrete. For example, the team drew multiple dashboards and visualized the structure of the dashboard with post-its. This was experienced as helpful by all members of the project team.
Investigating (1) how far the stakeholders' opinions of the concept match the previously formulated values and (2) if changes to the concept are required for it to optimally fit the stakeholders' preferences		Patients (n=10) and therapists (n=8) from all different locations of Transfore, the forensic hospital		Overall, the idea fits the values of the participant, mostly with regard to the unique added value to treatment. No major changes to the basic idea were necessary. In later stages, attention should be paid to the usability of the application, training, and protocols to successfully embed VR in treatment	

XSL•FO
RenderX

Research goal	Method	Target group	Main results	Conclusions	Lessons learned
	Interviews—idea finalization: In the first part, open-ended questions, based on an adapted version of the TAM[c] [41], were asked to check the attitudes toward the concept of the VR application. The second part focused on the participant's overall opinion of the VR application. The developed low-fidelity prototype and a scenario on its use in treatment were used		The first part was coded deductively using the constructs of the TAM (see Multimedia Appendix 1: 1.10), the second part was coded deductively with the 13 formulated values (see Multimedia Appendix 1: 1.11). Overall, the idea was evaluated positively, but there were some concerns about the ease of use of the application. All values were, to some extent, present in the participants' answers. Most positive remarks were about the added value for treatment, for example, fit with patient and new insights. Points of attention were related to the implementation in treatment		• This second set of interviews was considerably shorter than the first one: they only took about 15 to 20 minutes. It proved to be easier to include patients, which might be because of the relatively little time that was required to participate. • Using the values to code these interviews was useful to determine the positive and negative aspects of the idea in relation to the added value that it should have had. In this way, it became very clear what the points of improvements were, which might not have been the case with an inductive, bottom-up coding process. It also allowed the project team to check whether the idea was still in line with the values. • The TAM was used in the interview scheme and coding process. Although it helped to structurally ask about and analyze the participants' attitudes and intentions, it provided hardly any information about the treatment context and characteristics of (other) persons [42,43]. The second part, in which the added value in general was discussed, appeared to be necessary to paint a full picture of the participants' opinion. • Merely using the TAM would not have sufficed in this interview.

[a]VR: virtual reality.

[b]PII: personal involvement inventory.

[c]TAM: technology acceptance model.

Discussion

Reflection on Development Methods

The main goals of this study were to analyze the suitability of the development methods for participatory eHealth development in a complex context and reflect on the development model used: the CeHRes Roadmap. This study can contribute to the development of a broad tool kit from which researchers can choose appropriate methods for the stage of their development process, participants, and context. In hindsight, this study would have benefited from such a tool kit, as the results showed that all methods generated valuable information, but not each method proved to be very suitable for the target group and their context. Besides generating knowledge on suitable methods, this type of study can also facilitate reflection and accompanying improvements of the development model used. Although this study offers a contribution, more studies that pay attention to development methods and models are required to make generalizable statements about methods and models.

The first goal of this study was to reflect on the suitability of different development methods. The relevance of this goal became clear from the experiences of the project team, as a major challenge was to identify the suitable methods for the forensic psychiatric patient population. These types of vulnerable patient populations are often difficult to involve in research, and not much is known about the suitable methods for these types of population [27,44]. On the basis of the experiences with methods used in this study, several conclusions and recommendations can be drawn on the suitability of methods.

A first set of recommendations focuses on involving patients in research. First of all, working with concrete examples seemed to work better than merely asking patients for their opinion or ideas without much guidance or input [4]. Using existing or potential examples is also possible in the earliest stages of the process, when not much is known yet, and can be done by using methods derived from a human-centered design, such as scenarios, personas, or prototypes [45,46]. A second recommendation based on the findings of this study is to keep data collection as short as possible, because patients might have difficulties with concentration or are not motivated to invest a lot of time. This recommendation is also relevant for health care professionals, because although researchers often want to collect

XSL•FO
RenderX

as much data as possible, the professionals often not have a lot of time to participate [17]. The balance between how much in-depth information should be collected and the duration of data collection was experienced as difficult, so more research on this topic is needed. Finally, participating in research should be perceived as personally relevant or rewarding [47]. Although we used rewards such as VR goggles in the questionnaire and interviews, including participants for the focus groups proved to be easier. A reason for this might have been that participants could experience VR during the focus groups, which was perceived as new and exciting by both patients and therapists. Consequently, it appears to be worthwhile to spend time on identifying personally relevant rewards for participants.

The second set of recommendations centers on combining multiple methods and perspectives to paint a clear and complete picture of the context and stakeholder perspective. First, although involving patients proved to be very valuable, the development process also benefited from the perspectives of other types of stakeholders, such as therapists, managers, researchers, and technology developers, as they might have different needs or a more overarching view [28,48]. For example, the analysis of the first set of interviews showed that patients mostly mentioned the use of VR to observe situations and stimuli, whereas therapists also pointed out the importance of other possibilities, such as skill training, which was not mentioned by the patients. Second, involving participants via multiple methods enabled the project team to gain different types of information that supported them in getting a good grasp of all perspectives on VR in forensic mental health care. Finally, it can be concluded that more knowledge on suitable methods for involving patients, therapists, and other stakeholders in eHealth development is needed to be able to make more generalizable statements and create a tool kit [47].

Operationalization of the Development Model

Besides reflections on development methods, this study also aimed to reflect on the application of the development model that was used: the CeHRes Roadmap. It is of course not possible to conclude whether the development process guided by the CeHRes Roadmap resulted in better outcomes than another development method, partly because that would require 2 parallel development processes in identical settings [18], which is difficult both practically and conceptually. Nevertheless, based on the experiences of the project team, it can be concluded that the CeHRes Roadmap provided a valuable guidance for the development process. This process resulted in a concept for a VR application that is based on the wishes and preferences of the therapists and patients. The fit with their wishes became especially clear in the second round of interviews that showed that participants were enthusiastic about the concept and their opinions closely matched the previously formulated values.

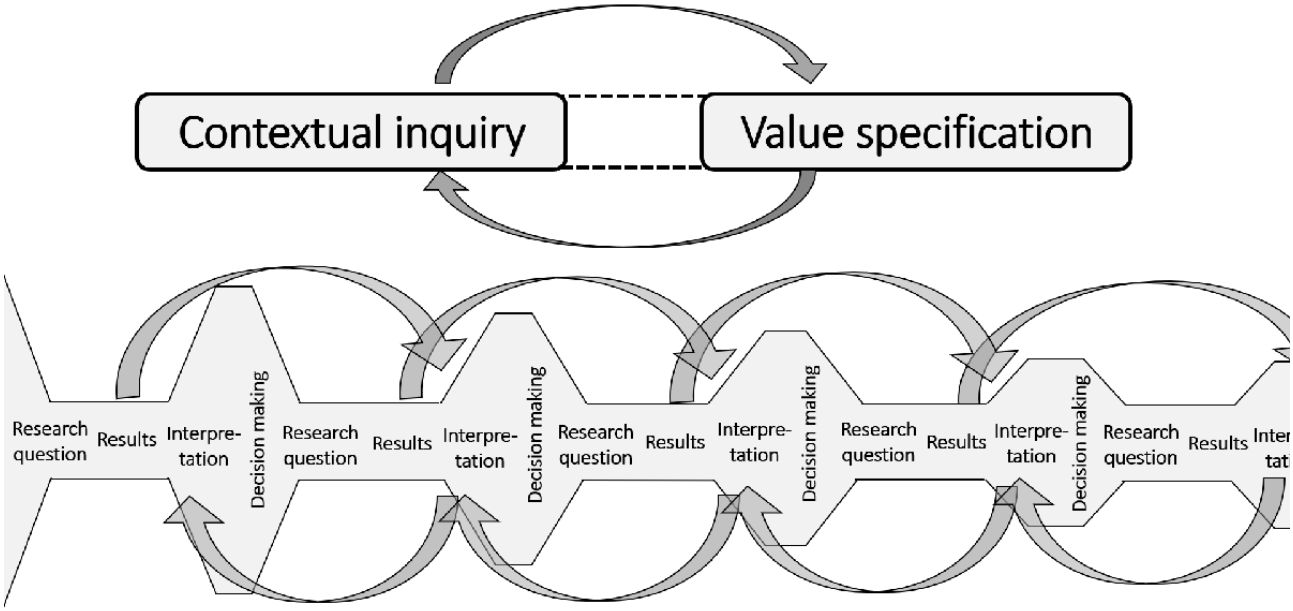
On the basis of the experiences of this study, several recommendations can be made on how to operationalize the CeHRes Roadmap and similar development models. First of all, an important principle of eHealth development is that it should comprise multiple formative evaluation cycles. The experiences of this study confirmed that the Roadmap should not be used as a linear, sequential approach with a fixed order of phases and accompanying activities [18]. To illustrate, the first set of interviews and focus groups provided information that was relevant for both the contextual inquiry and value specification phase. Also, during the value specification, activities from the design phase, such as prototyping and scenarios, were used to elicit opinions. Consequently, although the phases of the Roadmap are visualized as separate blocks (see Figure 1), they should be used as overlapping, interwoven sets of principles and methodologies. A thorough understanding of the principles of the Roadmap appeared to be more important than strictly following the order of separate phases.

A second important finding was that the formulation of clear, specific research goals was pivotal in structuring this development process. A pitfall of an elaborate development process in a complex setting is that it might become unstructured or vague [6,14]. We tried to prevent this by formulating multiple clear, specific research questions that were based on the goals of the Roadmap's phases and its 5 underlying principles [6]. To keep the process coherent, the project team carefully thought about how these research questions related to the outcomes of the previous development activities. Also, we added multiple formative evaluations to check whether the outcomes of different activities remained consistent with each other. This process is visualized in Figure 3.

Third, although constituting and managing an interdisciplinary project team was complex and time-consuming, the team was found to be an important part of the development process as it facilitated decision making from multiple perspectives [22,49,50]. Multidisciplinary teamwork in health care is often complex [51], so several measures were taken to increase the chances on a successful collaboration. Among other things, patients and therapists that participated in the project team were involved as active co-designers instead of passive informants [52,53] and thus took part in activities, such as designing studies, interpreting results, and creating and adapting ideas. To achieve this, the project leader ensured that each project team member had a clear task, as was, for example, done in the creation of scenarios, where each member actively participated in creating an idea and writing the script for 2 of the videos. Fourth, much attention was paid to the functioning of the team. Among other things, roles and tasks of all team members were made clear; regular, bimonthly meetings were held and there was ample communication in between meetings; individual members got the opportunity to be involved in activities of their own choice; there was a mix of skills and interests of members; there was a positive climate of trust and common respect; and, importantly, the team had a common, clear goal [51]. However, as these findings are based on only 1 development process, they are not generalizable. As the functioning of a project team seems to be a relevant topic in eHealth development, more studies on how to compose and organize multidisciplinary project teams should be conducted to be able to draw generalizable conclusions and recommendations. Finally, when operationalizing the CeHRes Roadmap—or any other development model—a thorough understanding of the model's underlying principles, continuous formative evaluations to prevent tunnel vision, clear research questions with suitable methods, and a well-functioning multidisciplinary team were found to be important.

XSL•FO
RenderX

Figure 3. The structure of the goal-driven development process with multiple formative evaluation cycles.



Reflection on the CeHRes Roadmap

While using the CeHRes Roadmap to shape the development process, we identified several strong points but also some points of improvement. First of all, the participatory development principle was used to determine what the main goal of the VR application should be in a bottom-up manner. According to this principle, it is important to involve users from the start to ensure that a technology addresses actual problems or points of improvement and is of added value for them [54]. However, in many cases, the goal of an eHealth technology is determined by researchers and/or developers, and stakeholders are involved as mere informants in later stages to provide feedback on concepts that were created in a top-down manner [55]. In this project, we tried to prevent this by actively involving stakeholders from the start, among other things, by asking them about points of improvement of the current situation and enabling them to come up with their own ideas about VR. Further along the process, values were formulated to specify the goal of the VR application. These values forced the project team to explicitly state the added value that a technology should have for patients and therapists. However, during the value specification, we noticed that there was a lack of clear guidelines on how to formulate these values and what topics they should cover. Although this value-driven approach was experienced as useful to keep an eye on people and their context, there is still much uncharted territory. We recommend that more studies using values in their development process should be conducted to be able to create clear guidelines.

Second, the Roadmap emphasizes the importance of formative evaluation and use of multiple methods. This indeed proved to be essential in this development process, especially because at the start of the project, there was no knowledge about the use of VR in forensic treatment. Consequently, much information had to be generated to make substantiated choices for the goal and content of the VR app. Just using 1 or 2 research methods would not have sufficed. This can be illustrated by the following example on personalization of VR. The first interviews and literature study indeed pointed out that personalization was important [35-37] but did not provide in-depth information about this topic. The results of the questionnaire offered more insights into what stakeholders wanted to be able to personalize: virtual people, environments, and scenarios. Throughout the process, the project team further specified these preferences and translated them into concepts for personalized VR applications via low-fidelity prototypes that were evaluated with stakeholders and fine-tuned accordingly. If only 1 interview study would have been conducted, the project team would not have had enough input to create a personalized VR application. A disadvantage of the multimethod, iterative approach was that it was very time-consuming. It might be possible that, if more would have been known about VR in forensic mental health care or suitable development methods, less research would have been required, which might have resulted in a shorter and more efficient development process. But again, more research on different types of development methods is required to draw more conclusions on this topic.

Finally, when reflecting on the development process, a more systematic approach toward involving domain-specific theories and models could have been used. Owing to the involvement of researchers and professionals with much knowledge on existing treatment models and theories on offending, this information was included but in an ad hoc manner. As other studies and models such as intervention mapping point out, it is important to incorporate theories that explain and change behavior in eHealth interventions [7,12,56,57]. In this project, this relates to models that explain delinquent behavior or theories that underpin treatment of forensic psychiatric patients, such as

XSL•FO
RenderX

the general theory of crime [58] or the risk-need-responsivity model [59]. Consequently, we recommend that the use of domain-specific theories and models to explain behavior and treatment can be explicitly integrated in the Roadmap. To do this, the pillar on persuasive design could be adapted. In its current state, it focuses on behavior change via persuasive design. We suggest a change to this pillar, so that, besides persuasive theory, it also entails the use of domain-specific theories and models throughout the entire development process. Goals and activities derived from this adapted principle could be added to the contextual inquiry and value specification phases to add more focus on domain-specific theories at the beginning of the development process.

Conclusions

This study described and reflected on the methods and development model used in a development process of a VR application for a complex setting: forensic mental health care. To take the domain of eHealth development to the next level, more studies need to report and reflect on the development processes in a standardized way to generate more knowledge on suitable methods. This might result in a tool kit that researchers can use to choose and operationalize methods. Based on this study, we conclude that eHealth development is much more than programing a technology or just going with the flow; it requires thorough research via methods that fit the participants, stage in the development process and context, structured project coordination by a multidisciplinary project team, a flexible and open mind-set, and the inclusion of multiple perspectives in every decision.

Acknowledgments

Funding for this study was provided by Stichting Vrienden van Oldenkotte. The authors would like to thank the members of the *VooRuit met VR* project team for their valuable contributions to this project: Dirk Dijkslag, Kirby Weerink, Ron Voorhuis, Jakob Visser, Kevin Krimmel, and Anne Marike Halma. A special thanks to Ankie Kuiper and Ines Brünninghoff for their work in analyzing the questionnaire and conducting the second set of interviews and to Kirby Weerink for conducting the first round of interviews.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Main results of the research methods.
[PDF File (Adobe PDF File), 1MB - jmir_v21i8e12972_app1.pdf]

References

1. van Gemert-Pijnen L, Kip H, Kelders SM, Sanderman R. Introducing ehealth. In: van Gemert-Pijnen L, Kelders SM, Kip H, Sanderman R, editors. eHealth Research, Theory and Development: A Multi-Disciplinary Approach. London: Routledge; 2018:23-46.
2. Mohr DC, Schueller SM, Montague E, Burns MN, Rashidi P. The behavioral intervention technology model: an integrated conceptual and technological framework for ehealth and mhealth interventions. J Med Internet Res 2014 Jun 5;16(6):e146 [FREE Full text] [doi: 10.2196/jmir.3077] [Medline: 24905070]
3. van Gemert-Pijnen L, Peters O, Ossebaard HC, editors. Improving eHealth. The Netherlands: Eleven International Publishing; 2013.
4. Beerlage-de Jong N. eHealth vs Infection: Participatory Development of Persuasive eHealth to Support Safe Care. The Netherlands: University Library/University of Twente; 2016.
5. Pieterse ME, Kip H, Cruz-Martínez RR. The complexity of ehealth implementation: a theoretical and practical perspective. In: van Gemert-Pijnen L, Kelders SM, Kip H, Sanderman R, editors. eHealth Research, Theory and Development: A Multi-Disciplinary Approach. London: Routledge; 2018:247-270.
6. Kip H, van Gemert-Pijnen L. Holistic development of ehealth technology. In: van Gemert-Pijnen L, Kelders SM, Kip H, Sanderman R, editors. eHealth Research, Theory and Development: A Multi-Disciplinary Approach. London: Routledge; 2018:151-186.
7. Michie S, Yardley L, West R, Patrick K, Greaves F. Developing and evaluating digital interventions to promote behavior change in health and health care: recommendations resulting from an international workshop. J Med Internet Res 2017 Dec 29;19(6):e232 [FREE Full text] [doi: 10.2196/jmir.7126] [Medline: 28663162]
8. 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]

XSL•FO
RenderX

9. van Gemert-Pijnen JE, Nijland N, van Limburg M, Ossebaard HC, Kelders SM, Eysenbach G, et al. A holistic framework to improve the uptake and impact of ehealth technologies. J Med Internet Res 2011 Dec 5;13(4):e111 [FREE Full text] [doi: 10.2196/jmir.1672] [Medline: 22155738]
10. Yardley L, Morrison L, Bradbury K, Muller I. The person-based approach to intervention development: application to digital health-related behavior change interventions. J Med Internet Res 2015 Jan 30;17(1):e30 [FREE Full text] [doi: 10.2196/jmir.4055] [Medline: 25639757]
11. Mohr DC, Lyon AR, Lattie EG, Reddy M, Schueller SM. Accelerating digital mental health research from early design and creation to successful implementation and sustainment. J Med Internet Res 2017 Dec 10;19(5):e153 [FREE Full text] [doi: 10.2196/jmir.7725] [Medline: 28490417]
12. Eldredge LK, Markham CM, Ruiter RA, Fernández ME, Kok G, Parcel GS. Planning Health Promotion Programs: An Intervention Mapping Approach. Hoboken, New Jersey: John Wiley & Sons; 2016.
13. Oinas-Kukkonen H, Harjumaa M. Persuasive systems design: key issues, process model, and system features. Commun Assoc Inf Syst 2009;24(1):28. [doi: 10.17705/1CAIS.02428]
14. Hekler EB, Klasnja P, Riley WT, Buman MP, Huberty J, Rivera DE, et al. Agile science: creating useful products for behavior change in the real world. Transl Behav Med 2016 Dec;6(2):317-328 [FREE Full text] [doi: 10.1007/s13142-016-0395-7] [Medline: 27357001]
15. van Gemert-Pijnen L, Kelders SM, Kip H, Sanderman R, editors. eHealth Research, Theory and Development: A Multi-Disciplinary Approach. London: Routledge; 2018.
16. Wentzel J, van Velsen L, van Limburg M, de Jong N, Karreman J, Hendrix R, et al. Participatory ehealth development to support nurses in antimicrobial stewardship. BMC Med Inform Decis Mak 2014 Jun 5;14:45 [FREE Full text] [doi: 10.1186/1472-6947-14-45] [Medline: 24898694]
17. van Velsen L, Wentzel J, van Gemert-Pijnen JE. Designing ehealth that matters via a multidisciplinary requirements development approach. JMIR Res Protoc 2013 Jun 24;2(1):e21 [FREE Full text] [doi: 10.2196/resprot.2547] [Medline: 23796508]
18. Kelders SM, Pots WT, Oskam MJ, Bohlmeijer ET, van Gemert-Pijnen JE. Development of a web-based intervention for the indicated prevention of depression. BMC Med Inform Decis Mak 2013 Feb 20;13:26 [FREE Full text] [doi: 10.1186/1472-6947-13-26] [Medline: 23425322]
19. van der Weegen S, Verwey R, Spreeuwenberg M, Tange H, van der Weijden T, de Witte L. The development of a mobile monitoring and feedback tool to stimulate physical activity of people with a chronic disease in primary care: a user-centered design. JMIR Mhealth Uhealth 2013 Jul 2;1(2):e8 [FREE Full text] [doi: 10.2196/mhealth.2526] [Medline: 25099556]
20. Wolpin SE, Halpenny B, Whitman G, McReynolds J, Stewart M, Lober WB, et al. Development and usability testing of a web-based cancer symptom and quality-of-life support intervention. Health Informatics J 2015 Mar;21(1):10-23 [FREE Full text] [doi: 10.1177/1460458213495744] [Medline: 24406906]
21. Beerlage-de Jong N, van Gemert-Pijnen L, Wentzel J, Hendrix R, Siemons L. Technology to support integrated antimicrobial stewardship programs: a user centered and stakeholder driven development approach. Infect Dis Rep 2017 Mar 30;9(1):6829 [FREE Full text] [doi: 10.4081/idr.2017.6829] [Medline: 28458796]
22. Geissbuhler A. Lessons learned implementing a regional health information exchange in Geneva as a pilot for the Swiss national ehealth strategy. Int J Med Inform 2013 May;82(5):e118-e124. [doi: 10.1016/j.ijmedinf.2012.11.002] [Medline: 23332387]
23. Covvey HD. Healthcare as a complex adaptive system. In: van Gemert-Pijnen L, Kelders SM, Kip H, Sanderman R, editors. eHealth Research, Theory and Development: A Multi-Disciplinary Approach. London: Routledge; 2018:89-110.
24. Greenhalgh T, Wherton J, Papoutsi C, Lynch J, Hughes G, A'Court C, et al. Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. J Med Internet Res 2017 Dec 1;19(11):e367 [FREE Full text] [doi: 10.2196/jmir.8775] [Medline: 29092808]
25. Ross J, Stevenson F, Lau R, Murray E. Factors that influence the implementation of e-health: a systematic review of systematic reviews (an update). Implement Sci 2016 Dec 26;11(1):146 [FREE Full text] [doi: 10.1186/s13012-016-0510-7] [Medline: 27782832]
26. Sieverink F, Köhle N, Cheung K, Roefs A, Trompetter H, Keizer J, et al. Evaluating ehealth. In: van Gemert-Pijnen L, Kelders SM, Kip H, Sanderman R, editors. eHealth Research, Theory and Development: A Multi-Disciplinary Approach. London: Routledge; 2018.
27. Flaskerud JH, Winslow BJ. Conceptualizing vulnerable populations health-related research. Nurs Res 1998;47(2):69-78. [doi: 10.1097/00006199-199803000-00005] [Medline: 9536190]
28. Kip H, Bouman YH, Kelders SM, van Gemert-Pijnen LJ. eHealth in treatment of offenders in forensic mental health: a review of the current state. Front Psychiatry 2018;9:42 [FREE Full text] [doi: 10.3389/fpsyt.2018.00042] [Medline: 29515468]
29. Kip H, Kelders SM, Weerink K, Kuiper A, Brüninghoff I, Bouman YH, et al. Identifying the added value of virtual reality for treatment in forensic mental health: a scenario-based, qualitative approach. Front Psychol 2019;10:406 [FREE Full text] [doi: 10.3389/fpsyg.2019.00406] [Medline: 30873093]

XSL•FO
RenderX

30. Kip H, Kelders SM, van Gemert-Pijnen LJ. Putting the Value in VR: How to Systematically and Iteratively Develop a Value-Based VR Application with a Complex Target Group. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. 2019 Presented at: CHI'19; May 4-9, 2019; Glasgow, Scotland UK. [doi: 10.1145/3290605.3300365]
31. van Woezik AF, Braakman-Jansen LM, Kulyk O, Siemons L, van Gemert-Pijnen JE. Tackling wicked problems in infection prevention and control: a guideline for co-creation with stakeholders. Antimicrob Resist Infect Control 2016;5:20 [FREE Full text] [doi: 10.1186/s13756-016-0119-2] [Medline: 27213040]
32. Renaud P, Trottier D, Rouleau J, Goyette M, Saumur C, Boukhalfi T, et al. Using immersive virtual reality and anatomically correct computer-generated characters in the forensic assessment of deviant sexual preferences. Virtual Real 2013 Oct 30;18(1):37-47. [doi: 10.1007/s10055-013-0235-8]
33. Fromberger P, Meyer S, Kempf C, Jordan K, Müller JL. Virtual viewing time: the relationship between presence and sexual interest in androphilic and gynephilic men. PLoS One 2015;10(5):e0127156 [FREE Full text] [doi: 10.1371/journal.pone.0127156] [Medline: 25992790]
34. Renaud P, Chartier S, Rouleau J, Proulx J, Goyette M, Trottier D, et al. Using immersive virtual reality and ecological psychology to probe into child molesters' phenomenology. J Sex Aggress 2011 Nov 8;19(1):102-120. [doi: 10.1080/13552600.2011.617014]
35. Renaud P, Rouleau JL, Proulx J, Trottier D, Goyette M, Bradford JP, et al. Virtual characters designed for forensic assessment and rehabilitation of sex offenders: standardized and made-to-measure. J Virtual Real Broadcast 2010;7(5):1-9. [doi: 10.20385/1860-2037/7.2010.5]
36. Fromberger P, Jordan K, Müller JL. [Use of virtual reality in forensic psychiatry. A new paradigm?]. Nervenarzt 2014 Mar;85(3):298-303. [doi: 10.1007/s00115-013-3904-7] [Medline: 24549691]
37. Benbouriche M, Nolet K, Trottier D, Renaud P. Virtual Reality Applications in Forensic Psychiatry. In: Proceedings of the 2014 Virtual Reality International Conference. 2014 Presented at: VRIC'14; April 9-11, 2014; Laval, France. [doi: 10.1145/2617841.2620692]
38. YouTube. 2018. Vooruit met VR URL: https://www.youtube.com/playlist?list=PLPPQsPUGLUhmgEKfTAr0wJ6UemosnAsiS [accessed 2018-01-04]
39. Kelders SM. Involvement as a Working Mechanism for Persuasive Technology. In: Proceedings of the International Conference on Persuasive Technology. 2015 Presented at: PERSUASIVE'15; June 3-5, 2015; Chicago, IL, USA p. 3-14. [doi: 10.1007/978-3-319-20306-5_1]
40. Zaichkowsky JL. The personal involvement inventory: reduction, revision, and application to advertising. J Advert 1994 Dec;23(4):59-70. [doi: 10.1080/00913367.1943.10673459]
41. Venkatesh V, Davis FD. A theoretical extension of the technology acceptance model: four longitudinal field studies. Manage Sci 2000 Feb;46(2):186-204. [doi: 10.1287/mnsc.46.2.186.11926]
42. van Velsen L, Evers M, Bara CD, den Akker HO, Boerema S, Hermens H. Manage sciunderstanding the acceptance of an ehealth technology in the early stages of development: an end-user walkthrough approach and two case studies. JMIR Form Res 2018 Jun 15;2(1):e10474 [FREE Full text] [doi: 10.2196/10474] [Medline: 30684434]
43. Legris P, Ingham J, Collerette P. Why do people use information technology? A critical review of the technology acceptance model. Inform Manage 2003 Jan;40(3):191-204. [doi: 10.1016/S0378-7206(01)00143-4]
44. Emergency Preparedness and Response | CDC. 2010. Public Health Workbook: To Define, Locate, and Reach Special, Vulnerable, and At-Risk Populations in an Emergency URL: https://emergency.cdc.gov/workbook/pdf/ph_workbookfinal.pdf
45. Beerlage-de Jong N, Wentzel J, Hendrix R, van Gemert-Pijnen L. The value of participatory development to support antimicrobial stewardship with a clinical decision support system. Am J Infect Control 2017 Apr 1;45(4):365-371. [doi: 10.1016/j.ajic.2016.12.001] [Medline: 28089673]
46. Burns C. Human-centred design. In: van Gemert-Pijnen L, Kelders SM, Kip H, Sanderman R, editors. eHealth Research, Theory and Development: A Multi-Disciplinary Approach. London: Routledge; 2018.
47. Dugas M, Trottier ME, Dansokho SC, Vaisson G, Provencher T, Colquhoun H, et al. Involving members of vulnerable populations in the development of patient decision aids: a mixed methods sequential explanatory study. BMC Med Inform Decis Mak 2017 Jan 19;17(1):12 [FREE Full text] [doi: 10.1186/s12911-016-0399-8] [Medline: 28103862]
48. van der Vaart R, Witting M, Riper H, Kooistra L, Bohlmeijer ET, van Gemert-Pijnen LJ. Blending online therapy into regular face-to-face therapy for depression: content, ratio and preconditions according to patients and therapists using a Delphi study. BMC Psychiatry 2014 Dec 14;14:355 [FREE Full text] [doi: 10.1186/s12888-014-0355-z] [Medline: 25496393]
49. Feldman SS, Schooley BL, Bhavsar GP. Health information exchange implementation: lessons learned and critical success factors from a case study. JMIR Med Inform 2014 Aug 15;2(2):e19 [FREE Full text] [doi: 10.2196/medinform.3455] [Medline: 25599991]
50. Nielsen JA, Mathiassen L. Interpretive flexibility in mobile health: lessons from a government-sponsored home care program. J Med Internet Res 2013 Oct 30;15(10):e236 [FREE Full text] [doi: 10.2196/jmir.2816] [Medline: 24172852]
51. Nancarrow SA, Booth A, Ariss S, Smith T, Enderby P, Roots A. Ten principles of good interdisciplinary team work. Hum Resour Health 2013 May 10;11:19 [FREE Full text] [doi: 10.1186/1478-4491-11-19] [Medline: 23663329]

XSL•FO
RenderX

52. DeSmet A, Thompson D, Baranowski T, Palmeira A, Verloigne M, de Bourdeaudhuij I. Is participatory design associated with the effectiveness of serious digital games for healthy lifestyle promotion? A meta-analysis. J Med Internet Res 2016 Apr 29;18(4):e94 [FREE Full text] [doi: 10.2196/jmir.4444] [Medline: 27129447]
53. Yip J, Clegg T, Bonsignore E, Gelderblom H, Rhodes E, Druin A. Brownies or Bags-of-Stuff?: Domain Expertise in Cooperative Inquiry With Children. In: Proceedings of the 12th International Conference on Interaction Design and Children. 2013 Presented at: IDC'13; June 24-27, 2013; New York, USA p. 201-210. [doi: 10.1145/2485760.2485763]
54. Kip H, Beerlage-de Jong N, Wentzel J. The contextual inquiry. In: van Gemert-Pijnen JE, Kelders SM, Kip H, Sanderman R, editors. eHealth Research, Theory and Development: A Multi-Disciplinary Approach. London: Routledge; 2018:187-206.
55. Scaife M, Rogers Y, Aldrich F, Davies M. Designing for or Designing With? Informant Design for Interactive Learning Environments. In: Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems. 1997 Presented at: CHI'97; March 22-27, 1997; Atlanta, Georgia, USA p. 343-350. [doi: 10.1145/258549.258789]
56. Moller AC, Merchant G, Conroy DE, West R, Hekler E, Kugler KC, 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 2017 Feb;40(1):85-98 [FREE Full text] [doi: 10.1007/s10865-016-9818-7] [Medline: 28058516]
57. Webb TL, Joseph J, Yardley L, Michie S. Using the internet to promote health behavior change: a systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. J Med Internet Res 2010 Feb 17;12(1):e4 [FREE Full text] [doi: 10.2196/jmir.1376] [Medline: 20164043]
58. Gottfredson MR. A General Theory Of Crime. Stanford: Stanford University Press; 1990.
59. Bonta J, Andrews DA. Risk-Need-Responsivity Model for Offender Assessment and Rehabilitation. Canada: Public Safety Canada; 2007.

Abbreviations

CeHRes: Centre for eHealth Research
eHealth: electronic health
PII: personal involvement inventory
TAM: technology acceptance model
VR: virtual reality

Edited by G Eysenbach; submitted 30.11.18; peer-reviewed by L van Velsen, C Prahm, E Børøsund; comments to author 31.03.19; revised version received 26.04.19; accepted 10.06.19; published 19.08.19.

Please cite as:
Kip H, Kelders SM, Bouman YHA, van Gemert-Pijnen LJEWC
The Importance of Systematically Reporting and Reflecting on eHealth Development: Participatory Development Process of a Virtual Reality Application for Forensic Mental Health Care
J Med Internet Res 2019;21(8):e12972
URL: http://www.jmir.org/2019/8/e12972/
doi:10.2196/12972
PMID:31429415

©Hanneke Kip, Saskia M Kelders, Yvonne HA Bouman, Lisette JEWC van Gemert-Pijnen. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 19.08.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Design and Evaluation of a Contextual Model for Information Retrieval From Web-Scale Discovery Services to Improve Evidence-Based Practice by Health Care Practitioners: Mixed Methods Study

Alvet Miranda[1], BSc, MBA, MAppFin; Shah Jahan Miah[1], BSc, MSc, MInfoTech, PhD

Victoria University Business School, Victoria University, Footscray Park Campus, Footscray, Australia

Corresponding Author:
Alvet Miranda, BSc, MBA, MAppFin
Victoria University Business School
Victoria University
Footscray Park Campus
Footscray,
Australia
Phone: 61 432498563
Email: contact@alvet.com.au

Abstract

Background: Practicing evidence-based health care is challenging because of overwhelming results presented to practitioners by Google-like Web-scale discovery (WSD) services that index millions of resources while retrieving information based on relevancy algorithms with limited consideration for user information need.

Objective: On the basis of the user-oriented theory of information need and following design science principles, this study aimed to develop and evaluate an innovative contextual model for information retrieval from WSD services to improve evidence-based practice (EBP) by health care practitioners.

Methods: We identified problems from literature to support real-world requirements for this study. We used design science research methodology to guide artefact design. We iteratively improved prototype of the context model using artificial formative evaluation. We performed naturalistic summative evaluation using convergent interviewing of health care practitioners and content analysis from a confirmatory focus group consisting of health researchers to evaluate the model's validity and utility.

Results: The study iteratively designed and applied the context model to a WSD service to meet 5 identified requirements. All 5 health care practitioners interviewed found the artefact satisfied the 5 requirements to successfully evaluate the model as having validity and utility. Content analysis results from the confirmatory focus group mapped top 5 descriptors per requirement to support a true hypothesis that there is significant discussion among participants to justify concluding that the artefact had validity and utility.

Conclusions: The context model for WSD satisfied all requirements and was evaluated successfully for information retrieval to improve EBP. Outcomes from this study justify further research into the model.

(J Med Internet Res 2019;21(8):e12621) doi:10.2196/12621

KEYWORDS

information retrieval; design science research; Web-scale discovery; evidence-based practice; libraries, digital; artefacts

Introduction

Background

Evidence-based practice (EBP) represents an amalgamation of research evidence, clinical experience and expertise, and patient values or preferences in the process of clinical patient care [1], as shown in Figure 1. This study introduces a contextual approach to research evidence modeled using design science research (DSR) [2] so health care practitioners can improve EBP.

A positive correlation was found between clinician's behavior and utilization of research evidence from information systems as identified by a study involving 439 nurses and physicians

XSL•FO
RenderX

from public and private hospitals [3]. The study, however, concluded that more support is needed for health care practitioners to use research evidence systems.

There is a myriad of evidence-based clinical information resources across various publisher platforms and use of these resources varies from practitioner to practitioner as illustrated in Figure 2.

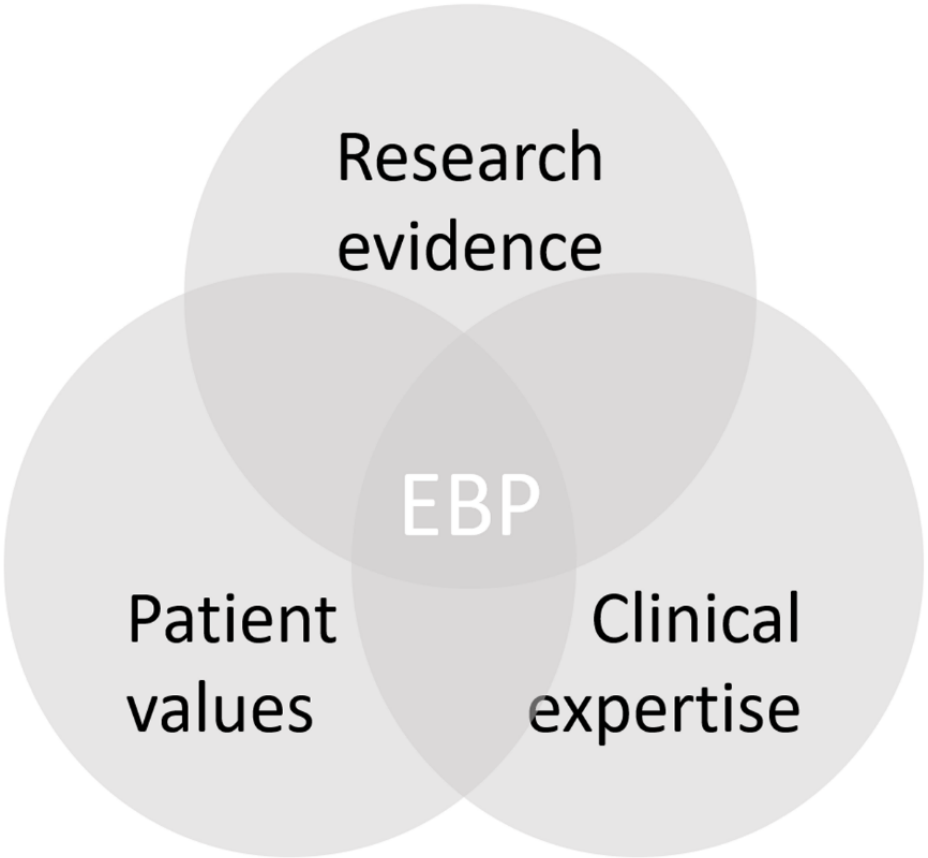
Electronic resources (e-resources) themselves are classified as journals, books, databases, and clinical decisions support references, and each classification has many associated publishers, vendors, and electronic service providers. Being aware of and retrieving information from all these e-resources was challenging for clinicians who needed information for patient care, medical research, and professional growth [4].

At the turn of the 21st century, this problem was solved to an extent by federated search systems; however, they were unreliable and slow as they connected with each publisher platform in real time, causing users to wait several minutes to see results or miss results entirely when connectors failed [5].

Evidence-based clinical information is traditionally licensed content inaccessible to Web search engines, so the information retrieval gap was satisfied by Web-scale discovery (WSD) services starting with OCLC's (Online Computer Library Center) WorldCat Local in 2007, Summon from Serial Solutions in mid-2009, EBSCO Discovery Service (EDS) in 2010, followed by Ex Libris Primo Central and Innovative interfaces Encore Synergy [6].

WSD service–based solutions offer user-friendly search interfaces, relevance ranking, and large, centralized indexes, allowing rapid, simultaneous searching [7]. Figure 3 draws upon Figure 2 and illustrates how a WSD solved the issues of fragmented and unknown e-resources by indexing licensed content metadata and making it retrievable by clinicians using a single interface on the screen. By pre-harvesting content from myriad databases into a single index, WSD tools improve on federated searching tools' speed, de-duplication abilities, relevancy rankings, and the amount of content that can be accessed [8]. WSD services are used worldwide; for example, the EDS is used in over 100 countries across 11,000 institutions [9].

Figure 1. Three components of evidence-based practice.



XSL•FO
RenderX

Figure 2. Health information.

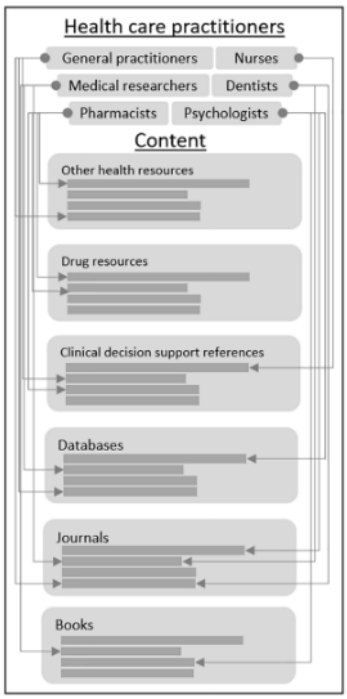
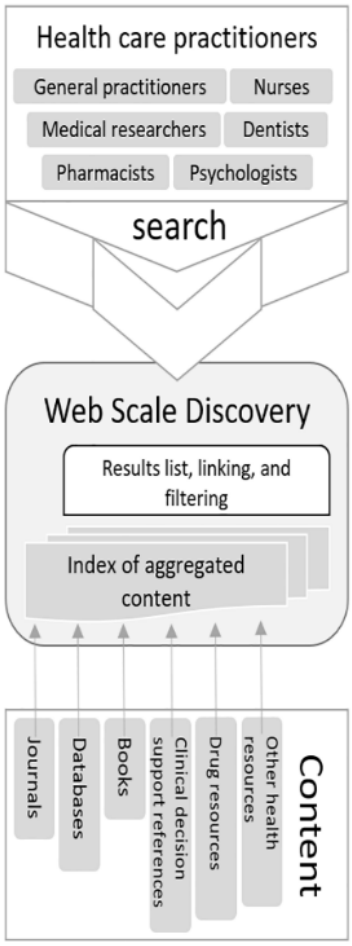


Figure 3. Health care practitioner's retrieving information from a Web-scale discovery.



XSL•FO
RenderX

Problem Identification and Motivation for Research

Focusing on the medical sector, a health science research comparing WSD from 3 vendors concluded that none of the services were overwhelmingly more effective in retrieving relevant information, placing them at 50% to 60% likelihood an average user would find adequate resources [7].

When introducing WSD, a study by Hoy uncovered various disadvantages that needed to be addressed further through new research [10], namely, (1) searches produced too many results overwhelming users, (2) object types and formats caused confusion, (3) users experienced noise, and (4) users needed help using limiters to retrieve specific information.

There are also challenges in offering the same set of e-resources to all clinician groups, as a study [11] comparing the information delivery needs of physicians and nurses found significant discrepancies with wide gaps in behavior and motivation as physicians accessed different online resources from nurses for clinical practice.

Articles from the *Journal of Medical Internet Research* since 2010 discussed research evidence and barriers faced by health care professional to find, appraise, and apply emerging evidence at the point of delivery [12]. As a means to improve EBP adoption rates, health care organizations adopt several strategies such as local consensus processes, distribution of education material, outreach visits, and reminders; however, lack of time, perceived difficulty, and nonintuitive platforms are cited as issues [13]. Web-based knowledge resources had significant impact on access to evidence-based information, which positively correlates with the growth in usage of WSD services over the 2010 decade [14].

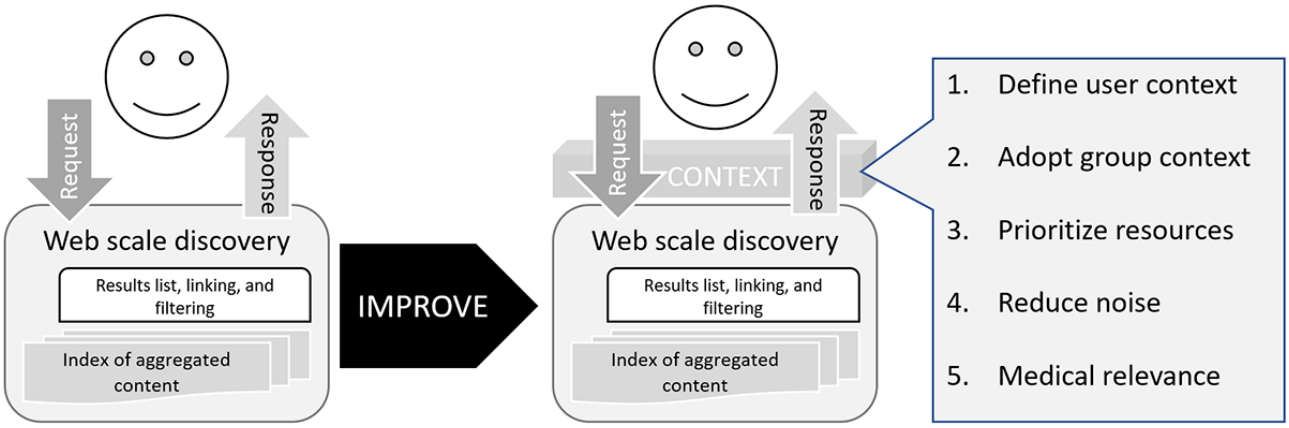
It can be deduced by comparing Figure 2 and Figure 3 that although WSDs gave health care practitioners a single point of entry for information retrieval, the array of aggregated content becomes problematic for individuals and clinical groups requiring context they would otherwise get by using the various content platforms directly.

Aims and Solution Objectives

This research aimed to develop and evaluate a model [15] to improve information retrieval from WSD services so health care practitioners can better adopt EBP by satisfying research evidence needs. The model introduces a context layer to existing WSD services to reduce the ambiguity of information by pre-capturing and integrating clinician's context during the information retrieval process [16].

The contextual layer improves WSDs having a computer science perspective to a knowledge formulation or acquisition system based on the theory of information need for information retrieval by Cole connecting information to knowledge [17]. Figure 4 focuses on the WSD part in Figure 3 and illustrates the improvement by adding a context layer between the user and the WSD that meets requirements listed in Table 1 derived from the literature referenced above.

Figure 4. Improving Web-scale discovery using contextual artefact or model.



XSL•FO
RenderX

Table 1. Context model requirements.

No.	Requirement	Description
R1[a]	Define user context	Allows a clinician to select settings and configuration to define their context to apply when querying a WSD[b]
R2[c]	Adopt group context	Adopt a clinical group context to receive similar settings to other health care practitioners. Adjust to user level if necessary
R3[d]	Prioritize resources	Display the resources that are most important to the clinician higher up the result list based on applied context
R4[e]	Reduce noise	Applying context to queries with WSD services should reduce noise from resources not applicable to the clinician
R5[f]	Medical relevance	Improved relevance of information presented to the clinician after the context is applied

[a]R1: Define user context.

[b]WSD: Web-scale discovery.

[c]R2: Adopt group context.

[d]R3: Prioritize resources.

[e]R4: Reduce noise.

[f]R5: Medical relevance.

Literature Review

This section reviews prior work including problems and findings from practice relevant to this research. It also explores prior knowledge and artefacts attempting to solve similar problems to scope knowledge gap. Finally, a literature-based conceptual framework is developed to guide this research.

Problems With Web-Scale Discovery Services for Evidence-Based Practice

Finding and retrieving research evidence is an essential part of EBP as shown in Figure 1, but a relevant study [4] found that the top 3 barriers to EBP besides time were unfamiliarity with bibliographic databases, difficulty accessing research material, and not seeing the value of research for improving practices.

To minimize these barriers, WSD services offer large centralized indexes of many resources, a relevance ranking algorithm to find information, and a user-friendly interface offering a Google-like experience. Although adoption of WSD services in academia and the public sector is high, their uptake in medical institutions has been slim [7].

Mayo clinic based in the United States is a health care institution that looked to adopt a WSD and published the results from their analysis as they evaluated the "promise to deliver quick, efficient and comprehensive search experience through a single-entry point." The Mayo Clinic WSD workgroup concluded to forgo purchasing a WSD service at the time as their collection and users are heavily concentrated within the medicine and health care disciplines. Necessary requirements to meet their user demands and expectations lacked; however, the study acknowledged any future evaluation should consider a trial because of the customizable nature of the services [18].

In addition to health care–specific information needs, a study into research information needs and barriers [19] found that there were significant differences in needs between a primary care physicians and hospital-based physicians as they have different types of consultations. To add to the complexities of information needs, researchers compared accessing online databases between physicians and nurses to conclude that the behavior and motivations varied significantly between the 2 groups [11]. This is a problem for WSD services as they offer a single point of entry for all health care practitioners and rely on the relevance algorithm, which is not sufficient as established by Hoy's research [10].

Prior Artefacts to Solve Problems

A study comparing 3 WSDs for health care sciences found that none of 3 WSD services—Ex Libris Primo, ProQuest's Summon, and EDS—were overwhelmingly more effective in returning relevant results for health sciences research specifically [7].

According to a study looking at information needs of health care practitioners caring for cancer patients, a range of information systems were developed, but most failed to meet the information needs. A possible reason in their opinion was the systems were developed without respecting and analyzing the information needs of the practitioners [20].

Narayanan, in an article published in December 2017, reviewed the current state of WSD services including the services mentioned before plus OCLC WorldCat Discovery services and concluded that although they have had a positive impact, user-centered requirements such as relevancy and personalization were not present [21].

In the past, however, user-centered or context-sensitive information seeking was attempted by systems with a smaller dataset compared with WSD services. For example, Saparova piloted a federated search system that factored in information needs of 51 clinicians. Clinicians found the search system easy and intuitive to use; however, a key piece of feedback was personalization features as the study concluded that successful adoption of a clinical information system depends on its human, technology, and organization fit [22].

In the book *Implementing Web-Scale Discovery Services—A Practical Guide for Librarians*, Thompson details the levels of customization and configuration capabilities available in the WSD services using the back-end administration tool [23]. This, however, does not satisfy the user and group context requirements, but it eludes to the possibility that a context model

XSL•FO
RenderX

could be implemented to extend or improve information retrieval from WSD.

Requirements Gap

The above literature demonstrates that significant research was conducted to determine the state of WSD services to facilitate EBP by health care practitioners and suggests that information retrieval systems lacked functions to meet the requirements of this research. Table 2 demonstrates the functional gap across systems identified in the literature.

Table 2. Requirements gap.

Previous studies to solve information retrieval	R1[a]	R2[b]	R3[c]	R4[d]	R5[e]
EBSCO Discovery Service	No	No	No	Yes	Yes
Proquest Summon	No	No	No	Yes	Yes
ExLibris Primo (ExLibris was acquired by Proquest)	No	No	No	Yes	Yes
OCLC[f] WorldCat Discovery	No	No	No	Yes	Yes
Federated Search System [22]	No	No	Yes	Yes	Yes

[a]R1: Define user context.

[b]R2: Adopt group context.

[c]R3: Prioritize resources.

[d]R4: Reduce noise.

[e]R5: Medical relevance.

[f]OCLC: Online Computer Library Center.

Conceptual Framework

Using a theory of information need, this research looks to approach developing the solution artefact differently compared with previous attempts discussed in the literature review. Health care practitioners use evidence-based clinical information to deliver patient care, conduct clinical research, and fill knowledge gaps. The WSDs are central to connecting clinicians' information needs to knowledge, and the literature reveals the 2 primary functions of WSDs, which are aggregation of resources or knowledge and facilitate searching using algorithms. A contextual layer is needed to improve information retrieval, and for this we look to Cole's theory of information use. Figure 5 shows how the following types are layered inward such as peels of an onion:

1. Prefocus: User borrows an existing frame for a new topic, and this frame can be based on adjacent topics, analogies, or knowledge stored in memory.
2. Focusing: Focus formulates, and the topic manifests its own frame.
3. The information need frame is fully developed to govern searching the topic.

The fulfillment of the prefocus, focusing, and postfocus inward journey of the user will help form the contextual framework necessary to improve information retrieval from WSDs. Applying these attributes to the context model is central to incorporating a user-oriented lens to this research as highlighted by Cole [17].

While reviewing implementation of WSD services, it is evident that these systems can be customized; however, the context is limited to the institutional level. A key aim of this research is to improve the current systems and offer personalization or context at the clinical group and user level as illustrated in Figure 6.

XSL•FO
RenderX

Figure 5. Theory of information need applied as context to Web-scale discovery. CDSS: WSD: Web-scale Discovery

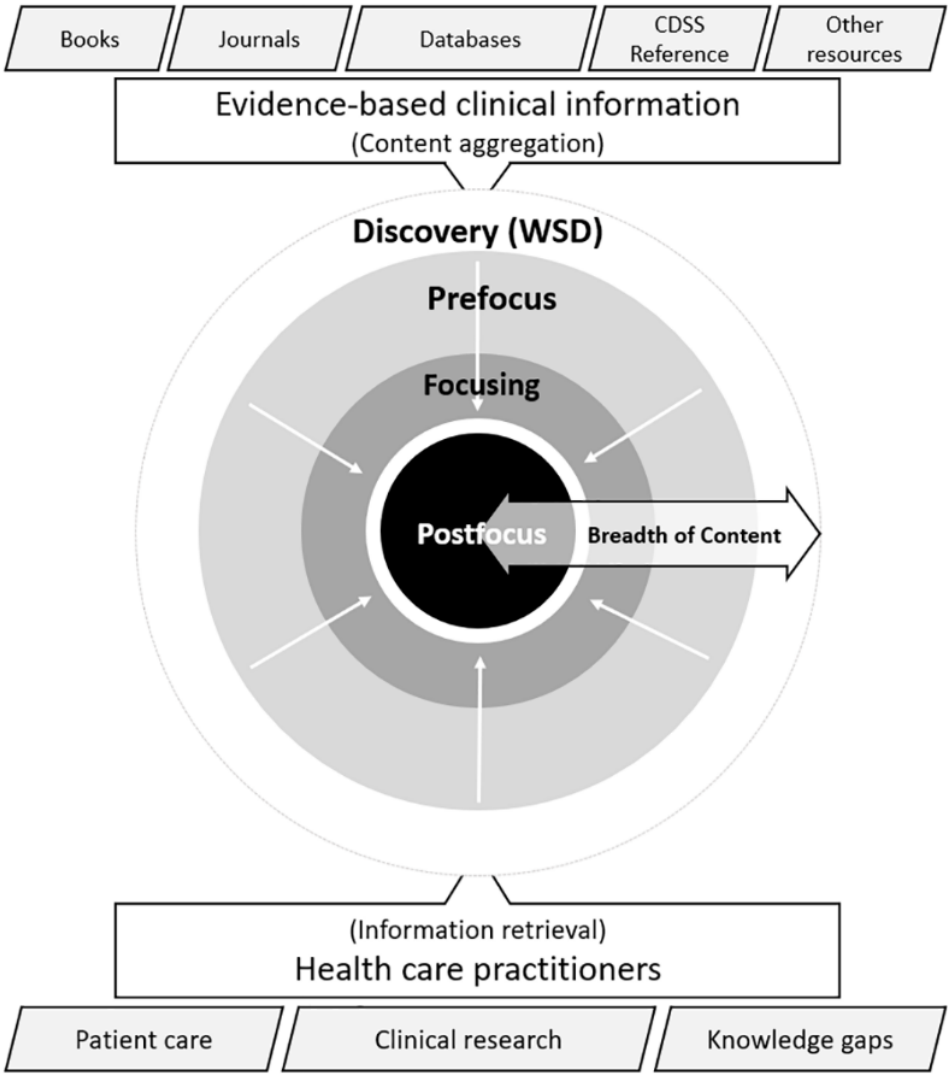
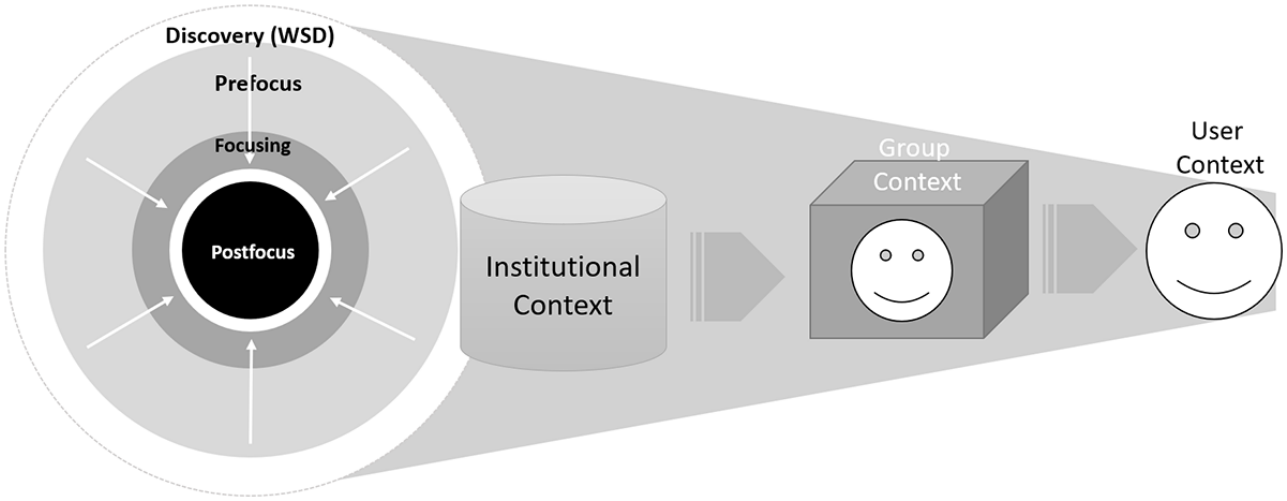


Figure 6. Nature of Web-scale discovery context at institution, group, and user level.



Methods

Paradigm

The study adopted a constructivism paradigm to maintain an appropriate research philosophy [24]. It is a paradigm that encourages intuitive thinking and guesswork from researchers who should discover and learn principles, facts, and concepts for themselves.

XSL•FO
RenderX

Design Science Research

This research will use DSR methodology (DSRM), which falls under the constructivist paradigm to develop a solution artefact model as part of a larger context [25] to solve the problems described earlier in this study. In the case of this research, the artefact or context model created to address the problem [26] is an innovative approach that applies context to information retrieval from WSDs to improve research evidence as part of EBP.

Referring to the knowledge contribution framework in Figure 7, this research falls under the improvement quadrant [27] as it addresses a known problem clinicians and health care professional face while retrieving evidence-based information from WSDs for EBP by developing an innovative artefact model prescribing a contextual approach to retrieving results from WSDs.

The use of DSRM for improvement is a proven principle based on several case studies [2].

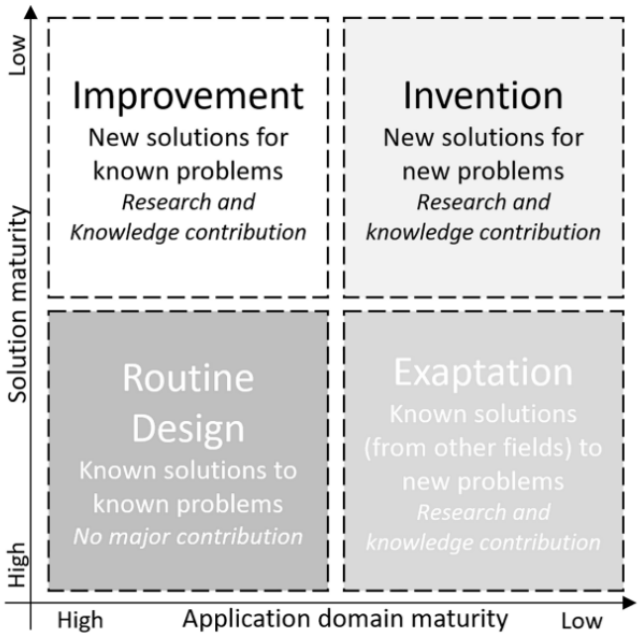
The DSRM illustrated in Figure 8 consists of 6-step nominal process sequences that interact iteratively based on findings from subsequent processes [2]. During step 1, the problem is identified along with motivation for a solution, described in the Introduction section of this study.

This is followed by defining solution objectives (Aims section of this study), which serves as inputs to the design and development for creating the artefact. Versions of the artefact are developed through iterative prototyping [28] and used to solve the problem in the demonstration step.

The evaluation step involves determining the artefact's usefulness based on validity, utility, quality, and efficacy. This research uses the technical risk and efficacy strategy [29] for iterative artificial formative evaluations. Naturalistic summative evaluation via convergent interviewing [30] was used to determine the validity and utility of the artefact with 5 health care practitioners who each were practicing actively for over 15 years. A confirmatory focus group [31] consisting of 5 doctoral level health researchers doing academic research within digital health care for at least 3 years was also conducted, and the transcribed content was analyzed quantitively [32] to ensure significant discussion to support the group's conclusion.

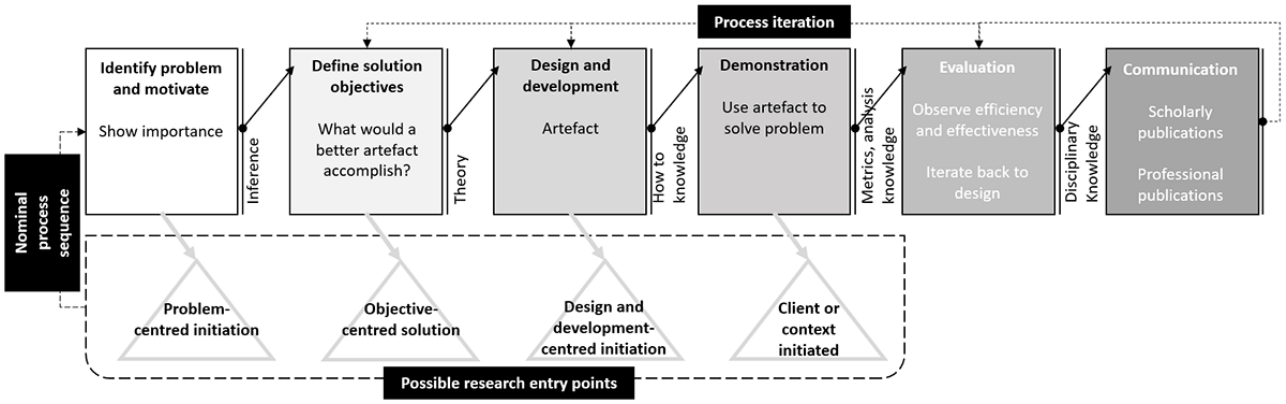
Findings from this study are communicated using recommendations by Gregor [27] to contribute to the existing body of knowledge.

Figure 7. Knowledge contribution framework.



XSL•FO
RenderX

Figure 8. Design Science Research Methodology derived from Peffers (2014).



Artefact Description

Up to this point, this research study has described the need for EBP and how WSD services solved the problem of accessing research evidence as illustrated in Figures 2 and 3, but there are problems with the current services. Figure 4 proposed an improvement by applying a contextual layer to Cole's model of a user-centered system instead of a computer science–focused system as illustrated in Figure 5. Figure 9 draws on gradual development of this theory to propose a contextual model to integrate with existing WSD services.

Starting from the rightmost end of the illustration, a user first borrows a frame for information retrieval for prefocus. For WSD services without the contextual model, the frame is an institutional level context that was configured as part of the WSD implementation. For WSD services that implement the contextual model, the frame can be further personalized to the user by allowing them to borrow a peer's frame. Focusing deeper, a user's focus formulates as part of the user's own information needs until an information need is fully formed as postfocus to capture the context and submit it with the query to retrieve results that are applicable to the user's context.

Developing the context, however, is a prefocus step where the model allows institutions to iteratively develop clinical group templates to borrow in the group template store by analyzing resource needs if a WSD was not available. These initial templates satisfy the adopt group context requirement of the model to help create a scope for information retrieval.

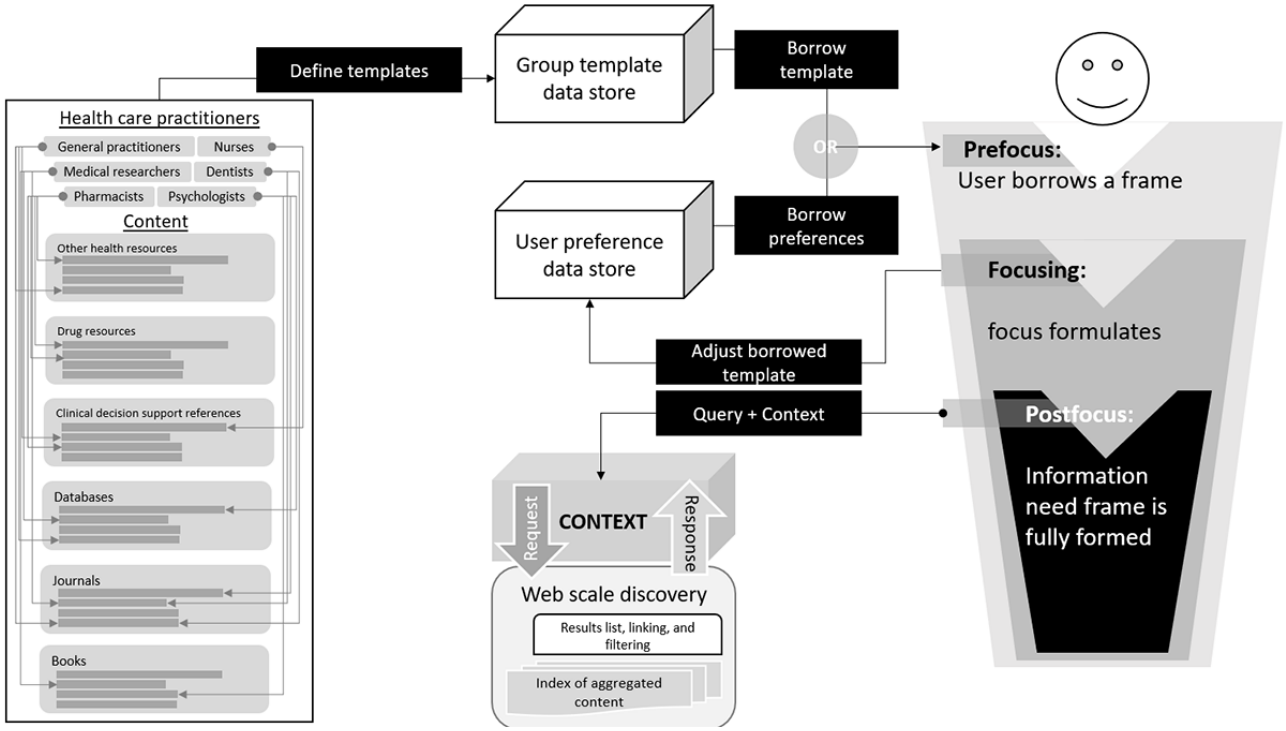
Individual users then have the option to personalize the template and define a more specific user context, which is stored in the user preference data store. This introduces persistence for context, so users do not have to define their context every time they initiate a new session on the WSD. The context is applied as soon as the system recognizes the user.

Data collection and analysis of the model was done through observation and refinements based on iterative prototyping [28]. Instantiations of the model were developed on the EDS [33] using its application programming interface (API) to iteratively prototype the context layer. A combination of Web technologies such as HTML, CSS, and JavaScript was used for the mobile first front-end. PHP was used for server-side, whereas Google Firebase [34] was used for authentication and storage.

On the basis of observations and comparison of results with and without the context layer, the model and subsequent instantiations were refined several times until it was observed that the model had an impact on information retrieved from the WSD to satisfy R1 to R5.

XSL•FO
RenderX

Figure 9. Artefact model to incorporate context when retrieving information from Web-scale discovery services.



Results

Evaluation Overview

Once observations showed sufficient evidence that the prototype as a context layer satisfied R1 to R5 for information retrieved from a WSD compared with its noncontextual state, the instantiation was evaluated in the real world with select health care practitioners using convergent interviewing [30] for a summative evaluation followed by a confirmatory focus group consisting of health researchers.

Demonstration

As per DSRM, demonstration was used to solve the problems, observe results, and develop an instantiation or prototype of the context model [2].

The first step is to recognize the user. The developed prototype allows users to sign in using providers such as GitHub, Twitter, or Facebook shown in Figure 10. Once the user authorizes the instantiation to recognize the user, it checks to see if the user has already borrowed a frame or if the default institutional frame should be applied.

The user is presented with all the filters that are available through the WSD API, and the user context is applied automatically if the frame was borrowed before. If the user logs into the instantiation for the first time, none of the filters are preselected, and the default frame is the institutional frame. User can choose to define their context or borrow a context from predefined frames. The prototype shown in Figure 11 includes frames for a general practitioner, nurse, physiotherapist, medical researcher, and psychologist. The model allows institutions to define and present any number of frames to borrow.

The settings enabled through the EBSCO WSD to define the context are databases, source type, subject, publication, and publisher. These settings will be different for WSD from other vendors.

It is not necessary for a user to define a full context right away, and this can be a gradual process as the prototype will store the selections for future reference allowing for gradual refinement of the context.

The screenshot in Figure 12 shows a comparison between information retrieved from the prototype with the context layer (left) versus results without the context layer (right). This example borrows the general practitioner frame to submit the query “fatty liver cure.” The example with the context layer (left) retrieved 5101 results, and the context is visible under the active facets section. The same query without the context layer (right) returned 28,000 results using the institutional frame. The type of information retrieved is presented on screen for qualitative evaluation.

XSL•FO
RenderX

Figure 10. Login screen to recognize user and load existing or create new context.

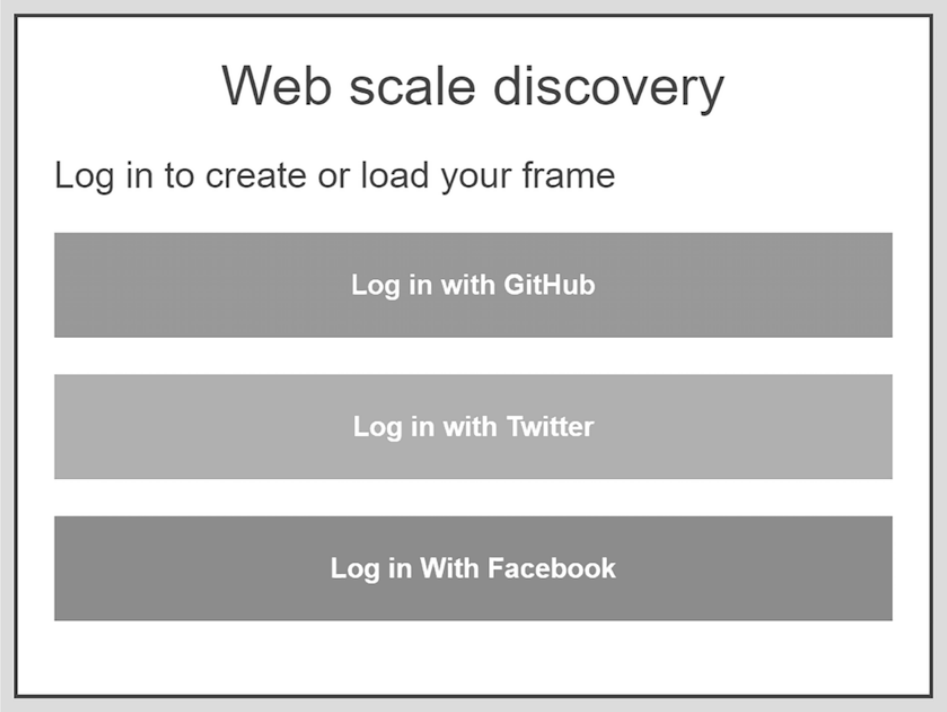
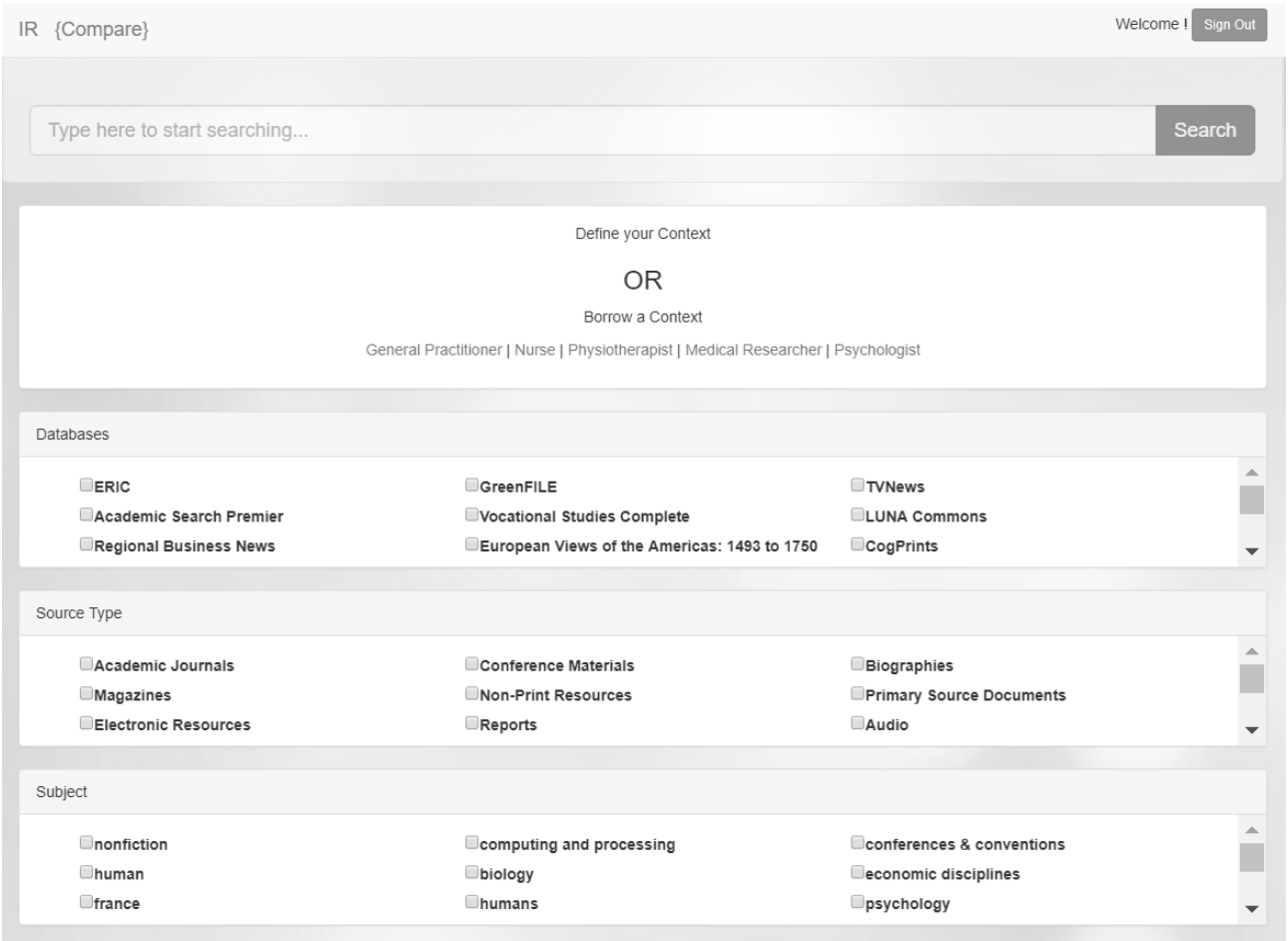
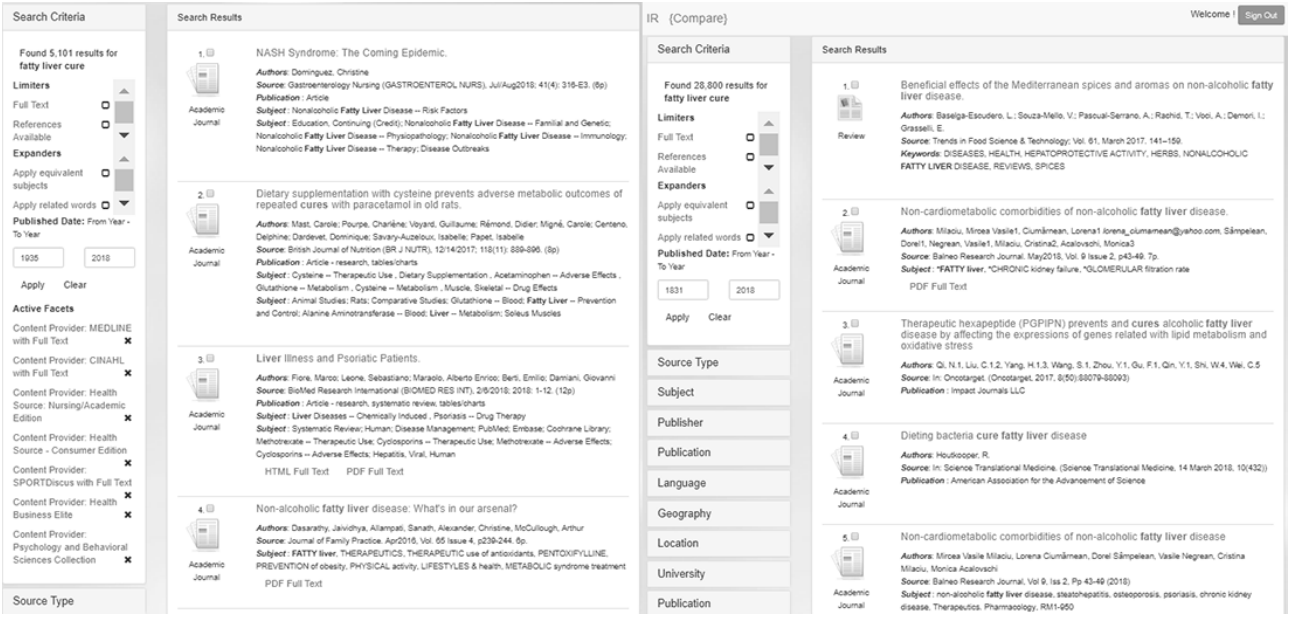


Figure 11. Screenshot of prototype showing options to define context or borrow an existing context.



XSL•FO
RenderX

Figure 12. Screenshot of prototype comparing results with context (left) and without context (right).



Results of Artefact Evaluation

"The technical risk and efficacy evaluation strategy emphasizes artificial formative evaluations iteratively early in the process but progressively moving toward summative artificial evaluations" [29].

A key risk in the evaluation was to determine if a contextual layer could technically be implemented to improve WSD services, so the initial version of the model was developed using the theory of iterative prototyping to allow observations and refinements. Once data collection and observations showed that WSD information retrieval using the context layer satisfied R1 to R5 as part of artefact formative evaluations, a naturalistic summative evaluation was conducted using qualitative and quantitative research methods with health care professionals and health researchers, respectively.

Convergent Interviewing

Convergent interviewing [30] allows researchers to identify and select participants to interview them for key issues and finalize results from qualitative analysis. This technique targeted health care practitioners with a significant length of experience in the health care sector as medical practices has progressed significantly over the years making EBP invaluable. Their practical experience with research evidence to provide patient care using current relevant practices qualifies for qualitative naturalistic summative evaluation.

During the one-on-one interview, the 5 health care practitioners were informed about the nature of this study complying with ethical requirements. Once consent was obtained, a live demonstration of the instantiation was provided to the health care practitioners followed by questions to evaluate the prototype for its validity and utility. They were also able to compare the results against a WSD implementation without the context layer. The results are shown in Table 3, which maps interview questions asked against the requirements identified for this study.

On all 5 counts R1 to R5, the health care practitioners saw benefits and confirmed that the context model improved the information retrieval process using WSD services compared with not using a contextual layer. One of the doctors expressed keenness to have this implemented across their medical practice, while a professor who is also a registered psychologist highlighted the value this would have for his students.

XSL•FO
RenderX

Table 3. Results of prototype evaluation through convergent interviewing.

Requirements, Evaluation criteria, Evaluated by	Satisfied (Yes/No)	Comments (if any)
Evaluating validity		
R1[a]: Does the instance allow you as a clinician to select and configure your own settings for a personalized profile?		
Doctor 1	Yes	"The system gives features for defining my context with preferences."
Doctor 2	Yes	"Profiling seems ok to me but can we add any new details?"
Doctor 3	Yes	"Professional preferences can be added that are supportive in finding quick resources."
Psychologist 1	Yes	"There are many settings for detailed selection."
Psychologist 2	Yes	"Several options not related to my area."
R2[b]: Are you as a clinician able to borrow a predefined context for a clinical group you belong to or adopt another clinical group to see difference in context?		
Doctor 1	Yes	"Yes, it is a good way of seeing a comparison."
Doctor 2	Yes	"The difference can lead to a good outcome."
Doctor 3	Yes	"It is a learning by seeing the difference, although we don't have time to check it during surgery time."
Psychologist 1	Yes	"Seeing different groups shows wide application across healthcare."
Psychologist 2	Yes	"I could select psychologist and other groups."
R3[c]: Does the instance prioritize the resources you as a clinician would prefer to search using a Web-scale discovery?		
Doctor 1	Yes	"It is a good system to see quick and concise searching result."
Doctor 2	Yes	"Searching process seems simple enough."
Doctor 3	Yes	"It can bring me a to-the-point answer."
Psychologist 1	Yes	"Psychologist group pre-selected all the resources I would select generally."
Psychologist 2	Yes	No comment
R4[d]: Does the instance apply your defined context to reduce noise or irrelevant results to your particular case?		
Doctor 1	Yes	"I can see only very relevant resources."
Doctor 2	Yes	"Outcome of searching makes sense to particular aspect."
Doctor 3	Yes	"The application removes unnecessary resources."
Psychologist 1	Yes	"Removing irrelevant databases shows improvement."
Psychologist 2	Yes	"Better compared to results without selecting a group."
Evaluating utility		
R5[e]: Does the instance improve the medical relevance of the Web-scale discovery compared with using it without clinical context?		
Doctor 1	Yes	"It offers better with context."
Doctor 2	Yes	"I'm happy to use this application as it is very work related."
Doctor 3	Yes	"The system supports with our work practices."
Psychologist 1	Yes	"Since only medical sources are selected."
Psychologist 2	Yes	"Saves a lot of time."
R5[e]: Dvidence-based practice or EBP consists of patient values, clinical expertise, and research evidence. Does this instance add value to the research evidence dimension of a clinician doing EPB?		
Doctor 1	Yes	No comment
Doctor 2	Yes	"I found it is helpful for our practice in the future please let us know we are very happy to provide more feedback."
Doctor 3	Yes	"It would be great if we could get the complete product for all of our GPs."

XSL•FO
RenderX

Requirements, Evaluation criteria, Evaluated by	Satisfied (Yes/No)	Comments (if any)
Psychologist 1	Yes	"Would benefit psychology students and new practitioners."
Psychologist 2	Yes	No comment

[a]R1: Define user context.

[b]R2: Adopt group context.

[c]R3: Prioritize resources.

[d]R4: Reduce noise.

[e]R5: Medical relevance.

Confirmatory Focus Group

Focus groups are used as an evaluation method in DSR and are an appropriate approach [35], which can either be exploratory in nature more aligned to formative evaluations or confirmatory used for summative evaluations [31]. In this case, a confirmatory focus group was formed to quantitatively analyze the discussion content to determine significant focus on requirements to justify the group's conclusion.Content analysis is a research technique that explores data obtained directly from sources such as human interactions and written documents for research quantitatively or qualitatively or both [36].

Forming a confirmatory focus group requires selection relevant to the area of research [32], so 5 doctoral level health researchers with hands-on exposure to using a discovery service were invited to take part in the discussion. An initial introduction was provided to the group along with the 5 requirements R1 to R5, as shown in Table 2, that assessed other discover services followed by a demonstration of the model's instantiation to discuss if it met all requirements.

Although there were comments about improving the detailed functionality of the prototype, such as the way the options screen could be more user-friendly and the limiting nature of the context to multidisciplinary content, the group consensus was that the instantiation satisfied all requirements for the artefact to have validity and utility.

Content analysis in the context of focus groups comprises forming a hypothesis to test through analysis. Semantical content analysis "which seeks to classify signs accordingly to their meaning," in particular the subclassification assertion analysis "which provides the frequency with which certain objects are characterized in a particular way" [32], was used to test if the hypothesis that all requirements R1 to R5 were discussed significantly for the focus group to justify the conclusion that the artefact instantiation satisfied all the requirements.

For this analysis, the requirements R1 to R5 were identified as objects positioned as columns and codes or words with frequency positioned as rows in the resulting matrix.

To rigorously analyze the discussion and its correlation to the conclusion, the session was recorded and later transcribed by converting the audio recording to a video file and uploading to YouTube to auto-generate the transcript. The text was copied into a raw HTML file to open in a browser removing new lines and saved as a continuous stream in a text file.

The text file was imported into Microsoft Excel as a space separated data file and transposed so each word appeared in a row. After this, a visual basic script was developed to analyze each word and determine its parts of speech (eg, verb and noun) and populate the cell adjacent to the word. This Excel file with the word, part of speech, and word length as columns was then imported into Tableau Software's Desktop tool for further analysis.

Inductive coding [32] was done using the imported data by creating a scatter plot using the number of times a word was mentioned in the discussion as x-axis and word length on y-axis for further interpretation. Select words belonging to parts of speech such as adverbs, conjunctions, interjections, pronouns, and other were removed leaving only adjectives, nouns, and verbs. Words with length less than 5 characters were ignored including words that were mentioned less than 6 times. The scatterplot was further analyzed word for word to selectively remove obscure words such as apply, being, and possible that did not apply to the research context, leaving the scatter plot with descriptors or codes to present outcomes shown in Figure 13.

Words with their frequency count (descriptors or codes) were analyzed for signs of association with 1 of the requirements R1 to R5 (objects) to form a matrix Table 4 consisting of top 5 descriptors per object to demonstrate that significant discussion for all requirements took place to confirm utility and validity of the artefact to declare the hypothesis as true.

XSL•FO
RenderX

Figure 13. Result of content analysis descriptors from confirmatory focus group.

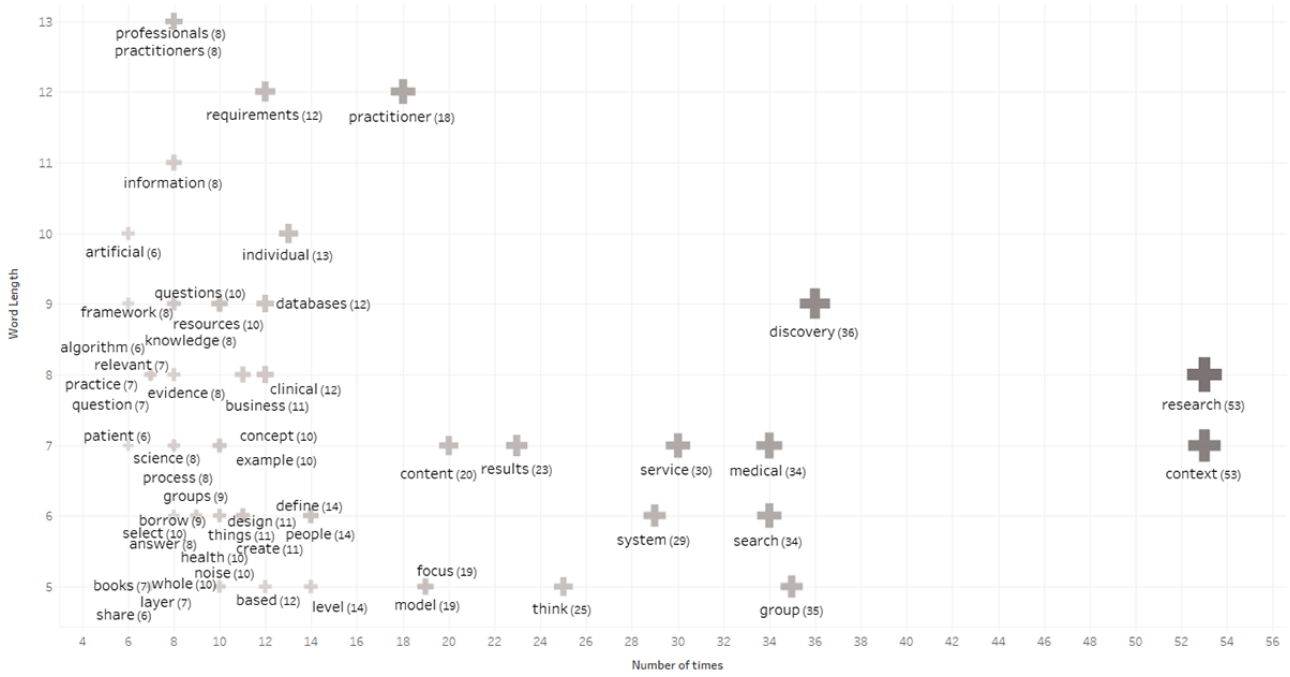


Table 4. Results of assertion analysis.

R1[a]	R2[b]	R3[c]	R4[d]	R5[e]
context (53)	group (35)	focus (19)	noise (10)	medical (34)
individual (13)	level (14)	content (20)	relevant (7)	health (10)
person (7)	select (10)	resources (10)	results (23)	clinical (12)
create (11)	borrow (9)	knowledge (8)	discovery (36)	practitioner (18)
Define (14)	share (6)	information (8)	research (53)	databases (12)

[a]R1: Define user context.

[b]R2: Adopt group context.

[c]R3: Prioritize resources.

[d]R4: Reduce noise.

[e]R5: Medical relevance.

Discussion

Meeting Requirements

Using a combination of artificial formative evaluations based on iterative prototyping followed by naturalistic summative evaluation using convergent interviewing and confirmatory focus group, the context model satisfied all requirements R1 to R5 of this research.

This research reviewed the state of EBP by health care practitioners and identified that the research evidence dimension was a significant barrier because of the vast myriad of resources. WSD services facilitated research evidence by indexing most licensed content while offering practitioners a single point of entry into research; this solved the problem to a certain degree.

The problems with this approach as identified in the literature was the lack of relevancy to the medical field, creation of noise, unprioritized resources, and missing user context. These problems were synonymous with information retrieval that were developed with a computer science approach focusing on systems and algorithms instead of user need.

On the basis of the theory of information need by Cole, this research designed and evaluated a contextual layer or artefact model to apply to WSD services to improve them. Using DSRM principles, this research used iterative prototyping to observe, design, and refine the model and its prototype against the EBSCO WSD using its APIs.

Once the prototype demonstrated that it satisfied requirements R1 to R5 during a series of artificial formative evaluations, 5 health care practitioners were selected for convergent interviewing for a qualitative naturalistic summative evaluation.

Principal Findings

Outcomes from this research provided very positive views about the artefact as all health care practitioners who participated in the field experiments found the proposed approach useful, compatible to their own practices, and added value to WSD services to improve EBP.

XSL•FO
RenderX

A quantitative naturalistic summative evaluation was also performed by forming a confirmatory focus group consisting of 5 health researchers who discussed and concluded that the artefact instantiation met all the requirements. The discussion was recorded and transcribed for content analysis to identify descriptors or codes and associate them to objects or requirements to form Table 4 for a positive hypothesis that there was evidence of significant discussion across the requirements to justify the conclusion.

Evaluation is one of the crucial elements of DSR as it provides feedback to further improve the artefact and bring it to a state where its utility, quality, and efficacy are validated [26].

Prior research [29] shared strategies for DSR evaluation, namely, quick and simple, human risk and efficacy, technical risk and efficacy, and the purely technical artefact strategy. The choice of evaluation depends on the functional purpose of the artefact, which in this research is formative as it aims to improve the outcomes of the process under evaluation. It is also summative as the outcomes will be judged to what extent they match expectations. The other factor for choice of strategy is the evaluation paradigm: if it is artificial or naturalistic.

In this dimension, the paradigm is both artificial and naturalistic as the context artefact is evaluated using iterative prototyping to determine if the context is feasible with a WSD as part of its development. This guided the development in an incremental and iterative fashion to address the research problem. Here, the prototype design and development effort is used to validate or invalidate the theory [28].

Conclusions

The evaluation concluded that the context layer or artefact model had validity and utility as an artefact to solve problems with WSD for EBP. Rigorous research in DSR [29] showed that formative evaluation is necessary to identify weaknesses and areas of improvement during artefact development, and summative evaluation is important to conduct the DSR study by outlining overall utilities or efficiencies and potential benefits of the proposed artefact. Considering artificial evaluation via conducting iterative prototyping and naturalistic summative evaluation, the technical risk and efficacy evaluation strategy was an appropriate fit for this research.

Further research into the contextual approach is recommended to evaluate the artefact using the human risk and effectiveness strategy.

Conflicts of Interest

None declared.

References

1. Samonte PR, Vallente RU. Evidence-Based Practice (EBP). Ipswich, Massachusetts: Salem Press; 2016.
2. Peffers K, Tuunanen T, Rothenberger MA, Chatterjee S. A design science research methodology for information systems research. J Manage Inform Syst 2014 Dec 8;24(3):45-77. [doi: 10.2753/MIS0742-1222240302]
3. Lialiou P, Pavlopoulou I, Mantas J. Health professionals' use of online information retrieval systems and online evidence. Stud Health Technol Inform 2016;228:801-803. [doi: 10.3233/978-1-61499-678-1-801] [Medline: 27577498]
4. Solomons NM, Spross JA. Evidence-based practice barriers and facilitators from a continuous quality improvement perspective: an integrative review. J Nurs Manag 2011 Jan;19(1):109-120. [doi: 10.1111/j.1365-2834.2010.01144.x] [Medline: 21223411]
5. Wisniewski J. Web scale discovery: the future's so bright I gotta wear shades. Online 2010 Jul;34(4):55-57 [FREE Full text]
6. Vaughan J. Web scale discovery what and why? Libr Technol Rep 2011;47(1):5-11 [FREE Full text] [doi: 10.5860/ltr.47n1]
7. Hanneke R, O'Brien KK. Comparison of three web-scale discovery services for health sciences research. J Med Libr Assoc 2016 Apr;104(2):109-117 [FREE Full text] [doi: 10.3163/1536-5050.104.2.004] [Medline: 27076797]
8. Asher AD, Duke LM, Wilson S. Paths of discovery: comparing the search effectiveness of EBSCO discovery service, summon, Google scholar, and conventional library resources. Coll Res Libr 2013 Sep 1;74(5):464-488. [doi: 10.5860/crl-374]
9. EBSCO Information Services. 2018. EBSCO Information Services Releases On Architecture URL: https://www.ebsco.com/news-center/press-releases/ebsco-information-services-releases-onarchitecture [accessed 2018-10-24] [WebCite Cache ID 73ODUR9wV]
10. Hoy MB. An introduction to web scale discovery systems. Med Ref Serv Q 2012;31(3):323-329. [doi: 10.1080/02763869.2012.698186] [Medline: 22853305]
11. Chiu YW, Weng YH, Lo HL, Shih YH, Hsu CC, Kuo KN. Comparison of accessing online databases between physicians and nurses in Taiwan. Inform Health Soc Care 2012 Dec;37(4):230-241. [doi: 10.3109/17538157.2012.654842] [Medline: 22657100]
12. Maloney S, Tunnecliff J, Morgan P, Gaida JE, Clearihan L, Sadasivan S, et al. Translating evidence into practice via social media: a mixed-methods study. J Med Internet Res 2015 Oct 26;17(10):e242 [FREE Full text] [doi: 10.2196/jmir.4763] [Medline: 26503129]
13. David I, Poissant L, Rochette A. Clinicians' expectations of web 2.0 as a mechanism for knowledge transfer of stroke best practices. J Med Internet Res 2012 Sep 13;14(5):e121 [FREE Full text] [doi: 10.2196/jmir.2016] [Medline: 23195753]

XSL•FO
RenderX

14. Eng JJ, Noonan VK, Townson AF, Higgins CE, Rogers J, Wolfe DL. Impact of an online medical internet site on knowledge and practice of health care providers: a mixed methods study of the spinal cord injury rehabilitation evidence project. J Med Internet Res 2014 Dec 23;16(12):e296 [FREE Full text] [doi: 10.2196/jmir.3453] [Medline: 25537167]
15. March ST, Smith GF. Design and natural science research on information technology. Decis Support Syst 1995 Dec;15(4):251-266. [doi: 10.1016/0167-9236(94)00041-2]
16. Tamine-Lechani L, Boughanem M, Daoud M. Evaluation of contextual information retrieval effectiveness: overview of issues and research. Knowl Inf Syst 2009 Jul 15;24(1):1-34. [doi: 10.1007/s10115-009-0231-1]
17. Cole C. A theory of information need for information retrieval that connects information to knowledge. J Assoc Inf Sci Technol 2011 Apr 19;62(7):1216-1231. [doi: 10.1002/asi.21541]
18. Brigham TJ, Farrell AM, Trzasko LC, Attwood CA, Wentz MW, Arp KA. Web-scale discovery service: is it right for your library? Mayo clinic libraries experience. J Hosp Librariansh 2016 Feb 3;16(1):25-39. [doi: 10.1080/15323269.2016.1118280]
19. Davies K. Information needs and barriers to accessing electronic information: hospital-based physicians compared to primary care physicians. J Hosp Librariansh 2011 Jul;11(3):249-260. [doi: 10.1080/15323269.2011.587103]
20. Ciarlo G, Liebl P, Zell J, Fessler J, Koester MJ, Ruetters D, et al. Information needs of oncologists, general practitioners and other professionals caring for patients with cancer. Eur J Cancer Care (Engl) 2016 Nov;25(6):1015-1023. [doi: 10.1111/ecc.12557] [Medline: 27550233]
21. Narayanan N, Byers DF. Improving web scale discovery services. Ann Libr Inf Stud 2017 Dec;64(4):276-279. [doi: 10.2139/ssrn.3139357]
22. Saparova D, Belden J, Williams J, Richardson B, Schuster K. Evaluating a federated medical search engine: tailoring the methodology and reporting the evaluation outcomes. Appl Clin Inform 2014;5(3):731-745 [FREE Full text] [doi: 10.4338/ACI-2014-03-RA-0021] [Medline: 25298813]
23. Thompson J. Configuring and branding the discovery layer. In: Implementing Web-Scale Discovery Services: A Practical Guide for Librarians. Volume 9. Lanham, Maryland: Rowman & Littlefield Publishers; 2014:87-95.
24. Chisholm B. Handbook of Constructivism (Concepts and Theories). Delhi, India: Learning Press; 2012.
25. Johannesson P, Perjons E. An Introduction to Design Science. New York: Springer; 2014.
26. Hevner AR, March ST, Park J, Ram S. Design science in information systems research. MIS Q 2004 Mar;28(1):75-105. [doi: 10.2307/25148625]
27. Gregor S, Hevner AR. Positioning and presenting design science research for maximum impact. MIS Q 2013 Feb 2;37(2):337-355. [doi: 10.25300/MISQ/2013/37.2.01]
28. Vaishnavi V, Kuechler W. Design Science Research Methods and Patterns: Innovating Information and Communication Technology. Boca Raton, Florida: CRC Press; 2015.
29. Venable J, Pries-Heje J, Baskerville R. FEDS: a framework for evaluation in design science research. Eur J Inf Syst 2017 Dec 19;25(1):77-89. [doi: 10.1057/ejis.2014.36]
30. Jepsen DM, Rodwell JJ. Convergent interviewing: a qualitative diagnostic technique for researchers. Manag Res News 2008 Jul 18;31(9):650-658. [doi: 10.1108/01409170810898545]
31. Miah SJ, Hasan N, Hasan R, Gammack J. Healthcare support for underserved communities using a mobile social media platform. Inform Syst 2017 Jun;66:1-12. [doi: 10.1016/j.is.2017.01.001]
32. Stewart DW, Shamdasani PN, Rook D. Analyzing focus group data. In: Focus Groups: Theory and Practice. Volume 20. Thousand Oaks, California: Sage Publications; 2006.
33. EDS Wiki. 2018. EDS API Documentation URL: http://edswiki.ebscohost.com/EDS_API_Documentation [accessed 2018-10-28] [WebCite Cache ID 73ULh91JI]
34. Firebase - Google. 2018. URL: http://firebase.google.com [accessed 2017-07-12] [WebCite Cache ID 6rsNaVwj9]
35. Miah SJ, Hasan J, Gammack JG. On-cloud healthcare clinic: an e-health consultancy approach for remote communities in a developing country. Telemat Inform 2017 Feb;34(1):311-322. [doi: 10.1016/j.tele.2016.05.008]
36. Miah SJ, Gammack J, Hasan N. Extending the framework for mobile health information systems Research: a content analysis. Inform Syst 2017 Sep;69:1-24. [doi: 10.1016/j.is.2017.04.001]

Abbreviations

API: application programming interface
DSR: design science research
DSRM: design science research methodology
EBP: evidence-based practice
EDS: EBSCO Discovery Service
e-resources: electronic resources
OCLC: Online Computer Library Center
WSD: Web-scale discovery

XSL•FO
RenderX

Edited by CL Parra-Calderón; submitted 27.10.18; peer-reviewed by A Zam, D Tao, M Alshehri; comments to author 12.02.19; revised version received 05.03.19; accepted 09.07.19; published 21.08.19.

Please cite as:
Miranda A, Miah SJ
Design and Evaluation of a Contextual Model for Information Retrieval From Web-Scale Discovery Services to Improve Evidence-Based Practice by Health Care Practitioners: Mixed Methods Study
J Med Internet Res 2019;21(8):e12621
URL: http://www.jmir.org/2019/8/e12621/
doi:10.2196/12621
PMID:31436167

©Alvet Miranda, Shah Jahan Miah. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 21.08.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Information Literacy in Food and Activity Tracking Among Parkrunners, People With Type 2 Diabetes, and People With Irritable Bowel Syndrome: Exploratory Study

Pamela McKinney[1*], BA, MSc, PhD; Andrew Martin Cox[1*], PhD; Laura Sbaffi[1*], PhD

Information School, University of Sheffield, Sheffield, United Kingdom

[*]all authors contributed equally

Corresponding Author:
Pamela McKinney, BA, MSc, PhD
Information School
University of Sheffield
Regent Court
211 Portobello Street
Sheffield, S1 4DP
United Kingdom
Phone: 44 0114 2222650
Email: p.mckinney@sheffield.ac.uk

Abstract

Background: The tracking, or logging, of food intake and physical activity is increasing among people, and as a result there is increasing evidence of a link to improvement in health and well-being. Crucial to the effective and safe use of logging is a user's information literacy.

Objective: The aim of this study was to analyze food and activity tracking from an information literacy perspective.

Methods: An online survey was distributed to three communities via parkrun, diabetes.co.uk and the Irritable Bowel Syndrome Network.

Results: The data showed that there were clear differences in the logging practices of the members of the three different communities, as well as differences in motivations for tracking and the extent of sharing of said tracked data. Respondents showed a good understanding of the importance of information accuracy and were confident in their ability to understand tracked data, however, there were differences in the extent to which food and activity data were shared and also a lack of understanding of the potential reuse and sharing of data by third parties.

Conclusions: Information literacy in this context involves developing awareness of the issues of accurate information recording, and how tracked information can be applied to support specific health goals. Developing awareness of how and when to share data, as well as of data ownership and privacy, are also important aspects of information literacy.

(J Med Internet Res 2019;21(8):e13652) doi:10.2196/13652

KEYWORDS

activity logging; food logging; information literacy; irritable bowel syndrome; personal informatics; quantified self; running; self-tracking; type 2 diabetes

Introduction

Background

Self-tracking has been defined as *practices in which people knowingly and purposively collect information about themselves, which they then review and consider applying to the conduct of their lives* [1].

While manual recording of personal data has been advocated for over many years, the potential of digital devices and apps to monitor and measure your own data and then share that information with others has a huge potential for improving personal health [2]. Mobile phones are ubiquitous, powerful, and connected devices that are highly valued by users, and they have the potential to support healthy behaviors through their built-in sensors and downloadable apps [3,4]. Smartphone penetration is high, with 83% of people in the United Kingdom

XSL•FO
RenderX

owning one [5] and 25% of smartphone app users regularly using a health or fitness app [6]. The act of tracking has been shown to be beneficial in terms of increasing desired behaviors in the health arena [7,8], and research has focused on the value of apps and technologies to support health goals in a variety of contexts (eg, menstrual tracking [9], management of migraines [10], diet and exercise [11,12] and chronic disease [13,14]).

Lifestyle changes supported by self-management for people with noncommunicable diseases are a key factor in their prevention and treatment [15]. It is accepted that to achieve health goals involving weight loss, people must address both diet (in terms of reducing calorific intake) and also increase physical exercise [16]. People who track both diet and physical activity are more likely to lose weight [11,17]. Wearable devices that automatically track physical activity, such as FitBit, are increasingly popular, with the market research organization Mintel estimating that 38% of UK consumers have an interest in wearable technology to monitor health and fitness [18].

However, there are a number of barriers to the effective and safe use of tracking, such as whether tracked data is accurate enough to be used by health professionals, ease of use of apps, and the information and digital literacies required to use them effectively. In addition, there is the threat to personal privacy from reuse of tracked data shared to third parties, and the problem that consumers still need to develop an understanding of the social norms of tracking data and of sharing said data [3,13,19]. Of particular interest to this paper is the way that levels of information literacy might be one important determinant of effective and safe use of tracking. Information literacy can be defined as *the ability to think critically and make balanced judgements about any information we find and use. It empowers us as citizens to develop informed views and to engage fully with society* [20].

Initially, research in information literacy focused on the educational context, but it has increasingly broadened towards developing an understanding of information literacy across a range of contexts, such as everyday life [21], the workplace [22] and health [23]. Information literacy is highly contextual, with the necessary set of skills, abilities and practices varying enormously depending on the setting [24]. Previous research into the information literacy aspects of diet and fitness tracking has shown a need for skills in a number of inter-related areas [25], such as: (1) understanding the importance of quality in data inputs; (2) interpreting tracking information outputs in the context of the limitations of the technology; (3) being aware of data privacy and ownership; and also (4) managing appropriate information sharing.

In order to investigate the role of information literacy in the safe and effective use of tracking in a range of contexts, we selected three contrasting populations to study: participants in parkrun free running events, people with type 2 diabetes, and people with irritable bowel syndrome (IBS). The populations were identified during a previous study as being inclined to want to support their health through tracking [25], and were selected to capture variations of underlying motivation, the need for tracking, and in tracking behaviors. An investigation across the three groups offers insights into the diversity of tracking practices.

Information Literacy and Health

There is increasing interest in the contextual nature of information literacy, which, in the health field, is often referred to as health information literacy [26]. The interest in the relationship between information and health is driven by the increasing demand for such information among the population, as well as the changing nature of the relationship between people and healthcare providers [27]. Self-tracking could also be understood as a response to a growing perception of individual responsibility for health [28].

Phenomenographic studies have demonstrated substantial variation in conceptions of health information literacy, which can mean: (1) striving for or reaffirming wellness; (2) knowing or protecting oneself; (3) screening, storing, or creating knowledge; (4) using information to choose a treatment path; (5) paying attention to the body; or (6) participating in learning communities [26,27]. This variation underlines the complexity of both the concept of health information literacy, as well as the multiple and distinctive ways in which people engage with and use health information in their lives. Health literacy has been defined specifically within an electronic context as electronic health (eHealth) literacy and is understood as a transactional model which focuses on people's ability to interact with technology, other users, and to apply information for improved health [29]. Understanding how people engage with and use health information, as well as develop their information literacy, is of interest to public health bodies as they attempt to design health messages that will have an impact on people's behaviors [30]. The view that health information literacy is an example of a contextual application of information literacy is adopted in this paper [24].

Food and Activity Tracking

Research has shown that use of apps can motivate people to adopt healthy behaviors, including a healthy diet, increased physical activity and weight loss [11,31]. Self-management of diet is seen to be a critical issue in some chronic disease management [32], and it has been found that mobile apps for dietary assessment are as valid and reliable as more traditional methods of food tracking [33]. The MyFitnessPal app, popular with both health professionals and the general public, has been found to promote positive changes to the lifestyles of people suffering from diabetes [16]. Tracking can give people a sense that they are taking control of aspects of their life, that they are developing enhanced self-knowledge and self-management, and that they have improved understanding of their own bodies [1,34]. Research has revealed different styles of personal tracking: (1) directive or goal driven tracking; (2) documentary tracking to simply record bodily information; (3) diagnostic tracking to link different aspects of behavior; (4) collecting rewards as a way to register achievement; and (5) fetishized tracking characterized by an interest in gadgets and technology [35]. People gain enjoyment from setting and achieving personal goals from tracked data [1], with use of multiple devices common among those who actively engage in tracking behavior [36].

XSL•FO
RenderX

However, there are a number of potential issues associated with tracking practices, including: (1) tracking can radically alter eating practices; (2) can cause a loss of pleasure in food [1]; (3) can remind people of the negative aspects of a chronic disease [13]; (4) users may fetishize data and develop unhealthy obsessions [1,12,13]; (5) apps tend to not to be based on any behavior change theory [37,38]; and (6) people can find the apps very time-consuming to use, leading to a culture of temporary use, which is particularly true if apps do not meet expectations [3,32,39]. In addition to these issues, there are a number revolving specifically around the information literacies required to make effective and safe use of tracking. Accuracy of data input in tracking is important, but people recognize that their own recording practices may not be sufficiently diligent [39]. People who use apps should have concerns around their ability to enter information accurately and avoid issues of self-deception [3], as understanding quality in data input is one key aspect of information literacy in tracking.

In addition, tracking devices are not necessarily scientifically reliable. They remain unregulated, and there has been considerable speculation about their accuracy [40] and the extent to which valuable bodily data cannot be recorded with apps and devices [12]. So, an information literate individual would be aware of these issues and either find ways to take them into account or not use them at all. Yang et al [41] investigated how people themselves attempted to test trackers' accuracy, though approaches to doing this were often flawed. Furthermore, the outputs of apps are not necessarily understandable by those who use them, since tracking demands the ability to interpret information outputs [19].

The extent to which people are aware of issues to do with the privacy of their personal data held in mobile apps or shared online is also an aspect of information literacy. Although market research in the United Kingdom has shown a majority of app users express concern about privacy and the extent to which apps share information about them, they are not always wary of using a social media account to access app functions, indicating a lack of awareness about potential reuse of data [6]. This parallels what has been dubbed the privacy paradox in social media, in that people are concerned about privacy but do risky things anyway [42]. This could be because they are not sure how to protect themselves, because they are not fully aware of the risk, or because of cynicism about having any privacy in a connected world. Further, research has found that many apps lack a data privacy statement, and often share data with third party organizations [43]. There have been several high-profile data breaches of consumer data, including in 2018, when 150 million users of the popular MyFitnessPal app were hacked [44]. A US study found that users were confident that apps kept their personal data secure [38], but other studies have found that users do have concerns about the privacy of their health data, particularly if data was sold to third parties [3,39]. A further area of concern is long-term access to data, whether because of the disappearance of platforms or the difficulty of exporting data when moving between devices. Thus, issues around data privacy and ownership constitute another area of information literacy relevant to tracking.

There are also aspects of information literacy bound up with appropriate data sharing. Research has shown that people are much more comfortable sharing activity and exercise data than they are sharing food and diet data [3,11,25]. Some studies have found positive perceptions of sharing exercise data, such as people enjoying a competitive relationship with friends and family relating to physical activity [11]. Digital health communities, where people share tracked data, have been identified as a motivating factor in increasing exercise [45]. Through these communities, it is possible for people to gain intimacy and social support through sharing, to benefit from crowd-sourced expertise, and to learn from others who have the same chronic condition [1,34]. However, other studies have shown that there are sensitivities with sharing tracked health data online, with some people considering sharing some health data (including diet information) to be completely unacceptable [3,25]. Unwanted automatic sharing of data with friends is a reason why people discontinue app use [38].

There seems to be a problematic relationship between people, their tracked health data and health professionals, as relatively few people report sharing data with a healthcare provider [36]. However, mobile apps that record diet have been identified as potentially useful, particularly for dietetic professionals [16], and people have seen value in being able to provide accurate data to healthcare professionals [10,38]. In particular, people with IBS are often advised to keep a diary of their symptoms and diet in order to share them with a doctor later [46,47]. However, healthcare providers often regard self-tracked data as unreliable, partly due to lack of diligence on the part of the patient and also due to their supposed unwillingness to admit to negative data [13]. There is also a perception among healthcare professionals that using apps in the context of managing a specific health problem could cause people to either undertake inappropriate or dangerous behaviors [19], or to promote obsessive or compulsive behaviors [13]. Patients feel that healthcare professionals are dismissive of their ability to collect accurate data or to know their own bodies [34].

In summary, the literature review identified that the adoption and use of tracking behaviors and technologies requires people to develop information literacy, both to understand the collection and interpretation of their own data, but also the social constraints around the sharing of that data. Understanding potential issues around privacy and security of data is also an aspect of information literacy in this context.

Thus, the research questions for this study were:

1. What do people in the three communities track and why?
2. What barriers to effective and safe use do they encounter, particularly in relation to information literacy?

This study is the first to investigate self-tracking for health and wellbeing in these specific communities and offers a novel comparative perspective on the attitudes and behaviors of people in regard to supporting their health. Framing tracking behaviors within an information literacy perspective focuses on users' levels of competence in using information to meet their goals, which contributes to an increased understanding of the way people engage with information in the health arena.

XSL•FO
RenderX

Methods

Research Design

A questionnaire-based survey was used to gather insights about food and activity logging habits of three different populations of potential app users. Survey-based research designs have been previously used with success in other studies on food and activity logging [48-50]. The survey was composed of three main sections in which were eleven questions. Ten of these questions were closed-ended (three included a text box for additional comments) and one of them was a fully open-ended question to allow respondents to elaborate more on their experience as food and activity loggers. Closed-ended questions included demographic questions (Section A) such as age, gender, education level and an indication of the onset of their medical condition or experience as parkrunners, and questions related to the respondents' views on logging (Section B) and information use (Section C) were included as both 5-point Likert scale statements and multiple-choice items. Prior to use in this study, the questionnaire had been pilot tested on a small sample of people representative of the three target populations, in order to guarantee consistency and improve readability. The study received ethical approval from the University of Sheffield Information School.

Participants

The survey was distributed online via the websites parkrun, diabetes.co.uk and IBS Network in early 2018, and produced 143 valid responses from parkrunners, 140 valid responses from diabetes.co.uk, and 45 valid responses from the IBS Network. Each community received a tailored version of the survey, with questions such as the length of time respondents had been engaged in running, how long they had suffered from IBS, or when they had been diagnosed with type 2 diabetes tailored for each community. Response rates are not available as the survey was distributed by moderators of the communities in lieu of the authors. No incentives were offered to participants for completing the survey.

Study Populations

The selection of these three specific communities is based on findings from a previous qualitative study [25], which highlighted how users with IBS and type 2 diabetes could benefit from food and activity tracking. In addition, that study identified a difference in tracking behaviors between diet and fitness tracking. Therefore, the present study aims to explore, in more detail, how very diverse groups of users make use of food and activity tracking functions.

Founded in the United Kingdom in 2004, parkrun is a not-for-profit organization that organizes weekly, timed, five kilometer runs in public spaces [51,52]. Events are free to enter and organized by volunteers, and parkrun's ethos emphasizes inclusivity, as shown by how most participants were not regular runners before registering for parkrun. Evidence suggests that running has positive impacts on physical health and well-being, so mobilizing an inclusive community around running has significant potential public health benefits [53].

Type 2 diabetes is a lifelong condition which occurs when the human body cannot use insulin effectively and blood glucose (sugar) levels rise to higher than normal values [54]. Even though type 2 diabetes is mostly diagnosed in adults, it can also develop from a young age. Type 2 diabetes can be controlled if treated properly in its early stages by adopting a healthy lifestyle and healthier habits, such as exercising regularly, maintaining a normal weight and following a low carbohydrate diet [55]. If not managed correctly, though, it can lead to health complications such as heart disease, stroke, blindness, kidney failure and foot or leg amputations [56].

IBS has been defined as "a functional bowel disorder characterized by symptoms of abdominal pain or discomfort and associated with disturbed defecation." [46] It is not understood as a single disease but instead as a range of physiological factors that contribute to commonly experienced symptoms [46]. While the causes are unknown, it is strongly linked to diet and stress, and diet changes are recommended as a way to control the symptoms that can vary enormously from person to person [47,57]. One commentator has estimated that around 11% of the global population has IBS [58]. Those self-identifying as suffering with IBS are predominately women [59].

Data Analysis

All numerical data were entered into IBM SPSS Statistics version 24 and analyzed using descriptive statistics. The results of the 5-point Likert scale statements were aggregated to produce overall figures for agreement and disagreement responses. In addition, two-tailed independent samples *t* tests were performed to identify potential differences in attitudes between men and women, and also to differentiate between the levels of education of the respondents. The qualitative responses were manually coded independently by two members of the research team using thematic analysis [60], and the central themes surfaced for discussion alongside the quantitative data.

Results

The demographic data obtained from the questionnaire (Section A) is reported in Table 1. A summary of the reported app usage (Section B) is reported in Figure 1. A summary of the responses to the Likert scale questions exploring opinions and behaviors relating to logging (Section C) are presented in Tables 2-4 and Figure 2. Tables 2-4 report on frequency of tracking behaviors in the three respondent groups, and Figure 2 presents opinions and views of respondents on their own tracking behaviors.

Selected quantitative and qualitative data are presented thematically below. Participants across the three communities tracked a variety of personal data related to exercise, food, and the body, and a variety of apps, automated devices and manual tracking procedures were used. To reflect this, the first three sections summarize the distinctive nature of tracking in the three groups, and this is followed by a thematic analysis of aspects of respondents' information literacy that synthesizes data from across the groups.

XSL•FO
RenderX

Table 1. Demographic data from the respondents of the questionnaire.

Demographic characteristics	Parkrun, n (%)	Diabetes, n (%)	IBS[a], n (%)
How long have you been running for/have had type 2 diabetes/IBS?			
Less than 2 years	54 (37.8)	41 (29.3)	3 (6.7)
2-5 years	51 (35.7)	48 (34.3)	12 (26.7)
6-10 years	20 (14.0)	19 (13.6)	6 (13.3)
More than 10 years	18 (12.6)	32 (22.9)	24 (53.3)
Gender			
Male	45 (31.5)	57 (40.7)	4 (8.9)
Female	97 (67.8)	83 (59.3)	41 (91.1)
Other	1 (0.7)	0 (0.0)	0 (0.0)
Age			
18-24 years	13 (9.1)	1 (0.7)	2 (4.4)
25-34 years	19 (13.3)	2 (1.4)	14 (31.1)
35-44 years	49 (34.3)	12 (8.6)	15 (33.3)
45-54 years	44 (30.8)	43 (30.7)	7 (15.6)
55-64 years	13 (9.1)	43 (30.7)	2 (4.4)
65+ years	5 (3.5)	38 (27.1)	4 (8.9)
Prefer not to say	0 (0.0)	0 (0.0)	1 (2.2)
Highest level of Education			
Below GCSE[b]	0 (0.0)	6 (4.3)	0 (0.0)
GCSE	11 (7.7)	24 (17.1)	4 (8.9)
A level	24 (16.8)	14 (10.0)	11 (24.4)
Undergraduate	62 (43.4)	44 (31.4)	15 (33.3)
Postgraduate	43 (30.1)	43 (30.7)	13 (28.9)
Prefer not to say	3 (2.1)	9 (6.4)	2 (4.4)

[a]IBS: irritable bowel syndrome.

[b]GCSE: General Certificate of Secondary Education.

XSL•FO
RenderX

Figure 1. Apps used, reasons for tracking, and who data is shared with. IBS: irritable bowel syndrome.

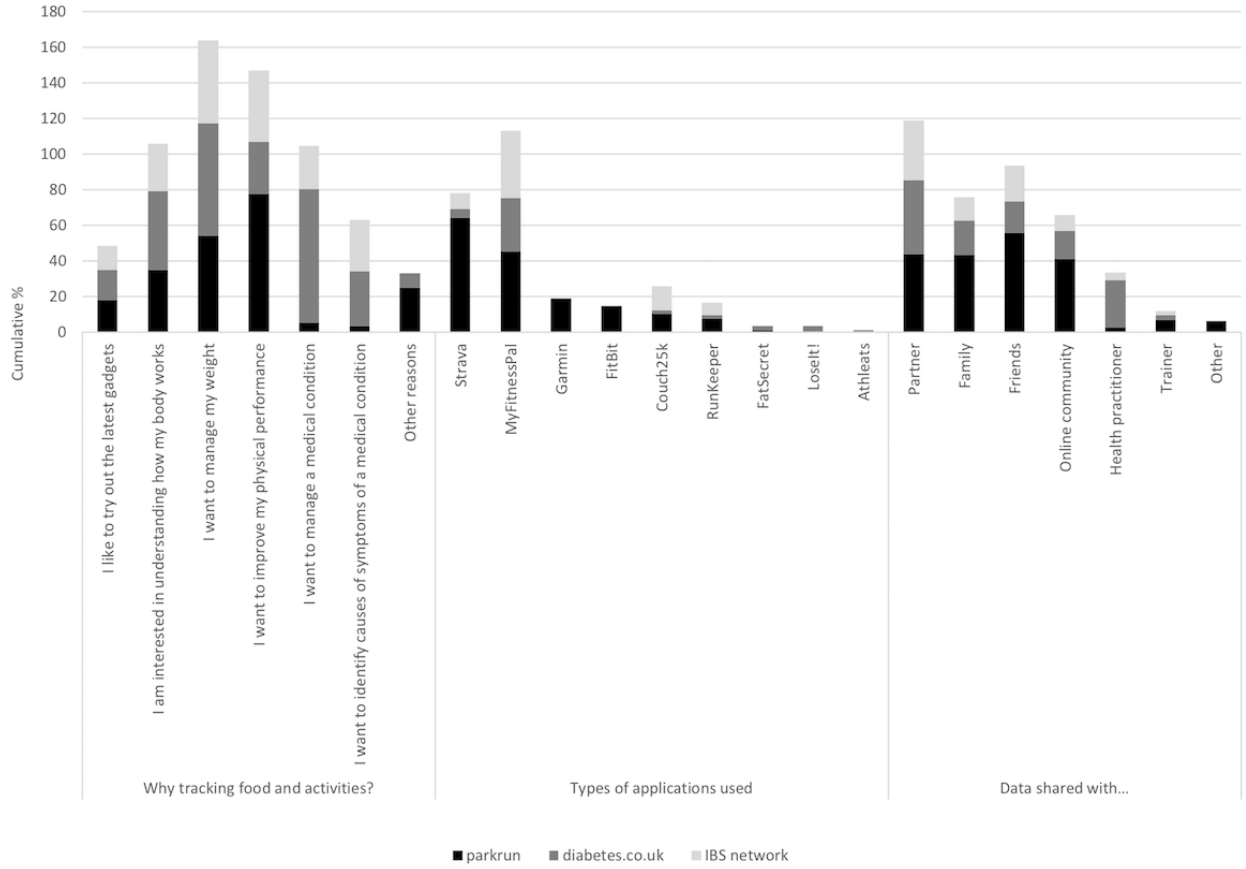


Table 2. Frequency of tracking behaviors in parkrun respondents.

Behavior	Every day, n (%)	2-3 times per week, n (%)	Once a week, n (%)	Less than once a week, n (%)	In the past but not at the moment, n (%)	Never, n (%)
I use a food logging app	45 (31.5)	4(2.8)	1 (0.7)	1 (0.7)	36 (25.2)	56 (39.2)
I use a step counter	84 (58.7)	6 (4.2)	2 (1.4)	4 (2.8)	20 (14.0)	27 (18.9)
I use a device that records running	51 (35.7)	79 (55.2)	6 (4.2)	3 (2.1)	3 (2.1)	1 (0.7)
I track my heart rate and other vital signs	47 (32.9)	27 (18.9)	0 (0.0)	11 (7.7)	15 (10.5)	43 (30.1)
I keep a manual food diary	11 (7.7)	0 (0.0)	1 (0.7)	1 (0.7)	27 (18.9)	103 (72.0)
I track my weight	27 (18.9)	13 (9.1)	44 (30.8)	30 (21.0)	12 (8.4)	17 (11.9)
I track my mood	10 (7.0)	3 (2.1)	1 (0.7)	5 (3.5)	4 (2.8)	120 (83.9)

Table 3. Frequency in tracking behaviours in diabetes.co.uk respondents.

Behavior	Every day, n (%)	2-3 times per week, n (%)	Once a week, n (%)	Less than once a week, n (%)	In the past but not at the moment, n (%)	Never, n (%)
I use a food logging app	40 (28.6)	9 (6.4)	1 (0.7)	0 (0.0)	28 (20.0)	62 (44.3)
I use a step counter	58 (41.4)	6 (4.3)	2 (1.4)	2 (1.4)	23 (16.4)	49 (35.0)
I use a device that records running	22 (15.7)	7 (5.0)	2 (1.4)	1 (0.7)	9 (6.4)	99 (70.7)
I track my heart rate and other vital signs	27 (19.3)	14 (10.0)	7 (5.0)	17 (12.1)	10 (7.1)	65 (46.4)
I keep a manual food diary	24 (17.1)	6 (4.3)	3 (2.1)	3 (2.1)	32 (22.9)	72 (51.4)
I track my weight	40 (28.6)	26 (18.6)	33 (23.6)	20 (14.3)	12 (8.6)	9 (6.4)
I track my mood	19 (13.6)	10 (7.1)	6 (4.3)	6 (4.3)	16 (11.4)	83 (59.3)

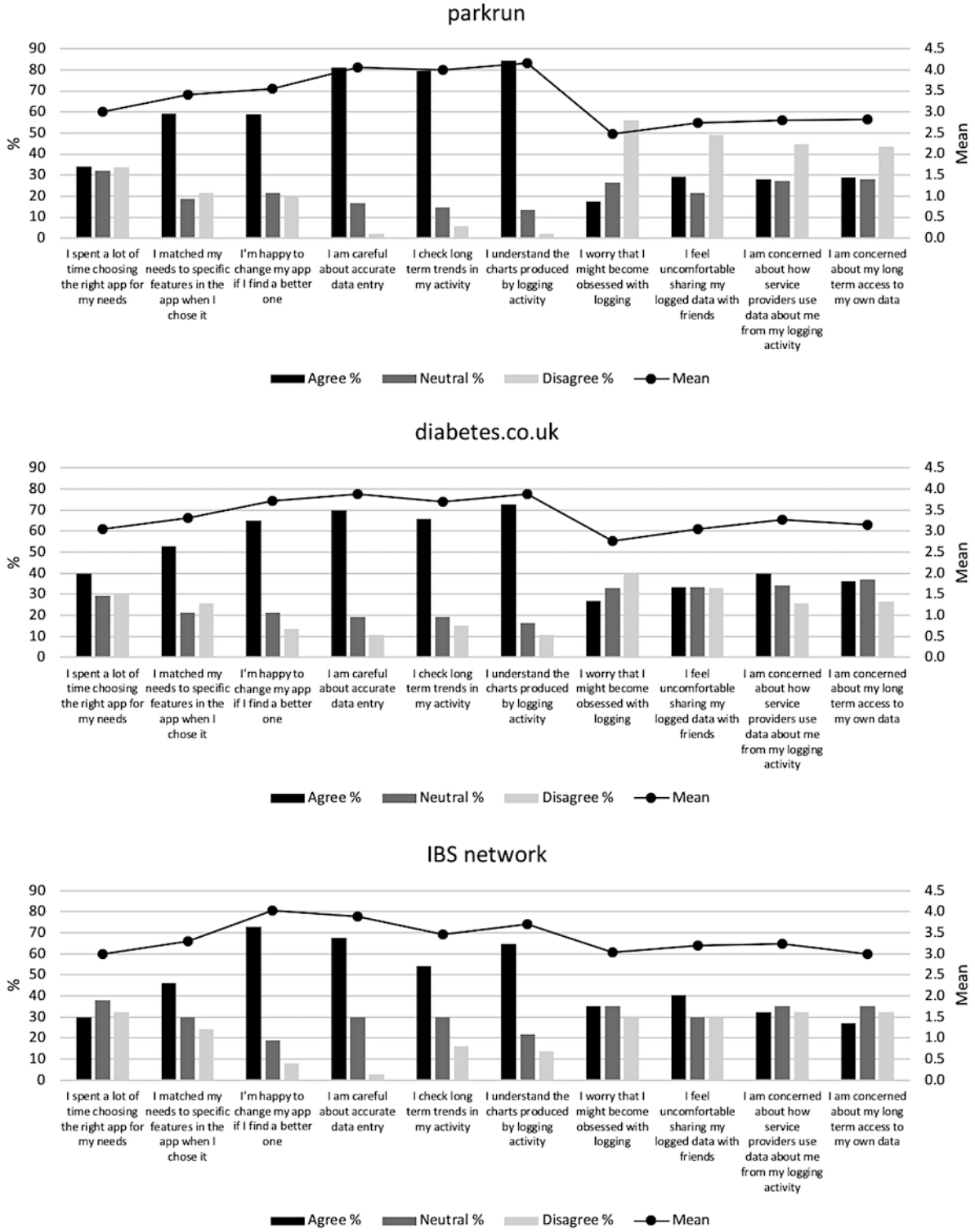
XSL•FO
RenderX

Table 4. Frequency in tracking behaviours in IBS Network respondents.

Behavior	Every day, n (%)	2-3 times per week, n (%)	Once a week, n (%)	Less than once a week, n (%)	In the past but not at the moment, n (%)	Never, n (%)
I use a food logging app	8 (17.8)	1 (2.2)	0 (0.0)	1 (2.2)	16 (35.6)	19 (42.2)
I use a step counter	22 (48.9)	1 (2.2)	1 (2.2)	0 (0.0)	7 (15.6)	14 (31.1)
I use a device that records running	6 (13.3)	2 (4.4)	3 (6.7)	1 (2.2)	8 (17.8)	25 (55.6)
I track my heart rate and other vital signs	5 (11.1)	0 (0.0)	0 (0.0)	3 (6.7)	7 (15.6)	30 (66.7)
I keep a manual food diary	4 (8.9)	0 (0.0)	1 (2.2)	1 (2.2)	17 (37.8)	22 (48.9)
I track my weight	4 (8.9)	4 (8.9)	8 (17.8)	15 (33.3)	5 (11.1)	9 (20.0)
I track my mood	6 (13.3)	1 (2.2)	1 (2.2)	8 (17.8)	4 (8.9)	25 (55.6)

XSL•FO
RenderX

Figure 2. Views on app usage in the three communities. IBS: irritable bowel syndrome.



Parkrunners

Parkrunners used more varied apps and devices to log than the other groups, reporting the use of least two apps on average, with some using as many as five. Unsurprisingly, parkrunners were the biggest users of devices that record running, with 35.7% (n=51) using one every day, and 55.2% (n=79) using one 2-3 times a week. Indeed, using or experimenting with one of these devices seems integral to the practice of running, as only 0.7% (n=1) of parkrunners had never used one. Recording of heart rate or other vital signs was also a popular aspect of tracking for parkrunners, with 32.9% (n=47) using one daily and 18.9% (n=27) using one 2-3 times a week. Parkrunners were primarily motivated by a desire to improve their performance (77.6%; n=111), and tracking was often used by parkrunners to compare their past performance to that of others:

> *I like to be able to track progress and have a goal because I tend to be results orientated.* [Parkrun]

Parkrunners reported tracking a variety of data related to their running practice, but this could be discontinuous and related to personal challenges:

XSL•FO
RenderX

> *I logged and referred to my steps daily as part of two challenges. One to do 10000 steps a day for one week for WI and another was to do 12,000 on average a day for the whole of Lent.* [Parkrun]

In addition, according to the Independent Samples *t* test results, among the parkrunner respondents, those with a higher level of formal education (undergraduate degree or above) reported the highest means of the whole sample in terms of checking long term trends of their activity (mean 4.09; SD 0.79) and understanding the charts produced by the logging activity (mean 4.25; SD 0.72).

Diabetes

Tracking specific aspects of diet (eg, sugar intake) was frequent among diabetes respondents, with over half (55.7%; n=78) engaging in this tracking on a daily basis, and only a quarter (24.3%; n=34) having never tracked a specific aspect of their diet. Some respondents tracked heart rate or other vital signs either daily (19.3%; n=27) or 2-3 times a week (10%; n=14). Manual tracking was a feature of logging among the diabetes respondents (eg, in spreadsheets), with 75% (n=105) indicating that managing their condition was a motivation for their tracking:

> *I use my own log via an excel spreadsheet, that includes blood glucose testing results for each meal, food eaten and exercise on a daily basis. Helps me monitor my condition, track foods and/or exercise that helps or hinders control of my health.* [Diabetes]

> *My logging has been in physical journals and in computer documents. I use data, graphs etc. of my results when I am participating in a particular experiment concerning diet and activity, and my blood glucose levels, HBA1c, waist height ratio, hips, weight.* [Diabetes]

Diabetes respondents displayed a technical knowledge of their condition and the factors that they could log in order to manage it:

> *It is the main cause [that] my HCA1b is now in the 34 area which is normal non-diabetic level, arb intake around 280 grams a day.* [Diabetes]

Logging provided an element of control over the condition:

> *The process of logging helps me stay focussed.* [Diabetes]

> *Generally, I enjoy logging my daily actives and food intake it gives me a better understanding of how my blood sugar levels are impacted by diet, exercise and medication* [Diabetes]

Irritable Bowel Syndrome

Although IBS is a condition that often involves sensitivity towards certain foods, surprisingly few IBS respondents were current users of food logging apps (17.8%; n=8). However, over a third of respondents (35.6%; n=16) had used one in the past, indicating that logging food could be valuable but possibly only over the short-term to identify triggering foods:

> *Great to start but cumbersome, especially if you have to log each ingredient every time. I tend to get bored and apps stop getting used.* [IBS]

IBS respondents were concerned about accuracy in data entry in common with the other groups, and the qualitative comments revealed a particular focus on perceived inaccuracies in food logging apps that could make the practice pointless:

> *I often feel that apps are lacking when it comes to logging food when you eat out or have a takeaway. I often find that many apps seem to suggest American based options so it can be difficult to find the right food. Sometimes it feels more like guess work than accurate tracking and logging.* [IBS]

Despite the interest in specialist diets (eg, FODMAP: fermentable oligo-, di-, monosaccharides and polyols) that have been shown to be effective in managing IBS symptoms [61], only 28.6% (n=40) of respondents (Table 4) used food logging for this purpose.

Although mood tracking was generally not a common aspect of tracking (see Tables 2 and 4), IBS respondents had the highest reported (35.5%; n=16) incidence of mood tracking from across the positive responses.

> *I tend to log my running activity so I can keep track of where I am with my progress. I also note in the tracking of how I felt on the day health / digestion wise so I can see if there is a link to anything in particular. I have had a good experience with tracking and will continue to do so in the future.* [IBS]

IBS respondents were motivated in their logging by a desire to manage weight (46.7%; n=21) and performance (40%; n=18). Surprisingly, managing a medical condition (24.2%; n=11) or identifying the cause of a symptom (28.9%; n=13) were not usually acknowledged as motives. The qualitative responses also underline the importance of weight management to logging practice for this group:

> *I started logging on and off in 2015. Logging my food intake has helped me to lose about 7kg and keep it off, taking me from borderline overweight to the middle of the healthy BMI range.* [IBS]

In contrast to the diabetes respondents, the logging practice for this group was not integral to their condition but more about maintaining general health through exercise and weight management.

Information Literacy: Data Entry Quality

Overall, many participants demonstrated a strong awareness of issues around data quality, with 81.1% (n=116) of parkrunners, 70% (n=100) of diabetes and 67.5% (n=30) of IBS respondents agreeing with the statement “I am careful about accurate data entry”, thus recognizing the critical nature of data quality in their own inputs. The nature of food logging in particular requires people to be precise, including recording everything and weighing and measuring a complex range of ingredients in recipes. Interestingly, parkrunners (65.1%; n=93) and diabetes (52.9%; n=74) respondents were more likely than IBS

XSL•FO
RenderX

respondents (68.2%; n=30) to be careful to log absolutely everything they ate if they used a food logging app.

The qualitative data revealed that people were well aware of the issues around accuracy of their own data input in the food logging context:

> *Difficult when local products are not in database and when item is scanned nothing is heard back. Recipes are tricky to enter.* [Diabetes]

> *Many apps are US based which means it's sometimes hard to find UK foods, but most of the time the barcode scanning works. Where it's less accurate is things like cherry tomatoes. I don't weigh them every time but I know an average weight that I use so I can go by quantity.* [Parkrun]

These complexities may explain why nearly all participants monitored their weight, yet the rate of food tracking was low. Only 17.8% (8) of IBS sufferers, 31.5% (45) of parkrunners and 28.6% (40) of diabetes respondents used a food logging app every day. Qualitative comments suggested why this was, as the practices of food logging and activity tracking had a very different feel. Food tracking was perceived to be worthy but time consuming, fiddly and potentially obsessive. Activity or running tracking was more automatic and seemed to be more inherently enjoyable, and often part of the enjoyment was data sharing. The nature of food logging meant it needed to be done multiple times in a day and would be checked frequently. Thus, food logging was more demanding, and as a consequence there were more complaints about the effort required:

> *Tedious but worthwhile. Any methods to make inputting information easier would be welcome.* [Parkrun]

A discourse of addiction or obsession was often used in relation to tracking, but it happened more often with food logging. It was perceived as more dangerous with food than activity, and thus, one person commented satirically on their obsession with recording running data:

> *I can be a bit obsessed with the "data" so much so that nothing happens until I have uploaded the info!!* [Parkrun]

The tone of the comment is lighthearted, but Independent Samples *t* tests conducted on gender among the parkrunner respondents show that females (mean 2.59; SD 1.06) are statistically significantly more worried about becoming obsessed with data logging than males (mean 2.22; SD 0.80). In addition, becoming obsessive about food emerged as a significant barrier to sustained use:

> *It's okay short term- long term tends to get obsessive and can, in my experience lead to disordered eating.* [Parkrun]

> *I try to balance keeping track of my numbers with not becoming obsessed by them.* [Diabetes]

The demanding requirement to gather accurate data throughout the day could be seen as creating this obsessive element. Thus, part of the information literacy of food tracking could be the management of risk around becoming obsessed with collecting data in a counterproductive way. Food tracking seemed often to be undertaken for short periods, probably for this reason. In contrast, comments on activity or running tracking often emphasized long term practice, enjoyment, how motivating it was, the online community element of it and how it was easy to do:

> *Run logging is fun and easy.* [Parkrun]

Where they did persevere with food logging, a number of solutions to data quality issues had been developed by participants, such as: (1) becoming a data creator (e.g. entering information from recipes into the app); (2) being particular about weighing food; and (3) modifying interpretations of the results to take into account any perceived inaccuracies.

> *I have created my own food entries in MyFitness Pal to be sure that my data is correct.* [Diabetes]

Information Literacy: Interpretation of Tracking Information Outputs

The questionnaire results showed that around half (51.8%; n=72) of diabetes respondents had concerns about the extent to which apps took account of their personal metabolism, and a similar number of diabetes respondents (53.1%; n=74) and parkrunners (52.2%; n=75) had concerns about the quality of data entered by other users in the app. This reveals a critical awareness of the reliability of tracking apps in terms of information literacy. Concerns about data entry and the assumptions built into the app were a barrier to this form of tracking:

> *Haven't started using a food logging app as I find it mind boggling and difficult to use when it comes to home made food, plus their general approach to diet seems to fall onto the calorie deficit thinking whereas I view it more as quality of food i.e. not all calories are equal.* [Parkrun]

Regarding the interpretation of the information output of tracking, 121 (84.6%) parkrunners, 102 (72.8%) diabetes respondents and 29 (64.8%) IBS respondents said that they understood the charts produced by their apps. They also engaged closely with the data: 114 (79.7%) parkrunners said that they checked their long-term trends in activity, and so did 91 (65%) diabetes respondents and 24 (54%) IBS respondents. Again, qualitative responses suggested quite sophisticated use of apps, such as combining multiple devices or tracking different data in parallel:

> *I initially used My fitness pal to see how many calories were in specific foods and also to see how the calories balanced against manually inputted exercise. Then I got a Fitbit and linked the two. I am type 1 diabetic and am interested in keeping my weight at a healthy BMI. I also use Endomondo for logging runs and the training plan in it for my first half marathon in September.* [Parkrun]

Indeed, at least some participants had a critical sense of the limits of current designs of the tracking devices themselves:

> *Logging can be negative if a device wants you to move and you cannot, due to medical or personal reasons.*

XSL•FO
RenderX

> *Interfaces need to evolve and become more personal, flexible and compassionate.* [Parkrun]
>
> *You need to decide exactly what you want out of the process and not let an app designer dictate to you. Also don't get fixated on completeness and logging history. Keep asking the question: why is this useful?* [Parkrun]

Information Literacy: Data Privacy and Ownership

Participants were asked about the extent to which they were concerned about how service providers used their logged data, with parkrunners most likely to be unconcerned (44.8%; n=64). Diabetes respondents were more worried about reuse of their data, with 24.3% (n=34) agreeing strongly and 40% (n=56) agreeing overall. The most common responses from IBS respondents were evenly distributed between agreeing, agreeing strongly or neutral, with 37.8% (n=17) choosing these options. Therefore, for each group, less than half of the respondents had concerns about potential reuse of their data.

Respondents were also relatively unconcerned about threats to long-term access of their data. Of the parkrunners, 43.4% (n=62) disagreed with the statement, "I am concerned about the long-term access to my data", with 28% (n=40) answering neutral. Diabetes respondents were slightly more concerned, with around a third (36.4%; n=51) agreeing overall with the statement, but the most popular answer for this group was neutral (37.1%; n=52). IBS respondents were evenly split across agreeing or agreeing strongly (32.4%; n=15), neutral (35.1%; n=16) and disagreeing or disagreeing strongly (32.4%; n=14).

Information Literacy: Information Sharing and Privacy

Different types of information seemed to be shared quite differently. Activity data was fairly freely shared, with the parkrunners as the greatest sharers of tracked data with friends and family, and by far the biggest sharers with online communities (41.3%; n=59). In the qualitative comments, data sharing was more commonly mentioned in relation to running activity, and seen as part of the enjoyment:

> *Seeing what my friends are doing (and knowing that they see what I do) is a major motivator for me in exercise and encourages me to get out and do things when I don't necessary feel like it. I also like statistics and tracking my performance.* [Parkrun]

Just a few qualitative comments revealed privacy concerns about running data:

> *I stopped using Strava because you could not hide runs from the public, which is a privacy concern as they could see or workout where I live and where I run on a regular basis.* [Parkrun]

IBS respondents shared the least data overall and were most likely to agree that data sharing made them feel uncomfortable, with 40.5% (n=18) agreeing or agreeing strongly and only 29.7% (n=13) disagreeing or disagreeing strongly. Specific to type 2 diabetes, women feel significantly more uncomfortable sharing data (mean 3.24; SD 1.27) than men (mean 2.75; SD 1.20). This probably reflects that, rather than activity data, they were collecting data about a medical condition or weight and diet, which was seen as more private. Sharing different types of data reflects an awareness of social norms surrounding tracked data.

A few strategies were mentioned as part of maintaining privacy, such as manual data tracking:

> *I strongly disagree with "cloud" based apps where I can't restrict data sharing. That's both for privacy, and also risk of losing access.* [Diabetes]

Similar sorts of sensitivities were reflected in who data was shared with. Partners were the most popular people to share data with across diabetes (41.4%; n=58) and IBS (33.3%; n=14) respondents, but friends were the most popular for parkrunners (55.9%; n=80). Diabetes respondents were the most likely to share data with a health practitioner, but the numbers were still quite low (26.4%; n=37). Less than 10% of the other two groups shared their data with an expert such as a trainer or doctor.

In summary, the results present the different varieties of tracking practice among the three communities studied and demonstrate that there are significant differences in motivations for tracking and uses for the data gathered. People who engage in self-tracking show evidence of information literacy through: ensuring data quality, understanding the information produced by tracking technologies and how this relates to their particular situation or medical condition, developing awareness of when and how to share their data, and developing an understanding of who has access to their data and the potential for sharing and reuse without their explicit consent.

Discussion

Overview

Tracking is used in different ways by different groups, but in all contexts, it is an information intense activity based on gathering, interpreting and managing data mediated by various devices and apps. The question of how information literate trackers are (how good their critical understanding of the information they are using is) thus becomes central to effective and safe tracking. This is one of the first papers to bring this perspective on tracking explicitly to the fore and complements research that has examined self-tracking from a Human-computer interaction perspective [12,62,63], a health behavior change perspective, [3] and a sociological perspective [13].

Respondents showed an understanding of the importance of their own accurate data entry, but also a skeptical awareness of its limits, especially in the context of food logging. In some cases, it was this critical understanding that led to nonuse, and in others, people found approaches to ensuring data quality or only used it intermittently. This is consistent with previous studies that show that simplifying diet and nutrition apps to make data entry less time consuming and more automatic was a key improvement desired by users [38]. It would make food logging much easier and also remove one aspect that created a fear of obsession, which is a common issue identified in self-tracking research [1,13].

XSL•FO
RenderX

While data accuracy is an important aspect of successful tracking, previous research into self-tracking has highlighted a tension between trusting data or trusting bodily sensations [12], with speculation regarding the relationship and potential value of each. In information literacy research, the role of corporeal information as a valuable source of information alongside social, epistemic, or formal sources is widely understood [22,25,64]. Conceptions of health information literacy indicate that assessing and evaluating information is a key activity, and that paying attention to the body and developing self-awareness supports the interpretation of other health information [26]. Diabetes respondents actively used information, often manually recorded, to manage their condition, which could be seen as an example of diagnostic tracking [35]. This extends conceptions of self-tracking beyond simply understanding a person's relationship with technology to a broader understanding of their relationship with information [63]. Previous research into the information behavior of people with type 2 diabetes found that connecting information gathered from different objective and subjective bodily observations was an aspect of effectively managing the condition [34]. Integrating bodily information with app-related information has also been shown to be an important aspect of elite runners' personal informatics practices [12]. Becoming information literate with regard to self-tracking, therefore, involves developing an understanding of how to integrate app data with corporeal information in order to achieve specific health goals.

Although the app MyFitnessPal was popular with participants in this study, perceived inaccuracies in either the app or one's own measurements were also barriers to food logging. This is exacerbated by the acknowledged US bias of the food and measurements in the app's database [16]. Many people, therefore, log food for only as long as it takes to either learn better food habits or to learn which foods trigger aspects of their condition. Discontinuing use can also occur due to the burdensome nature of tracking [38]. Information literacy therefore revolves around learning at what point the information needs have been met, and when to modify or discard the logging practice.

As regards interpreting information from tracking, respondents were confident in their own information literacy in interpreting the data output by the apps, and often used multiple devices in rather sophisticated tracking practices. They also made some critical comments on the questionable assumptions or expectations built into app design. This is consistent with previous research that has also found that users can be very capable of taking critical stance towards apps [65]. An area of rather more concern, consistent with some previous studies [40], is that many respondents were not worried about the use of their data by the platform or about long-term access to it, particularly in the current climate where app data is often widely shared with third parties without the express consent of the user [1,43]. Fortunately, some studies have found that data reuse is a serious issue for participants [3], and opinions may start shifting because of recent cases in the news.

Data sharing with friends and even in online communities was found to be central to activity tracking for many participants, which is consistent with previous research that has found sharing to be a fun aspect of tracking [12]. However, participants in this study were more reluctant to share data about health conditions, diet and weight, a point also identified by previous authors [3]. Information literacy research has found that sharing information about a chronic disease is a way for people to draw friends and family into their landscape and create a narrative about a disease [24], but this does not seem to be the case for participants in this study. Consistent with previous research [36,38], despite the potential benefits data was not often being shared with a trainer or doctor, which might reflect lack of interest by practitioners rather than trackers' willingness to share data since they were already sharing with others, such as partners.

In summary, from an information literacy perspective, users seemed to be very literate in many aspects of logging practice. The relatively low use of food apps seems to reflect a critical perspective on the effort required to use them, their accuracy and their potentially obsessive effects.

Limitations

This is a small-scale exploratory study, which only begins to identify the information literacy aspects of tracking behaviors for the three participant groups. All three groups of respondents reported a high level of prior education, which may not be representative of the populations as a whole. This may reflect a higher use of logging by higher socioeconomic groups [66], as more educated users seemed to have more confidence in their information literacy capabilities. In several respects we do not know how well the respondents represent the wider target population, partly because we do not have data about the demographics of the wider population and partly because respondents were self-selecting. The skew towards women among parkrunners, for example, may reflect greater willingness to participate in surveys rather than the actual proportions in the population [53]. The response rate from IBS sufferers was lower than for the other groups and so the results for this group should be treated with additional caution.

By definition, respondents to the questionnaire were current or past users of tracking apps. Many had not used or lapsed from use of particular forms of tracking, but the sample did not include those who had never tracked at all. This places a further limitation on the data as a means of understanding barriers to tracking in all contexts, however, studies of nonusers are inherently difficult.

The survey was based on asking participants to self-evaluate some of their information literacy skills (eg, their ability to understand charts produced by apps), so this may differ from actual competence. Overconfidence in information literacy is a known phenomenon [67]; however, levels of information literacy were implicit in many of the qualitative comments, which reflected complex, personalized practices of use.

Conclusions and Implications

An information literacy perspective is of value because tracking is an information intensive activity, involving the user in entering data, in interpreting the information outputs of the device, and then managing access to said data. Effective and safe use of tracking thus depends on information literacy. This study showed that in three very different domains devices were

XSL•FO
RenderX

used quite differently and levels of information literacy were also variable. In terms of understanding data entry quality, interpreting information, and appropriate sharing, respondents seemed to demonstrate good information literacy; however, a greater area of concern is around people's lack of awareness of risks around platform use of data and continuity of access. This implies the need for much better public awareness around data ownership, and simplified privacy statements might assist in this. Organizations such as parkrun, Diabetes UK and the IBS network should consider this issue for their communities when they are providing advice and support about using mobile apps. The European Union's General Data Protection Regulation is a move in a favorable direction in increasing protection of trackers' privacy, but simple tools to extract data and maintain access to personal tracking data in the long-term are also needed. Additionally, there seems to be a gap in the market regarding mobile apps to support both the management of type 2 diabetes and IBS, given the reported manual tracking of one community and the pattern of app use and nonuse of the other.

Acknowledgments

This research was funded by the Information School at the University of Sheffield, and was facilitated by the diabetes.co.uk community, the IBS Network charity and the parkrun organization, who contributed to the design and distribution of the survey.

Conflicts of Interest

None declared.

References

1. Lupton D. The Quantified Self. Cambridge: Polity Press; 2016.
2. Lupton D. Health promotion in the digital era: a critical commentary. Health Promot Int 2015 Mar;30(1):174-183. [doi: 10.1093/heapro/dau091] [Medline: 25320120]
3. 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):1-12 [FREE Full text] [doi: 10.2196/jmir.2583] [Medline: 23598614]
4. Dute D, Bemelmans W, Breda J. Using mobile apps to promote a healthy lifestyle among adolescents and students: A review of the theoretical basis and lessons learned. JMIR mHealth uHealth 2016;4(2):e39. [Medline: 23598614]
5. Mintel. Mintel Reports. 2018. Mobile phones-UK-April 2018 URL: http://reports.mintel.com/display/859075/?__cc=1
6. Mintel. Mintel Academic. 2017. Mobile device apps-UK-November 2017 URL: http://academic.mintel.com/display/793809/
7. Klasnja P, Pratt W. Healthcare in the pocket: mapping the space of mobile-phone health interventions. J Biomed Inform 2012 Feb;45(1):184-198 [FREE Full text] [doi: 10.1016/j.jbi.2011.08.017] [Medline: 21925288]
8. Schoeppe S, Alley S, Van Lippevelde W, Bray NA, Williams SL, Duncan MJ, et al. Efficacy of interventions that use apps to improve diet, physical activity and sedentary behaviour: a systematic review. Int J Behav Nutr Phys Act 2016 Dec 07;13:127 [FREE Full text] [doi: 10.1186/s12966-016-0454-y] [Medline: 27927218]
9. Epstein D, Lee N, Kang J, Agapie E, Schroeder J, Pina L, et al. Examining Menstrual Tracking to Inform the Design of Personal Informatics Tools. In: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. 2017 Presented at: CHI Conference on Human Factors in Computing Systems; May 06 - 11, 2017; Denver, Colorado p. 6876-6888 URL: http://dl.acm.org/citation.cfm?doid=3025453.3025635
10. Schroeder J, Chung C, Epstein D, Karkar R, Parsons A, Murinova N, et al. Examining Self-Tracking by People with Migraine. In: DIS '18 Proceedings of the 2018 Designing Interactive Systems Conference. 2018 Presented at: DIS 18: Designing Interactive Systems Conference; June 09 - 13, 2018; Hong Kong- China p. 135-148 URL: https://dl.acm.org/citation.cfm?id=3196738 [doi: 10.1145/3196709.3196738]
11. Wang Q, Egelandsdal B, Amdam GV, Almli VL, Oostindjer M. Diet and Physical Activity Apps: Perceived Effectiveness by App Users. JMIR Mhealth Uhealth 2016;4(2):e33 [FREE Full text] [doi: 10.2196/mhealth.5114] [Medline: 27056639]
12. Rapp A, Tirabeni L. Personal Informatics for Sport. ACM Trans Comput-Hum Interact 2018 Jun 28;25(3):1-30. [doi: 10.1145/3196829]
13. Ancker JS, Witteman HO, Hafeez B, Provencher T, Van de Graaf M, Wei E. "You Get Reminded You're a Sick Person": Personal Data Tracking and Patients With Multiple Chronic Conditions. J Med Internet Res 2015;17(8):e202 [FREE Full text] [doi: 10.2196/jmir.4209] [Medline: 26290186]
14. Mentis H, Komlodi A, Schrader K, Phipps M, Gruber-Baldini A, Yarbrough K, et al. Crafting a View of Self-Tracking Data in the Clinical Visit. In: CHI '17 Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. 2017 Presented at: 2017 CHI Conference on Human Factors in Computing Systems; May 06 - 11, 2017; Denver, Colorado p. 5800-5812 URL: https://dl.acm.org/citation.cfm?id=3025589 [doi: 10.1145/3025453.3025589]
15. Lunde P, Nilsson B, Bergland A, Kværner KJ, Bye A. The Effectiveness of Smartphone Apps for Lifestyle Improvement in Noncommunicable Diseases: Systematic Review and Meta-Analyses. J Med Internet Res 2018 May 04;20(5):e162-e112 [FREE Full text] [doi: 10.2196/jmir.9751] [Medline: 29728346]

XSL•FO
RenderX

16. Chen J, Lieffers J, Bauman A, Hanning R, Allman-Farinelli M. The use of smartphone health apps and other mobile health (mHealth) technologies in dietetic practice: a three country study. J Hum Nutr Diet 2017 Aug;30(4):439-452. [doi: 10.1111/jhn.12446] [Medline: 28116773]
17. Flores Mateo G, Granado-Font E, Ferré-Grau C, Montaña-Carreras X. Mobile Phone Apps to Promote Weight Loss and Increase Physical Activity: A Systematic Review and Meta-Analysis. J Med Internet Res 2015;17(11):e253 [FREE Full text] [doi: 10.2196/jmir.4836] [Medline: 26554314]
18. Mintel. Mintel Reports. 2017. Wearable technology-UK-December 2017 URL: http://reports.mintel.com/display/794377/
19. Bert F, Giacometti M, Gualano M, Siliquini R. Smartphones and health promotion: A review of the evidence. J Med Syst 2014;38(1):9995. [doi: 10.1007/s10916-013-9995-7]
20. Chartered Institute of Library and Information Professionals. CILIP. 2018. CILIP definition of Information Literacy 2018 URL: https://infolit.org.uk/ILdefinitionCILIP2018.pdf
21. Papen U. Conceptualising information literacy as social practice: a study of pregnant women's information practices. Inf Res 2013;18(2):1 [FREE Full text]
22. Lloyd A. Informing practice: information experiences of ambulance officers in training and on‐road practice. Journal of Documentation 2009 Apr 24;65(3):396-419. [doi: 10.1108/00220410910952401] [Medline: 23490161]
23. Marshall A, Henwood F, Carlin L, Guy E, Sinozic T, Smith H. Information to fight the flab: findings from the Net.Weight Study. JIL 2009 Dec 08;3(2):39-52. [doi: 10.11645/3.2.218]
24. Lloyd A, Bonner A, Dawson-Rose C. The health information practices of people living with chronic health conditions: Implications for health literacy. Journal of Librarianship and Information Science 2013 May 17;46(3):207-216. [doi: 10.1177/0961000613486825]
25. Cox A, McKinney P, Goodale P. Food logging: an information literacy perspective. Aslib Journal of Info Mgmt 2017 Mar 20;69(2):184-200. [doi: 10.1108/AJIM-12-2016-0208]
26. Yates C. Exploring variation in the ways of experiencing health information literacy: A phenomenographic study. Library & Information Science Research 2015 Jul;37(3):220-227. [doi: 10.1016/j.lisr.2015.04.003]
27. Yates C, Partridge H, Bruce C. Learning wellness: how ageing Australians experience health information literacy. The Australian Library Journal 2009 Aug;58(3):269-285. [doi: 10.1080/00049670.2009.10735905]
28. Sharon T. Self-Tracking for Health and the Quantified Self: Re-Articulating Autonomy, Solidarity, and Authenticity in an Age of Personalized Healthcare. Philos. Technol 2016 Apr 18;30(1):93-121. [doi: 10.1007/s13347-016-0215-5]
29. Paige S, Stellefson M, Krieger J, Anderson-Lewis C, Cheong J, Stopka C. Proposing a Transactional Model of eHealth Literacy: Concept Analysis. J Med Internet Res 2018;20(10):e10175. [Medline: 30279155]
30. Yates C, Stoodley I, Partridge H, Bruce C, Cooper H, Day G, et al. Exploring Health Information Use by Older Australians within Everyday Life. Library Trends 2012;60(3):460-478. [doi: 10.1353/lib.2012.0004]
31. Ernsting C, Dombrowski SU, Oedekoven M, Kanzler M, Kuhlmey A, Gellert P. Using Smartphones and Health Apps to Change and Manage Health Behaviors: A Population-Based Survey. J Med Internet Res 2017 Apr 05;19(4):e101 [FREE Full text] [doi: 10.2196/jmir.6838] [Medline: 28381394]
32. Rusin M, Arsand E, Hartvigsen G. Functionalities and input methods for recording food intake: a systematic review. Int J Med Inform 2013 Aug;82(8):653-664. [doi: 10.1016/j.ijmedinf.2013.01.007] [Medline: 23415822]
33. Sharp DB, Allman-Farinelli M. Feasibility and validity of mobile phones to assess dietary intake. Nutrition 2014;30(11-12):1257-1266. [doi: 10.1016/j.nut.2014.02.020] [Medline: 24976425]
34. St. Jean B, Jindal G, Chan K. “You Have to Know Your Body!”: The Role of the Body in Influencing the Information Behaviors of People with Type 2 Diabetes. Library Trends 2018;66(3):289-314. [doi: 10.1353/lib.2018.0004]
35. Rooksby J, Rost M, Morrison A, Chalmers M. Personal tracking as lived informatics. In: CHI '14 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2014 Presented at: SIGCHI Conference on Human Factors in Computing Systems; April 26 - May 01, 2014; Toronto, Canada p. 1163-1172 URL: https://dl.acm.org/citation.cfm?id=2557039 [doi: 10.1145/2556288.2557039]
36. Paré G, Leaver C, Bourget C. Diffusion of the Digital Health Self-Tracking Movement in Canada: Results of a National Survey. J Med Internet Res 2018 May 02;20(5):e177 [FREE Full text] [doi: 10.2196/jmir.9388] [Medline: 29720359]
37. Azar KMJ, Lesser LI, Laing BY, Stephens J, Aurora MS, Burke LE, et al. Mobile applications for weight management: theory-based content analysis. Am J Prev Med 2013 Nov;45(5):583-589. [doi: 10.1016/j.amepre.2013.07.005] [Medline: 24139771]
38. Krebs P, Duncan DT. Health App Use Among US Mobile Phone Owners: A National Survey. JMIR Mhealth Uhealth 2015;3(4):e101 [FREE Full text] [doi: 10.2196/mhealth.4924] [Medline: 26537656]
39. Rapp A, Cena F. Personal informatics for everyday life: How users without prior self-tracking experience engage with personal data. International Journal of Human-Computer Studies 2016 Oct;94:1-17. [doi: 10.1016/j.ijhcs.2016.05.006]
40. Hoy MB. Personal Activity Trackers and the Quantified Self. Med Ref Serv Q 2016;35(1):94-100. [doi: 10.1080/02763869.2016.1117300] [Medline: 26794199]
41. Yang R, Shin E, Newman M, Ackerman M. When fitness trackers don't 'fit': end-user difficulties in the assessment of personal tracking device accuracy. In: Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing. 2015 Presented at: the 2015 ACM International Joint Conference on Pervasive and Ubiquitous

XSL•FO
RenderX

Computing; September 07 - 11, 2015; Osaka, Japan p. 623-634 URL: https://dl.acm.org/citation.cfm?id=2804269 [doi: 10.1145/2750858.2804269]
42. Vitak J, Liao Y, Kumar P, Zimmer M, Kritikos K. Privacy attitudes and data valuation among fitness tracker users. In: Transforming Digital Worlds: 13th International Conference, iConference 2018. Sheffield: Springer International Publishing; 2018 Presented at: Transforming Digital Worlds: 13th International Conference, iConference 2018; March 25-28, 2018; Sheffield p. 229-239. [doi: 10.1007/978-3-319-78105-1]
43. Ackerman L. Mobile health and fitness applications and information privacy: report to California Consumer Protection Foundation. San Diego, California: Privacy Rights Clearinghouse; 2013. URL: https://www.privacyrights.org/mobile-medical-apps-privacy-consumer-report.pdf
44. The Guardian. 2018 Mar 30. Hackers steal data of 150 million MyFitnessPal app users URL: https://www.theguardian.com/technology/2018/mar/30/hackers-steal-data-150m-myfitnesspal-app-users-under-armour
45. Ba S, Wang L. Digital health communities: The effect of their motivation mechanisms. Decision Support Systems 2013 Nov;55(4):941-947. [doi: 10.1016/j.dss.2013.01.003]
46. American Gastroenterological Association. American Gastroenterological Association medical position statement: Irritable Bowel Syndrome. Gastroenterology 2002 Nov;123(6):2105-2107. [doi: 10.1053/gast.2002.37095b]
47. National Health Service. 2017. What is IBS? Irritable Bowel Syndrome (IBS) URL: https://www.nhs.uk/conditions/irritable-bowel-syndrome-ibs/ [accessed 2018-08-02]
48. Burke LE, Wang J, Sevick MA. Self-monitoring in weight loss: a systematic review of the literature. J Am Diet Assoc 2011 Jan;111(1):92-102 [FREE Full text] [doi: 10.1016/j.jada.2010.10.008] [Medline: 21185970]
49. Shim J, Oh K, Kim H. Dietary assessment methods in epidemiologic studies. Epidemiol Health 2014;36:e2014009. [Medline: 5078382]
50. James P, Weissman J, Wolf J, Mumford K, Contant C, Hwang W, et al. Comparing GPS, Log, Survey, and Accelerometry to Measure Physical Activity. Am J Health Behav 2016;40(1):123-131 [FREE Full text] [doi: 10.5993/AJHB.40.1.14] [Medline: 26685821]
51. Hindley D. “More Than Just a Run in the Park”: An Exploration of Parkrun as a Shared Leisure Space. Leisure Sciences 2018 Jan 10;1:1-21. [doi: 10.1080/01490400.2017.1410741]
52. Stevinson C, Wiltshire G, Hickson M. Facilitating participation in health-enhancing physical activity: a qualitative study of parkrun. Int J Behav Med 2015 Apr;22(2):170-177. [doi: 10.1007/s12529-014-9431-5] [Medline: 25096794]
53. Stevinson C, Hickson M. Exploring the public health potential of a mass community participation event. J Public Health (Oxf) 2014 Jun;36(2):268-274. [doi: 10.1093/pubmed/fdt082] [Medline: 23954885]
54. WebMD. 2018. Types of diabetes mellitus URL: https://www.webmd.com/diabetes/guide/types-of-diabetes-mellitus#1 [accessed 2018-10-11]
55. Boden G, Sargrad K, Homko C, Mozzoli M. Short-term effects of low-carbohydrate diet compared with usual diet in obese patients with type 2 diabetes. Ann Intern Med 2005 Mar 15;142(6):I44. [doi: 10.7326/0003-4819-142-6-200503150-00002] [Medline: 15767614]
56. Turner R, Holman R, Frighi V, Manley S, Matthews D, Neil A, et al. Tight blood pressure control risk of macrovascular and microvascular complications in type 2 diabetes: UKPDS 38. Br Med J 1998;317(7160):703. [Medline: 9732337]
57. The IBS Network. 2018. What is IBS? URL: https://www.theibsnetwork.org [accessed 2018-08-02]
58. Canavan C, West J, Card T. The epidemiology of irritable bowel syndrome. Clin Epidemiol 2014;6:71-80 [FREE Full text] [doi: 10.2147/CLEP.S40245] [Medline: 24523597]
59. Drossman Z, Li Z, Andruzzi E, Temple RD, Talley NJ, Thompson WG, et al. U.S. householder survey of functional gastrointestinal disorders. Prevalence, sociodemography, and health impact. Dig Dis Sci 1993 Sep;38(9):1569-1580. [doi: 10.1007/bf01303162] [Medline: 8359066]
60. Braun V, Clarke V. Using thematic analysis in psychology. Qualitative Research in Psychology 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
61. Halmos E, Power V, Shepherd S, Gibson P, Muir J. A diet low in FODMAPs reduces symptoms of irritable bowel syndrome. Gastroenterology 2014 Jan;146(1):67-75.e5. [doi: 10.1053/j.gastro.2013.09.046] [Medline: 24076059]
62. Rapp A, Tirassa M. Know Thyself: A Theory of the Self for Personal Informatics. Human–Computer Interaction 2017 Jan 27;32(5-6):335-380. [doi: 10.1080/07370024.2017.1285704]
63. Ohlin F, Olsson C. Beyond a utility view of personal informatics: A postphenomenological framework. In: Adjunct Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers.: ACM; 2015 Presented at: the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers; September 07 - 11, 2015; Osaka, Japan p. 1087-1092. [doi: 10.1145/2800835.2800965]
64. Lloyd A. Framing information literacy as information practice: site ontology and practice theory. Journal of Documentation 2010 Mar 09;66(2):245-258. [doi: 10.1108/00220411011023643]
65. Jimoh F, Lund EK, Harvey LJ, Frost C, Lay WJ, Roe MA, et al. Comparing Diet and Exercise Monitoring Using Smartphone App and Paper Diary: A Two-Phase Intervention Study. JMIR Mhealth Uhealth 2018 Jan 15;6(1):e17 [FREE Full text] [doi: 10.2196/mhealth.7702] [Medline: 29335239]

XSL•FO
RenderX

66. Régnier F, Chauvel L. Digital Inequalities in the Use of Self-Tracking Diet and Fitness Apps: Interview Study on the Influence of Social, Economic, and Cultural Factors. JMIR Mhealth Uhealth 2018 Apr 20;6(4):e101 [FREE Full text] [doi: 10.2196/mhealth.9189] [Medline: 29678807]
67. Mahmood K. Do People overestimate their information literacy skills? A systematic review of empirical evidence on the Dunning-Kruger effect. Commun Inf Lit 2016;10(2):199-213. [doi: 10.7548/cil.v10i2.385]

Abbreviations

eHealth: electronic health
FODMAP: fermentable oligo-, di-, monosaccharides and polyols
IBS: irritable bowel syndrome

Edited by G Eysenbach; submitted 07.02.19; peer-reviewed by A Rapp, X Fan; comments to author 28.03.19; revised version received 04.06.19; accepted 29.06.19; published 01.08.19.

Please cite as:
McKinney P, Cox AM, Sbaffi L
Information Literacy in Food and Activity Tracking Among Parkrunners, People With Type 2 Diabetes, and People With Irritable Bowel Syndrome: Exploratory Study
J Med Internet Res 2019;21(8):e13652
URL: https://www.jmir.org/2019/8/e13652/
doi:10.2196/13652
PMID:31373277

©Pamela McKinney, Andrew Martin Cox, Laura Sbaffi. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 01.08.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Electronic Health Literacy Among Magnetic Resonance Imaging and Computed Tomography Medical Imaging Outpatients: Cluster Analysis

Lisa Lynne Hyde[1,2,3,4], BPsych (Hons); Allison W Boyes[1,2,3,4], PhD; Lisa J Mackenzie[1,2,3,4], PhD; Lucy Leigh[3], PhD; Christopher Oldmeadow[1,3], PhD; Carlos Riveros[1,3], PhD; Rob Sanson-Fisher[1,2,3,4], PhD

[1]School of Medicine and Public Health, Faculty of Health and Medicine, University of Newcastle, Callaghan, Australia

[2]Priority Research Centre for Health Behaviour, University of Newcastle, Callaghan, Australia

[3]Hunter Medical Research Institute, New Lambton Heights, Australia

[4]Hunter Cancer Research Alliance, Newcastle, Australia

Corresponding Author:
Lisa Lynne Hyde, BPsych (Hons)
School of Medicine and Public Health
Faculty of Health and Medicine
University of Newcastle
University Drive
Callaghan, 2308
Australia
Phone: 61 2 4913 8799
Fax: 61 2 4042 0040
Email: Lisa.L.Hyde@uon.edu.au

Abstract

Background: Variations in an individual's electronic health (eHealth) literacy may influence the degree to which health consumers can benefit from eHealth. The eHealth Literacy Scale (eHEALS) is a common measure of eHealth literacy. However, the lack of guidelines for the standardized interpretation of eHEALS scores limits its research and clinical utility. Cut points are often arbitrarily applied at the eHEALS item or global level, which assumes a dichotomy of high and low eHealth literacy. This approach disregards scale constructs and results in inaccurate and inconsistent conclusions. Cluster analysis is an exploratory technique, which can be used to overcome these issues, by identifying classes of patients reporting similar eHealth literacy without imposing data cut points.

Objective: The aim of this cross-sectional study was to identify classes of patients reporting similar eHealth literacy and assess characteristics associated with class membership.

Methods: Medical imaging outpatients were recruited consecutively in the waiting room of one major public hospital in New South Wales, Australia. Participants completed a self-report questionnaire assessing their sociodemographic characteristics and eHealth literacy, using the eHEALS. Latent class analysis was used to explore eHealth literacy clusters identified by a distance-based cluster analysis, and to identify characteristics associated with class membership.

Results: Of the 268 eligible and consenting participants, 256 (95.5%) completed the eHEALS. Consistent with distance-based findings, 4 latent classes were identified, which were labeled as low (21.1%, 54/256), moderate (26.2%, 67/256), high (32.8%, 84/256), and very high (19.9%, 51/256) eHealth literacy. Compared with the low class, participants who preferred to receive a lot of health information reported significantly higher odds of moderate eHealth literacy (odds ratio 16.67, 95% CI 1.67-100.00; *P*=.02), and those who used the internet at least daily reported significantly higher odds of high eHealth literacy (odds ratio 4.76, 95% CI 1.59-14.29; *P*=.007).

Conclusions: The identification of multiple classes of eHealth literacy, using both distance-based and latent class analyses, highlights the limitations of using the eHEALS global score as a dichotomous measurement tool. The findings suggest that eHealth literacy support needs vary in this population. The identification of low and moderate eHealth literacy classes indicate that the design of eHealth resources should be tailored to patients' varying levels of eHealth literacy. eHealth literacy improvement interventions are needed, and these should be targeted based on individuals' internet use frequency and health information amount preferences.

XSL•FO
RenderX

(J Med Internet Res 2019;21(8):e13423) doi:10.2196/13423

KEYWORDS

internet; health; literacy; cluster analysis; medical imaging

Introduction

Electronic Health Literacy Is Important for the Use and Receipt of Benefits From Electronic Health Programs

Web-based interventions have been reported to be consistently more effective than non-Web-based modalities in changing patient health behaviors and health-related knowledge [1]. Information and communication technology is also recognized as a promising enabler of safe, integrated, and high-quality health care, yet more scientifically rigorous research is needed [2,3]. Accordingly, internet-enabled health care is a strategic priority globally [4-7]. Electronic health (eHealth) literacy is one important factor influencing the use and receipt of benefits from Web-based health resources [8-10]. eHealth literacy refers to an individual's ability to seek, find, understand, and appraise health information from electronic sources, and apply the knowledge gained to addressing or solving a health problem [11]. The concept is derived from 6 literacy types (ie, health, computer, media, science, information, traditional literacy, and numeracy), which play an important role in facilitating engagement with Web-based health resources [11]. Inadequate eHealth literacy has been self-reported as a barrier to use of the internet for health information seeking purposes among the chronically ill [12]. Furthermore, descriptive research indicates that eHealth literacy is associated with positive cognitive (eg, understanding of health status) [8], instrumental (eg, self-management, physical exercise, and dieting) [8-10], and interpersonal (eg, physician interaction) [8] outcomes from Web-based health information searches. Individuals with lower eHealth literacy have been suggested to be older [8,13,14], less educated [8,14,15], have lower access to, or use of, the internet [15-17], and have poorer health [8].

Interpretations of Electronic Health Literacy Data are Inconsistent

Approaches used to assess eHealth literacy have included objective performance testing [18,19] and self-reported measurement [20-23]. The most commonly used self-reported measure is the 8-item, eHealth Literacy Scale (eHEALS) [20]. Compared with other self-report measures of eHealth literacy, strengths of the eHEALS include its psychometric rigor, brevity, ease of administration, and availability in a number of languages [17,19,20,24-26]. One of the key issues limiting the utility of the eHEALS is the lack of information about interpretation of these data. Although there is a convention that higher scores represent a higher level of eHealth literacy [20], there is an absence of guidance for the standardized interpretation of these scores. This guidance is needed to inform decision-making and follow-up actions [27]. eHEALS mean and median scores [8,13,14,28], as well as item response frequencies [14,29,30], are typically reported. Cut points have been arbitrarily applied at the item level [15], which disregards scale constructs. Furthermore, the common use of a single cut point to the global scale [8,16,28] implies a dichotomy of high versus low eHealth literacy and does not account for respondent self-perceived competency across the multiple eHEALS factors (ie, awareness, skills, and evaluation) [24,31]. These factors have only recently been identified [24,31], demonstrating that our understanding of the eHEALS and its psychometric properties is continuing to evolve more than a decade after the scale was published.

A Robust Approach to Analyzing Electronic Health Literacy Data Is Required

Shortcomings in the interpretation of eHEALS scores highlight the need for a robust approach to analyzing and interpreting eHealth literacy data. In line with the principles of scale development [27,32], measures should be refined as new data about a scale's properties accumulates. This includes retesting a scale when it is used in new populations and as new analytical techniques become available [27,32]. Cluster analysis is a sophisticated analytical approach, which has not previously been applied to eHealth literacy research. This powerful technique is used to identify natural groupings or structures within data and can therefore classify individuals who score similarly on an outcome measure, such as the eHEALS [33]. It has several strengths including: First, it is a data-driven exploratory technique and therefore not dependent on scoring thresholds, which are arbitrarily imposed by the author(s). Second, being able to observe and characterize natural structures or groupings means that researchers have a better understanding of subgroups of eHealth literacy in the sample population. If classes (or clusters) exist, ignoring their presence by analyzing the data as a single group could lead to an averaging out of any effects of interest [34]. Third, this approach allows for the multiple eHEALS domains (ie, skill, awareness, and evaluate) to be considered simultaneously across subgroups. For example, it can be known if one subgroup self-rates their *awareness* as highest, whereas another subgroup self-rates their *skills* as highest. Finally, regression analyses can be completed to examine patient characteristics associated with assignment to each eHealth literacy class.

By understanding the number and characteristics of groupings, it can be known whether a one size fits all approach to eHealth literacy improvement is appropriate, or whether more tailored interventions are required. If tailoring is needed, understanding how different classes scored across the eHEALS factors allows researchers and clinicians to ensure interventions are designed to specifically address the needs of that subgroup. Furthermore, understanding patient characteristics associated with class membership allows the identification of individuals who should be targeted for interventions, or who will require more intensive support throughout periods of eHealth delivery. A cluster analysis of eHEALS data is therefore an important next step to better understand the multicomponent nature of eHealth literacy and how these eHEALS factors coexist in subgroups of patients.

XSL•FO
RenderX

This study aimed to determine (1) whether there are identifiable eHealth literacy classes among magnetic resonance imaging (MRI) and computed tomography (CT) medical imaging outpatients; and (2) sociodemographic and internet use characteristics associated with each eHealth literacy class.

Methods

Design and Setting

This cross-sectional study was completed with MRI and CT medical imaging outpatients attending the imaging department of a large, tertiary hospital, located within New South Wales, Australia. The results of this study have been reported in accordance with the Strengthening the Reporting of Observational Studies in Epidemiology checklist [35] and the Checklist for Reporting Results of Internet E-Surveys [36].

Participants

Eligible participants were: (1) attending for an outpatient MRI or CT scan; (2) 18 years or older; and (3) reported having access to the internet for personal use. Participants were excluded if they were: (1) non-English speaking; (2) deemed by reception staff to be cognitively or physically unable to consent or complete the survey; or (3) identified as having completed the survey previously. MRI and CT medical imaging outpatients were the focus of this research because they have high unmet information preferences, which could potentially be met by eHealth capabilities [37].

Procedure

Medical imaging department receptionists identified potentially eligible participants when they presented for their outpatient appointment. Potentially eligible participants were informed about the research and invited to speak with a trained research assistant. Interested patients were provided with a written information sheet and introduced to the research assistant, who gave an overview of the study and obtained the patient's verbal consent to participate. During this overview, interested patients were told that the Web-based questionnaire would take approximately 10-15 mins to complete, participation was voluntary, and responses would remain confidential. The age, gender, and scan type of noninterested and nonconsenting patients were recorded. Consenting patients were provided with a tablet computer and asked to complete a Web-based questionnaire before their scan. Participants' study identification number, assigned by the receptionist and entered by the research assistant, provided access to the questionnaire. Each participant could move freely through each screen using next and back buttons. The questionnaire was pilot tested with MRI and CT medical imaging outpatients 2 weeks before study commencement, which confirmed the acceptability and feasibility of electronic survey administration in this study setting. A paper-and-pen version of the questionnaire was available to participants who requested it. If the patient was called for their procedure before finishing the questionnaire, only those questions that had been completed were used for data analysis. Electronic responses were deidentified, collected using the QuON platform [38], and stored securely on an access-restricted part of the University of Newcastle server. Ethics approval was obtained from the Human Research Ethics Committees of the Hunter New England Local Health District (16/10/19/5.11) and University of Newcastle (H-2016-0386).

Measure

eHealth literacy was assessed using the 8-item eHEALS. All 8 eHEALS items were administered on 1 screen within the Web-based questionnaire, and the presentation of these items was not random. Respondents indicated their level of agreement with each statement on a 5-point Likert scale from 1 *strongly disagree* to 5 *strongly agree*. Responses were summed to give a final score ranging from 8 to 40, with higher scores indicating higher eHealth literacy. The tool has demonstrated test-retest reliability [17], internal consistency [17,19,28], and measurement invariance across English speaking countries [24]. Previous studies, largely employing exploratory factor analysis, have suggested that the scale measures a single factor [8,17,19,20]. Emerging research using confirmatory factor analysis and based on the theoretical underpinnings of eHealth literacy suggests that the scale measures 3 factors: awareness, skills, and evaluate [24,31]. This 3-factor eHEALS structure has been identified in the medical imaging study setting (standardized root mean residual=0.038; confirmatory fit index=0.944; and root mean square error of approximation=0.156) [31]. As such, self-rated awareness, skills, and evaluate competencies of patients within each subgroup were explored within this study.

Study Factors

On the basis of previous research indicating an association with eHealth literacy, standard self-report items assessed participant gender, age, marital status, education, internet use frequency, and overall health status [8,13-17]. Remoteness of residence, health information amount preference (no information; some information; and a lot of information), and internet use for scan preparation (yes; no; and don't know) were hypothesized to influence eHealth literacy and were, therefore, included as covariates. Participant postcodes were mapped to the Accessibility/Remoteness Index of Australia Plus to categorize participant remoteness as metropolitan (major cities of Australia) or nonmetropolitan (inner regional, outer regional, remote, or very remote Australia) [39].

Data Analysis

Participant characteristics were summarized as frequencies and percentages or means and standard deviations. Consent bias was assessed for gender, scan type, and age group using Chi-square tests. Given the high completion rate (98.1%, 256/261 for individuals starting eHEALS items), only complete eHEALS data were included in the analyses. Items relating to each eHEALS factor were summed to generate separate awareness, skill, and evaluate factor scores.

Identification of Electronic Health Literacy Classes

Cluster analysis was completed using a 2-phased approach. Distance-based unsupervised clustering was undertaken as an initial exploratory knowledge discovery technique, to identify natural clusters of patients according to their responses (refer Multimedia Appendix 1 for methods and results). Secondary clustering of patients, using latent class analysis (LCA) as a

XSL•FO
RenderX

statistical modeling approach, was to be completed as a follow-up if distance-based cluster structures were observed. LCA was subsequently performed to verify the 4-cluster structure identified. LCA is less sensitive to choice of parameters (eg, distance metric), allows for uncertainty in class membership, and has greater power and lower type 1 error rates when compared with other clustering techniques [34], and was, therefore, selected as the primary analysis technique. Latent class membership probabilities were calculated to determine the proportion of the sample that belonged to each of the classes. Item response probabilities were calculated to determine the probability of endorsing each response option, conditional on class membership. The Bayesian Information Criterion (BIC) and G^2-statistic were computed to aid in determining the optimal number of classes (with plateauing indicating no improvements to model fit) [40], as were overall class interpretability and model parsimony. Model entropy was computed, with values closer to 1 representing clear class delineation [41]. The maximum posterior probability of class membership was also calculated for each participant, based on the optimal number of classes, with values greater than .5 indicating adequate probability for class assignment [42].

Characteristics Associated With Class Membership

An LCA regression analysis was performed to identify participant sociodemographic and internet use characteristics associated with class membership. Given the exploratory nature of data analysis, all covariates were initially cross-tabulated with class membership (assigned according to maximum posterior probability) to identify model sparseness, and then analyzed using univariate LCA regression: gender; age (<65 years vs 65+ years); geographic location of residence (major city vs regional or rural); marital status (married or living with partner vs not married); education (high school or less vs more than high school); overall health (fair or worse; good or better than good); information amount preference (a lot of information vs not a lot of information); internet use for scan preparation; and internet use frequency (daily vs less than daily). Likelihood ratio tests (based on the univariate results) were performed to determine whether each predictor significantly improved the fit of the model. Covariates with a statistically significant likelihood ratio test (P<.05) were included in the final multivariable LCA regression. Distance-based and latent class analyses were performed in R 3.4 [43]. Descriptive statistics were computed in STATA v13.

Sample Size

Sample sizes of at least 200 have been suggested as adequate for LCA, dependent on subsequent model fit and number of classes [40,44]. As such, a sample of at least 200 was deemed appropriate for this study.

Results

Sample

A total of 405 potentially eligible patients were invited to discuss the study with a research assistant during the 7-week recruitment period, of which 354 (87.4%) were interested in participating. Of 268 eligible participants, 261 (97.4%) started the eHEALS, 256 (95.5%) completed all eHEALS items, and 222 (82.8%) completed all eHEALS and study factor items. There were no significant differences between patients who were and were not interested in participating in the study based on gender, scan type, or age group. Table 1 provides a summary of the sociodemographic, scan, and internet characteristics of the study sample.

XSL•FO
RenderX

Table 1. Participant sociodemographic, scan, and internet characteristics (N=256). Number of observations for each characteristic may not total 256 because of missing data.

Characteristic	Value
Age (years), mean (SD)	53 (15.0)
Electronic Health Scale (eHEALS) domain score, mean (SD)	
Awareness (possible total=10)	6.9 (2.0)
Skills (possible total=15)	10.9 (2.9)
Evaluate (possible total=15)	10.0 (3.1)
Gender, n (%)	
Male	112 (43.8)
Female	144 (56.3)
Marital status, n (%)	
Married or living with partner	146 (64.6)
Not married or living with partner	80 (35.4)
Education completed, n (%)	
High school or less	128 (56.6)
More than high school	98 (43.4)
Geographic location, n (%)	
Metropolitan	200 (78.1)
Nonmetropolitan	56 (21.9)
Overall health, n (%)	
Poor	17 (7.7)
Fair	75 (34.1)
Good	94 (42.7)
Very good	34 (15.5)
Scan type, n (%)	
Computed tomography	101 (39.4)
Magnetic resonance imaging	152 (59.4)
Don’t know	3 (1.2)
Used internet for scan, n (%)	
Yes	27 (10.5)
No	228 (89.1)
Don’t know	1 (0.4)
Frequency of internet use, n (%)	
Less than once a month	11 (4.3)
Once a month	5 (1.9)
A few times a month	14 (5.5)
A few times a week	33 (12.9)
About once a day	47 (18.4)
Several times a day	146 (57.0)
Information amount preference, n (%)	
No information	2 (0.8)
Some information	58 (25.9)
A lot of information	165 (73.3)

XSL•FO
RenderX

Identification of Electronic Health Literacy Classes

The BIC and G^2-statistic continued to decrease as the number of classes (K) increased, but the improvement was progressively smaller after 3 classes (see Table 2). On the basis of the interpretability of the latent classes, the reduction in class size beyond K=4, and parsimony, the 4 class model was selected as the optimal class structure. The lowest maximum posterior probability under this 4 class model was .516. As such, all participants exceeded the threshold of .5 for maximum posterior probability and were assigned to a class. Hence, LCA findings on number of classes were consistent with that of distance-based clustering (see Multimedia Appendix 1).

Multimedia Appendix 2 shows the unconditional item response probabilities of each eHEALS response option based on class assignment. Classes were named according to likely level of eHealth literacy, with respect to that of other classes identified in the analysis:

- Class 1—*low eHealth literacy* (21.1% of respondents, 54/256): when compared with other classes, class 1 had the highest probability of responding *disagree* and *strongly disagree* across all eHEALS items. The probability of this group responding either *disagree* or *strongly disagree* was highest for awareness items (0.88 and 0.89), followed by evaluate items (0.79, 0.81, and 0.88) and skills items (0.66, 0.75, and 0.84).
- Class 2—*moderate eHealth literacy* (26.2% of respondents, 67/256): when compared with other classes, class 2 had the highest probability of responding *undecided* across all eHEALS items, and the second highest probability of responding *agree* across awareness and skills items. This group was most likely to respond *undecided* to awareness items (0.56 and 0.59), either agree (0.54 and 0.58) or undecided (0.48) to skills items, and undecided to evaluate items (0.55, 0.61, and 0.63).
- Class 3—*high eHealth literacy* (32.8% of respondents, 84/256): when compared with other classes, class 3 had the highest probability of responding *agree* across all eHEALS items. The probability of this class responding *agree* was greatest for skills items (0.97, 0.97, and 1.00), followed by awareness (0.80 and 0.91), and evaluate items (0.68, 0.71, and 0.81).
- Class 4—*very high eHealth literacy* (19.9% of respondents, 51/256): when compared with other classes, class 4 had the highest probability of responding *strongly agree* across all eHEALS items. The probability of this class responding *strongly agree* was greatest for skills items (0.71, 0.79 and 0.90), followed by evaluate (0.57, 0.74, and 0.86), and awareness items (0.53 and 0.61).

Characteristics Associated With Class Membership

Internet use for scan preparation was not included in regression analyses because of sparseness (ie, 10.5%, 27/256 of participants responded *yes* to internet use for scan preparation). Following univariate analyses, likelihood ratio difference tests indicated that age; education, marital status, overall health status, information amount preference, and internet use frequency all significantly improved the fit of the model ($P<.05$; see Multimedia Appendix 3) and were included in the multivariable regression analysis (see Table 3).

Class 1 (low eHealth literacy) was selected as a reference class for multivariable regression. This was because these participants likely need additional support to engage with eHealth, making identification of the characteristics of participants in this subgroup a priority. As shown in Table 3, participants who indicated that they preferred not to receive a lot of information about their health had 0.06 times the odds of belonging to class 2 (moderate eHealth literacy), compared with class 1 (low eHealth literacy), and this difference was statistically significant. Furthermore, participants who reported using the internet less than daily had 0.21 times the odds of belonging to class 3 (high eHealth literacy), compared with class 1 (low eHealth literacy), and this difference was statistically significant. There were no other significant differences in sociodemographic or internet use attributes between participants in class 1 (low eHealth literacy) and classes 2, 3, and 4 (moderate, high, and very high eHealth literacy, respectively).

Table 2. Goodness of fit indices for 1 to 5 class structures.

Class structure	BIC[a]	G^2-statistic	Entropy
1 class structure	5893.74	3402.83	1.00
2 class structure	5148.66	2474.76	0.97
3 class structure	4651.68	1794.79	0.98
4 class structure	4556.81	1516.93	0.92
5 class structure	4545.21	1322.34	0.90

[a]BIC: Bayesian Information Criterion.

XSL•FO
RenderX

Table 3. Adjusted odds ratios associated with membership of classes 2, 3, and 4, compared with class 1.

Variable	Class 1 versus class 2 (low vs moderate)		Class 1 versus class 3 (low vs high)		Class 1 versus class 4 (low vs very high)	
	Odds ratio (95% CI)	*P* value	Odds ratio (95% CI)	*P* value	Odds ratio (95% CI)	*P* value
Age						
<65 years	Ref[a]	Ref	Ref	Ref	Ref	Ref
65 years or older	0.37 (0.06-2.11)	.26	0.32 (0.10-1.03)	.06	0.37 (0.07-2.00)	.25
Education						
High school or less	Ref	Ref	Ref	Ref	Ref	Ref
More than high school	1.09 (0.15-7.65)	.93	2.21 (0.52-9.47)	.29	3.89 (0.67-22.76)	.14
Marital status						
Married or living with partner	Ref	Ref	Ref	Ref	Ref	Ref
Not married or living with partner	1.63 (0.26-10.23)	.60	0.96 (0.27-3.41)	.96	0.91 (0.14-6.01)	.92
Information amount preference						
A lot of information	Ref	Ref	Ref	Ref	Ref	Ref
Not a lot of information	0.06 (0.01-0.60)	.02[b]	0.61 (0.18-2.04)	.43	0.23 (0.04-1.29)	.10
Overall health						
Fair or worse	Ref	Ref	Ref	Ref	Ref	Ref
Good or better than good	1.10 (0.24-5.02)	.91	1.16 (0.35-3.87)	.81	1.48 (0.33-6.68)	.61
Internet use frequency						
Daily	Ref	Ref	Ref	Ref	Ref	Ref
Less than once a day	0.62 (0.14-2.67)	.52	0.21 (0.07-0.63)	.007[b]	0.17 (0.02-1.76)	.14

[a]Ref: reference category.

[b]Statistically significant.

Discussion

Principal Findings

This study was the first to identify classes of patients based on eHealth literacy, and to assess characteristics associated with class membership. The identification of multiple classes, using both distance-based and latent class analyses, highlights issues with using the eHEALS global score as a dichotomous measurement tool. In particular, these findings suggest that it may be important to account for multiple eHealth literacy subgroups when developing standardized guidance for the interpretation of eHEALS scores. Furthermore, the identification of multiple classes suggests that the design and delivery of eHealth resources may need to be tailored based on eHealth literacy. Patient characteristics, such as internet use frequency and health-related information amount preferences, may provide an indication of eHealth literacy, and related support needs.

Multiple Electronic Health Literacy Subgroups Were Identified

In total, 4 eHealth literacy classes were identified, and the probabilities of belonging to each of the 4 classes were similar (ie, range 19.9%-32.8%). The finding that eHealth literacy varied substantially in this population suggests that MRI and CT medical imaging outpatients may have differing support needs relating to the use of eHealth technology. Subgroups of patients were characterized by having either very high, high, moderate, or low eHealth literacy. Within the very high eHealth literacy subgroup, awareness was the lowest scoring competency. This may be because consumers who are familiar with eHealth also understand the masses of Web-based information that is available and the common difficulty of locating valid and reliable information sources [12]. Across all classes, participants reported being most competent in their skills using eHealth resources. Such skills may be perceived highly because they align to the computer and media literacy types, which comprise eHealth literacy [11]. These literacy types are increasingly used in the digital era, with 87% of Australians being identified as internet users in 2016-2017 [45].

In total, 2 out of 4 classes, comprising 52.7% of respondents, had the highest probability of responding either agree or strongly agree to eHEALS items, reflecting high and very high eHealth literacy. Despite this, there was room for improvement in awareness, skills, and evaluation competencies for the remaining 2 classes, comprising 47.3% of respondents and reflecting low and moderate eHealth literacy. This approximately even split in eHealth literacy capabilities is also apparent in other studies completed with cardiovascular disease patients [16] and chronic disease patients [46], which used arbitrary cut points to dichotomize high versus low eHealth literacy. It is possible that the application of dichotomous cut points prevented the identification of such diverse eHealth literacy subgroups. Further

XSL•FO
RenderX

research using cluster analyses should be conducted to determine whether multiple eHealth literacy subgroups exist across other health consumer populations. This information may inform the development of more targeted eHealth literacy improvement interventions.

Internet Use Frequency and Health Information Amount Preferences Predicted Class Membership

Those who had used the internet less than daily had approximately 5 times the odds of belonging to the low eHealth literacy class compared with the high eHealth literacy class. Although mixed findings exist [19], an association between internet use and eHealth literacy has been reported in studies with chronically ill patients and the general public [15-17]. Our findings may suggest that frequent internet users do use the internet for health, and this may result in greater self-reported eHealth literacy. Alternatively, they may indicate that frequent internet users self-perceive that their ability to engage with and evaluate general internet resources is transferable to health-related content.

Those with a preference not to receive a lot of information about their health had over 16 times the odds of belonging to the low eHealth literacy class, compared with the moderate eHealth literacy class. To the authors' knowledge, this study is the first to explore the association between preferred amount of information and eHealth literacy. It is possible that the inclusion of an *undecided* response option resulted in imposter syndrome for those in the moderate class [47]. In this case, participants underestimate their competency, opting for a neutral response option, to prevent being perceived as overconfident. Therefore, those in the moderate class may be more eHealth literate than findings suggest, which could contribute to a significant finding when comparing low and moderate classes. It may also be possible that those who prefer to receive a lot of information about their health are Web-based health-related information seekers, hence requiring eHealth literacy. An evidence review completed by the Australian Commission on Quality and Safety in Health Care found that patients typically use the internet as a supplement to advice from a health professional [48]. It is therefore likely that those who have greater preferences for health-related information require and develop the awareness, skills, and evaluation abilities needed to use this Web-based supplementary information. An analysis of the potentially moderating effects of Web-based health-related information seeking on the association between information amount preference and eHealth literacy should be explored in the future. This analysis should include an examination of the types of eHealth resources being accessed and used.

The technology acceptance model provides a theoretical justification for the characteristics related to a subgroup assignment [49]. Under this model, technology acceptance is influenced by perceived ease of use, and usefulness of the internet [49]. Accordingly, those who use the internet more frequently may be more likely to perceive ease of use of Web-based health resources. Similarly, those who prefer to receive a lot of health-related information may be more likely to deem eHealth as useful. Such perceived acceptability may result in greater self-rated eHealth literacy. Continued studies are needed to investigate this association and determine whether other factors not explored in this study, which promote perceived ease of use and usefulness of eHealth (eg, speed and availability of the internet, and self-management of chronic conditions, respectively), are associated with eHealth literacy. Contrary to expectations and inconsistent with previous studies [8,13-15], no other examined sociodemographic characteristics significantly influenced class membership. Inconsistencies with existing literature may indicate that the predictors of eHealth literacy differ across populations, settings, or cut points applied.

Practice Implications

The identification of low and moderate eHealth literacy classes suggests that eHealth literacy improvement interventions may be warranted within this population. However, there is minimal high-quality research investigating the effectiveness of such interventions, highlighting a need for continued research in this area [50]. Given their association with low class membership, those who use the internet less than daily and prefer not to receive a lot of health information should be the focus of such eHealth literacy improvement interventions. In the interim, researchers and clinicians should tailor the design and delivery of eHealth resources to patients' eHealth literacy, to maximize engagement and potential receipt of benefits. As skills were the highest rated competency across all classes within this study population, future eHealth interventions should be designed with a focus on promoting awareness and reducing the need to evaluate eHealth resources within the imaging setting. A written provider recommendation, which directs consumers toward credible eHealth resources, may be one scalable strategy to do this [31,51]. In cases where skills are low, alternative strategies may be needed, such as clear instructions on how to appropriately navigate Web-based content, reduced click-through requirements to retrieve Web-based materials, and the use of persuasive system design elements to enhance usability and maintain engagement [52].

Limitations and Future Research

To aid in the interpretation of findings, labels (ie, very high, high, moderate, and low) were arbitrarily assigned to eHealth literacy classes. It is therefore unclear whether, for example, those classified as very high eHealth literacy were indeed very high. As this study applied a novel approach to data analysis and interpretation, the generalizability of findings across medical imaging settings and to other patient groups is unknown. This class structure and the predictors of class membership should be studied and replicated in other populations. Furthermore, it is possible that the setting influenced responses as participants may have assumed that eHEALS questions related to scan-specific information on the internet rather than general eHealth resources.

The eHEALS was selected because of its established psychometric properties, emerging research proposing a 3-factor structure, and wide application [17,19,20,24,28,31]. However, it has been criticized for not measuring health 2.0. (ie, user-generated content and interactivity) and, therefore, lacking relevance to modern technology [21,24,53]. Some studies have adapted the scale to address this limitation, yet the body of research is small and as a result, the impacts on scale

XSL•FO
RenderX

psychometric properties remain unclear [21,24]. The generation of new Web-based content is, however, not highly relevant within the context of preparatory information provision for medical imaging procedures and this limitation is, therefore, not expected to influence our study.

Conclusions

This study used sophisticated analytical techniques to add to evidence about the nature of eHEALS scores within a clinical population. Cluster analyses were used to identify 4 classes of patients with differing eHealth literacy within this sample of MRI and CT medical imaging outpatients. The proportion of participants assigned to each latent class was similar, suggesting that eHealth literacy varies within this study setting. Across all classes, skills were perceived as the highest rated competency followed by either awareness or evaluation. The frequency of participants' personal internet use and their health-related information preferences predicted class membership. Tools such as the eHEALS may need to be administered to identify class assignment, and inform eHealth literacy improvement interventions, as well as the design and delivery of eHealth resources. Findings from this study should also contribute to the development of guidance for eHEALS scoring interpretation, which is a necessary next step to improve scale utility [27]. Study findings should be replicated in other populations and settings to increase the generalizability of results.

Acknowledgments

The authors thank the patients for their involvement in this study, as well as the administrative and clinical staff at Hunter New England Imaging for assistance with recruitment and data collection. This paper is supported by a Priority Research Centre for Health Behaviour Small Grant, and Hunter Cancer Research Alliance Implementation Science Flagship 2018 Student Award. LH is supported by an Australian Government Research Training Program Scholarship. AB is supported by a National Health and Medical Research Council Early Career Fellowship (APP1073317) and Cancer Institute New South Wales Early Career Fellowship (13/ECF/1-37). LM is supported by a postdoctoral fellowship (PF-16-011) from the Australian National Breast Cancer Foundation.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Distance-based cluster analysis.
[PDF File (Adobe PDF File), 330KB - jmir_v21i8e13423_app1.pdf]

Multimedia Appendix 2
Unconditional item response probabilities for a 4-class model of electronic health literacy.
[PDF File (Adobe PDF File), 159KB - jmir_v21i8e13423_app2.pdf]

Multimedia Appendix 3
Log likelihood difference tests.
[PDF File (Adobe PDF File), 12KB - jmir_v21i8e13423_app3.pdf]

References

1. 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]
2. Car J, Black A, Anandan C, Cresswell K, Pagliari C, McKinstry B, et al. University of Birmingham. 2008. The Impact of eHealth on the Quality & Safety of Healthcare: A Systemic Overview & Synthesis of the Literature. Report for the NHS Connecting for Health Evaluation Programme URL: https://www.birmingham.ac.uk/Documents/college-mds/haps/projects/cfhep/projects/001/NHSCFHEP001FinalReport.pdf [accessed 2019-08-08]
3. Black AD, Car J, Pagliari C, Anandan C, Cresswell K, Bokun T, et al. The impact of ehealth on the quality and safety of health care: a systematic overview. PLoS Med 2011 Jan 18;8(1):e1000387 [FREE Full text] [doi: 10.1371/journal.pmed.1000387] [Medline: 21267058]
4. Queensland Health. 2015. eHealth Investment Strategy URL: https://www.health.qld.gov.au/system-governance/strategic-direction/plans/ehealth-investment-strategy [accessed 2018-12-13] [WebCite Cache ID 74d1NO9Lh]
5. Health and Social Care Board. 2016. eHealth and Care Strategy for Northern Ireland: Improving Health and Wealth Through the Use of Information and Communication Technology URL: http://www.hscboard.hscni.net/download/Consultations/2014-15%20eHealth_and_Care_Strategy_Consultation/Consultation_Document-e-Health_and_Care-PDF_2mb.pdf [accessed 2019-08-08]

XSL•FO
RenderX

6. NHS Scotland. 2015. eHealth Strategy 2014-2017 URL: https://www.gov.scot/publications/ehealth-strategy-2014-2017/ [accessed 2019-08-08]
7. United States Department of Health and Human Services. 2015. Federal Health IT Strategic Plan: 2015-2020 URL: https://www.healthit.gov/sites/default/files/9-5-federalhealthitstratplanfinal_0.pdf [accessed 2019-08-08]
8. Neter E, Brainin E. eHealth literacy: extending the digital divide to the realm of health information. J Med Internet Res 2012 Jan 27;14(1):e19 [FREE Full text] [doi: 10.2196/jmir.1619] [Medline: 22357448]
9. Kim SH, Son YJ. Relationships between ehealth literacy and health behaviors in Korean adults. Comput Inform Nurs 2017 Feb;35(2):84-90. [doi: 10.1097/CIN.0000000000000255] [Medline: 27258808]
10. Mitsutake S, Shibata A, Ishii K, Oka K. Associations of ehealth literacy with health behavior among adult internet users. J Med Internet Res 2016 Jul 18;18(7):e192 [FREE Full text] [doi: 10.2196/jmir.5413] [Medline: 27432783]
11. Norman CD, Skinner HA. eHealth literacy: essential skills for consumer health in a networked world. J Med Internet Res 2006 Jun 16;8(2):e9 [FREE Full text] [doi: 10.2196/jmir.8.2.e9] [Medline: 16867972]
12. Lee K, Hoti K, Hughes JD, Emmerton L. Dr Google and the consumer: a qualitative study exploring the navigational needs and online health information-seeking behaviors of consumers with chronic health conditions. J Med Internet Res 2014 Dec 2;16(12):e262 [FREE Full text] [doi: 10.2196/jmir.3706] [Medline: 25470306]
13. 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 2;15(5):e93 [FREE Full text] [doi: 10.2196/jmir.2645] [Medline: 23639979]
14. Tennant B, Stellefson M, Dodd V, Chaney B, Chaney D, Paige S, et al. eHealth literacy and web 2.0 health information seeking behaviors among baby boomers and older adults. J Med Internet Res 2015 Mar 17;17(3):e70 [FREE Full text] [doi: 10.2196/jmir.3992] [Medline: 25783036]
15. 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]
16. Richtering SS, Hyun K, Neubeck L, Coorey G, Chalmers J, Usherwood T, et al. eHealth literacy: predictors in a population with moderate-to-high cardiovascular risk. JMIR Hum Factors 2017 Jan 27;4(1):e4. [doi: 10.2196/humanfactors.6217] [Medline: 28130203]
17. Chung SY, Nahm ES. Testing reliability and validity of the ehealth literacy scale (eHEALS) for older adults recruited online. Comput Inform Nurs 2015 Apr;33(4):150-156 [FREE Full text] [doi: 10.1097/CIN.0000000000000146] [Medline: 25783223]
18. Neter E, Brainin E. Perceived and performed ehealth literacy: survey and simulated performance test. JMIR Hum Factors 2017 Jan 17;4(1):e2 [FREE Full text] [doi: 10.2196/humanfactors.6523] [Medline: 28096068]
19. van der Vaart R, van Deursen A, Drossaert C, Taal E, van Dijk J, van de Laar M. 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 Nov 9;13(4):e86 [FREE Full text] [doi: 10.2196/jmir.1840] [Medline: 22071338]
20. Norman CD, Skinner HA. eHEALS: the ehealth literacy scale. J Med Internet Res 2006 Nov 14;8(4):e27 [FREE Full text] [doi: 10.2196/jmir.8.4.e27] [Medline: 17213046]
21. van der Vaart R, Drossaert C. Development of the digital health literacy instrument: measuring a broad spectrum of Health 1.0 and Health 2.0 skills. J Med Internet Res 2017 Jan 24;19(1):e27 [FREE Full text] [doi: 10.2196/jmir.6709] [Medline: 28119275]
22. 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 Dec 12;16(12):e287 [FREE Full text] [doi: 10.2196/jmir.3542] [Medline: 25499086]
23. Koopman RJ, Petroski GF, Canfield SM, Stuppy JA, Mehr DR. Development of the PRE-HIT instrument: patient readiness to engage in health information technology. BMC Fam Pract 2014 Jan 28;15:18 [FREE Full text] [doi: 10.1186/1471-2296-15-18] [Medline: 24472182]
24. Sudbury-Riley L, FitzPatrick M, Schulz PJ. Exploring the measurement properties of the ehealth literacy scale (eHEALS) among baby boomers: a multinational test of measurement invariance. J Med Internet Res 2017 Feb 27;19(2):e53 [FREE Full text] [doi: 10.2196/jmir.5998] [Medline: 28242590]
25. Diviani N, Dima AL, Schulz PJ. A psychometric analysis of the Italian version of the ehealth literacy scale using item response and classical test theory methods. J Med Internet Res 2017 Apr 11;19(4):e114 [FREE Full text] [doi: 10.2196/jmir.6749] [Medline: 28400356]
26. Chung S, Park BK, Nahm ES. The Korean ehealth literacy scale (K-eHEALS): reliability and validity testing in younger adults recruited online. J Med Internet Res 2018 Apr 20;20(4):e138 [FREE Full text] [doi: 10.2196/jmir.8759] [Medline: 29678800]
27. HealthMeasures. 2013. PROMIS® Instrument Maturity Model URL: http://www.healthmeasures.net/images/PROMIS/PROMISStandards_Vers_2_0_MaturityModelOnly_508.pdf [accessed 2019-06-07] [WebCite Cache ID 78wZrhm3Z]
28. 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 Nov 13;14(6):e153 [FREE Full text] [doi: 10.2196/jmir.1927] [Medline: 23149453]

XSL•FO
RenderX

29. 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 Sep 29;13(3):e75 [FREE Full text] [doi: 10.2196/jmir.1697] [Medline: 21960017]
30. 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]
31. Hyde LL, Boyes AW, Evans TJ, Mackenzie LJ, Sanson-Fisher R. Three-factor structure of the ehealth literacy scale among magnetic resonance imaging and computed tomography outpatients: a confirmatory factor analysis. JMIR Hum Factors 2018 Feb 19;5(1):e6 [FREE Full text] [doi: 10.2196/humanfactors.9039] [Medline: 29459356]
32. McDowell I. The theoretical and technical foundations of health measurement. In: Measuring Health: A Guide to Rating Scales and Questionnaires. Third Edition. New York: Oxford University Press; 2006:10-46.
33. Hastie T, Tibshirani R, Friedman J. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Second Edition. New York: Springer-Verlag; 2009.
34. Lanza ST, Rhoades BL. Latent class analysis: an alternative perspective on subgroup analysis in prevention and treatment. Prev Sci 2013 Apr;14(2):157-168 [FREE Full text] [doi: 10.1007/s11121-011-0201-1] [Medline: 21318625]
35. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP, STROBE Initiative. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. PLoS Med 2007 Oct 16;4(10):e296 [FREE Full text] [doi: 10.1371/journal.pmed.0040296] [Medline: 17941714]
36. Eysenbach G. Improving the quality of web surveys: the checklist for reporting results of internet e-surveys (CHERRIES). J Med Internet Res 2004 Sep 29;6(3):e34 [FREE Full text] [doi: 10.2196/jmir.6.3.e34] [Medline: 15471760]
37. Hyde L, Mackenzie L, Boyes AW, Evans TJ, Symonds M, Sanson-Fisher R. Prevalence and correlates of patient-centred preparatory information provision to computed tomography and magnetic resonance imaging outpatients: a cross-sectional study. Patient Educ Couns 2018 Oct;101(10):1814-1822. [doi: 10.1016/j.pec.2018.05.025] [Medline: 29884532]
38. Paul D, Wallis M, Henskens F, Nolan K. QuON: A Generic Platform for the Collation and Sharing of Web Survey Data. In: Proceedings of the 9th International Conference on Web Information Systems and Technologies. 2013 Presented at: WEBIST'13; May 8-10, 2013; Auchen, Germany p. 111-116 URL: https://ogma.newcastle.edu.au/vital/access/manager/Repository/uon:15474
39. The University of Adelaide. 2005. Accessibility/Remoteness Index of Australia (ARIA) URL: https://www.adelaide.edu.au/hugo-centre/services/aria [accessed 2019-08-08]
40. Nylund KL, Asparouhov T, Muthén BO. Deciding on the number of classes in latent class analysis and growth mixture modeling: a Monte Carlo simulation study. Struct Equ Modeling 2007 Dec 5;14(4):535-569. [doi: 10.1080/10705510701575396]
41. Celeux G, Soromenho G. An entropy criterion for assessing the number of clusters in a mixture model. J Classif 1996 Sep;13(2):195-212. [doi: 10.1007/BF01246098]
42. Clark S, Muthén B. Muthén & Muthén, Mplus. 2009. Relating Latent Class Analysis Results to Variables not Included in the Analysis URL: https://www.statmodel.com/download/relatinglca.pdf [accessed 2019-06-07] [WebCite Cache ID 78wb4hBuw]
43. R Project. 2018. The R Project for Statistical Computing URL: https://www.r-project.org/ [accessed 2019-08-09]
44. Tein JY, Coxe S, Cham H. Statistical power to detect the correct number of classes in latent profile analysis. Struct Equ Modeling 2013 Oct 1;20(4):640-657 [FREE Full text] [doi: 10.1080/10705511.2013.824781] [Medline: 24489457]
45. Australian Bureau of Statistics, Australian Government. 2018. 8146.0 - Household Use of Information Technology, Australia, 2016-17 URL: http://www.abs.gov.au/ausstats/abs@.nsf/mf/8146.0 [accessed 2019-06-24] [WebCite Cache ID 74d1a0s3q]
46. Paige SR, Krieger JL, Stellefson M, Alber JM. eHealth literacy in chronic disease patients: an item response theory analysis of the ehealth literacy scale (eHEALS). Patient Educ Couns 2017 Feb;100(2):320-326 [FREE Full text] [doi: 10.1016/j.pec.2016.09.008] [Medline: 27658660]
47. Clance PR, Imes SA. The imposter phenomenon in high achieving women: dynamics and therapeutic intervention. Psychotherapy 1978;15(3):241-247. [doi: 10.1037/h0086006]
48. Ramsay I, Peters M, Corsini N, Eckert M. Australia Commission on Safety and Quality in Healthcare. 2017. Consumer Health Information Needs and Preferences: A Rapid Evidence Review URL: https://www.safetyandquality.gov.au/sites/default/files/migrated/Consumer-health-information-needs-and-preferences-A-rapid-evidence-review.pdf [accessed 2019-08-08]
49. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep;13(3):319-340. [doi: 10.2307/249008]
50. Car J, Lang B, Colledge A, Ung C, Majeed A. Interventions for enhancing consumers' online health literacy. Cochrane Database Syst Rev 2011 Jun 15(6):CD007092 [FREE Full text] [doi: 10.1002/14651858.CD007092.pub2] [Medline: 21678364]
51. Usher W. Gold Coast general practitioners' recommendations of health websites to their patients. Med J Aust 2007;187(2):82-83. [doi: 10.5694/j.1326-5377.2007.tb01145.x]

XSL•FO
RenderX

52. Kelders SM, Kok RN, Ossebaard HC, van Gemert-Pijnen JE. Persuasive system design does matter: a systematic review of adherence to web-based interventions. J Med Internet Res 2012 Nov 14;14(6):e152 [FREE Full text] [doi: 10.2196/jmir.2104] [Medline: 23151820]
53. van de Belt TH, Engelen LJ, Berben SA, Schoonhoven L. Definition of Health 2.0 and Medicine 2.0: a systematic review. J Med Internet Res 2010 Jun 11;12(2):e18 [FREE Full text] [doi: 10.2196/jmir.1350] [Medline: 20542857]

Abbreviations

BIC: Bayesian Information Criterion
CT: computed tomography
eHEALS: eHealth Literacy Scale
eHealth: electronic health
LCA: latent class analysis
MRI: magnetic resonance imaging

Edited by G Eysenbach; submitted 16.01.19; peer-reviewed by PJ Schulz, S Antani; comments to author 01.05.19; revised version received 20.06.19; accepted 19.07.19; published 28.08.19.

Please cite as:
Hyde LL, Boyes AW, Mackenzie LJ, Leigh L, Oldmeadow C, Riveros C, Sanson-Fisher R
Electronic Health Literacy Among Magnetic Resonance Imaging and Computed Tomography Medical Imaging Outpatients: Cluster Analysis
J Med Internet Res 2019;21(8):e13423
URL: http://www.jmir.org/2019/8/e13423/
doi:10.2196/13423
PMID:31464188

©Lisa Lynne Hyde, Allison W Boyes, Lisa J Mackenzie, Lucy Leigh, Christopher Oldmeadow, Carlos Riveros, Rob Sanson-Fisher. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 28.08.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

What Do Patients Complain About Online: A Systematic Review and Taxonomy Framework Based on Patient Centeredness

Jing Liu[1], MS; Shengchao Hou[1,2], MS; Richard Evans[3], PhD; Chenxi Xia[1], PhD; Weidong Xia[4], PhD; Jingdong Ma[1], MD, PhD

[1]School of Medicine and Health Management, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China

[2]Library, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China

[3]College of Engineering, Design and Physical Sciences, Brunel University London, London, United Kingdom

[4]Department of Information Systems and Business Analytics, College of Business, Florida International University, Miami, FL, United States

Corresponding Author:
Jingdong Ma, MD, PhD
School of Medicine and Health Management
Tongji Medical College
Huazhong University of Science and Technology
No 13 Hangkong Road
Qiaokou District
Wuhan,
China
Phone: 86 27 83692826
Email: jdma@hust.edu.cn

Abstract

Background: Complaints made online by patients about their health care experiences are becoming prevalent because of widespread worldwide internet connectivity. An a priori framework, based on patient centeredness, may be useful in identifying the types of issues patients complain about online across multiple settings. It may also assist in examining whether the determinants of patient-centered care (PCC) mirror the determinants of patient experiences.

Objective: The objective of our study was to develop a taxonomy framework for patient complaints online based on patient centeredness and to examine whether the determinants of PCC mirror the determinants of patient experiences.

Methods: First, the best fit framework synthesis technique was applied to develop the proposed a priori framework. Second, electronic databases, including Web of Science, Scopus, and PubMed, were searched for articles published between 2000 and June 2018. Studies were only included if they collected primary quantitative data on patients' online complaints. Third, a deductive and inductive thematic analysis approach was adopted to code the themes of recognized complaints into the framework.

Results: In total, 17 studies from 5 countries were included in this study. Patient complaint online taxonomies and theme terms varied. According to our framework, patients expressed most dissatisfaction with *patient-centered processes* (101,586/204,363, 49.71%), followed by *prerequisites* (appropriate skills and knowledge of physicians; 50,563, 24.74%) and *the care environment* (48,563/204,363, 23.76%). The least dissatisfied theme was *expected outcomes* (3651/204,363, 1.79%). People expressed little dissatisfaction with *expanded* PCC dimensions, such as *involvement of family and friends* (591/204,363, 0.29%). Variation in the concerns across different countries' patients were also observed.

Conclusions: Online complaints made by patients are of major value to health care providers, regulatory bodies, and patients themselves. Our PCC framework can be applied to analyze them under a wide range of conditions, treatments, and countries. This review has shown significant heterogeneity of patients' online complaints across different countries.

(J Med Internet Res 2019;21(8):e14634) doi:10.2196/14634

KEYWORDS

patient-centered care; delivery of health care; systematic review; taxonomy

XSL•FO
RenderX

Introduction

As internet availability and usage grows worldwide, patients are spontaneously rating their experiences with physicians and hospitals by sharing their opinions about health encounters on the World Wide Web via mediums such as social media websites, Web-based consumer opinion platforms, and physician rating websites (PRWs) [1-5]. Previous research has demonstrated that patients are often influenced by peer-submitted comments posted on opinion and rating websites when making health care decisions [6,7]. On the basis of this notion, medical providers are able to leverage the information posted on such platforms to better comprehend patient experiences and engagement levels [4] and increase the understanding of patient frustrations and joy points during hospital visits [8-16]. By capturing patient data in real time, health care providers can use them as a quality metric to highlight insufficient physician performances or irregular events [5,17]. On the basis of extensive circumstantial evidence [18-20], countries such as the United Kingdom systematically collect data relating to patient experiences from their quality-reporting website (National Health Service, NHS choices) to support the further development of patient-centered care (PCC) [8,21-23].

Given the intrinsic value of comments posted online by patients, it is important that health care providers make efficient use of the information collected. Practices observed from the adverse event taxonomy proposed by Harrison et al [24], which facilitated the collection and aggregation of data to compare findings, identify priorities, and develop wide-reaching patient safety solutions, demonstrated the momentousness of a unified, agreed framework with standardized concepts and terms. Despite the large volume of work published in this domain (eg, the studies by Reader et al [25] and Li et al [26]), currently available taxonomies for analyzing online complaints made by patients often lack standardized themes, terminology, and underlying unifying theory, creating difficulties in making sense of data that cannot be used to compare against other services, organizations, or countries. An operational and rigorous framework that classifies complaints made by patients, containing standardized concepts with agreed definitions and preferred terminology and establishing the relationships between concepts based on an explicit and nonoverlapping domain ontology, is required [24].

When we take into consideration the well-developed *PCC framework*, which forms the basis for patient experience measurement systems in the United States, the United Kingdom, and other parts of Europe [27,28], we can see that it includes clear and proven terminologies with standardized dimensions and concepts. To create a patient-centric health care system that meets the needs and preferences of patients is one of the primary goals of numerous countries [17,21,27,29-32]. Therefore, it is feasible that we use the principles of PCC to guide the analysis of online complaints and examine whether the determinants of PCC are the same as the patient experience. To confirm this approach, a literature search was completed using a combination of keywords and subject headings, based on the defined concepts of patient complaints and PCC. Through analysis of the search results, it was identified that no study is yet to be completed that categorizes issues based on patient centeredness and that a generic taxonomy is required that appropriately analyzes issues against a wide range of conditions and in the context of different health care settings. These findings led to the following research questions (RQs) being posed:

- RQ1: Have previous studies formed or adopted a credible taxonomy framework?
- RQ2: Are available frameworks based on patient centeredness commensurate with what patients currently complain about online?
- RQ3: Which dimensions of PCC constitute the focus of online complaints made by patients?
- RQ4: Could a taxonomy framework allow us to identify the differences in patient complaints in a multicountry context?

To answer these questions, a systematic review approach was used. First, we followed a process of synthesis [29] to propose our a priori framework. Then, comprehensive searches were conducted to systematically identify qualified studies relating to patient complaints online; at this point, data were extracted to match with the a priori framework; those that matched were compared between countries.

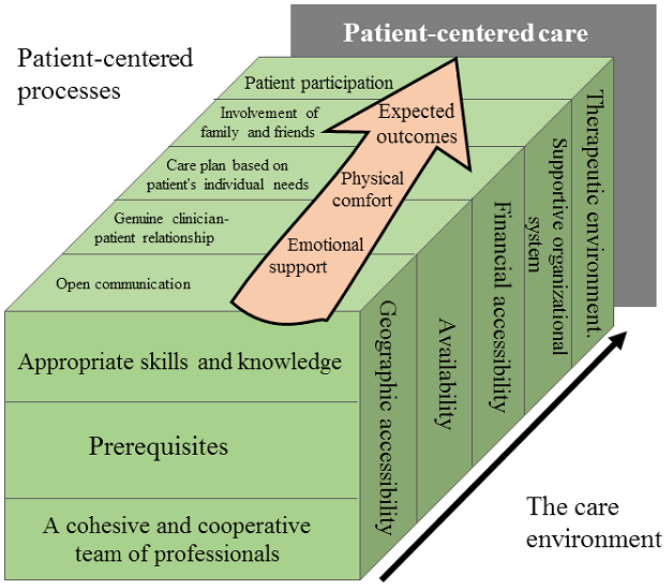
To create the proposed framework, we synthesized all studies relating to PCC using the best fit framework synthesis technique proposed by Booth and Carroll [29]. First, we identified all the relevant frameworks or conceptual models that related to PCC, which are published in academic literature. At the forefront is the widely understood *Picker Principles of Care* framework, an internationally renowned approach used for measuring quality improvement in health care in the United States and the United Kingdom. Designed by the UK-based Institute for Healthcare Improvement, the PCC framework includes 8 dimensions: (1) respect for patient values, preferences, and expressed needs; (2) coordination and integration of care; (3) information and education; (4) physical comfort; (5) emotional support and alleviation of fear and anxiety; (6) involvement of family and friends; (7) continuity and transition; and (8) access to care. Although well adopted, the framework is considered a single-layer structure, which may lead to inefficiency in identifying homogeneous underlying problems. Brendan McCormack et al [33] developed a patient-centered framework comprising 4 constructs—prerequisites, the care environment, patient-centered processes, and expected outcomes—which was derived from Donabedian's [34] structure-process-outcome assessment model. The proposed framework has been rigorously developed and tested in acute hospital settings [33] and is comprehensive enough to incorporate PCC dimensions. Second, we identified all relevant publications relating to the dimensions of PCC. Aside from the frameworks mentioned above, Kitson et al [35] and Rathert et al [36] completed systematic reviews of the PCC field and synthesized the common core elements of PCC. These 2 studies are highly cited and have been validated by a variety of follow-up studies. We compared and synthesized the dimensions of these 5 models. Third, we conducted a framework synthesis using thematic analysis [37]. Finally, grounded in the above, we developed an a priori framework with 4 domains, 8 categories, and 25 subcategories of online

XSL•FO
RenderX

health care complaints based on patient centeredness, as shown in Figure 1.

In the first layer of the framework, the prerequisites focus on health care professionals having appropriate skills and knowledge and the team of professionals being cohesive and cooperative. The care environment refers to the context in which care is delivered and includes supportive organizational systems and accessibility, in terms of geography, financial affordability, and availability. Patient-centered processes focus on delivering care through a range of activities that operationalize person-centered care. Expected outcomes relate to the results expected from effective PCC, addressing a patients' physical and emotional needs.

Figure 1. Proposed an a priori framework of online complaints based on patient centeredness.



Methods

The systematic review reported on hereafter was conducted according to the Preferred Reporting Items for Systematic reviews and Meta-Analyses guidelines[38]. As Reader et al [25] have previously systematically reviewed patient complaints, this study adopted some of their reporting items.

Search Strategy

The electronic databases of ISI Web of Science, Scopus, and PubMed were searched for articles published between 2000 (because of the explosive growth of internet usage around 2000 [30]) and June 2018. A medical librarian (SCH) developed a Boolean search strategy. Then, a doctoral student (JL) carried out the search strategy, which was revised, if required, by SCH. The keywords searched in the *Title* or *Abstract* fields related to (1) *complaints* (eg, comments OR ratings OR suggestions OR reviews OR feedback) and (2) *online* (eg, free text OR social media OR e-health OR virtual OR internet OR Facebook OR twitter), which were subject to inquiry (see Multimedia Appendix 1).

Inclusion Criteria

Studies were considered eligible if they were (a) related to the collection of primary quantitative data about patient complaints; (b) submitted by patients or third parties on their behalf; (c) comments uploaded to PRWs, organization's websites, or any other online channel; or (d) conveyed in English to facilitate cross-country comparison.

Exclusion Criteria

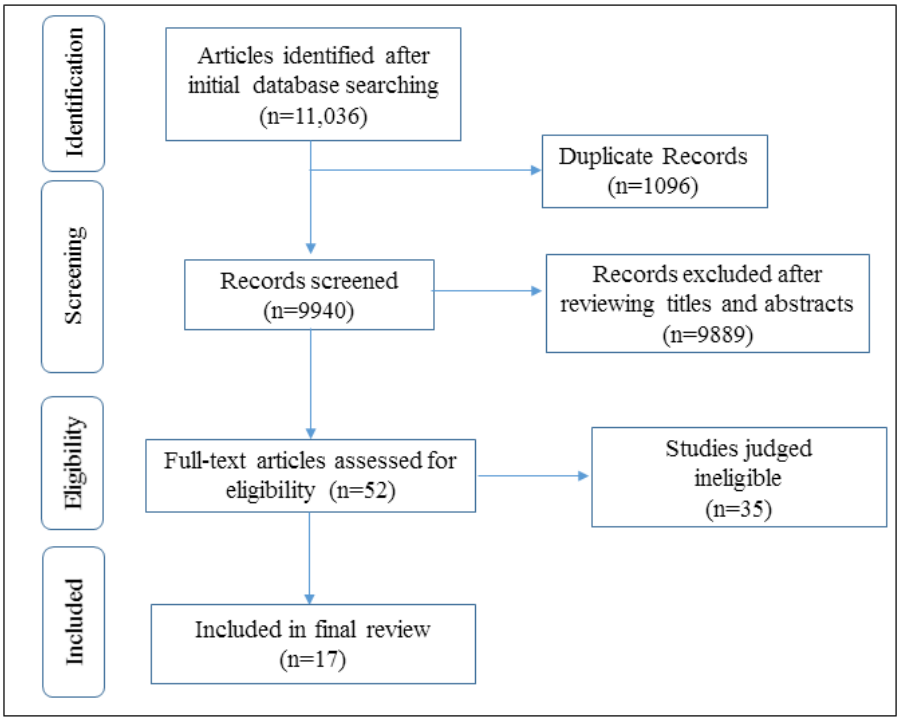
Studies considered ineligible from this research included those that only referred to (a) physician ratings, (b) satisfaction questionnaires, (c) complaints made to non–health care–related organizations, or (d) qualitative analysis without quantifiable themes.

Study Selection

After removal of duplicates, JL screened the titles and abstracts of all the remaining records for relevance. In the next stage, the full-article text of the retrieved results was independently examined by JL and SCH for inclusion. Discrepancies were adjudicated by a senior researcher (JDM). A total of 17 papers were included in this systematic analysis, as illustrated in Figure 2.

XSL•FO
RenderX

Figure 2. Preferred Reporting Items for Systematic reviews and Meta-Analyses flow chart.



Data Extraction and Analysis

Data were initially extracted by JL and subsequently checked by SCH, JL, and JDM, who together carried out the coding phase. The process consisted of the following phases: (1) descriptive and methodological data included in the study were extracted based on the items listed in the People-Centered and Integrated Health Services [33], (2) the number and categories of patient complaints were extracted, and (3) all themes or categories in selected studies were traversed and classified to match the initial categories. Data was analyzed using a thematic analysis approach adapted from the procedure outlined by Braun et al [37]. JL and SCH co-coded the issues of 2 papers to ensure that the coding framework and themes were commonly understood by the research team. JL and SCH independently coded the issues of the remaining 15 papers. In this paper, we achieved high interrater reliability (kappa=.82). Throughout the coding stage, to ensure consistency in categorizing the issue of complaints, 3 researchers (JL, SCH, and JDM) discussed to eliminate divergence.

Results

Search Results

During the process of synthesizing identified literature, 35 studies were excluded for various reasons, including inability to distinguish between positive and negative comments [2,39], fuzzy quantity [11,16,41], channels of complaints [40,42], or only the number of high-frequency words of complaints mentioned [15,43]. Ultimately, 17 publications were identified as eligible. A wide range of data sources were represented, including (1) PRWs (10/17, 59%), such as RateMDs and China's Good Doctor website; (2) government-managed health websites, such as NHS choices (2/17, 12%); (3) social networking sites (4/17, 24%); and (4) national online surveys (1/17, 6%). Of the 4 articles referring to social networking sites, 2 included data captured from tweets, whereas one was from Google+ reviews and one from Facebook.

Descriptive and Methodological Data

Through our analysis, we found that pertinent articles have emerged since 2012, with a steady increase observed ever since. Most of the research reported in the analyzed studies focused on PRWs or tweets examining country contexts, such as the United States, the United Kingdom, Germany, China, and Canada. From the data collected, it was identified that 10 articles focused on the United States (59%); 4 on the United Kingdom (24%); and 1 article each on China, Germany, and Canada. Most studies did not screen the departments in which the complainant arose (13/17, 76%), whereas 4 articles (24%) paid attention to complaints received from those dealing with a specific illness or encountering specific medical services. The number of complaints reported (or listed in the thematic analysis) in each study varied widely (average 6543, SD 15,547, range 36-57,028, and median 480).

It was noted that the 17 articles included in our sample had different classification criteria, theme terms, and granularity of complaints. Among them, 6 articles (35%) classified complaints based on published categorization schemes, whereas 9 articles (53%) generated their own coding framework from scratch. The coding framework of 2 articles (12%) came from data source organizations. For example, Zhang et al [9] classified complaints according to the stages of medical consultation, which resulted in a 3-layer classification of stages, including medical consultation, diagnosis and treatment processes, and specific complaint attributions. Emmert et al [44] classified complaints

XSL•FO
RenderX

according to the object being complained about and generated a 2-layer taxonomy, referring to both the object and specific complaint attributions. We assessed the quality of the included studies against interrater reliability/performance measures; number of layers; number of codes used, and if there were any definitions or descriptions; and examples provided, as shown in Table 1. The results indicated that the quality of studies varied: 8 did not report any classification measurement results and 4 studies used less than 6 codes, overall, whereas 15 did not provide definitions. A total of 13 studies did not provide examples of each theme/class. Thus, the taxonomy for patients' online complaints is unstandardized, and it is deemed difficult to identify consistent problems arising in patient care.

XSL•FO
RenderX

Table 1. A breakdown of descriptive and methodological data.

Article	Country	Health care settings	Data sources	Complaints reported, n	Source of coding frame	Complaints coded by	Classification quality			
							IRR[a]/performance measure	Layers, n	Codes used, n	Definitions or descriptions present	Examples present
Alemi et al [45] (2012)	United States	Multiple	PRW[b]	307	Survey items	Machine learning algorithms	precision, recall, *F* measure, and area under ROC curve	2	32	No	No
López et al [46] (2012)	United States	Primary care	PRW	263	Developed	Authors	κ[c]	2	24	No	Yes
Lagu et al [8] (2013)	United Kingdom	Multiple	Health website	200	Literature and NHS[d] Choices prompts	Authors	κ	2	22	No	No
Detz et al [47] (2013)	United States	Primary care	PRW	36	Literature	Authors	NR[e]	2	18	No	No
Emmert et al [44] (2014)	Germany	Multiple	PRW	480	Literature	Authors	κ	2	49	No	No
Greaves et al [48] (2014)	United Kingdom	Multiple	SNS[f] (tweets)	60	Literature	Authors	κ	3	17	No	No
Macdonald et al [49] (2015)	Canada	Multiple	Dental services	15	Developed	Authors	NR	2	16	No	No
Hawkins et al [39] (2015)	United States	Hospitals	SNS (tweets)	814	Developed	Amazon Mechanical Turk workers/curators	κ	1	10	No	No
Lagu et al [50] (2015)	United States	Multiple	SNS (Facebook)	37	Developed	Two investigators	r_s[g]	1	4	No	Yes
Trehan et al [51] (2016)	United States	Multiple	PRW	533	Literature	Authors	NR	2	5	No	No
Cunningham and Wells [52] (2017)	United Kingdom	Multiple	Official online survey	1969	Developed	Authors	NR	2	22	No	Yes
James et al [53] (2017)	United States	Multiple	PRW	10,992	Developed	Machine learning algorithms	NR	1	3	No	No
Xu et al [54] (2017)	United States	Multiple	PRW	125	Developed	Authors	NR	1	9	Yes	No
King et al [55] (2017)	United States	Multiple	SNS (Google+ reviews)	34,748	Developed	Customized software	NR	1	2	Yes	No
Brookes and Baker [23] (2017)	United Kingdom	Multiple	Health website (NHS choices)	57,028	Literature	Computer-assisted methods (CQPweb)	NR	2	23	No	Yes
Zhang et al [9] (2018)	China	Multiple	PRW	3012	Developed	Authors	α[h]	3	50	No	No
Emmert et al [56] (2018)	United States	Multiple	PRW	618	Literature	Authors	κ	1	20	No	No

[a]IRR: Interrater reliability.

XSL•FO
RenderX

[b]PRW: physician rating website.

[c]κ: Cohen kappa coefficient.

[d]NHS: National Health Service.

[e]NR: not retrievable.

[f]SNS: social networking service.

[g]r_s: Spearman correlation.

[h]α: Cronbach alpha.

All papers specified the coders of the complaints; among them, 71% (n=12) were coded by the authors, 12% (n=2) were coded by data curators or investigators, and 18% (n=3) were coded automatically via machine learning techniques. In addition, 9 articles reported intercoder reliability (6 with the Cohen kappa coefficient, 2 with the Spearman correlation or Cronbach alpha, and 1 unspecified).

Coding Results

In total, across the 17 papers, 326 issue codes were used to code 154,762 complaints. Among them, 36 issue codes incorporating 9602 complaints were not classified into our classification framework for their ambiguous meaning of the category (such as *others*). When classifying all complaint codes, identified from the literature, into our complaint classification code system, we determined that a single code may include a number of new complaint codes that have been assigned individual type codes. The coding results containing concept explanations and issue numbers are provided in Table 2. Patients' online complaints were seen to fit into the 4 domains, proposed in the a priori framework: (1) prerequisites, (2) patient-centered processes, (3) care environment, and (4) expected outcomes. From the complaints analyzed, it was concluded that patients have the most dissatisfaction with the *patient-centered processes* (101,586/204,363, 49.71%), followed by *prerequisites* (50,563/204,363, 24.74%), and *care environment* (48,563/204,363, 23.76%), with the least satisfied being *expected outcomes* (3651/204,363, 1.79%).

The *prerequisites* domain referred to dissatisfaction with the professional skills and knowledge of the health care provider (48251/204,363, 23.61%) and cooperation between professionals in the medical team (2312/204,363, 1.13%). Among the 4 subcategories, the most common referred to comments about *attributes of the patient-centered professional* (38,314/204,363, 18.75%), which represents the explicit patient-centered personality traits of professionals.

The domain with the most online complaints, *patient-centered processes*, was represented by a number of categories, as shown in Table 2. Within this domain, the greatest number of complaints related to a lack of *open communication of knowledge, personal expertise, and clinical expertise between the patient and the professional* (47,385/204,363, 23.19%). The second category *care plan based on patient's individual needs* collected 40,722 issues (40,722/204,363, 19.93%). The remaining 3 categories contained a small number of complaints, for example, *patient participation as a respected and autonomous individual* was represented by 4.00% (8186/204,363) of total issues. Among the 14 subcategories in this domain, *information, communication, and education* accounted for the majority of complaints, representing 22.80% (46,596/204,363) of total issues reported, whereas no complaints were reported on *patient autonomy*. The most frequently mentioned subtheme in *the care environment* domain was *availability* (28,784/204,363, 14.08%), which represented the timeliness of service and the accessibility of medical staff, facilities, and materials. Common problems mentioned in several articles were lengthy telephone calls made by the physician during consultation and difficulties in patients booking an appointment or seeing a clinician. It is worth noting that issues of *therapeutic environment* emerged in 12 articles. The *expected outcomes* domain contained complaints relating to physical comfort and physical care (3009/204,363, 1.47%), and emotional support for alleviation of anxiety issues (642/204,363, 0.31%).

XSL•FO
RenderX

Table 2. Main results of the coding.

Domains, categories, and subcategories	Definition
Prerequisites (50,563 /204,363, 24.74%)	
Health professionals have appropriate skills and knowledge (48,251/204, 363, 23.61%)	
Professional competence (9937/204,363, 4.86%)	Professional competence focuses on the knowledge and skills of the professionals to make decisions and prioritize care and includes competence in physical or technical aspects of care.
Attributes of the patient-centered professional (38,314/204,363, 18.75%)	The following care attributes are important in professionals' approach to patients: respect, good manners, being polite, good etiquette, sensitive, welcoming, and empathetic.
A cohesive and co-operative team of professionals (2312/204,363, 1.13%)	
Cooperation among clinicians a priority (2312/204,363, 1.13%)	Patient-centered clinicians are described as being committed and cooperative in an effective team that draws on individuals from different disciplines to complement one another in patient care.
Differences in perception of role between doctors, nurses, and patients (0/204,363, 0.00%)	Members of the team know exactly the differences in the roles of doctors, nurses, and patients.
Patient-centered processes (101,586/204,363, 49.71%)	
Participation of the patient as a respected and autonomous individual (8186/204,363, 4.01%)	
Respect for patients' values, preferences, and expressed needs (7446/204,363, 3.64%)	Patient-centered care (PCC) responds precisely to each patient's wants, needs, and preferences.
Patient as a source of control (370/204,363, 0.18%)	Patients should be given the necessary information and the opportunity to exercise the degree of control they choose over health care decisions that affect them. The health system should be able to accommodate differences in patient preferences and encourage shared decision making.
Patient's active involvement and participation (370/204,363, 0.18%)	It gives patients abundant opportunities to be informed and involved in medical decision making and guides and supports those providing care in attending to their patients' physical and emotional needs and maintaining or improving their quality of life as far as possible.
Patient autonomy (0/204,363, 0.00%)	Patients direct their lives according to their personal convictions and individual reasons and goals, ultimately to achieve self-governance and self-care.
Involvement of family and friends (591/204,363, 0.29%)	
Family and friends supported as caregivers (591/204,363, 0.29%)	This dimension of patient-centeredness focuses on accommodating family and friends on whom patients may rely, involving them as appropriate in decision making, supporting them as caregivers, making them welcomed and comfortable in the care delivery setting, and recognizing their needs and contributions.
Care plan based on patient's individual needs (40,722/204,363, 19.93%)	
Care customized according to patient needs and values (3101/204,363, 1.52%)	PCC is highly customized, incorporates cultural competence and empowers patient decision making
Needs are anticipated (462/204,363, 0.23%)	Care plan meets the future needs of patients.
Coordination and integration of care (35,923/204,363, 17.58%)	The extent to which patient care services are coordinated across people, functions, activities, and sites in a timely manner to maximize the value of services delivered to patients. Patients identified 3 areas in which care coordination can reduce feelings of vulnerability: coordination of clinical care, coordination of ancillary and support services, and coordination of frontline patient care.
Transition and continuity of care (1236/204,363, 0.60%)	Support patients with their ability to care for themselves after discharge. Meeting patient needs in this respect requires the following: understandable, detailed information regarding medications, physical limitations, dietary needs, etc; coordinate and plan ongoing treatment and services after discharge; and provide information regarding access to clinical, social, physical, and financial support on a continuing basis.
Genuine clinician-patient relationship (4702/204,363, 2.30%)	
Care based on a continuous healing relationship (4385/204,363, 2.15%)	Patients should receive care whenever they need it and in many forms, not just face-to-face visits. This rule implies that the health care system should be responsive round the clock (24×7) and that access to care should be provided over the internet, by telephone, and by other means in addition to face-to-face visits.
Clinician-patient relationship (317/204,363, 0.16%)	The effective clinician-patient relationships coming from a healing model, with education and disease management information delivered within the context of the healing relationship.

XSL•FO
RenderX

Domains, categories, and subcategories	Definition
Open communication of knowledge, personal expertise, and clinical expertise between the patient and the professional (47,385/204,363, 23.19%)	
Knowledge shared and information flows freely (752/204,363, 0.37%)	Patients should have unfettered access to their own medical information and to clinical knowledge. Clinicians and patients should communicate effectively and share information.
Information, communication, and education (46,596/204,363, 22.80%)	With respect to their health, people tend to wonder (1) what is wrong (diagnosis) or how to stay well, (2) what is likely to happen and how it will affect them (prognosis), and (3) what can be done to change or manage their prognosis. Common to all such interactions is the desire for trustworthy information (often from an individual clinician that is attentive, responsive, and tailored to an individual's needs).
Feedback mechanisms to measure patient experience (37/204,363, 0.02%)	Clinicians can move beyond their individual patients and use survey instruments and other tools that invite patients to report collectively about their clinical experiences.
The care environment (48,563/204,363, 23.76%)	
System issues (48,563/204,363, 23.76%)	
Geographic accessibility (536/204,363, 0.26%)	The physical distance, travel time, and cost from service delivery point to the patient.
Availability (28,784/204,363, 14.08%)	Having the right type of care available to those who need it, such as hours of operation and waiting times that meet the demands of those who would use care, as well as having the appropriate type of service providers, materials and facilities such as parking, food, and hand hygiene.
Financial accessibility (2964/204,363, 1.45%)	The relationship between the price of services (in part affected by their costs) and the willingness and ability of users to pay for those services as well as be protected from the economic consequences of health costs.
Supportive organizational system (8624/204,363, 4.22%)	A system that promotes a philosophy conducive to PCC. Specifically, the system's managers and employees (usually not clinical experts) create and maintain a responsive, secure, and orderly system on their own or via information systems.
Therapeutic environment (7655/204,363, 3.75%)	It is the context in which care is delivered. A place quiet, peaceful, neat, clean, and private, if necessary.
Expected outcomes (3651/204,363, 1.79%)	
Addressing a patient's physical and emotional needs (3651/204.363, 1.79%)	
Physical comfort (3009/204,363, 1.47%)	Attention to physical comfort implies timely, tailored, and expert management of pain, shortness of breath, or other discomforts, with the best possible curative effect. Try the best to avoid unexpected patient events and actively deal with them once they occur.
Emotional support—alleviation of anxiety (642/204,363, 0.31%)	PCC attends to the anxiety that accompanies every injury and illness, whether because of uncertainty, fear of pain, disability or disfigurement, loneliness, financial impact, or the disease burden on one's family.

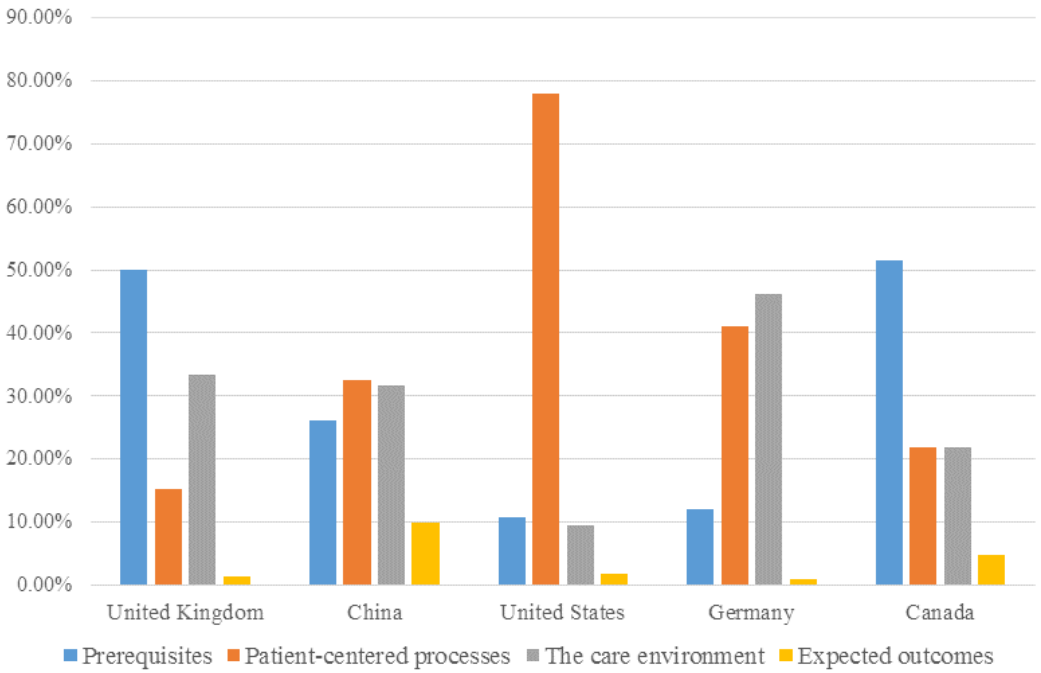
Country-Specific Differences in Complaints

We analyzed online complaints from patients based in the United Kingdom, China, the United States, Germany, and Canada because of these being the focus of the studies identified in the literature. The distribution of the 4 domains of online health care complaints across different countries is displayed in Figure 3.

With regard to care aspects that were complained about most frequently, Canadian (33/64, 52%) and British patients expressed the greatest dissatisfaction with *prerequisites* (34,828/69,746/, 49.94%), as compared with those of other nationalities. Referring to Canadian patient experiences of dental practice, dentists' professional competence caused the greatest dissatisfaction (20/64, 31%), whereas the grievances of British patients to prerequisites were largely related to patients' discontent over the attributes of the patient-centered professionals, which contributed to 42.74% (29,810 /69,746) of all complaint issues reviewed from British patients. Patients based in the United States attached significant importance to *patient-centered processes* (76,349/97,937, 77.96%), especially relating to information, communication, and education (41,133/97,937, 42.00%). German patients expressed negative comments to *systemic problems* (14,337/31,095), accounting for 46.11% of all complaints identified from German patients. In particular, poor therapeutic environments led to the most-complained-about topic. Chinese patients' complaints relating to *expected outcomes* represented 9.82% (542/5521) of the total complaints made, far exceeding the sample average (3651/204,363, 1.79%).

XSL•FO
RenderX

Figure 3. Distribution of 4 domains of online health care complaints across different countries.



In addition to the aforementioned observations, it was identified that UK patients showed special features. First, they expressed dissatisfaction with the dimensions of care that patients in other countries did not complain about. Such dimensions included *family and friends supported as caregivers*, *needs are anticipated*, and *transition and continuity of care*. Second, UK patients expressed far-less-than-average complaints about certain dimensions of care than those in other countries, such as *financial accessibility* (0% vs 1.45%) and *information, communication, and education* (1.94% vs 22.80%). The proportion of complaints from Chinese patients to experts' professional competence and physical comfort was significantly higher than that of patients from other countries.

Discussion

Principal Findings

The results of this systematic review show that there is not yet a widely adopted taxonomy framework for classifying patient complaints made online in various settings. This means that studies have defined and classified complaints differently, resulting in limited comparability between available studies and research contexts [24,25,57]. To eliminate this gap, we have developed an a priori framework of PCC, which can incorporate patient complaints online. The NHS Institute for Innovation and Improvement found that PCC frameworks (eg, Picker framework) are broadly appropriate for "what matters most" to patients [58]. We have further validated the scope of the PCC frameworks using data from patient complaints online. Our research is considered beneficial for identifying gaps in the evidence base for patient experience expressed online, which has been identified as *domains of PCC*.

From the analysis of investigated studies, we have identified that the generic themes of *prerequisites*, *the care environment*, and *the processes* were complained about the most or, put it another way, greatly valued by the majority of patients, regardless of where they came from, and should be deemed priorities for PCC. To be more specific, a practicing clinician should be trained more on these dimensions: attributes of patient-centered professional, information, communication, and education.

Although involvement of family and friends is increasingly viewed as an important component of PCC [1,59-61], patients seldom expressed dissatisfaction with it in our research; apparently far less complained about the low participation of themselves. Compared with technical competence, which constituted a fundamental aspect of health care provision [23], interpersonal attributes of professionals were much more likely to be evaluated by patients; the potential implications of this are twofold: first, consistent with the study by Jia Li et al [26], patients' needs have different hierarchy, which were firstly stated by Maslow [62], and second, patients have uninformed expectation—"patients are not capable or are reluctant to communicate their expectations" [61], which was validated by Rothenfluh and Schulz [63]. Although there were common concerns across different countries' patients, variation existed in this study as well. Overall, 5 countries included in our study have different health care systems in terms of health care insurance, drug pricing, physician compensation, etc, and lead to disparities in quality of care, care coordination, and physician education: for example, health care in countries such as the United Kingdom and Canada is publicly financed and the coverage is universal; however, the health coverage of America remains fragmented, with numerous private and public sources, as well as wide gaps in insured rates across the US population [64].

In general, UK patients contributed most of the complaints about the attributes of patient-centered professionals. This finding correlates with that of the NHS Institute for Innovation and Improvement [58]. Besides, British patients had a higher pursuit

XSL•FO
RenderX

as they voiced high-level needs such as *involvement of family and friends* (591 issues), *anticipated needs* (462 issues), *feedback mechanisms to measure patient experience* (37 issues), and *emotional support* (642 issues). British patients were also the only nationality to convey discontent over these aspects of care. Given that the NHS has been collecting data on patients' experience of care for over 10 years [58] and professionals are occasionally accused of being incompetent in satisfying patients' needs to be treated as a person, eliminating the gap between knowing and doing is crucial. From another perspective, UK patients are assumed to be more *informed* about PCC by virtue of a variety of regular national health and social care surveys carried out in the United Kingdom, with frameworks of several surveys adopting Picker's PCC principles being available [58,65]; however, this hypothesis awaits further confirmation.

Despite nearly none of the British patients in this study expressing discontent over financial accessibility (0/69,746, 0.00%) and physical outcomes (376/69,746, 0.54%), approximately 10% (542/5521, 9.82%; 441/5521, 7.99%) of Chinese patients' complaints referred to these issues. Although the coverage by publicly financed health insurance in China is near-universal, out-of-pocket spending per capita represented approximately 32% of total health expenditures in 2014 [64]. Aside from this, high registration fees, the formidable markup by ticket touts operating in health care locations [9], fees related to excessive tests and treatments [26,66,67], and insurance reimbursement obstacles (cross-regional medical treatment) [67] constitute Chinese patients' financial barriers. It is likely that physician-dominated decision-making [68], inadequate communication, and patient distrust of doctors [66,68,69] have led to dissatisfaction with their physical outcomes.

Looking into the aspects of care that deeply concerned German patients, we find it necessary to improve the therapeutic environment of health care provision, in terms of privacy and entertainment, as well as maintaining a continuous healing relationship through telecommunication and house visits [44]. Given that studies on patient experience in Germany have not taken these vulnerable care aspects into account [70-72], these dissatisfaction factors should be tested in future research.

In this study, patients based in the United States conveyed great dissatisfaction for information, communication, and education (41,133/97,937, 42.00%), as well as coordination and integration of care (33,520/97,937, 34.23%). Previous work on patient-centered communication demonstrated a positive correlation between skilled physician communication and patient satisfaction [73-75]. Utilizing patient-centered communication guidelines and codes of conduct, such as physicians' humility and communication training for physicians and medical students, may bring better patient experience and diminish patient complaints [70,73-76]. As for coordination and integration of care, as necessitated when patients encounter long waiting times in hospitals or disorganized operations, previous studies have focused on hospital-level care coordination strategies associated with better patient experience [77]. Besides, it is envisaged that information technologies can reduce the need to craft laborious, case-by-case strategies for coordinating patient care [30,78,79].

Our findings suggest that it is feasible to identify gaps in evidence bases for patient experiences, which have been identified as *domains of PCC*. It was observed that differences and commonalities coexist across countries, after applying the proposed taxonomy framework, and we found that there is much leeway for the countries of interest to seek improvement in patient-centeredness.

Conclusions

Patient complaints online can indicate weaknesses in the health care system through the eyes of the patients' themselves. The proposed PCC framework can be applied to analyze the complaints under a wide PCC range of conditions, treatments, and countries. This review has shown significant heterogeneity of patients' online complaints in different countries, attributable to the diversity in culture, health care institution, and health literacy. Further work is required to apply the framework, using a plethora of data sources, to compare with other services, organizations, and countries or within the health care service over time, that is, a longitudinal study.

All RQs, proposed in the Introduction of this paper, were answered through conduction of the systematic review. Despite certain studies classifying patient complaints online, none were found to include or adopt a credible taxonomy framework. The proposed PCC framework aligns with what patients currently complain about online. By applying the taxonomy, results show that health professionals' skills and knowledge, open communication of knowledge, and system issues of PCC constitute the focus of online complaints made by patients. In addition, the differences in patient complaints in a multicountry context are discussed.

Limitations

As always, there are several limitations to this study. First, it was based on searches in merely 3 databases and focused only on currently available peer-reviewed literature; for this reason, we may have missed information in the gray literature. Second, regarding the small number of included articles, because of artificial screening, and the resulting relatively small sample size, our conclusions, especially the country-specific ones, may not be free of overgeneralizations and missing targets, even by taking into account the "community of common destiny." Finally, our interpretation of the concepts and scope of the various categories of complaints included in the article may not be fully consistent with the authors of the included papers, especially if the explanations or quotes were not given. Considering the limits to the time, space and researcher resources of this study, it is, nonetheless, a worthy trial that merits further exploration.

Acknowledgments

This work was supported by *Fundamental Research Funds for the Central Universities*, HUST: 2015AE017.

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1

Search strategy.

[PDF File (Adobe PDF File), 57KB - jmir_v21i8e14634_app1.pdf]

References

1. Drevs F, Hinz V. Who chooses, who uses, who rates: the impact of agency on electronic word-of-mouth about hospitals stays. Health Care Manage Rev 2014;39(3):223-233. [doi: 10.1097/HMR.0b013e3182993b6a] [Medline: 23727786]
2. Griffiths A, Leaver MP. Wisdom of patients: predicting the quality of care using aggregated patient feedback. BMJ Qual Saf 2018 Dec;27(2):110-118 [FREE Full text] [doi: 10.1136/bmjqs-2017-006847] [Medline: 28971881]
3. Hopkins B, Yamaguchi JT, Cloney MB, Shlobin NA, Dahdaleh NS. Effects of the physician payments sunshine act on the patient experience and perception of care amongst neurosurgeons: a comparative study of online PRW ratings and industry payments. Clin Neurol Neurosurg 2019 Jan;176:127-132. [doi: 10.1016/j.clineuro.2018.12.008] [Medline: 30557766]
4. Rozenblum R, Greaves F, Bates DW. The role of social media around patient experience and engagement. BMJ Qual Saf 2017 Dec;26(10):845-848. [doi: 10.1136/bmjqs-2017-006457] [Medline: 28428244]
5. Greaves F, Ramirez-Cano D, Millett C, Darzi A, Donaldson L. Harnessing the cloud of patient experience: using social media to detect poor quality healthcare. BMJ Qual Saf 2013 Mar;22(3):251-255. [doi: 10.1136/bmjqs-2012-001527] [Medline: 23349387]
6. Burkle CM, Keegan MT. Popularity of internet physician rating sites and their apparent influence on patients' choices of physicians. BMC Health Serv Res 2015 Sep 26;15:416 [FREE Full text] [doi: 10.1186/s12913-015-1099-2] [Medline: 26410383]
7. Hanauer DA, Zheng K, Singer DC, Gebremariam A, Davis MM. Parental awareness and use of online physician rating sites. Pediatrics 2014 Oct;134(4):e966-e975 [FREE Full text] [doi: 10.1542/peds.2014-0681] [Medline: 25246629]
8. Lagu T, Goff SL, Hannon NS, Shatz A, Lindenauer PK. A mixed-methods analysis of patient reviews of hospital care in England: implications for public reporting of health care quality data in the United States. Jt Comm J Qual Patient Saf 2013 Jan;39(1):7-15. [doi: 10.1016/S1553-7250(13)39003-5] [Medline: 23367647]
9. Zhang W, Deng Z, Hong Z, Evans R, Ma J, Zhang H. Unhappy patients are not alike: content analysis of the negative comments from China's good doctor website. J Med Internet Res 2018 Dec 25;20(1):e35 [FREE Full text] [doi: 10.2196/jmir.8223] [Medline: 29371176]
10. Lagu T, Hannon NS, Rothberg MB, Lindenauer PK. Patients' evaluations of health care providers in the era of social networking: an analysis of physician-rating websites. J Gen Intern Med 2010 Sep;25(9):942-946 [FREE Full text] [doi: 10.1007/s11606-010-1383-0] [Medline: 20464523]
11. Asanad K, Parameshwar PS, Houman J, Spiegel BM, Daskivich TJ, Anger JT. Online physician reviews in female pelvic medicine and reconstructive surgery: what do patients really want? Female Pelvic Med Reconstr Surg 2018;24(2):109-114. [doi: 10.1097/SPV.0000000000000503] [Medline: 29474282]
12. Bardach NS, Lyndon A, Asteria-Peñaloza R, Goldman LE, Lin GA, Dudley RA. From the closest observers of patient care: a thematic analysis of online narrative reviews of hospitals. BMJ Qual Saf 2016 Dec;25(11):889-897 [FREE Full text] [doi: 10.1136/bmjqs-2015-004515] [Medline: 26677215]
13. Wallace BC, Paul MJ, Sarkar U, Trikalinos TA, Dredze M. A large-scale quantitative analysis of latent factors and sentiment in online doctor reviews. J Am Med Inform Assoc 2014;21(6):1098-1103 [FREE Full text] [doi: 10.1136/amiajnl-2014-002711] [Medline: 24918109]
14. Rastegar-Mojarad M, Ye Z, Wall D, Murali N, Lin S. Collecting and analyzing patient experiences of health care from social media. JMIR Res Protoc 2015 Jul 2;4(3):e78 [FREE Full text] [doi: 10.2196/resprot.3433] [Medline: 26137885]
15. Greaves F, Ramirez-Cano D, Millett C, Darzi A, Donaldson L. Use of sentiment analysis for capturing patient experience from free-text comments posted online. J Med Internet Res 2013 Nov 1;15(11):e239 [FREE Full text] [doi: 10.2196/jmir.2721] [Medline: 24184993]
16. Noll D, Mahon B, Shroff B, Carrico C, Lindauer SJ. Twitter analysis of the orthodontic patient experience with braces vs invisalign. Angle Orthod 2017 May;87(3):377-383. [doi: 10.2319/062816-508.1] [Medline: 28059576]
17. Rozenblum R, Bates DW. Patient-centred healthcare, social media and the internet: the perfect storm? BMJ Qual Saf 2013 Mar;22(3):183-186. [doi: 10.1136/bmjqs-2012-001744] [Medline: 23378660]
18. Manary MP, Boulding W, Staelin R, Glickman SW. The patient experience and health outcomes. N Engl J Med 2013 Jan 17;368(3):201-203. [doi: 10.1056/NEJMp1211775] [Medline: 23268647]
19. Aiken LH, Sermeus W, van den Heede K, Sloane DM, Busse R, McKee M, et al. Patient safety, satisfaction, and quality of hospital care: cross sectional surveys of nurses and patients in 12 countries in Europe and the United States. Br Med J 2012 Mar 20;344:e1717 [FREE Full text] [doi: 10.1136/bmj.e1717] [Medline: 22434089]

XSL•FO
RenderX

20. GOV.UK. 2008. High Quality Care for All: NHS Next Stage Review Final Report URL: https://www.gov.uk/government/publications/high-quality-care-for-all-nhs-next-stage-review-final-report [accessed 2018-09-15]
21. Greaves F, Millett C, Nuki P. England's experience incorporating 'anecdotal' reports from consumers into their national reporting system: lessons for the United States of what to do or not to do? Med Care Res Rev 2014 Oct;71(5 Suppl):65S-80S. [doi: 10.1177/1077558714535470] [Medline: 24836765]
22. Asprey A, Campbell JL, Newbould J, Cohn S, Carter M, Davey A, et al. Challenges to the credibility of patient feedback in primary healthcare settings: a qualitative study. Br J Gen Pract 2013 Mar;63(608):e200-e208 [FREE Full text] [doi: 10.3399/bjgp13X664252] [Medline: 23561787]
23. Brookes G, Baker P. What does patient feedback reveal about the NHS? A mixed methods study of comments posted to the NHS choices online service. BMJ Open 2017 Dec 27;7(4):e013821 [FREE Full text] [doi: 10.1136/bmjopen-2016-013821] [Medline: 28450463]
24. Harrison R, Walton M, Healy J, Smith-Merry J, Hobbs C. Patient complaints about hospital services: applying a complaint taxonomy to analyse and respond to complaints. Int J Qual Health Care 2016 Apr;28(2):240-245. [doi: 10.1093/intqhc/mzw003] [Medline: 26826722]
25. Reader TW, Gillespie A, Roberts J. Patient complaints in healthcare systems: a systematic review and coding taxonomy. BMJ Qual Saf 2014 Aug;23(8):678-689 [FREE Full text] [doi: 10.1136/bmjqs-2013-002437] [Medline: 24876289]
26. Li J, Liu M, Li X, Liu X, Liu J. Developing embedded taxonomy and mining patients' interests from web-based physician reviews: mixed-methods approach. J Med Internet Res 2018 Dec 16;20(8):e254 [FREE Full text] [doi: 10.2196/jmir.8868] [Medline: 30115610]
27. Giuseppe P. Picker: Patient Centred Care Improvement. 2016. Person-centred care in Europe: a cross-country comparison of health system performance, strategies and structures URL: https://www.picker.org/wp-content/uploads/2016/02/12-02-16-Policy-briefing-on-patient-centred-care-in-Europe.pdf
28. Redding D, Hutchinson S. National Voices. 2017. Person-Centred Care in 2017 URL: https://www.nationalvoices.org.uk/publications/our-publications/person-centred-care-2017 [accessed 2019-04-10]
29. Booth A, Carroll C. How to build up the actionable knowledge base: the role of 'best fit' framework synthesis for studies of improvement in healthcare. BMJ Qual Saf 2015 Nov;24(11):700-708 [FREE Full text] [doi: 10.1136/bmjqs-2014-003642] [Medline: 26306609]
30. 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.
31. Ministry of Solidarity and Health. 2017. [The 2018-2022 National Health Strategy] URL: https://solidarites-sante.gouv.fr/systeme-de-sante-et-medico-social/strategie-nationale-de-sante/article/la-strategie-nationale-de-sante-2018-2022 [accessed 2018-11-18]
32. NHS England. Creating the Conditions for Person-Centred Care URL: https://www.england.nhs.uk/ourwork/patient-participation/patient-centred/ [accessed 2018-11-18]
33. McCormack B, McCance TV. Development of a framework for person-centred nursing. J Adv Nurs 2006 Dec;56(5):472-479. [doi: 10.1111/j.1365-2648.2006.04042.x] [Medline: 17078823]
34. Donabedian A. The quality of care. How can it be assessed? J Am Med Assoc 1988;260(12):1743-1748. [doi: 10.1001/jama.1988.03410120089033] [Medline: 3045356]
35. Kitson A, Marshall A, Bassett K, Zeitz K. What are the core elements of patient-centred care? A narrative review and synthesis of the literature from health policy, medicine and nursing. J Adv Nurs 2013 Jan;69(1):4-15. [doi: 10.1111/j.1365-2648.2012.06064.x] [Medline: 22709336]
36. Rathert C, Wyrwich MD, Boren SA. Patient-centered care and outcomes: a systematic review of the literature. Med Care Res Rev 2013 Aug;70(4):351-379. [doi: 10.1177/1077558712465774] [Medline: 23169897]
37. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
38. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009 Jul 21;6(7):e1000097 [FREE Full text] [doi: 10.1371/journal.pmed.1000097] [Medline: 19621072]
39. Hawkins JB, Brownstein JS, Tuli G, Runels T, Broecker K, Nsoesie EO, et al. Measuring patient-perceived quality of care in US hospitals using Twitter. BMJ Qual Saf 2016 Dec;25(6):404-413 [FREE Full text] [doi: 10.1136/bmjqs-2015-004309] [Medline: 26464518]
40. Wagland R, Recio-Saucedo A, Simon M, Bracher M, Hunt K, Foster C, et al. Development and testing of a text-mining approach to analyse patients' comments on their experiences of colorectal cancer care. BMJ Qual Saf 2016 Dec;25(8):604-614. [doi: 10.1136/bmjqs-2015-004063] [Medline: 26512131]
41. Kowalski R. Patients' Written Reviews as a Resource for Public Healthcare Management in England. 2017 Presented at: EICN'17; September 18-20, 2017; Lund, Sweden p. 545-550. [doi: 10.1016/j.procs.2017.08.275]
42. Corner J, Wagland R, Glaser A, Richards SM. Qualitative analysis of patients' feedback from a PROMs survey of cancer patients in England. BMJ Open 2013 Apr 10;3(4):e002316 [FREE Full text] [doi: 10.1136/bmjopen-2012-002316] [Medline: 23578681]

XSL•FO
RenderX

43. Paul MJ, Wallace BC, Dredze M. What Affects Patient (Dis)satisfaction? Analyzing Online Doctor Ratings With a Joint Topic-Sentiment Model. In: Expanding the Boundaries of Health Informatics Using AI. 2013 Presented at: AAAI'13; July 14-15, 2013; Bellevue, Washington.
44. Emmert M, Meier F, Heider AK, Dürr C, Sander U. What do patients say about their physicians? An analysis of 3000 narrative comments posted on a German physician rating website. Health Policy 2014 Oct;118(1):66-73. [doi: 10.1016/j.healthpol.2014.04.015] [Medline: 24836021]
45. Alemi F, Torii M, Clementz L, Aron DC. Feasibility of real-time satisfaction surveys through automated analysis of patients' unstructured comments and sentiments. Qual Manag Health Care 2012;21(1):9-19. [doi: 10.1097/QMH.0b013e3182417fc4] [Medline: 22207014]
46. López A, Detz A, Ratanawongsa N, Sarkar U. What patients say about their doctors online: a qualitative content analysis. J Gen Intern Med 2012 Jun;27(6):685-692 [FREE Full text] [doi: 10.1007/s11606-011-1958-4] [Medline: 22215270]
47. Detz A, López A, Sarkar U. Long-term doctor-patient relationships: patient perspective from online reviews. J Med Internet Res 2013 Jul;15(7):e131 [FREE Full text] [doi: 10.2196/jmir.2552] [Medline: 23819959]
48. Greaves F, Laverty AA, Cano DR, Moilanen K, Pulman S, Darzi A, et al. Tweets about hospital quality: a mixed methods study. BMJ Qual Saf 2014 Oct;23(10):838-846 [FREE Full text] [doi: 10.1136/bmjqs-2014-002875] [Medline: 24748372]
49. Macdonald ME, Beaudin A, Pineda C. What do patients think about dental services in Quebec? Analysis of a dentist rating website. J Can Dent Assoc 2015;81:f3 [FREE Full text] [Medline: 26030597]
50. Lagu T, Goff SL, Craft B, Calcasola S, Benjamin EM, Priya A, et al. Can social media be used as a hospital quality improvement tool? J Hosp Med 2016 Jan;11(1):52-55 [FREE Full text] [doi: 10.1002/jhm.2486] [Medline: 26390277]
51. Trehan SK, DeFrancesco CJ, Nguyen JT, Charalel RA, Daluiski A. Online Patient Ratings of Hand Surgeons. J Hand Surg 2016;41:98-103. [doi: 10.1016/j.jhsa.2015.10.006] [Medline: 26710742]
52. Cunningham M, Wells M. Qualitative analysis of 6961 free-text comments from the first National Cancer Patient Experience Survey in Scotland. BMJ Open 2017 Dec 15;7(6):e015726 [FREE Full text] [doi: 10.1136/bmjopen-2016-015726] [Medline: 28619780]
53. James TL, Calderon ED, Cook DF. Exploring patient perceptions of healthcare service quality through analysis of unstructured feedback. Expert Syst Appl 2017 Apr;71:479-492. [doi: 10.1016/j.eswa.2016.11.004]
54. Xu S, Atanelov Z, Bhatia AC. Online patient-reported reviews of Mohs micrographic surgery: qualitative analysis of positive and negative experiences. Cutis 2017 Feb;99(2):E25-E29. [Medline: 28319637]
55. King RH, Stanley J, Baum N. Hard internet truths: 34,748 online reviews reveal what patients really want from doctors. J Med Pract Manage 2016;31(5):309-312. [Medline: 27249884]
56. Emmert M, Meszmer N, Schlesinger M. A cross-sectional study assessing the association between online ratings and clinical quality of care measures for US hospitals: results from an observational study. BMC Health Serv Res 2018 Dec 5;18(1):82 [FREE Full text] [doi: 10.1186/s12913-018-2886-3] [Medline: 29402321]
57. Runciman W, Hibbert P, Thomson R, van der Schaaf T, Sherman H, Lewalle P. Towards an international classification for patient safety: key concepts and terms. Int J Qual Health Care 2009 Feb;21(1):18-26 [FREE Full text] [doi: 10.1093/intqhc/mzn057] [Medline: 19147597]
58. Robert G, Cornwell J. Wales NHS. 2011 Nov. What Matters to Patients? Developing the Evidence Base for Measuring and Improving Patient Experience URL: http://www.wales.nhs.uk/sites3/documents/420/Final%20Project%20Report%20pdf%20doc%20january%202012%20(2).pdf [accessed 2018-10-09]
59. National Center for Interprofessional Practice and Education. 2015. Picker Institute's Eight Principles of Patient-Centered Care URL: https://nexusipe.org/informing/resource-center/picker-institute%E2%80%99s-eight-principles-patient-centered-care [accessed 2018-12-11]
60. Gerteis M, Delbanco TL, Daley J, Edgman-Levitan S. Through the Patient's Eyes: Understanding and Promoting Patient-Centered Care. San Francisco, CA: Jossey-Bass; 1993.
61. Sofaer S, Firminger K. Patient perceptions of the quality of health services. Annu Rev Public Health 2005;26:513-559. [doi: 10.1146/annurev.publhealth.25.050503.153958] [Medline: 15760300]
62. Maslow AH. A theory of human motivation. Psychol Rev 1943;50(4):370-396. [doi: 10.1037/h0054346]
63. Rothenfluh F, Schulz PJ. Physician rating websites: what aspects are important to identify a good doctor, and are patients capable of assessing them? A mixed-methods approach including physicians' and health care consumers' perspectives. J Med Internet Res 2017 Dec 1;19(5):e127 [FREE Full text] [doi: 10.2196/jmir.6875] [Medline: 28461285]
64. International Health Care System Profiles. Country Profiles URL: https://international.commonwealthfund.org/countries/ [accessed 2019-05-30]
65. Coulter A, Locock L, Ziebland S, Calabrese J. Collecting data on patient experience is not enough: they must be used to improve care. Br Med J 2014 Mar 26;348:g2225. [doi: 10.1136/bmj.g2225] [Medline: 24671966]
66. Wang X, Jiang R, Li J, Chen J, Burström B, Burström K. What do patients care most about in China's public hospitals? Interviews with patients in Jiangsu Province. BMC Health Serv Res 2018 Dec 8;18(1):97 [FREE Full text] [doi: 10.1186/s12913-018-2903-6] [Medline: 29422095]
67. Gong X. The health care gap between China and America. Ann Transl Med 2014 Apr;2(4):39 [FREE Full text] [doi: 10.3978/j.issn.2305-5839.2014.04.04] [Medline: 25333015]

XSL•FO
RenderX

68. Ting X, Yong B, Yin L, Mi T. Patient perception and the barriers to practicing patient-centered communication: a survey and in-depth interview of Chinese patients and physicians. Patient Educ Couns 2016 Mar;99(3):364-369. [doi: 10.1016/j.pec.2015.07.019] [Medline: 26776708]
69. Hung LM, Shi L, Wang H, Nie X, Meng Q. Chinese primary care providers and motivating factors on performance. Fam Pract 2013 Oct;30(5):576-586. [doi: 10.1093/fampra/cmt026] [Medline: 23788201]
70. Breuning M, Lucius-Hoene G, Burbaum C, Himmel W, Bengel J. [Patient experiences and patient centeredness: the website project DIPEx Germany]. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2017 Apr;60(4):453-461. [doi: 10.1007/s00103-017-2524-y] [Medline: 28251275]
71. Kraska RA, Weigand M, Geraedts M. Associations between hospital characteristics and patient satisfaction in Germany. Health Expect 2017 Dec;20(4):593-600 [FREE Full text] [doi: 10.1111/hex.12485] [Medline: 27447595]
72. Schoenfelder T, Klewer J, Kugler J. Determinants of patient satisfaction: a study among 39 hospitals in an in-patient setting in Germany. Int J Qual Health Care 2011 Oct;23(5):503-509. [doi: 10.1093/intqhc/mzr038] [Medline: 21715557]
73. King A, Hoppe RB. 'Best practice' for patient-centered communication: a narrative review. J Grad Med Educ 2013 Sep;5(3):385-393 [FREE Full text] [doi: 10.4300/JGME-D-13-00072.1] [Medline: 24404300]
74. Brueck MK, Salib AM. Strategies for acing the fundamentals and mitigating legal and ethical consequences of poor physician-patient communication. AMA J Ethics 2017 Mar 1;19(3):289-295 [FREE Full text] [doi: 10.1001/journalofethics.2017.19.3.hlaw1-1703] [Medline: 28323610]
75. Ruberton PM, Huynh HP, Miller TA, Kruse E, Chancellor J, Lyubomirsky S. The relationship between physician humility, physician-patient communication, and patient health. Patient Educ Couns 2016 Dec;99(7):1138-1145. [doi: 10.1016/j.pec.2016.01.012] [Medline: 26830544]
76. Boissy A, Windover AK, Bokar D, Karafa M, Neuendorf K, Frankel RM, et al. Communication skills training for physicians improves patient satisfaction. J Gen Intern Med 2016 Dec;31(7):755-761 [FREE Full text] [doi: 10.1007/s11606-016-3597-2] [Medline: 26921153]
77. Figueroa JF, Feyman Y, Zhou X, Joynt Maddox K. Hospital-level care coordination strategies associated with better patient experience. BMJ Qual Saf 2018 Oct;27(10):844-851. [doi: 10.1136/bmjqs-2017-007597] [Medline: 29618639]
78. McKay C, Vanaskie K. Partnering for success: the role of the nurse leader in health information technology implementation for coordination of care. Nurse Lead 2018 Dec;16(6):385-388. [doi: 10.1016/j.mnl.2018.07.012]
79. Chang L, Wanner KJ, Kovalsky D, Smith KL, Rhodes KV. 'It's really overwhelming': patient perspectives on care coordination. J Am Board Fam Med 2018;31(5):682-690 [FREE Full text] [doi: 10.3122/jabfm.2018.05.180034] [Medline: 30201664]

Abbreviations

NHS: National Health Service
PCC: patient-centered care
PRW: physician rating website
RQ: research question

Edited by G Eysenbach; submitted 07.05.19; peer-reviewed by S Lin, R Druz; comments to author 29.05.19; revised version received 05.06.19; accepted 09.06.19; published 07.08.19.

Please cite as:
Liu J, Hou S, Evans R, Xia C, Xia W, Ma J
What Do Patients Complain About Online: A Systematic Review and Taxonomy Framework Based on Patient Centeredness
J Med Internet Res 2019;21(8):e14634
URL: https://www.jmir.org/2019/8/e14634/
doi:10.2196/14634
PMID:31392961

©Jing Liu, Shengchao Hou, Richard Evans, Chenxi Xia, Weidong Xia, Jingdong Ma. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 07.08.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Association of Social Media Presence with Online Physician Ratings and Surgical Volume Among California Urologists: Observational Study

Justin Houman[1*], MD; James Weinberger[2*], BS; Ashley Caron[1*], BS; Alex Hannemann[1*], BS; Michael Zaliznyak[1*], BA; Devin Patel[1*], MD; Ariel Moradzadeh[1*], MD; Timothy J Daskivich[1*], MD

[1]Cedars-Sinai Medical Center, Los Angeles, CA, United States

[2]David Geffen School of Medicine, University of California Los Angeles, Los Angeles, CA, United States

[*]all authors contributed equally

Corresponding Author:
Justin Houman, MD
Cedars-Sinai Medical Center
8700 Beverly Blvd
Los Angeles, CA, 90048
United States
Phone: 1 7149287950
Email: justin.houman@cshs.org

Abstract

Background: Urologists are increasingly using various forms of social media to promote their professional practice and attract patients. Currently, the association of social media on a urologists' practice is unknown.

Objectives: We aimed to determine whether social media presence is associated with higher online physician ratings and surgical volume among California urologists.

Methods: We sampled 195 California urologists who were rated on the ProPublica Surgeon Scorecard website. We obtained information on professional use of online social media (Facebook, Instagram, Twitter, blog, and YouTube) in 2014 and defined social media presence as a binary variable (yes/no) for use of an individual platform or any platform. We collected data on online physician ratings across websites (Yelp, Healthgrades, Vitals, RateMD, and UCompareHealthcare) and calculated the mean physician ratings across all websites as an average weighted by the number of reviews. We then collected data on surgical volume for radical prostatectomy from the ProPublica Surgeon Scorecard website. We used multivariable linear regression to determine the association of social media presence with physician ratings and surgical volume.

Results: Among our sample of 195 urologists, 62 (32%) were active on some form of social media. Social media presence on any platform was associated with a slightly higher mean physician rating (β coefficient: .3; 95% CI 0.03-0.5; *P*=.05). However, only YouTube was associated with higher physician ratings (β coefficient: .3; 95% CI 0.2-0.5; *P*=.04). Social media presence on YouTube was strongly associated with increased radical prostatectomy volume (β coefficient: 7.4; 95% CI 0.3-14.5; *P*=.04). Social media presence on any platform was associated with increased radical prostatectomy volume (β coefficient: 7.1; 95% CI –0.7 to 14.2; *P*=.05).

Conclusions: Urologists' use of social media, especially YouTube, is associated with a modest increase in physician ratings and prostatectomy volume. Although a majority of urologists are not currently active on social media, patients may be more inclined to endorse and choose subspecialist urologists who post videos of their surgical technique.

(J Med Internet Res 2019;21(8):e10195) doi:10.2196/10195

KEYWORDS

social media; surgical volume; physician ratings

XSL•FO
RenderX

Introduction

Social media use is becoming increasingly common among both health care consumers and urologists. A recent Pew Research Center study showed that the number of US internet users active on social media has increased from 8% in 2005 to 74% in 2014 [1]. Urologists are also part of this trend; in 2014, more than 70% of urologists were reported to be active on some form of social media [2,3]. Urologists currently use social media for a variety of reasons, including discussing patient cases, sharing patient education materials, creating forums to discuss journal articles, and connecting attendees at large academic conferences [4,5]. Beside these reasons, there are numerous academic advantages to social media use to expand professional networks, create new opportunities for academic collaboration, and increase citation potential for papers [6]. However, another presumed intent of social media among physicians is improvement in business productivity in terms of case volume and public perception. Physicians who are active on social media are undoubtedly trying to promote themselves or their practice. Steinmetz et al [7] highlighted how various forms of advertising, such as social media, can improve sales, profits, and reputation of a business.

Despite this, to date, there has been no published literature on the impact of social media on the public perception of providers or business productivity.

Although surgical volume is a clear indicator of business productivity, public perception is a more intangible concept that is difficult to measure, especially given the lack of published literature in this area. However, online physician ratings may be a reasonable proxy for the public's perception of a physician, based on the scope of use and impact of online ratings on health care consumers' behavior. This is evidenced by the fact that in the United States, 47% of patients performed online searches of their physicians in 2010 [8], and a recent population-based analysis of 600 physicians showed a median of seven reviews per provider [9]. In 2010, one in six practicing US physicians had received an online review [10]. These ratings are frequently used as a key source of information through which patients choose a physician. This is further supported by a survey of 1000 surgical patients at the Mayo clinic, which found that 81% of patients would seek consultation from a physician based on positive reviews alone, and 77% would not seek consultation from a physician based solely on negative reviews [11]. Similar data have been reported in Europe [12]. Because of the strong association of online ratings with health care consumers' choices, it may be reasonable to use online physician ratings as a proxy measure of reputation, if this is defined as the likelihood that a patient will choose to consult with a given practice based on community opinions.

In this study, we aimed to determine whether social media presence among urologists impacts their reputation (vis-à-vis their online consumer rating) and surgical volume. To address this question, we sampled 195 California urologists rated on the ProPublica Surgeon Scorecard website to determine whether professional use of social media platforms (Facebook, Instagram, Twitter, blog, and YouTube) was associated with average numeric physician rating across five popular websites and radical prostatectomy surgical volume in 2014 as determined by the Medicare Physician and Other Supplier Public Use File. Recognizing the potential for confounding by institutional branding and practice setting, we corrected for whether the urologist was affiliated with an academic or private practice and performed a subgroup analysis to determine if the effect size was consistent.

We hypothesized that physicians with a more active social media presence would have higher online physicians ratings and surgical volume.

Methods

Data Source and Participants

The Cedars-Sinai Institutional Review Board provided an exemption certification for this study (IRB #00050328)**.** This study was conducted in accordance with all relevant guidelines and procedures of Cedars-Sinai Medical Center. We sampled all California urologists rated on the ProPublica Surgeon Scorecard website (n=195). The ProPublica Surgeon Scorecard website was used because it provided the most comprehensive list of currently practicing urologist. These urologists completed at least 20 radical prostatectomy or transurethral resection of the prostate procedures in the calendar year 2014 according to Medicare claims data [13]. Physicians were excluded from the online review portion of the analysis if they had no online reviews (n=12). Physicians were excluded from the surgical volume portion of the analysis if they performed less than 20 radical prostatectomies in the calendar year 2014 (n=110).

Variables

Primary Predictor

The primary predictor was physician social media presence. One of the investigators (JH) collected data on physicians' professional social media presence in calendar year 2014 on five popular social media platforms: Facebook, Instagram, Twitter, YouTube, and professional blog. Social media presence was considered a binary variable (yes/no), defined as any social media posts promoting their medical practice. We coded social media presence both at the individual platform level and across all platforms. We also collected data on frequency of posts, but elected not to subdivide our sample based on this characteristic given the uniformity of frequency (90% posted information on social media once a month, and only 10% posted information with greater or lesser frequency).

Covariates

We gathered demographic data on physicians in our sample, including practice setting (academic or private), years since medical school graduation, and location of medical school (domestic or international) from the California Medical Board website.

Outcomes

Online Physician Ratings

We collected online ratings for each physician across the 5 most popular online physician-rating platforms according to Google

XSL•FO
RenderX

Trends: HealthGrades, Vitals, Yelp, RateMD, and UCompareHealth. Each of these websites asks consumers to rate physicians using a "5-star" scale. Using these data, we calculated each physician's "average rating" as a weighted average of scores across the five websites, weighted by the number of reviews on each website.

Physician Surgical Volume

We collected data on radical prostatectomy surgical volume in 2014 using Medicare claims from the Medicare Physician and Other Supplier Public Use File. We linked these data to other data sources using National Physician Identifier numbers. We defined radical prostatectomy as Current Procedural Terminology claims codes 55840, 55842, and 55845.

Statistical Analysis

We compared characteristics of our sample population by activity on social media, using the Chi-square test for categorical variables and the Wilcoxon-Mann-Whitney test for nonparametrically distributed dependent variables.

We used multivariable linear regression analysis to assess the association of physician social media presence with online physician ratings. Our primary predictor in these models was social media presence, and the outcome was average online physician rating across the five websites. Covariates included practice setting, years since medical school graduation, and location of medical school. We created separate models to analyze the impact of social media presence on any platform and individual social media platforms on online physician ratings. We also performed sensitivity analyses of our aggregate model in subgroups of academic and private physicians.

We assessed the association of physician social media presence on physician surgical volume in a similar fashion, using multivariable linear regression analysis and identical predictor and covariate structure. We created separate models to analyze the impact of social media presence on any platform and of individual social media platforms on radical prostatectomy volume. We also performed sensitivity analyses of our aggregate model in subgroups of academic and private physicians.

Results

Overview

Characteristics of our sample across those active on social media versus those not active were recorded. Of the 195 California urologists, 62 (32%) were active professionally on some form of social media in 2014, including 53 (27%) on YouTube, 15 (8%) on Facebook, 14 (7%) on Twitter, 10 (5%) on blogs, and 6 (3%) on Instagram. Of the total, 159 urologists were in the private practice setting and 36 were in the academic setting. The average number of years since medical school graduation was 32.6 years. In addition, 163 attended medical school in the United States and 32 attended foreign medical schools (Multimedia Appendix 1).

Association of Social Media Presence With Online Physician Ratings

Multivariable linear regression models predicting the weighted average of online physician ratings showed that social media presence on any platform was associated with a significantly higher mean physician rating (β coefficient: .3; 95% CI 0.03-0.5; *P*=.05; Multimedia Appendix 2). A similar magnitude and direction of the effect persisted among subgroups of private (β coefficient: .2; 95% CI –0.1 to 0.5; *P*=.2) and academic physicians (β coefficient: .6; 95% CI 0.15-1.0; *P*=.01) in sensitivity analyses. However, in multivariable models assessing the association of individual social media platforms and online ratings, only presence on YouTube was associated with a significantly higher mean physician rating (β coefficient: .3; 95% CI 0.2-0.5; *P*=.04). There were no meaningful or statistically significant differences in online physician ratings with regard to the use of other social media platforms.

Association of Social Media Presence With Physician Surgical Volume

Multivariable linear regression models predicting surgical volume showed that social media presence on any platform was associated with a trend toward higher annual radical prostatectomy volume (β coefficient: 7.1; 95% CI –0.7 to 14.2; *P*=.05; Multimedia Appendix 3). A similar magnitude of effect was observed among subgroups of private (β coefficient: 7.1; 95% CI –0.9 to 15.2; *P*=.08) and academic (β coefficient: 6.7; 95% CI –11.3 to 24.8; *P*=.4) physicians in sensitivity analyses. In multivariable models assessing the association of individual social media platforms with surgical volume, presence on YouTube was significantly associated with higher annual radical prostatectomy volume (β coefficient: 7.4; 95% CI 0.3-14.5; *P*=.04; Multimedia Appendix 3). There was no statistically significant difference in surgical volume with regard to the use of other social media platforms.

Discussion

Although most California urologists were not active on social media, we found that professional use of social media was associated with higher online physician ratings and increased prostatectomy volume.

Although physicians are increasingly using online social media to interact with their patients and promote their practices, to date, it is unclear whether this activity has any demonstrable effect on productivity outcomes. Our study suggests that professional activity on social media sites may positively impact both the physician's reputation as well as their surgical volume. We found that social media activity on any of the five social media outlets studied (and YouTube specifically in subgroup analysis) had a statistically significant association with online physician ratings, with an average increase of 0.3 over physicians who were not active on social media. Although a 0.3 increase on a 5-point scale seems small, it represents a difference of 0.4 SDs from the mean, indicating a meaningful difference. We also found that presence on any one of the five social media sites was associated with a trend toward increased prostatectomy volume, with a statistically significant increase

XSL•FO
RenderX

noted among those active on You Tube. A urologist who posts videos on social media is likely to perform roughly seven more radical prostatectomies per year than a urologist who does not post videos (representing an average increase of roughly 25% over the mean annual rate of 27 prostatectomies).

To our knowledge, this is the first study to directly test the association between social media use and clinical productivity outcomes. Other surgical specialties have speculated on the advantages of social media use in clinical efficiency and productivity afforded by social media outreach, but have not engaged in formal hypothesis testing. Orthopedic surgeons have postulated that directing patients to social media sites focused on patient education may enable more efficient and effective communications with their patients, reducing patient phone call volume and increasing clinical efficiency [14]. Vascular surgeons have suggested the utility of social media outreach in circumventing traditional physician referral patterns, providing an advantage in gaining market presence [15]. Despite clear interest in the association between social media use and real-world outcomes, there remains a lack of evidence-based testing of these associations. Our study, while limited due to its retrospective design, provides some evidence for the purported utility of professional use of social media.

Our finding that YouTube is the social media form with the strongest association with a physician's online reputation and surgical volume is consistent with recent data showing that health care consumers can accurately identify quality of surgery by watching online samples of a surgeon's technique. A recent study showed that medically trained reviewers were able to identify surgeons with higher complication rates by watching videos of their technique in the context of laparoscopic bariatric surgery [16]. Surprisingly, health care consumers (via crowdsourcing among the general population) were also able to identify surgeons with higher complication rates by watching videos of their operative technique in the context of robotic radical prostatectomy [17]. Based on this finding, patients may be justified in choosing surgeons who post videos online, since they appear to be able discern good from bad surgeons by viewing examples of their best work.

Professional use of social media is one of a number of sources of information that contribute to a physician's "online dossier," and both health care consumers and physicians should be aware of the relative worth of each component. The comprehensive online data available to health care consumers for choosing a physician include social media activity, online physician ratings, quality metrics, surgical volume, office ratings, and publicly available personal information. Although this study highlights the importance of social media presence in cultivating an online persona, the other components undoubtedly can (and should) contribute to a patient's overall perception of a physician. For example, we recently argued that online rating data should not be used as the sole criterion to select physicians by health care consumers (as data suggests they are), since they have no association with quality or value of care [18]. Similar to online ratings, social media presence should not be used by patients in isolation to select physicians. Ideally, physicians and patients will consider social media activity in the context of other sources of information that provide independent information about the physician, such as data that capture quality of care (surgical volume, quality metrics, and complexity of case volume), value of care, and the patient experience (online ratings).

This study has several limitations that may affect our findings. First, the association of social media presence with reputation and productivity outcomes may be confounded by academic institutional branding. However, the magnitude of association was virtually identical for subsets of private and academic physicians, which suggests that the findings are robust across different practice settings. Second, we are unable to rule out reverse causality as an explanation for our findings (ie, whether being a high-volume, highly rated surgeon is predictive of activity on social media). However, even if some degree of reverse causality exists, the policy implication is still the same: More surgeons should be posting videos online to prove the adequacy of their skills compared to their peers. Finally, the observational nature of this study may incur selection bias, since it only includes California urologists who are performing a minimum of 20 urologic procedures (transurethral resection of the prostate and radical prostatectomy) per year. Considerations for future prospective study design could include an interrupted time series or a randomized controlled study.

In conclusion, our findings suggest that urologists should consider being active on social media to promote and build their professional practice. Given the trajectory of use of online resources such as online ratings in selecting providers, we believe that the use of social media will become an increasingly important outreach tool for clinicians to interact with their patients in a meaningful way.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Sample characteristics (n=195).
[PDF File (Adobe PDF File), 23KB - jmir_v21i8e10195_app1.pdf]

Multimedia Appendix 2
Association between social media presence and online physician ratings.
[PDF File (Adobe PDF File), 19KB - jmir_v21i8e10195_app2.pdf]

XSL•FO
RenderX

Multimedia Appendix 3
Association between social media presence and prostatectomy volume.
[PDF File (Adobe PDF File), 18KB - jmir_v21i8e10195_app3.pdf]

References

1. Greenwood S, Perrin A, Duggan M. Pew Research Center: Internet & Technology. 2016. Social Media Update 2016 URL: https://www.pewinternet.org/2016/11/11/social-media-update-2016/ [accessed 2019-07-24]
2. Salem J, Borgmann H, Murphy DG. Integrating Social Media into Urologic Health care: What Can We Learn from Other Disciplines? Curr Urol Rep 2016 Feb;17(2):13. [doi: 10.1007/s11934-015-0570-2] [Medline: 26757907]
3. Loeb SB, Bayne CE, Frey C, Davies BJ, Averch TD, Woo HH, American Urological Association Social Media Work Group. Use of social media in urology: data from the American Urological Association (AUA). BJU Int 2014 Jun;113(6):993-998 [FREE Full text] [doi: 10.1111/bju.12586] [Medline: 24274744]
4. Sternberg K, Loeb SL, Canes D, Donnelly L, Tsai MH. The use of Twitter to facilitate sharing of clinical expertise in urology. J Am Med Inform Assoc 2018 Dec 01;25(2):183-186. [doi: 10.1093/jamia/ocx055] [Medline: 28591771]
5. Filson C, Eggener S. AUA University. 2016. Update Series (2016) Lesson 1: The Role of Social Media in Urology URL: https://auau.auanet.org/content/update-series-2016-lesson-1-role-social-media-urology [accessed 2019-07-24]
6. Leveridge M. The emerging role of social media in urology. Rev Urol 2014;16(3):110-117. [Medline: 25337040]
7. Steinmetz L. Critical stages of small business growth. Business Horizons 1969 Feb;12(1):29-36. [doi: 10.1016/0007-6813(69)90107-4]
8. Kadry B, Chu LF, Kadry B, Gammas D, Macario A. Analysis of 4999 online physician ratings indicates that most patients give physicians a favorable rating. J Med Internet Res 2011 Nov 16;13(4):e95 [FREE Full text] [doi: 10.2196/jmir.1960] [Medline: 22088924]
9. Lagu T, Metayer K, Moran M, Ortiz L, Priya A, Goff SL, et al. Website Characteristics and Physician Reviews on Commercial Physician-Rating Websites. JAMA 2017 Dec 21;317(7):766-768 [FREE Full text] [doi: 10.1001/jama.2016.18553] [Medline: 28241346]
10. Gao G, McCullough JS, Agarwal R, Jha AK. A changing landscape of physician quality reporting: analysis of patients' online ratings of their physicians over a 5-year period. J Med Internet Res 2012 Feb 24;14(1):e38 [FREE Full text] [doi: 10.2196/jmir.2003] [Medline: 22366336]
11. Burkle C, Keegan T. Popularity of internet physician rating sites and their apparent influence on patients' choices of physicians. BMC Health Serv Res 2015 Sep 26;15:416 [FREE Full text] [doi: 10.1186/s12913-015-1099-2] [Medline: 26410383]
12. Emmert M, Meier F, Pisch F, Sander U. Physician choice making and characteristics associated with using physician-rating websites: cross-sectional study. J Med Internet Res 2013 Aug 28;15(8):e187 [FREE Full text] [doi: 10.2196/jmir.2702] [Medline: 23985220]
13. Wei S, Allen M. ProPublica. 2015. Surgeon Scorecard URL: https://projects.propublica.org/surgeons/ [accessed 2019-07-24] [WebCite Cache ID 71fkbdG53]
14. Curry E, Li X, Nguyen J, Matzkin E. Prevalence of internet and social media usage in orthopedic surgery. Orthop Rev (Pavia) 2014 Aug 08;6(3):5483 [FREE Full text] [doi: 10.4081/or.2014.5483] [Medline: 25317312]
15. Turnipseed WD. The social media: its impact on a vascular surgery practice. Vasc Endovascular Surg 2013 Apr;47(3):169-171. [doi: 10.1177/1538574413478471] [Medline: 23427284]
16. Birkmeyer JD, Finks JF, O'Reilly A, Oerline M, Carlin AM, Nunn AR, Michigan Bariatric Surgery Collaborative. Surgical skill and complication rates after bariatric surgery. N Engl J Med 2013 Oct 10;369(15):1434-1442. [doi: 10.1056/NEJMsa1300625] [Medline: 24106936]
17. Ghani KR, Miller DC, Linsell S, Brachulis A, Lane B, Sarle R, Michigan Urological Surgery Improvement Collaborative. Measuring to Improve: Peer and Crowd-sourced Assessments of Technical Skill with Robot-assisted Radical Prostatectomy. Eur Urol 2016 Apr;69(4):547-550. [doi: 10.1016/j.eururo.2015.11.028] [Medline: 26755338]
18. Daskivich TH, Houman J, Fuller G, Black JT, Kim HL, Spiegel B. Online physician ratings fail to predict actual performance on measures of quality, value, and peer review. J Am Med Inform Assoc 2018 Dec 01;25(4):401-407. [doi: 10.1093/jamia/ocx083] [Medline: 29025145]

XSL•FO
RenderX

Edited by G Eysenbach; submitted 23.02.18; peer-reviewed by S Bidmon, M Tkáč; comments to author 03.08.18; revised version received 27.09.18; accepted 02.04.19; published 13.08.19.

Please cite as:
Houman J, Weinberger J, Caron A, Hannemann A, Zaliznyak M, Patel D, Moradzadeh A, Daskivich TJ
Association of Social Media Presence with Online Physician Ratings and Surgical Volume Among California Urologists: Observational Study
J Med Internet Res 2019;21(8):e10195
URL: https://www.jmir.org/2019/8/e10195/
doi:10.2196/10195
PMID:31411141

©Justin Houman, James Weinberger, Ashley Caron, Alex Hannemann, Michael Zaliznyak, Devin Patel, Ariel Moradzadeh, Timothy J Daskivich. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 13.08.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Identifying Influence Agents That Promote Physical Activity Through the Simulation of Social Network Interventions: Agent-Based Modeling Study

Thabo J van Woudenberg[1*], MSc; Bojan Simoski[2*], MSc; Eric Fernandes de Mello Araújo[2], PhD; Kirsten E Bevelander[1,3], PhD; William J Burk[1], PhD; Crystal R Smit[1*], MSc; Laura Buijs[1], MSc; Michel Klein[2], PhD; Moniek Buijzen[1], PhD

[1]Behavioural Science Institute, Radboud University, Nijmegen, Netherlands

[2]Social AI Group, Vrije Universiteit, Amsterdam, Netherlands

[3]Radboud Institute for Health Sciences, Radboud University and Medical Centre, Nijmegen, Netherlands

[*]these authors contributed equally

Abstract

Background: Social network interventions targeted at children and adolescents can have a substantial effect on their health behaviors, including physical activity. However, designing successful social network interventions is a considerable research challenge. In this study, we rely on social network analysis and agent-based simulations to better understand and capitalize on the complex interplay of social networks and health behaviors. More specifically, we investigate criteria for selecting influence agents that can be expected to produce the most successful social network health interventions.

Objective: The aim of this study was to test which selection criterion to determine influence agents in a social network intervention resulted in the biggest increase in physical activity in the social network. To test the differences among the selection criteria, a computational model was used to simulate different social network interventions and observe the intervention's effect on the physical activity of primary and secondary school children within their school classes. As a next step, this study relied on the outcomes of the simulated interventions to investigate whether social network interventions are more effective in some classes than others based on network characteristics.

Methods: We used a previously validated agent-based model to understand how physical activity spreads in social networks and who was influencing the spread of behavior. From the observed data of 460 participants collected in 26 school classes, we simulated multiple social network interventions with different selection criteria for the influence agents (ie, in-degree centrality, betweenness centrality, closeness centrality, and random influence agents) and a control condition (ie, no intervention). Subsequently, we investigated whether the detected variation of an intervention's success within school classes could be explained by structural characteristics of the social networks (ie, network density and network centralization).

Results: The 1-year simulations showed that social network interventions were more effective compared with the control condition (beta=.30; t100=3.23; *P*=.001). In addition, the social network interventions that used a measure of centrality to select influence agents outperformed the random influence agent intervention (beta=.46; t100=3.86; *P*<.001). Also, the closeness centrality condition outperformed the betweenness centrality condition (beta=.59; t100=2.02; *P*=.046). The anticipated interaction effects of the network characteristics were not observed.

Conclusions: Social network intervention can be considered as a viable and promising intervention method to promote physical activity. We demonstrated the usefulness of applying social network analysis and agent-based modeling as part of the social network interventions' design process. We emphasize the importance of selecting the most successful influence agents and provide a better understanding of the role of network characteristics on the effectiveness of social network interventions.

(J Med Internet Res 2019;21(8):e12914) doi:10.2196/12914

KEYWORDS

physical activity; adolescent health; peer group; computer simulation; social network analysis

XSL•FO
RenderX

Introduction

Background

There has been an increasing interest in the use of *social network interventions* to promote health behaviors. Social network interventions are based on the diffusion of innovations theory [1] and capitalize on interpersonal influence to promote and catalyze desired behavioral changes [2]. A few studies have used social network interventions to promote health behaviors in school settings [3]. For example, the A Stop Smoking In Schools Trial study trained influence agents to encourage peers not to smoke in secondary schools [4]. Other studies have trained influence agents to stimulate peers to increase health behaviors, such as drinking more water [5] or being more physically active [6,7].

One of the most important assumptions of social network interventions is that some peers act as role models and can be important determinants of the behavior of the group [8]. Involving these important peers in the intervention can prove beneficial, as they can be used as an example for the rest of the social network; they can help ensure that the intervention message spreads among the individuals in the social network. In such an intervention, the health behavior is disseminated among the classmates through their network ties [1] and will lead to less resistance. Therefore, in most social network interventions, a subset of participants is selected as *influence agents* to initiate the diffusion of an idea or behavior. The influence agents can volunteer or be appointed by researchers, but many social network interventions rely on peer nominations within a social network to determine the influence agents [2]. Participants nominate peers on a number of questions (eg, "Who are your friends?"). On the basis of these nominations, 10% to 17.5% of individuals are approached to become influence agents [2]. The influence agents are trained to adopt and spread a new or improved health behavior or informally diffuse the intervention messages within their social network. However, it is not yet clear which individuals in a network make the most effective influence agents. In other words, what is the best selection criterion to determine influence agents?

An ideal solution to this question would be to run a large-scale field experiment with different criteria for selecting the influence agents. However, this would be a costly undertaking, which is probably the reason why this question has remained unanswered. Fortunately, advancements in computer science have enabled us to simulate hypothetical social network interventions by using computational models [9,10]. This contemporary approach is a big step forward in the intervention studies' design process. Computational models can be a promising method to understand the complex interplay between social influences and other factors that are driving certain health behaviors [11]. For example, researchers can collect baseline data, simulate a wide range of interventions, and opt for the intervention strategy with the biggest changes in behavior or the one that is most cost-effective. In addition, computational models could be used in consultation with key stakeholders to determine priorities, create expectations about the interventions, and tackle issues regarding implementation early on. Finally, simulations enable researchers to formulate data-driven hypotheses that can be tested in vivo. Therefore, computational models are a valuable addition to the toolbox of researchers and practitioners who aim to change behaviors.

Agent-based models (ABMs) are used to model interactions among individuals within a social network and, therefore, fit the theoretical underlying mechanisms of social network interventions. The behavior of an influence agent has an effect on the individuals with whom the influence agent shares a connection. To develop effective social network interventions, it is essential to understand how behavior spreads in a social network and what affects the spread of the desired behavior. ABMs are a helpful tool for this, as they enable researchers to experiment in simulated environments. In previous research, ABMs were used to ascertain effective ways of identifying important influencers [10,12-14]. In addition, ABM simulations have been increasingly explored as an alternative approach for addressing health research questions. Furthermore, previous studies have shown that ABMs can be used to model physical activity behavior [15-17] or obesity [18,19] in a social network.

The aim of this study was to test which selection criterion to determine influence agents in a social network intervention resulted in the biggest increase in physical activity in the social network. An ABM was used to test different selection criteria for influence agents by simulating social network interventions and observing the intervention's effect on the physical activity within school classes. In this study, we relied on the methods and model specifications of our previous study [20] to build the social networks and implement the computational model. Drawing on a previously validated model developed by Beheshti [12] and Giabbanelli et al [21], the computational model employed in this study was applied to the observed data of primary and secondary school children collected in the *MyMovez* project [22]. The model considered 2 factors as determinants for an individual's behavioral change: the class's social influence and the individual's social environment (for more information see [20]). In the model, the behavior of influence agents has an effect on those with whom he or she shares a relationship: the effect of the influence agents spreads from connection to connection. This is referred to as *social network influence*. In addition, the ABM used in this study considered the influence of the physical environment as a potentially important factor for promoting health behavior.

To further investigate the applicability of ABMs for social network interventions, this study examined whether the simulated effectiveness of social network interventions was dependent on several network characteristics. We built upon Valente's idea that the interventionist should not only use the networks as an intervention instrument but also learn from the available social network information to create better, meaningful interventions [20]. In addition, Giabbinelli et al [21] concluded that there are microlevel network structures to be investigated, which are involved in making the agents more resilient to change. Other studies also state that interventions might be less effective if they neglect the impact of social networks [23]. Therefore, we investigated if characteristics of social networks (classes) could affect the effectiveness of network-based health interventions.

XSL•FO
RenderX

The analysis presented in this study is based on the *MyMovez* dataset of 26 school classes. This provided a rare opportunity to simulate social network health interventions in school classes based on a comprehensive real-world dataset with real social networks and physical activity data. Multiple sociometric nominations submitted by the participants were used to define weighted relationships among the class peers and thereby build the social networks. Physical activity and social environment data were used to define the state of each agent per day.

In this study, we identified 2 sets of hypotheses. First, we compared the outcomes of different conditions to determine the selection criteria for most effective influence agents in social network interventions. The effectiveness of the interventions was measured by the difference in physical activity between the baseline and after 1 year of simulation. Second, this study investigated whether different network characteristics (ie, density and centralization) of the classes could affect the effectiveness of social network interventions.

This study is a product of collaborative research between social and computer scientists, with the motivation to translate the findings into applicable advice for preparing network-based interventions. The social and computer science research communities examine social networks and network-based health interventions from fairly contrasting angles. Significant improvements could ensue with respect to the way social network health interventions are designed and implemented owing to strong collaborations between social and computer science research communities.

Selecting Influence Agents

To assess the predictive validity of the computational model, the simulated interventions were compared with the no intervention condition. On the basis of social network theory and the overall positive outcomes of previous social network interventions [4-6,24], we expected a bigger increase in physical activity in the intervention conditions than in the no intervention condition.

Subsequently, we looked at selecting strategically placed influence agents, compared with having a random allocation of influence agents. Scholars have elaborated on different roles and positions of individuals within social networks (for an overview see [25]). Influence agents are often defined as individuals who are most *central* in the network [3]. This means that those individuals hold a prominent place in the network. Centrality is a measure of an individual's position relative to their social network, but there are a handful of definitions and algorithms used to define and measure centrality [3,25]. These definitions all assume that in one way or another, being central in the social network means that an individual is more influential. Therefore, we assumed that having central individuals as influence agents (regardless of the used definition) should increase the effectiveness of a social network intervention. Thus, we defined our first hypothesis as follows:

> *H2: The increase in physical activity will be higher in the simulated social network interventions based on centrality than in the simulated random influence agent intervention.*

As Freeman [25] discussed, there is no consensus on a common definition of centrality or how it should be measured. There are 3 widely used definitions of centrality: in-degree centrality, betweenness centrality, and closeness centrality [3,26].

In-Degree Centrality

The most often used centrality measure in the social network interventions literature is *in-degree centrality*. In-degree centrality is based on the number of peer nominations an individual receives (notably, this is referred to as out-degree centrality when researchers use influence models instead of nomination models). The more the incoming peer nominations, the higher the in-degree centrality. So, individuals with high in-degree centrality can be seen as an important channel of information [25]. In school settings, most often the in-degree central influence agents are the most popular children or adolescents and are clustered together in the network. Therefore, the intervention could affect that small cluster of individuals and not reach the important subgroups or peripheral nodes in the network (who might benefit the most from the intervention). In addition, popular peers may be reluctant to change their behavior or perform the role of an influence agent [27]. The popular peers have a large contribution to the social norms within the network, and deviating from the established social norm could have a negative effect on their social status. Therefore, Borgatti [26,28] argues that 2 other types of centrality are likely to be more important for the promotion of health behaviors: betweenness centrality and closeness centrality.

Betweenness Centrality

Betweenness centrality focuses on the role of influence agents as a gatekeeper of information within social networks. These influence agents are important for linking different individuals, groups, or subgroups together and are referred to as being a *bridge*. More specifically, betweenness centrality is based on the frequency with which an individual is a link in the shortest path between 2 other peers. This means that this individual controls the flow of information among other peers in the network. Such an individual can influence the network by withholding or distorting information in the diffusion. If the betweenness central agents are not selected to disseminate the intervention message, entire subgroups could be withheld from the intervention [25]. In particular, Borgatti argues that betweenness central agents should be used when the goal is to disrupt the network's ability to spread unhealthy behavior [28]. By removing these individuals from the social network, the residual network has the least possible cohesion and, therefore, will decrease the spread of negative behaviors in the network the most. In practice, it is not feasible to remove those individuals from a network but increase their physical activity to prevent a potential negative behavior (low physical activity) from spreading in the social network.

Closeness Centrality

Closeness centrality focuses on the reach of the influence agents within networks and dissemination speed of the intervention in the network. Closeness centrality represents the distance between the individuals and all other peers in a network. More specifically, closeness central individuals have on average the

XSL•FO
RenderX

shortest path to all other peers in a network. This means that the intervention will reach the entire network in the least amount of links, and it makes the intervention message most efficient. Therefore, Borgatti argues that closeness central influence agents should be used when the goal is to promote positive health behaviors [28]. The positive intervention message will reach all members of the social network in the most efficient way and will not exclude clusters of or subgroups from the intervention message. This approach fits within the notion that to reduce weight, it is more effective to promote a healthy behavior (eg, physical activity) than to discourage negative behaviors (eg, watching television) [29]. Because the simulated social network interventions entail the promotion of physical activity (ie, a positive behavior), we defined our second hypothesis as follows:

> *H2: The increase of physical activity will be higher for simulated social network intervention based on closeness centrality than in simulated social network intervention based on in-degree and betweenness centrality.*

Network Characteristics

Next to the measurement of network properties at the individual level, social network analysis can also be used to describe network properties at the group level. It is important to understand group-level network information to create better and more meaningful interventions [30,31]. Because all classes are unique in their network properties, social network interventions should keep the structure of the network in mind. *Density* and *centralization* are 2 of the most important network characteristics that could influence the effectiveness of a social network intervention [32].

Density

The density of a social network is a measure of the cohesion in a network and can be defined as a ratio between the number of ties between participants and the number of all possible ties in a network. This means that dense classes have a relatively high number of connections among the individuals and thus have a high degree of cohesion. Figure 1 provides examples of social networks of 2 classes. The node color refers to the individual's in-degree centrality. Red means a higher in-degree and blue means a low in-degree centrality. Ties between nodes are weighted based on 6 nomination questions, and participants could nominate an unlimited number of peers. The left network in Figure 1 is a classroom with high density as 90% of all possible ties are connected. The right network scores low on centrality, as only 46% of all possible ties are connected.

Networks with high density imply more peer interactions, therefore maximizing the opportunities for spreading an intervention within a social network [33]. We expected that this would also apply to social network interventions that promote physical activity; therefore, we defined our third hypothesis as follows:

> *H3: The effect of the simulated social network interventions will be higher in classes with high density than in classes with low density.*

Figure 1. Examples of density in social networks.

(a) Class 129, a class with high density

(b) Class 135, a class with low density

Centralization

Centralization of a network describes the distribution of the individual *centrality* measures of the participants in a network. In contrast to centrality, centralization is a network-level measure. Freeman describes centralization as the skewness of the distribution of nominations in a social network [25]. Centralization defines the extent to which interactions are concentrated in a small number of individuals rather than distributed equally among all peers [32]. This means that in highly centralized networks, there is a pronounced subgroup of central individuals. Network centralization can be calculated for all the centrality measures (ie, in-degree, betweenness, and closeness centrality).

Figure 2 is an example of in-degree centralization in 2 of the classes. The node color is proportional to the individual's in-degree centrality. Red means a higher in-degree and blue means lower in-degree centrality. The left network in Figure 2 is an example of a class with high in-degree centralization. As can be visually observed, there is 1 individual (ID 2892) in the

XSL•FO
RenderX

left social network who received proportionally more nominations than the rest of the class. Therefore, this class has a high score in in-degree centralization, and ID 2892 should be an effective influence agent in this class. In contrast, the network on the right has low in-degree centralization because it has a large subgroup of individuals who are high in in-degree centrality. The same principle applies to betweenness centralization and closeness centralization.

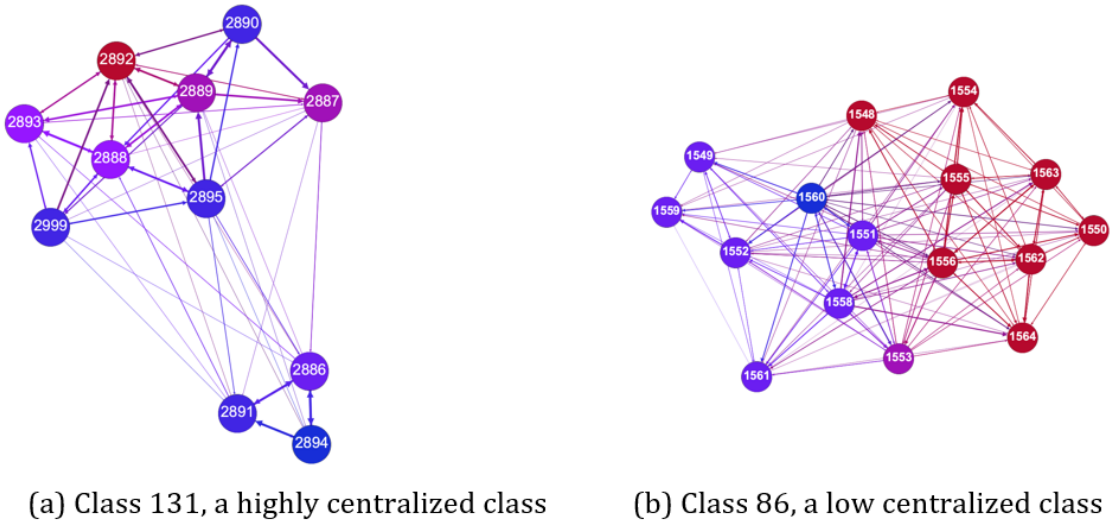
Previous research has shown the moderating role of centralization in the relationship between friendship networks and bullying in children [34]. More specifically, the centralization of the class predicted whether popularity related to aggressive behavior in boys. However, it has not been studied before whether social network interventions have more effect in centralized classes than in classes in which the nominations are spread evenly. We argue that classes with high centralization lend themselves better for social network interventions because the influence agents are more pronounced and, therefore, easier to detect by the researcher. These influence agents in centralized classes will have relatively more influence on the network than the influence agents in noncentralized classes. Therefore, we defined our last 3 hypotheses as follows:

H4a: The effectiveness of the simulated social network interventions based on in-degree centrality will be greater in classes with high in-degree centralization than in classes with low in-degree centralization.

H4b: The effectiveness of the simulated social network interventions based on betweenness centrality will be greater in classes with high betweenness centralization than in classes with low betweenness centralization.

H4c: The effectiveness of the simulated social network interventions based on closeness centrality will be greater in classes with high closeness centralization than in classes with low closeness centralization.

Figure 2. Examples of in-degree centralization in social networks.



(a) Class 131, a highly centralized class

(b) Class 86, a low centralized class

Methods

Participants and Procedure

The study used data from *MyMovez* project [22], a large-scale cross-sequential cohort study among children and adolescents (aged 8-12 and 12-15 years) from 21 primary and secondary schools [22]. In the project, participants received a smartphone with a research app on which they received daily questionnaires and a wrist-worn accelerometer (Fitbit Flex). This accelerometer has been shown to be a reliable measure of physical activity [35,36]. For this study, the first 4 waves of the *MyMovez* project were used: February to March 2016 (Wave 1), April to May 2016 (Wave 2), June to July 2016 (Wave 3), and February to March 2017 (Wave 4). To ensure that the influence agents are identified from a representative sample within each classroom, only classes with more than 60% of students participating were included. This resulted in 26 classes, with 460 participants (mean age 10.81 years, SD 1.28; male=51.5% [237/460]) in total.

Measures

Physical Activity

In each wave, participants wore the accelerometer on their nondominant hand for 7 consecutive days. The first and the last day were excluded because these were partial days (handing out and giving back the accelerometer), resulting in 5 complete days of data. In addition, days that did not add up to 1440 min (24 hours) and days with less than 1000 steps were excluded because these were partial days of data (eg, caused by empty battery or nonwear time).

The average physical activity per wave was calculated by taking the average steps per day of at least 3 days of valid data. If participants had less than 3 days of valid data per wave, daily step count was imputed with the same strategy as in the study by van Woudenberg et al [7], by using single multilevel (predictive mean matching) imputation [37]. Missing data were imputed based on other physical activity data of the same participant, day of the week, measurement period, sex, and age. On average, participants accumulated 10,505 steps per day (SD 5.730).

XSL•FO
RenderX

The physical activity measure had to be scaled to fit the ABM. In the previous study with the same ABM, the mean value of physical activity was set at 1.53 [12]. Therefore, we computed a new variable named the physical activity level (PAL) by dividing the steps by 10,000 and multiplying by 1.53. The mean PAL value in our dataset was 1.50, with a minimum of 0.45 and a maximum of 4.27.

Family Affluence

A measurement of the influences of the social environment was needed as a second input parameter of the ABM. The Family Affluence Scale (FAS) was used as a measure of socioeconomic status [23]. The FAS is a self-reported measure of family affluence and is an effective tool for assessing socioeconomic status in adolescents [38]. The participants were asked sets of questions (eg, "How many cars does your family own?" and "How often do you go on a holiday outside of the Netherlands?"). All answers (range 0-13) were summed ($mean_{FAS}$ 4.01, SD 1.52), reflected, and divided by the number of items (alpha=.41) to fit the model. This resulted in an environmental variable (*env*) with a value between 0 and 2 in which a higher *env* value reflects a lower family affluence.

Sociometric Nominations

In each wave, participants nominated peers from the same class by 6 sociometric questions based on the study by Starkey et al [24]. Participants received the questions at a random time during the day and nominated peers by clicking on their names in a list on the research smartphone. They were required to nominate at least one other peer, and no maximum on the peers nominated was given (note that self-nominations were not possible). For an overview of the questions, see Multimedia Appendix 1 [4,39,40].

Centrality

The social network characteristics at the individual level were calculated with the Python3 [41] package NetworkX 2.1 [42]. For an overview of the centrality measures, see Table 1. The individual's betweenness centrality did not correlate with in-degree centrality or closeness centrality, but in-degree centrality did correlate with closeness centrality (r_{457}=0.58; P<.001).

Table 1. Descriptive statistics for the individual- and group-level variables.

Variable name	Mean (SD)	Minimum	Maximum
Individual characteristics (centrality; N=451)			
In-degree	12.27 (4.15)	4.00	27.00
Betweenness	0.01 (0.02)	0.00	0.12
Closeness	0.78 (0.11)	0.49	1.00
Network characteristics (N=26)			
Density	0.72 (0.11)	0.46	0.90
Centralization			
In-degree	0.20 (0.08)	0.07	0.40
Betweennes	0.04 (0.03)	0.01	0.09
Closeness	0.22 (0.08)	0.09	0.39

Density and Centralization

Density and 3 centralization measures were calculated for each class. The density was calculated by taking the number of ties present in a social network and dividing this by the number of all possible ties, resulting in a number ranging from 0 (noncohesive network) to 1 (very cohesive network). In-degree centralization, betweenness centralization, and closeness centralization were calculated with the *igraph package* in the statistical computing package RStudio [43], resulting in a number ranging from 0 (*noncentralized network*) to 1 (*very centralized network*). The density and centralization scores were normalized given the different network sizes. For an overview of the density and centralization scores, see Table 1.

Design

Social Networks

On the basis of the sociometric nominations, a directed social network was constructed for each classroom. A directional social network comprises nodes that represent the participants within a class and edges representing (weighted) connection between 2 nodes (referred to as *edge*).

The weight is defined as the sum of nominations of a participant toward another, divided by the total number of nomination questions. Because 2 participants could nominate each other, the edges in the network are directional (represented by the arrow of the edge). As participants nominated peers on multiple sociometric questions, each edge was associated with a *connection weight* ranging from 0 (zero nominations) to 1 (all 6 nominations). The more nominations a participant gave to another peer, the stronger the edge's connection weight. Duplicate nominations were omitted (as a participant could nominate the same peers on the same items across waves), resulting in a maximum of 6 nominations toward another peer within all 4 waves.

XSL•FO
RenderX

Agent-Based Model

Computational models can be defined "as an abstract and simplified representation of a given reality, either already existing or just planned. Models are commonly defined to study and explain observed phenomena or to foresee future phenomena" [44].

ABMs are a particular category of computational models for simulating the communication among the agents in a common environment to understand their behavior. For this study, we relied on a previously validated ABM developed by Giabbanelli et al [21] and enriched by several adaptations [12,20].

Giabbanelli's [21] model was used as it accounted for the interaction of social networks with environmental factors, unlike earlier related computational models for social network interventions. In this model, individuals influence each other with respect to physical activity that might change also depending on the agent's physical environment. Their factor analysis on synthetic and real-world social networks showed that the environment was a crucial parameter for changes in bodyweight (their health behavior of interest). This particular model was favored as it was a good complement to the collected *MyMovez* data. Many previous studies used more complex models incorporating multiple parameters but based them on synthetic datasets. The purpose of this study was to use data collected from real human relations and behaviors, and this model was a good fit for the observed data.

The ABM simulates the spread of physical activity within social networks (classes), that is, simulating the spread of the intervention's effect through the classes. We assumed that physical activity spreads throughout the relationships and depends on the physical environment. Each agent, in our case participants within a class, was assigned 2 input parameters before running the simulations—the PAL and the *env* parameter. Yearlong simulations were run for each of the social network intervention strategies and for each class. During each step (represented by a single day) of the simulation, a PAL value was derived for each agent based on the social influence and the environmental influence. The social influence comes from all the peers who are connected to the agent. It is based on the connection weights between agent's peers and the associated peers' PAL. The environmental influence is the effect of the agent's family affluence, represented by *env*. The ABM does not make assumptions regarding probability of diffusion across ties.

Each simulation step potentially updates the agent's PAL and was calculated in 3 phases, similarly as presented by Giabbanelli et al [21]. First, the social influence parameter was calculated, coming from the adolescent's peers (dependent on peers' PAL and connection weights). Second, the social influence with the agent's environmental influence (ie, *env*) was combined in a single parameter, called the socioenvironmental influence. Third, the socioenvironmental influence parameter was compared with a predefined threshold to decide if agent's PAL will be modified or remain the same.

See Multimedia Appendix 2 for a detailed description of the used ABMs. Multimedia Appendix 2 presents the ABM's mathematical representation and gives more information about the model's thresholds [45].

Interventions

A total of 5 conditions were created based on 4 social network intervention strategies and a control condition (no intervention). In the centrality-based intervention conditions (ie, in-degree, betweenness, and closeness centrality), the top 15% of participants with the highest centrality were assigned as influence agents. When participants above and below the cutoff score had the same centrality scores, random participants from these cases were assigned as influence agents. In the *random agent* intervention condition, 15% of influence agents were randomly selected out of all participants in a classroom. To diminish the possible effect of selecting a particular set of influence agents in the random agent condition, 100 interventions were simulated and averaged afterward to provide a single outcome value. In the *control condition*, no intervention was simulated.

All interventions were based on the assumption that the training sessions of the social network interventions were able to increase the physical activity of the influence agents at the start of the intervention. Therefore, all influence agents received an artificial increase of 17% in their initial PAL based on the outcomes of a previous behavioral intervention [12,46]. After the increase in PAL of the influence agents, the intervention simulations were run for 365 days (day 0-364). The effectiveness of the health interventions was expressed as the *success rate*, the percentage of increase in a class's PAL from the start (day 0) to the end (day 364) of the simulation.

Ethics Approval and Consent to Participate

Informed consent was obtained from 1 of the parents of the participants in the *MyMovez* project. Study procedures were approved by the Ethics Committee of the Radboud University (ECSW2014–100614-222).

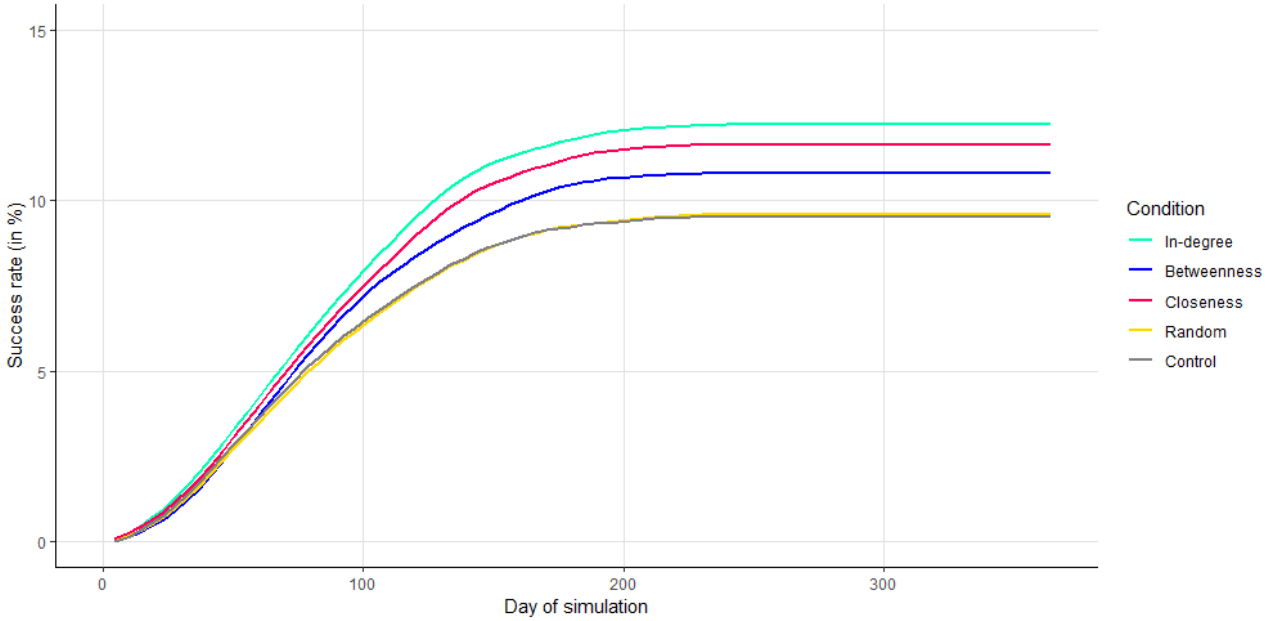
Results

Simulating Physical Activity

The simulations were used to observe the spread of physical activity among peers in the classes and determine the success rate of the different interventions. Figure 3 illustrates the trajectory of the averaged PAL of the classes for all the different simulated interventions for the 1-year period. What stands out from Figure 3 is that all simulated interventions increase the average physical activity of the networks. However, there are differences between the interventions in the increase of physical activity. A detailed overview of the interventions' success rates for all conditions can be found in Multimedia Appendix 3.

XSL•FO
RenderX

Figure 3. Intervention outcomes. Average success rate for the conditions over one-year simulation.



As a first step, we tested the overall differences among all conditions. A linear mixed-effects model was run [47], with success rate as the dependent variable, condition as the predictor, and random intercepts per class. Mauchly test indicated that the assumption of sphericity was not met (W=0.00, P<.001, ε=.41). Therefore, all degrees of freedom were corrected by using the Huynh-Feldt estimation of sphericity. The repeated measures analysis of variance showed that the simulated interventions differed from one another ($F_{1.66,41.42}$=7.72, P=.002, ε=.01). To investigate differences among the conditions as proposed in the hypotheses, planned contrasts (Helmert coding scheme) were used. In addition, all P values were corrected by using the Satterthwaite method as suggested by Luke [48].

Selecting Influence Agents

For checking model validity, the first planned contrast was used to compare the 4 social network intervention conditions with the control condition (no intervention). The contrast revealed that the success rates of the social network interventions (11.28%) were higher than the control condition (9.76%; beta=.30; t_{100}=3.30; P=.001). This means that the interventions were more successful in increasing physical activity than in the absence of interventions. Therefore, we presumed that ABM is a valid tool to simulate social network interventions.

To test the first hypothesis (H1), the second planned contrast compared the 3 centrality social network intervention conditions (ie, in-degree, betweenness, and closeness centrality conditions) with the random agent condition. The averaged success rate of the centrality social network intervention conditions (11.74%) was higher than the success rate in the random agent condition (9.90%; beta=.46; t_{100}=3.86; P<.001). This means that having central influence agents is more effective in increasing physical activity than having randomly sampled individuals in a network.

To test the second hypothesis (H2), the third and fourth planned contrasts compared the differences within the 3 centrality social network intervention conditions. The third contrast compared the betweenness and closeness centrality conditions (11.57%) with the in-degree condition (12.08%). The success rates did not differ from each other (beta=−.17; t_{100}=−1.00; P=.32). The fourth contrast compared the closeness centrality condition with the betweenness centrality condition. The success rates of the closeness centrality condition (12.16%) were higher than the betweenness centrality condition (10.98%; beta=.59; t_{100}=2.02; P=.046). This means that we did not find evidence that the closeness centrality condition outperformed the in-degree centrality condition, but the betweenness centrality condition was less effective in increasing physical activity in the networks compared with the in-degree and the closeness centrality conditions.

Network Characteristics

The success rates of the social network interventions varied among classes, as can be seen in Multimedia Appendix 3. A few classes stayed neutral to the interventions or even encountered negative effects (classes 101 and 129 showed PAL decrease), whereas other classes showed >30% average PAL increase over a year of simulation of particular intervention condition. Therefore, we investigated the effect of structural properties of the classes (ie, density, in-degree centralization, betweenness centralization, and closeness centralization) on the success rates of the interventions. More specifically, we added the different structural properties as moderators to the mixed-effects model. Table 2 displays the correlation coefficients of the 4 social network interventions and the 4 structural network properties. For a detailed overview of the structural properties per class, see Multimedia Appendix 4.

XSL•FO
RenderX

Table 2. Correlations between social network interventions and network structures.

Network structures	Network interventions			
	In-degree	Betweenness	Closeness	Random agent
Density	−0.37	−0.33	−0.35	−0.34
In-degree centralization	0.58[a]	0.57[a]	0.58[a]	0.56[a]
Betweenness centralization	0.26	0.26	0.26	0.21
Closeness centralization	0.35	0.30	0.33	0.30

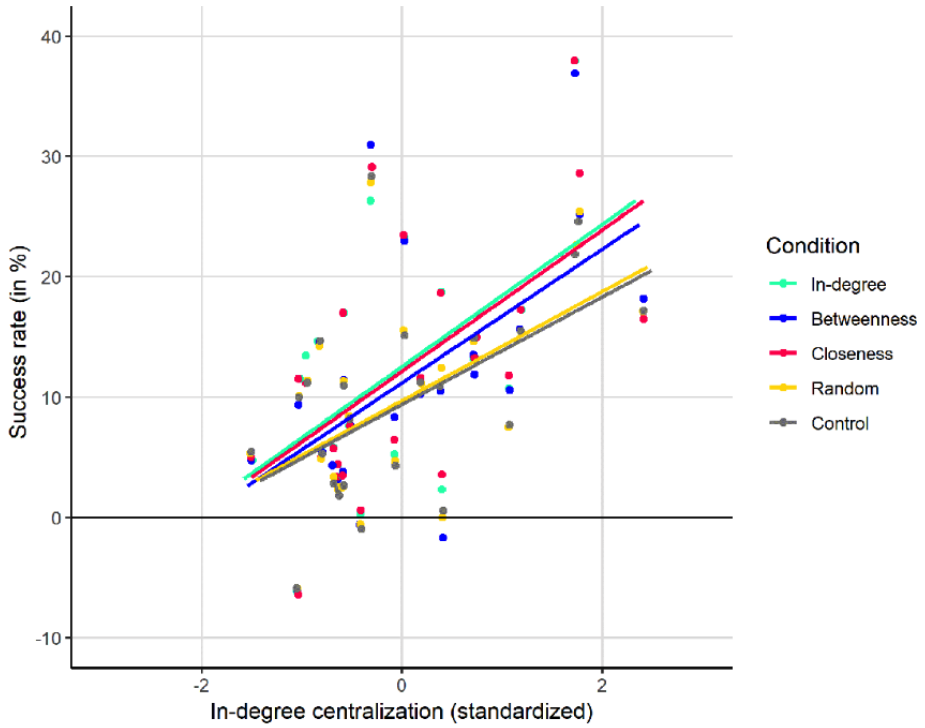
[a]$P<.05$.

The third hypothesis (H3) predicted that the interventions would be more effective in classes with high density. To test whether the density of the class moderated the effectiveness of the different interventions, the same mixed-effects model was run with the addition of the interaction effect of density (standardized). The analysis showed that there was no significant direct effect of density on the success rate (beta=−3.17; t_{24}=−1.58; P=.08). This means that the success rates were not higher in classes with high density than in classes with low density. In addition, no significant interaction effects of the planned contrasts of the social network conditions and the density of class were observed. This means that we did not find evidence to support the hypothesis (H3) that social network interventions were more effective in classes with high density compared with classes with low density.

The last 3 hypotheses (H4a, H4b and H4c) predicted that the interventions would be more effective in classes with high centralization based on the centrality measure that was used. For these analyses, the contrasts were changed per hypothesis, so that the centrality measure in focus was contrasted with the other social network interventions. For these 3 hypotheses, the same mixed-effects model was used, with the addition of the interaction effect of centralization.

The first linear mixed-effects model investigated in-degree centralization (H4a) and showed that there was a direct effect of in-degree centralization on the success rate (beta=5.27; $t_{9.94}$=3.55; P=.002). As can be seen in Figure 4, the social network interventions were more effective in classes with high in-degree centralization. This means that social network interventions are more effective when the class is more centralized around some in-degree central individuals. In addition, we looked at the interaction of in-degree centralization and the planned contrast of the in-degree centrality condition versus the other social network interventions. This interaction effect was nonsignificant (beta=.15; $t_{39.76}$=1.26; P=.21). This means the effect of in-degree centralization on the success rates was not stronger in the in-degree centrality condition than in the other social network conditions.

Figure 4. Effect of in-degree centralization on the success rates per condition.



XSL•FO
RenderX

The second linear mixed-effects model investigated betweenness centralization (H4b) and showed that there was no direct effect of betweenness centralization on the success (beta=2.22; $t_{9.94}$=1.25; *P*=.22). This means that the social network interventions were not more effective in classes with high betweenness centralization compared with classes with low betweenness centralization. In addition, the interaction effect was nonsignificant (beta=.09; $t_{39.76}$=0.73; *P*=.45). This means that the effect of betweenness centralization on the success rates was not stronger in the betweenness centrality condition than in the other social network conditions.

The last linear mixed-effects model investigated closeness centralization (H4c) and showed that there was no direct effect of closeness centralization on the success rate (beta=2.88; $t_{9.94}$=6.66; *P*=.11). This means that the interventions were not more effective in classes with high closeness centralization compared with low closeness centralization. In addition, the interaction effect was nonsignificant (beta=.11; $t_{39.76}$=0.93; *P*=.36). This means that the effect of closeness centralization on the success rates was not stronger in the closeness centrality condition than in the other social network conditions.

Given these results, our hypotheses that the effectiveness of the simulated social network interventions would be greater in classes with high centralization than classes with low centralization were rejected. We only found evidence that social network interventions were more effective in high in-degree centralized classrooms, irrespective of the type of social network intervention used.

Discussion

Principal Findings

The aim of this study was to test which selection criterion to determine influence agents in a social network intervention resulted in the biggest increase in physical activity in the social network. To test different selection criteria for influence agents, an ABM was used to simulate different selection criteria for social network interventions and observe the intervention's effect on the average physical activity of the classroom. In addition, the study investigated whether social network interventions were more effective in some classes than others based on their particular network characteristics.

The general effectiveness of social network interventions was compared with the control condition. The results showed that the increase in physical activity was of greater magnitude in social network interventions than in the control condition. This demonstrates that an increase in physical activity of a small group of individuals has the potential to spread to peers in the social network. Therefore, the ABM produced results in line with the social network theory, which predicts that behaviors spread in social networks [1,2,3]. We, therefore, assumed that our model was a valid tool to test our hypotheses.

In addition, the effect was stronger for the centrality-based social network intervention conditions compared with the random influence agent condition. This is not in line with the results of the first model of El-Sayed et al [13] who concluded (also based on simulations of literature-based parameters) that well-connected influence agents had little or no added value compared with random influence agents. This difference may be a result of the different model specifications in the 2 studies. In addition, the outcome variable in the study by El-Sayed et al [13] was the prevalence of obesity. On the contrary, the results of this study are in line with the second set of simulations of artificially high parameter models of El-Sayed et al [13] and Zhang et al [10]. These results corroborate the idea that central individuals hold an important position within their social networks [25]. Taking a random subsample of the participants as influence agents is not as effective as strategically located influence agents. Therefore, researchers should carefully select influence agents based on their position in the social network, as suggested by Borgatti [26] and Valente and Pumpuang [30]. When researchers are unable to strategically select the influence agents, Bahr et al recommend increasing the percentage of random influence agents to obtain the same success rates as the centrality conditions with 15% of the class as influence agents [49].

Contrary to expectations, no difference was observed between the in-degree centrality condition and the closeness centrality condition, as suggested by Borgatti [26] and Valente [30]. An explanation could be that Valente's argument [27] that in-degree agents are most often the popular individuals and not willing to change their behavior does not hold for simulated intervention. In the simulations, the artificial increase of physical activity of the influence agents was the same for the in-degree centrality condition and the closeness centrality condition. In contrast, a difference between the closeness centrality condition and the betweenness centrality condition was observed. In accordance with Borgatti's [28] reasoning that positive behavior should be promoted via closeness central agents, we observed that the closeness centrality condition had a higher success rate than the betweenness centrality condition. This corroborates the idea that when researchers want to increase a positive behavior, closeness centrality influence agents should be selected.

Finally, this study looked at the moderating role of structural characteristics of the class on the effectiveness of the social network interventions. The results showed that the density of the class did not affect the success rates of the social network interventions. This is not in line with social network theory, which argues that innovations spread quicker through highly connected networks [32]. We also anticipated that the specific centrality conditions were most effective as the classes were more centralized on the relevant centrality measure. However, the results indicated that only in-degree centralization had a direct effect on the success rates. This means that social network interventions are more effective when classes have a small number of individuals who receive the most nominations. The subsequent analyses showed that this effect was not stronger in the in-degree centrality condition than in the other social network intervention conditions. Therefore, we can conclude that social network interventions work well in classes with high in-degree centralization irrespective of the selection criterion used.

XSL•FO
RenderX

This study advanced the field of social network interventions and the use of ABMs for better understanding interventions in numerous ways. This study was one of the first to use simulations to test the difference among the selection criteria for the influence agents in social network health interventions. In addition, this study used empirical data as input for the model. The next step in the interplay between health interventions and computational models will be to replicate these simulated results with empirical data of social network health interventions.

The study provides implications for future research and can advise social network researchers. First, this study supports the idea that social network interventions can be an effective strategy to increase physical activity in the classroom. Second, it stresses the importance of strategically selecting the most central individuals as influence agents. Finally, the composition of the class can influence the effectiveness of social network interventions. In addition, this study shows the applicability of simulations to help researchers design the most effective interventions.

Comparison With Previous Studies

ABMs have been used previously to study the spread of health behaviors in simulated social environments after hypothetical interventions. For example, an ABM was used to investigate the spread of obesity in artificial participants after multiple obesity prevention campaigns [13]. The use of ABMs to investigate the spread of obesity was further refined by using the body mass index of an observed sample of participants and the addition of a socioenvironmental factor [21]. However, no behavioral data were available, so physical activity was imputed based on a random distribution. A subsequent study improved the previously mentioned ABM by incorporating individual thresholds for the change in health behaviors [12]. Our previous study used this model, but here we applied it to observed behavioral and sociometric data [20]. The previously mentioned ABMs [12,20,21] showed the same results as this study in that the simulations of interventions showed an increase that attenuated over time.

On the basis of different ABMs, 2 other studies have used agent-based simulations to investigate the effectiveness of different types of influence agents in social network interventions [9,10] but both with a slightly different aim. The study by Zhang et al [10] examined only the difference between randomly selected and in-degree central influence agents. Their conclusion aligns with the findings from this study in that physical activity increases more in the intervention that uses influence agents based on centrality compared with the intervention that uses random influence agents.

The study by Badham et al [9] matches the research question of this study more closely, that is, the study looked at the 3 different types of centrality measures. However, the outcome of the simulations was the amount of time (number of iterations) before the entire network adopted a behavior. In other words, the study by Badham et al [9] focused on the speed of adoption of the intervention and not on the magnitude of the behavior change after the simulated interventions. Despite the different outcome variable, the studies showed comparable outcomes to the findings of this study. More specifically, the most effective interventions are those with influence agents based on centrality (ie, in-degree, betweenness, and closeness centrality). Although the study by Badham et al [9] did not formally test the differences among the centrality measures, the observed steps to saturation did not indicate that there was a difference among them.

Limitations

To interpret the results of the simulation of social network interventions, a number of limitations have to be discussed. First, this study was based on the assumption that researchers were able to increase the amount of physical activity of the influence agents. However, it could be that this does not reflect the field experiments that train influence agents to become more active. In addition, increasing the targeted health behavior is only part of the influence agents' training. For example, most training sessions in social network interventions also focus on how the influence agent could communicate the health message in an informal way. This type of health promotion was not a part of the ABM that we used. Future studies could also imitate other aspects of a successful training. For example, researchers could consider increasing the number or the weight of the connections to reflect the communication component of the influence agents' training. Along the same lines, the success rates of the intervention are based on the embedded assumptions in the model of how people influence each other. In our model, the assumption was that the increase in physical activity diffuses over time. However, adopting a contagion framework, which looks at how many peers should increase in physical activity before the individual's physical activity increases, might lead to different success rates of the interventions.

Second, the employed ABM comes with a set of limitations. For example, based on the mathematical characteristics of the model, the ABM's outcome has an initial increase and reaches an equilibrium state after a particular time in the simulations, as shown in Figure 3. Consequently, the control condition also increased in physical activity, contrary to the usually observed decrease among the youth [50]. Therefore, caution is warranted in interpreting the absolute increase in classes' physical activity. Rather, we want to emphasize that the results focused on the relative differences among the selection criteria. In addition, the ABM outcomes enabled us to discuss the effects of *simulated* health interventions. Although the ABM has been validated and tuned to the empirical data, the presented simulation effects should be interpreted with caution. Following this limitation, in our next study we intend to perform similar statistical analyses on the empirical data when the intervention outcomes of the *MyMovez* project are available.

Third, the applied analyses were all based on data aggregated on a classroom level. However, we realize the importance of conducting more elaborate individual-level analyses by including personal characteristics, such as sex, personality traits, individual physical activity, or role in the social network. These personal characteristics can moderate the effect of the health intervention. By including more personal information, the ABM can be better specified. Adopting personality traits could help us understand how an individual perceives and reacts to peer

XSL•FO
RenderX

behaviors as well as learn about individuals' contributions to the class behavior.

Conclusions

In conclusion, we demonstrated the advantages of applying social network analyses and simulations to understanding social networks' characteristics and performing detailed simulations on peer influences. We advise future researchers to perform such simulations on peer influences, whenever possible, *before* doing real-world interventions to maximize the success rate of their interventions. This information can help in designing more effective social network health interventions.

Acknowledgments

The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007–2013)/ERC grant agreement number (61725).

Authors' Contributions

TJW and BS are the 2 primary researchers in the study and are both corresponding authors. All authors were involved in conceptualizing the design of the study. TJW, KEB, CRS, and LB provided the instructions at the location and collected the data. BS and TJW analyzed the data and wrote the concept version of the paper. EFMA, KEB, WJB, CRS, LB, MK, and MB critically reviewed the paper. All authors have given the final approval for the paper and agreed to be accountable for the accuracy and integrity of any part of the study.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Peer nomination questions.

[PDF File (Adobe PDF File), 15KB - jmir_v21i8e12914_app1.pdf]

Multimedia Appendix 2

Model Description.

[PDF File (Adobe PDF File), 28KB - jmir_v21i8e12914_app2.pdf]

Multimedia Appendix 3

Success rates per class of one year simulations of the interventions (in percentages).

[PDF File (Adobe PDF File), 19KB - jmir_v21i8e12914_app3.pdf]

Multimedia Appendix 4

Structural network parameters per class based on the weighted ties.

[PDF File (Adobe PDF File), 23KB - jmir_v21i8e12914_app4.pdf]

References

1. Rogers EM. Diffusion of Innovations. Fifth Edition. New York: Free Press; 2003.
2. Valente TW, Davis RL. Accelerating the diffusion of innovations using opinion leaders. Ann Am Acad Pol Soc Sci 2016 Sep 8;566(1):55-67. [doi: 10.1177/000271629956600105]
3. Valente TW. Network interventions. Science 2012 Jul 6;337(6090):49-53. [doi: 10.1126/science.1217330] [Medline: 22767921]
4. Campbell R, Starkey F, Holliday J, Audrey S, Bloor MJ, Parry-Langdon N, et al. An informal school-based peer-led intervention for smoking prevention in adolescence (ASSIST): a cluster randomised trial. Lancet 2008 May 10;371(9624):1595-1602 [FREE Full text] [doi: 10.1016/S0140-6736(08)60692-3] [Medline: 18468543]
5. Smit CR, de Leeuw RN, Bevelander KE, Burk WJ, Buijzen M. A social network-based intervention stimulating peer influence on children's self-reported water consumption: a randomized control trial. Appetite 2016 Dec 1;103:294-301. [doi: 10.1016/j.appet.2016.04.011] [Medline: 27085637]
6. Sebire S, Jago R, Edwards M, Banfield K, Campbell R, Kipping R, et al. The feasibility and acceptability of PLAN-A: A school-based peer-led physical activity intervention for adolescents girls in English secondary school. 2017 Jun 07 Presented at: The International Society of Behavioral Nutrition and Physical Activity; 07/06/2017; Victoria.

XSL•FO
RenderX

7. van Woudenberg TJ, Bevelander KE, Burk WJ, Smit CR, Buijs L, Buijzen M. A randomized controlled trial testing a social network intervention to promote physical activity among adolescents. BMC Public Health 2018 Dec 23;18(1):542 [FREE Full text] [doi: 10.1186/s12889-018-5451-4] [Medline: 29685112]
8. Valente TW, Fosados R. Diffusion of innovations and network segmentation: the part played by people in promoting health. Sex Transm Dis 2006 Jul;33(7 Suppl):S23-S31. [doi: 10.1097/01.olq.0000221018.32533.6d] [Medline: 16794552]
9. Badham J, Kee F, Hunter RF. Simulating network intervention strategies: implications for adoption of behaviour. Netw Sci 2018 May 16;6(2):265-280. [doi: 10.1017/nws.2018.4]
10. Zhang J, Shoham DA, Tesdahl E, Gesell SB. Network interventions on physical activity in an afterschool program: an agent-based social network study. Am J Public Health 2015 Apr;105(Suppl 2):S236-S243. [doi: 10.2105/AJPH.2014.302277] [Medline: 25689202]
11. Hammond RA. Social influence and obesity. Curr Opin Endocrinol Diabetes Obes 2010 Oct;17(5):467-471. [doi: 10.1097/MED.0b013e32833d4687] [Medline: 20689421]
12. Beheshti R, Jalalpour M, Glass TA. Comparing methods of targeting obesity interventions in populations: an agent-based simulation. SSM Popul Health 2017 Dec;3:211-218 [FREE Full text] [doi: 10.1016/j.ssmph.2017.01.006] [Medline: 29349218]
13. El-Sayed AM, Seemann L, Scarborough P, Galea S. Are network-based interventions a useful antiobesity strategy? An application of simulation models for causal inference in epidemiology. Am J Epidemiol 2013 Jul 15;178(2):287-295 [FREE Full text] [doi: 10.1093/aje/kws455] [Medline: 23696107]
14. Trogdon JG, Nonnemaker J, Pais J. Peer effects in adolescent overweight. J Health Econ 2008 Sep;27(5):1388-1399. [doi: 10.1016/j.jhealeco.2008.05.003] [Medline: 18565605]
15. Yang Y, Roux AV, Auchincloss AH, Rodriguez DA, Brown DG. A spatial agent-based model for the simulation of adults' daily walking within a city. Am J Prev Med 2011 Mar;40(3):353-361 [FREE Full text] [doi: 10.1016/j.amepre.2010.11.017] [Medline: 21335269]
16. Yang Y, Diez-Roux AV. Using an agent-based model to simulate children's active travel to school. Int J Behav Nutr Phys Act 2013 May 26;10:67 [FREE Full text] [doi: 10.1186/1479-5868-10-67] [Medline: 23705953]
17. Baker CW, Little TD, Brownell KD. Predicting adolescent eating and activity behaviors: the role of social norms and personal agency. Health Psychol 2003 Mar;22(2):189-198. [doi: 10.1037/0278-6133.22.2.189] [Medline: 12683739]
18. Widener MJ, Metcalf SS, Bar-Yam Y. Agent-based modeling of policies to improve urban food access for low-income populations. Appl Geogr 2013 Jun;40(2013):1-10. [doi: 10.1016/j.apgeog.2013.01.003]
19. Zhang J, Tong L, Lamberson PJ, Durazo-Arvizu RA, Luke A, Shoham DA. Leveraging social influence to address overweight and obesity using agent-based models: the role of adolescent social networks. Soc Sci Med 2015 Jan;125:203-213 [FREE Full text] [doi: 10.1016/j.socscimed.2014.05.049] [Medline: 24951404]
20. Araujo EF, Simoski B, van Woudenberg TJ, Bevelander KB, Smit CR, Buijs L, et al. Using Simulations for Exploring Interventions in Social Networks Modeling Physical Activity Behaviour in Dutch School Classes. In: Proceedings of the 8th International Conference on Simulation and Modeling Methodologies, Technologies and Applications. 2018 Presented at: SIMULTECH'18; July 29-31, 2018; Porto, Portugal URL: http://www.insticc.org/Primoris/Resources/PaperPdf.ashx?idPaper=68577 [doi: 10.5220/0006857704140425]
21. Giabbanelli PJ, Alimadad A, Dabbaghian V, Finegood DT. Modeling the influence of social networks and environment on energy balance and obesity. J Comput Sci 2012 Jan;3(1-2):17-27. [doi: 10.1016/j.jocs.2012.01.004]
22. Bevelander KE, Smit CR, van Woudenberg TJ, Buijs L, Burk WJ, Buijzen M. Youth's social network structures and peer influences: study protocol MyMovez project - phase I. BMC Public Health 2018 Dec 16;18(1):504 [FREE Full text] [doi: 10.1186/s12889-018-5353-5] [Medline: 29661223]
23. Torsheim T, Cavallo F, Levin KA, Schnohr C, Mazur J, Niclasen B, FAS Development Study Group. Psychometric validation of the revised family affluence scale: a latent variable approach. Child Indic Res 2016;9(3):771-784 [FREE Full text] [doi: 10.1007/s12187-015-9339-x] [Medline: 27489572]
24. Starkey F, Audrey S, Holliday J, Moore L, Campbell R. Identifying influential young people to undertake effective peer-led health promotion: the example of a stop smoking in schools trial (ASSIST). Health Educ Res 2009 Dec;24(6):977-988. [doi: 10.1093/her/cyp045] [Medline: 19684123]
25. Freeman LC. Centrality in social networks conceptual clarification. Soc Netw 1978 Jan;1(3):215-239. [doi: 10.1016/0378-8733(78)90021-7]
26. Borgatti SP. Centrality and network flow. Soc Netw 2005 Jan;27(1):55-71. [doi: 10.1016/j.socnet.2004.11.008]
27. Valente TW. Network Models of the Diffusion of Innovations. New York: Hampton Press; 1995.
28. Borgatti SP. Identifying sets of key players in a social network. Comput Math Organ Theory 2006 Apr;12(1):21-34. [doi: 10.1007/s10588-006-7084-x]
29. Salmon J, Ball K, Hume C, Booth M, Crawford D. Outcomes of a group-randomized trial to prevent excess weight gain, reduce screen behaviours and promote physical activity in 10-year-old children: switch-play. Int J Obes (Lond) 2008 Apr;32(4):601-612. [doi: 10.1038/sj.ijo.0803805] [Medline: 18253162]
30. Valente TW, Pumpuang P. Identifying opinion leaders to promote behavior change. Health Educ Behav 2007 Dec;34(6):881-896. [doi: 10.1177/1090198106297855] [Medline: 17602096]

XSL•FO
RenderX

31. Gesell SB, Barkin SL, Valente TW. Social network diagnostics: a tool for monitoring group interventions. Implement Sci 2013 Oct 1;8:116 [FREE Full text] [doi: 10.1186/1748-5908-8-116] [Medline: 24083343]
32. Sparrowe RT, Liden RC, Wayne SJ, Kraimer ML. Social networks and the performance of individuals and groups. Acad Manag J 2001 Apr 1;44(2):316-325. [doi: 10.2307/3069458]
33. Giordano PC, Cernkovich SA, Pugh MD. Friendships and delinquency. Am J Sociol 1986 Mar;91(5):1170-1202. [doi: 10.1086/228390]
34. Ahn HJ, Rodkin PC. Classroom-level predictors of the social status of aggression: friendship centralization, friendship density, teacher–student attunement, and gender. J Educ Psychol 2014 Nov;106(4):1144-1155. [doi: 10.1037/a0036091]
35. Alharbi M, Bauman A, Neubeck L, Gallagher R. Validation of Fitbit-Flex as a measure of free-living physical activity in a community-based phase III cardiac rehabilitation population. Eur J Prev Cardiol 2016 Dec;23(14):1476-1485. [doi: 10.1177/2047487316634883] [Medline: 26907794]
36. Diaz KM, Krupka DJ, Chang MJ, Peacock J, Ma Y, Goldsmith J, et al. Fitbit®: an accurate and reliable device for wireless physical activity tracking. Int J Cardiol 2015 Apr 15;185:138-140 [FREE Full text] [doi: 10.1016/j.ijcard.2015.03.038] [Medline: 25795203]
37. van Buuren S. Multiple imputation of multilevel data. In: Hox JJ, Roberts KJ, editors. Handbook of Advanced Multilevel Analysis. New York: Routledge Academic; 2010:173-196.
38. Boyce W, Torsheim T, Currie C, Zambon A. The family affluence scale as a measure of national wealth: validation of an adolescent self-report measure. Soc Indic Res 2006 Apr 20;78(3):473-487. [doi: 10.1007/s11205-005-1607-6]
39. Brechwald WA, Prinstein MJ. Beyond homophily: a decade of advances in understanding peer influence processes. J Res Adolesc 2011 Mar 1;21(1):166-179 [FREE Full text] [doi: 10.1111/j.1532-7795.2010.00721.x] [Medline: 23730122]
40. Salvy SJ, de la Haye K, Bowker JC, Hermans RC. Influence of peers and friends on children's and adolescents' eating and activity behaviors. Physiol Behav 2012 Jun 6;106(3):369-378 [FREE Full text] [doi: 10.1016/j.physbeh.2012.03.022] [Medline: 22480733]
41. van Rossum G, Drake FL. Python.org. 2003. The Python Language Reference URL: https://docs.python.org/3/reference/
42. Hagberg AA, Schult DA, Swart PJ. Exploring Network Structure, Dynamics, and Function using NetworkX. In: Proceedings of the 7th Python in Science Conference. 2008 Presented at: SciPy'08; August 19-24, 2008; Pasadena, USA p. 11-15 URL: https://conference.scipy.org/proceedings/scipy2008/paper_2/full_text.pdf
43. RStudio: Integrated Development for R. 2015. URL: https://www.rstudio.com/
44. Bandini S, Manzoni S, Vizzari G. Agent based modeling and simulation: an informatics perspective. J Artif Soc Soc Simul 2009;12(4):4 [FREE Full text]
45. McCormack GR, Hawe P, Perry R, Blackstaffe A. Associations between familial affluence and obesity risk behaviours among children. Paediatr Child Health 2011 Jan;16(1):19-24 [FREE Full text] [doi: 10.1093/pch/16.1.19] [Medline: 22211068]
46. Logue E, Sutton K, Jarjoura D, Smucker W, Baughman K, Capers C. Transtheoretical model-chronic disease care for obesity in primary care: a randomized trial. Obes Res 2005 May;13(5):917-927 [FREE Full text] [doi: 10.1038/oby.2005.106] [Medline: 15919846]
47. Bates DM. The Comprehensive R Archive Network. 2010. Lme4: Linear Mixed-effects Models Using 'Eigen' and S4 URL: https://cran.r-project.org/web/packages/lme4/index.html [accessed 2019-07-11]
48. Luke SG. Evaluating significance in linear mixed-effects models in R. Behav Res Methods 2017 Dec;49(4):1494-1502. [doi: 10.3758/s13428-016-0809-y] [Medline: 27620283]
49. Bahr DB, Browning RC, Wyatt HR, Hill JO. Exploiting social networks to mitigate the obesity epidemic. Obesity (Silver Spring) 2009 Apr;17(4):723-728 [FREE Full text] [doi: 10.1038/oby.2008.615] [Medline: 19148124]
50. Cooper AR, Andersen LB, Wedderkopp N, Page AS, Froberg K. Physical activity levels of children who walk, cycle, or are driven to school. Am J Prev Med 2005 Oct;29(3):179-184. [doi: 10.1016/j.amepre.2005.05.009] [Medline: 16168866]

Abbreviations

ABM: agent-based model
FAS: Family Affluence Scale
PAL: physical activity level

XSL•FO
RenderX

Edited by G Eysenbach; submitted 26.11.18; peer-reviewed by R Hunter, K de la Haye; comments to author 31.03.19; revised version received 15.05.19; accepted 31.05.19; published 05.08.19.

Please cite as:
van Woudenberg TJ, Simoski B, Fernandes de Mello Araújo E, Bevelander KE, Burk WJ, Smit CR, Buijs L, Klein M, Buijzen M Identifying Influence Agents That Promote Physical Activity Through the Simulation of Social Network Interventions: Agent-Based Modeling Study
J Med Internet Res 2019;21(8):e12914
URL: https://www.jmir.org/2019/8/e12914/
doi:10.2196/12914
PMID:31381504

©Thabo J van Woudenberg, Bojan Simoski, Eric Fernandes de Mello Araújo, Kirsten E Bevelander, William J Burk, Crystal R Smit, Laura Buijs, Michel Klein, Moniek Buijzen. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 05.08.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

A Web-Based and Print-Delivered Computer-Tailored Physical Activity Intervention for Older Adults: Pretest-Posttest Intervention Study Comparing Delivery Mode Preference and Attrition

Janet Maria Boekhout[1], MSc; Denise Astrid Peels[1], PhD; Brenda Angela Juliette Berendsen[1], PhD; Catherine Bolman[1], PhD; Lilian Lechner[1], PhD

Department of Psychology and Educational Science, Open University of the Netherlands, Heerlen, Netherlands

Corresponding Author:
Janet Maria Boekhout, MSc
Department of Psychology and Educational Science
Open University of the Netherlands
Valkenburgerweg 177
Heerlen, 6401 DL
Netherlands
Phone: 31 455762448
Email: janet.boekhout@ou.nl

Abstract

Background: Web-based interventions can play an important role in promoting physical activity (PA) behavior among older adults. Although the effectiveness of these interventions is promising, they are often characterized by low reach and high attrition, which considerably hampers their potential impact on public health.

Objective: The aim of this study was to identify the participant characteristics associated with the preference for a Web-based or a printed delivery mode and to determine whether an association exists between delivery modes or participant characteristics and attrition in an intervention. This knowledge may enhance implementation, sustainability of participation, and effectiveness of future interventions for older adults.

Methods: A real-life pretest-posttest intervention study was performed (N=409) among community-living single adults who were older than 65 years, with physical impairments caused by chronic diseases. Measurements were taken at baseline and 3 months after the start of the intervention. Hierarchical logistic regression was used to assess demographic and behavioral characteristics (age, gender, body mass index, educational attainment, degree of loneliness, and PA level), as well as psychosocial characteristics (social support for PA, modeling, self-efficacy, attitude, and intention) related to delivery mode preference at baseline and attrition after 3 months.

Results: The printed delivery mode achieved higher participation (58.9%, 241/409) than the Web-based delivery mode (41.1%, 168/409). Participation in the Web-based delivery mode was associated with younger age (B=–0.10; SE 0.02; Exp (B)=0.91; *P*<.001) and higher levels of social support for PA (B=0.38; SE 0.14; Exp (B)=1.46; *P*=.01); attrition was associated with participation in the Web-based delivery mode (B=1.28; SE 0.28; Exp (B)=3.58; *P*<.001) and low educational attainment (B=–0.53; SE 0.28; Exp (B)=0.59; *P*=.049).

Conclusions: A total of 41% of the participants chose the Web-based delivery mode, thus demonstrating a potential interest of single older adults with physical impairments in Web-based delivered interventions. However, attrition was demonstrated to be higher in the Web-based delivery mode, and lower educational attainment was found to be a predictor for attrition. Characteristics predicting a preference for the printed delivery mode included being older and receiving less social support. Although Web-based delivery modes are generally less expensive and easier to distribute, it may be advisable to offer a printed delivery mode alongside a Web-based delivery mode to prevent exclusion of a large part of the target population.

Trial Registration: Netherlands Trial Register NTR2297; https://www.trialregister.nl/trial/2173

International Registered Report Identifier (IRRID): RR2-DOI: 10.2196/resprot.8093

(J Med Internet Res 2019;21(8):e13416) doi:10.2196/13416

XSL•FO
RenderX

KEYWORDS

older adults; physical activity; Web-based intervention; print-delivered intervention; attrition

Introduction

Background

A majority of older adults in Western societies are not sufficiently physically active [1]. With societies ageing rapidly this poses a major public health concern, as insufficient physical activity (PA) is regarded as a major health risk. [2]. PA interventions that are cost effective in targeting this health risk are thus needed. Computer-tailored interventions designed to improve PA, delivered either via the Web or a printed delivery mode, have therefore become increasingly popular in the last decade, with substantial evidence supporting their effectiveness and cost-effectiveness [3-5]. However, computer-tailored interventions delivered on the Web may suffer from low reach (ie, the proportion of people participating in an intervention), low long-term adherence (ie, proportion of participants not following the intervention as intended), and high attrition (ie, the proportion of participants not completing the intervention) [6], thereby impeding the effectiveness of the interventions [7-9] and potentially decreasing their intended impact on public health. Web-based delivery modes can be especially challenging among older adults (older than 65 years). Current data on internet use among older adults show that despite the increasing popularity of Web-based interventions, a digital divide still exists [10], with only 50% of older adults regularly using the internet; a total of 34% of this age group express that they have little-to-no confidence to use the internet properly. Moreover, only 15% of adults older than 70 years use the internet for health applications [11,12]. Thus, alternative delivery modes should be provided for this population. Not only a higher age but also a lower educational attainment are characteristics of those on the disadvantaged side of the digital divide. Therefore, the individuals who are the most in need of health applications may not be reachable with interventions delivered via the Web [10]. Some physical limitations that older adults are confronted with, such as impaired eyesight, hearing, and dexterity, are inherent in this age group, and these may interfere with internet use in future generations of older adults [13]. Because offering interventions on the Web alone may exclude a vulnerable group, providing insight on how the (demographic and psychosocial) characteristics of older adults are associated with a preference for a printed or Web-based delivery mode and with attrition can contribute to stimulating reach and sustainability of such interventions. Moreover, it may help to explain the effectiveness of interventions and the impact on the public health of certain subpopulations [7,9,14].

Previously, the computer-tailored Active Plus65 intervention was developed with a primary aim to stimulate PA and a secondary aim to decrease loneliness [15]. The target population of Active Plus65 was single older adults with a physical impairment caused by a chronic disease, considering the high prevalence of insufficient PA and loneliness among this population [16-18]. Active Plus65 is available in 2 delivery modes, that is, Web-based (with both questionnaire and advice delivered through the internet) and printed (with a paper questionnaire and advice sent by paper post). Because interventions are often delivered in a single delivery mode, such as either on the Web or printed, it may be difficult to establish whether reach and attrition are related to the intervention itself or to its delivery mode. Apart from some technical applications, the content of the Web-based and printed versions of the Active Plus65 intervention was the same. This allowed a comparison of the Web-based and printed delivery mode of the computer-tailored Active Plus65 intervention.

Objectives

There is a paucity of research determining the preferred and actually used delivery mode by really offering participants a choice in a real-life intervention setting. Previously, preferences have predominantly been researched by asking participants, hypothetically, what delivery mode they would prefer if they were to join an intervention [19-22]. As far as could be determined, there have been only 2 studies pertaining to delivery mode preference, where participants could choose between participating in a printed and Web-based delivery mode of an intervention [23,24]. In these studies, participants with lower education and older age chose the printed delivery mode. However, no straightforward comparisons can be made because the interventions in both studies were designed for different target groups compared with that of Active Plus65. However, the preference of a printed delivery mode in older and lower educated adults is also consistent with data on the present digital divide [25].

In 2012, the delivery mode preference and attrition of Active Plus50, a previous version of Active Plus65, were studied [26]. In Active Plus65, participants were free to choose between delivery modes, whereas participants in Active Plus50 were randomly assigned to either the printed or Web-based delivery mode. Active Plus50 participants in the Web-based delivery mode were younger, more often male, had a higher body mass index (BMI), and a lower intention to be physically active. Moreover, a low intention to be physically active was also found to be a predictor for attrition in both delivery modes. Because intention had a predictive value for delivery mode preference and attrition in Active Plus50, it may be interesting to repeat these analyses in Active Plus65: because participants could not freely choose their delivery mode in Active Plus50, the role of intention in Active Plus65, where participants can freely choose the delivery mode, may be different. Considering intention had a certain predictive value for delivery mode preference and attrition, other psychosocial variables, which according to several behavior change models are important determinants of intention [27-29], could also have a predictive value. Because low intention was encountered in the Web-based delivery mode, and it predicted attrition in both delivery modes, it can be argued that a negative attitude toward PA, low social support, modeling, and self-efficacy (psychosocial determinants associated with behavior change) may also be associated with a preference for the Web-based delivery mode and with attrition. After all, in behavior change theories, these are often the major determinants of behavioral intention that are strongly related to predicting

XSL•FO
RenderX

behavior change. Such knowledge may be useful in the enrollment process of the intervention to ensure that participants join a delivery mode that is best suited for them; if, for example, our findings would establish a preference for a printed delivery mode in participants with low levels of social support, emphasizing on how easy it is to use the printed delivery mode or to point out that a helpdesk is available for Web-based participants at enrollment would be useful. Thus, participants with little social support who are hesitant about taking part in an intervention could be assessed with a short pre-enrollment questionnaire, thereby preventing them from not starting an intervention at all. Support for the incorporation of a broad range of (psychosocial) individual characteristics in the study can also be found in the Persuasion-Communication Matrix of McGuire [30]. This model describes several factors that predict whether an individual will use an intervention, among which the broad characteristics of the user are also included. For example, the approach required to stimulate a potential user with low self-efficacy and low social support to be physically active to join an intervention will vary from that for an individual with high self-efficacy and social support for health behavior. When trying to identify the characteristics that predict preference for a delivery mode, it may be interesting to focus on a broad range of determinants, including psychosocial determinants.

As Active Plus65 targets both PA and loneliness, it presents an opportunity to analyze whether the baseline level of PA and loneliness can predict delivery mode preference and attrition because this could link target subgroups to the most appropriate delivery mode. Only 2 studies have been found that considered PA to analyze delivery mode preference [20] or attrition [31]. In these studies, a higher level of PA was associated with a lower likelihood of preferring Web-based delivery modes and lower attrition. Comparisons may be difficult to make: either the Web-based delivery mode was not compared with a printed delivery mode but compared with face-to-face or group interventions instead [20] or only daily steps were measured [31] rather than a broad range of PA, which Active Plus65 does. Moreover, these studies were performed among a general population of adults. As far as could be determined, no previous research has considered loneliness when researching delivery mode and attrition of PA interventions for older adults; however, some assumptions may be made. In general, social isolation and loneliness are more prevalent among older adults [32,33]. Some studies have shown that social support is a prerequisite to adopt new technologies [34,35], which would suggest that participants who are lonelier will not choose a Web-based delivery mode. Considering the lack of studies for comparison, the analyses regarding PA and loneliness in our study will thus have a more exploratory character.

The aim of this study was to determine (1) which individual characteristics predict differences in delivery mode preference between the printed or Web-based delivery mode and (2) which user characteristics and delivery mode predict attrition. We hypothesized that a higher age, lower educational status, and lower presence of psychosocial determinants are predictors of a preference for a printed delivery mode and of attrition. Identifying the factors related to delivery mode preference and attrition could be of substantial use for researchers when optimizing the reach and sustainability of interventions because this could increase their impact on public health and prevent an important target population from being excluded when switching to only Web-based delivery prematurely.

Methods

Study Design

This study was part of a pretest-posttest trial, evaluating the Active Plus65 intervention [15,36,37]. The trial was executed in a real-life setting, without a control group. As the delivery mode preference and attrition of the participants in the printed and Web-based delivery mode of the intervention were compared, no control group was required for this study. This study was approved by the Research Ethics Committee of the Open University of the Netherlands (reference number U2016/02373/HVM). The original Active Plus50 studies were registered at the Netherlands Trial Register (NTR2297). All participants gave their informed consent before participation.

Intervention

Active Plus65 is a computer-tailored intervention, with a primary aim to stimulate PA and secondary aim to decrease loneliness among single older adults with physical impairments caused by a chronic disease. Active Plus65 was systematically developed [15], and changes in PA [36] and loneliness [37] have been demonstrated.

The advice is generated by computer tailoring; in the Web-based delivery mode, participants fill the questionnaire themselves on the intervention website, which in the printed delivery mode is done by the intervention providers after receiving the questionnaire by mail from the participant. The method and degree of tailoring in both delivery modes is identical; therefore, the advice in both delivery modes has identical content, with only some practical differences. For example, in the printed version, modeling texts and pictures are used versus modeling videos in the Web-based version: the role models that are portrayed are the same persons delivering exactly the same message. Moreover, the design and format, such as images, typeface, and layout, of both questionnaires and advice are identical. A screenshot of the intervention website is provided in Figure 1.

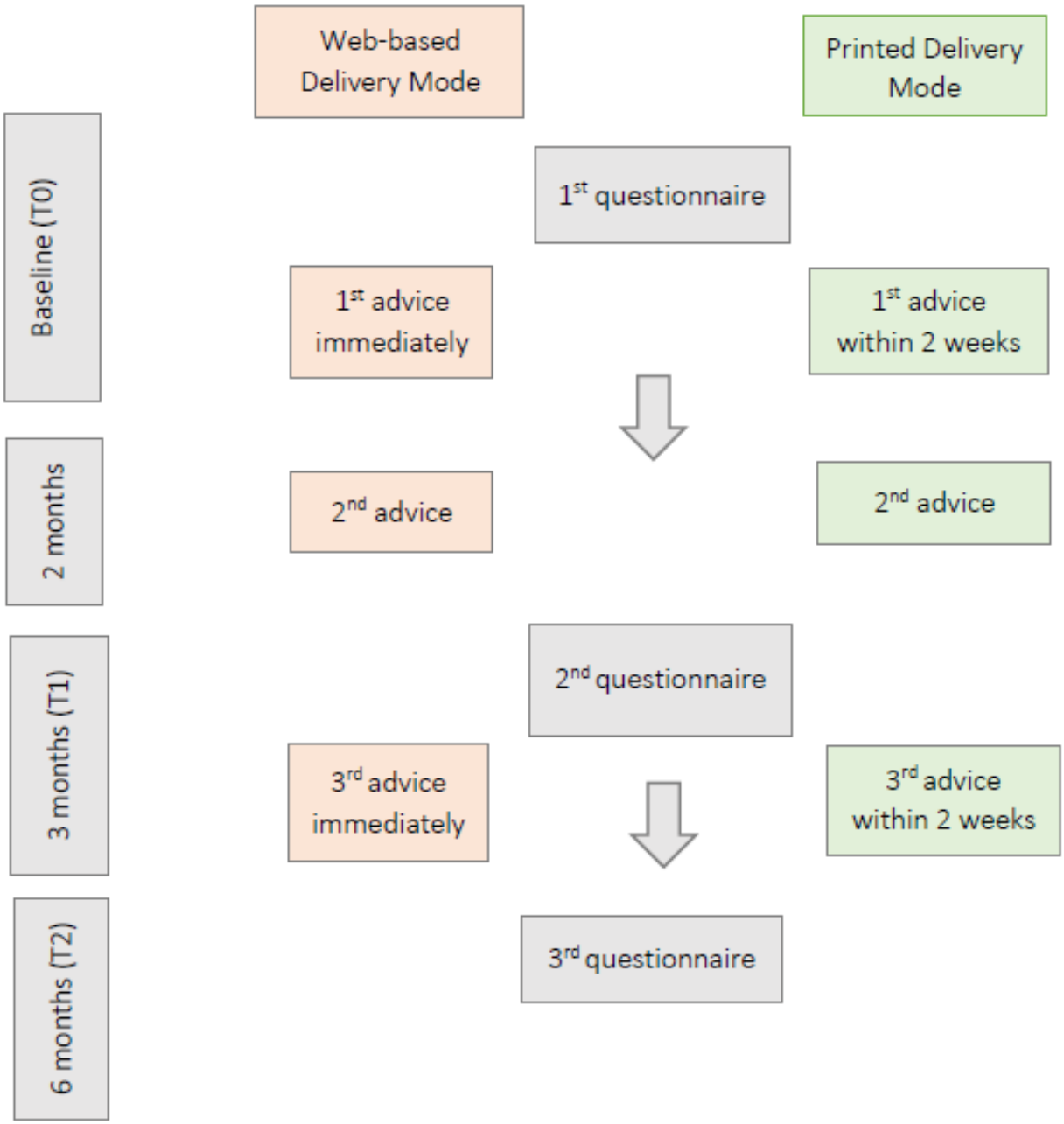
Both delivery modes provide a tailored advice at 3 time points (at the start of the intervention, 2 months after the start, and 3 months after the start), on the basis of 2 questionnaires, the first one at the start (T0) of the intervention (on which advice 1 and 2 are based) and the second one after 3 months (T1; for advice 3). A third questionnaire, 6 months after the start of the intervention (T2), does not result in an advice, but it serves as a follow-up measurement. The time needed to fill out the questionnaire is identical for both delivery modes, that is, about 15 min for the first and second questionnaire and 5 min for the third questionnaire. Figure 2 provides a schematic view of the intervention timelines.

XSL•FO
RenderX

Figure 1. The intervention website.

XSL•FO
RenderX

Figure 2. Intervention timelines per delivery mode.



Each advice is tailored to the characteristics that are assessed in the questionnaire, including demographic and psychosocial determinants, the amount of PA, and the degree of loneliness. The first advice aims to increase insights into the present level of PA, which is achieved by targeting premotivational psychosocial determinants, such as knowledge and awareness. The second advice motivates participants to become more physically active and focuses on the benefits of PA, especially when done with others: in this advice, motivational psychosocial determinants, such as attitude, intention, and social influence, are targeted. Participants are also stimulated to prepare for difficult situations that might hinder them from becoming more active. The intervention also stimulates participants to transfer their motivation into sustainable behavior: depending on how active participants already are, this is done in the second or third advice by targeting postmotivational determinants, such as strategic planning and coping planning. Depending on the assessment, the advice comprises 7 to 12 pages (A4 format) of text, pictures, diagrams, etc, for (passive) reading (Multimedia Appendix 1) and elements that require participants to actively contribute, such as planning sheets that have to be filled, schemes for handling difficult situations and formulating implementation intentions, and basic PA exercises to be performed at home (Multimedia Appendix 2). The organization that implements the intervention (usually a local council) provides information on PA in social meeting opportunities that are available in the area where the participant lives. A more extensive description is provided elsewhere [15,36,37].

XSL•FO
RenderX

Participants and Procedures

All citizens of a Dutch municipality, in the southern part of the Netherlands, who were single, older than 65 years, and living independently in the community (n=6751) were recruited by direct mailing. In this mail communication, it was explained that Active Plus65 is specifically suited for participants who have physical impairments caused by chronic diseases. For this study, only participants with a physical impairment were included. Invitations were sent by personalized letter and contained information about the intervention. Both log-in details for those wishing to participate via internet and a prepaid response card for requesting a paper questionnaire were included. It was mentioned that personal assistance with using internet when filling the questionnaire on the Web was available upon request. Only participants who completed the baseline questionnaire were enrolled. For the second assessment after 3 months, a printed invitation letter with the questionnaire and prepaid response envelope was sent to all participants who had completed the first questionnaire by the printed delivery mode. The participants in the Web-based delivery mode received an invitation by email for the follow-up questionnaires, with a direct link to the Web-based questionnaire. After 6 months, a third questionnaire was sent following the same procedure as for the second questionnaire.

Measurements

Demographic characteristics, psychosocial determinants, the amount of PA, and the degree of loneliness were assessed at baseline. The second questionnaire assessed the same variables, except for the demographic characteristics considering their stable qualities. The third questionnaire assessed only the amount and type of PA and the degree of loneliness. Other variables were also assessed, but as they are outside the scope of this study, they are not discussed here.

Demographic Characteristics

The assessed demographic characteristics were age, gender, height, weight, educational attainment, and presence of physical impairments caused by chronic diseases. Height and weight were used to calculate the BMI by dividing weight in kilograms by height in meters squared. Educational attainment was categorized into *low* (lower vocational education, medium general secondary education, secondary vocational education, and higher general secondary education) and *high* (higher vocational education and university education). The presence of physical impairments was categorized into *yes* or *no*.

Psychosocial Determinants

The assessed psychosocial determinants were attitude, modeling, social support, self-efficacy, and intention to be sufficiently physically active. Attitude to be sufficiently physically active was measured by 17 items (eg, *PA gives me a satisfied feeling*) on a 5-point scale (1=totally disagree to 5=totally agree). Modeling was measured by asking *Are the following persons physically active for at least 30 min per day on at least 5 days per week?* in 2 items, 1 for family and 1 for friends, on a 5-point scale (1=never to 5=always). Social support for PA was also measured by 2 items (1 relating to family, 1 relating to friends) by asking *To what degree do you expect to get support to be sufficiently physically active?* —with answers on a 5-point scale (1=never to 5=always). Self-efficacy was measured by asking to what degree one would manage to be physically active for at least 30 minutes per day for different situations (eg, *when the weather is bad*), with 11 items on a 5-point scale (1=definitely not to 5=definitely sure). The intention for performing sufficient PA was measured by 3 items on a 10-point scale (eg, *How likely do you think it is that you will stay or become sufficiently physically active?*). The scales to assess the psychosocial variables were based on validated questionnaires [38-42], and their usability has been demonstrated by pilot tests among the target population [15].

Physical Activity

The amount of PA was assessed with the Short Questionnaire to Assess Health Enhancing Physical Activity (SQUASH) [43]. This questionnaire assesses the amount and intensity of different types of PA performed during (volunteering) work, commuting, household, and leisure time. It allows for calculating the total minutes per week of PA performed with moderate to vigorous intensity (MVPA) [2]. The psychometric properties of the SQUASH have been found to be acceptable [43-45].

Loneliness

Loneliness was assessed with the De Jong Gierveld 6-item Loneliness Scale, whose psychometric properties have been found to be acceptable [46]. This scale has 6 items (eg, *I often feel rejected*), on an originally 6-point scale, but it was adapted to a 10-point scale (1=*absolutely not* and 10=*absolutely sure*), as this was deemed more suitable for older adults [47,48]. Items with answer ranges from 6 to 10 (indicating loneliness) are summed, resulting in a potential score of loneliness between 0 (not lonely) and 6 (extremely lonely).

Statistical Analyses

Differences in Delivery Mode Preference

Statistical analyses were performed on those participants who completed the baseline questionnaire. Univariate one-way analyses of variance and chi-square tests were performed on age, gender, educational attainment, BMI, PA, loneliness, modeling, social support, attitude, self-efficacy, and intention to assess whether participants differed at baseline between the printed and Web-based delivery mode. Hierarchical logistic regression was performed on the T0 data to identify the user characteristics that predict differences in the preference for the printed or Web-based delivery mode. Outcome measure was the dichotomous variable of delivery mode. Step 1 of the analyses contained the demographic variables and the weekly minutes of MVPA and loneliness. In step 2, the psychosocial determinants were added. In step 3, interaction terms were added: as previous research provided no directions for formulating hypotheses on potential interaction effects, only interaction effects for the determinants that were significant in step 2 were included (eg, the interaction between age and social support).

Differences in Attrition

Differences in attrition between the printed and Web-based delivery mode were analyzed with a chi-square test. To identify

XSL•FO
RenderX

potential factors related to attrition, hierarchical logistic regression analysis was performed on the T1 data, as this is the time when participants have to fill the second questionnaire, which will provide them with the third advice. Attrition occurs when participants do not fill this questionnaire. The outcome measure is the dichotomous variable of attrition. Demographic characteristics, weekly minutes of MVPA, and degree of loneliness were added in step 1; delivery mode in step 2; and psychosocial determinants in step 3 of the analysis. In step 4, interaction terms of user characteristics and delivery mode were added to assess whether dropout is associated with certain combinations of user characteristics and delivery mode: gender, age, and educational attainment were selected based on previous research [20,23,24].

All significance levels were set at P=.05, except for the final step of the regression analyses where significance was set at P=.10 because interaction terms are known to have less power [49]. SPSS version 24 (IBM Statistical Package for Social Sciences) was used to perform all analyses.

Ethics Approval and Consent to Participate

The study was reviewed and approved by the Committee for Ethics and Consent in Research of the Open University (Commissie Ethische Toetsing Onderzoek, reference number: U2016/02373/HVM). Participants provided written informed consent to participate in the study.

Results

Delivery Mode Distribution

Of all eligible participants of Active Plus65, 241 (58.9%, 241/409) participants took part in the printed delivery mode, and 168 (41.1%, 168/409) took part in the Web-based delivery mode. In total, a response rate of 6% of invited participants was realized. Figure 3 provides an overview.

Delivery Mode Preference

Several of the baseline characteristics of the printed and Web-based group differed significantly (Table 1).

Participants in the printed delivery mode were older (P<.001), had a lower educational attainment (P<.001), were more often female (P=.02), were less physically active (P<.001), had fewer family and friends who are sufficiently physically active as modeling roles (P<.001), received less social support for PA (P<.001), and had a lower intention to be sufficiently physically active (P=.01).

Baseline user characteristics related to delivery mode preference are presented in Table 2.

In step 1, age and degree of loneliness were significant predictors of delivery mode preference. Participants in the printed delivery mode were older (B=–0.10; SE 0.02; Exp (B)=0.91; P<.001) and lonelier (B=–0.14; SE 0.06; Exp (B)=0.87; P=.03) than the participants who chose the Web-based delivery mode. When entering the psychosocial variables to the analyses, age was still a significant predictor (B=–0.10; SE 0.02; Exp (B)=0.91; P<.001), but loneliness became nonsignificant (P=.16) and social support for PA then emerged as a significant predictor (B=0.38; SE 0.14; Exp (B)=1.46; P=.01) with participants in the Web-based delivery mode having higher levels of social support than those in the printed delivery mode. Explained variance (R^2) in the steps ranged between 0.15 and 0.19. The interaction in step 3 between age and social support was not significant (P=.34), indicating that the effect of age on delivery mode preference did not differ depending on the degree of social support.

XSL•FO
RenderX

Figure 3. Flow chart of reach and attrition in printed and Web-based delivery mode. "a" reported as percentage of invited participants; "b" reported as percentage of all participants completing T0 questionnaire; "c" reported as percentage of all participants starting in either printed or Web-based delivery mode; asterisk indicates that eligible participants are those who meet all requirements of being single, over the age of 65, and chronically impaired in physical activity.

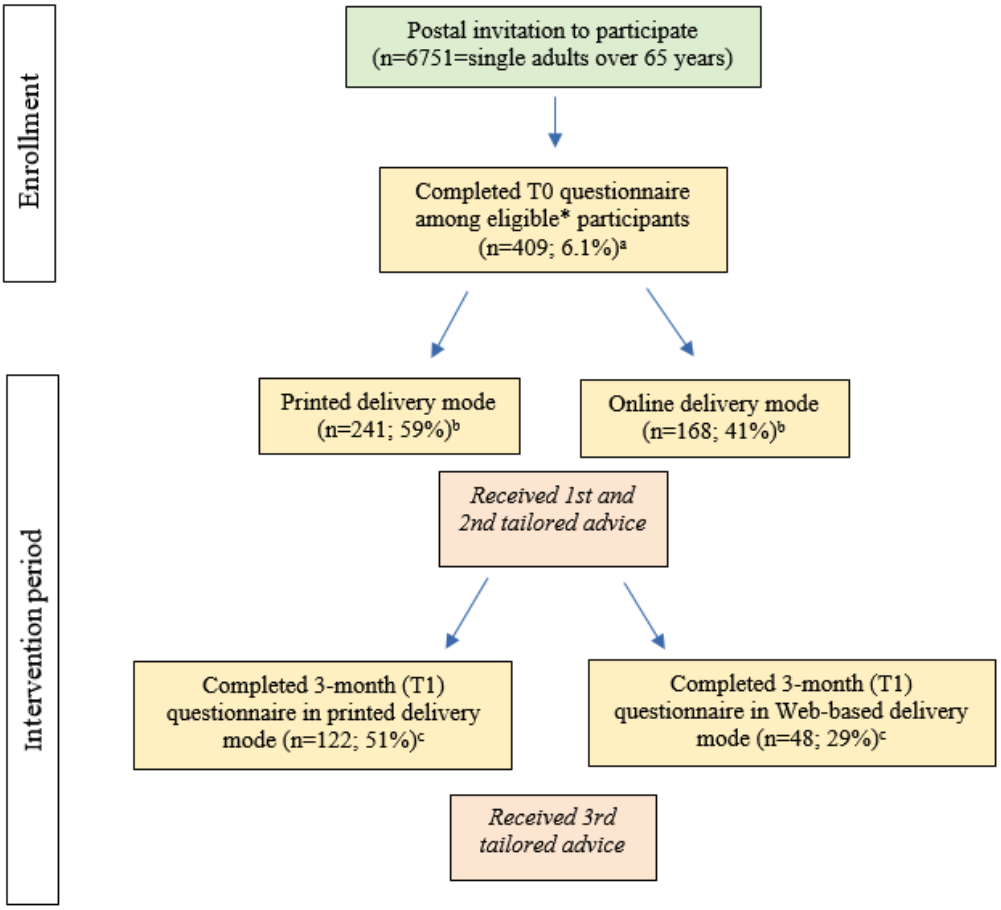


Table 1. Baseline characteristics of participants in printed and Web-based delivery mode.

Determinants	Printed (n=241)	Web-based (n=168)	*P* value
Age (years), mean (SD)	79.22 (7.59)	73.29 (6.61)	<.001
Body mass index (kg/m^2), mean (SD)	27.01 (4.89)	27.82 (5.27)	.12
Educational attainment (low), n (%)	167 (70.8)	83 (51.2)	<.001
Gender (men), n (%)	74 (30.7)	71 (42.3)	.02
Moderate-to-vigorous physical activity, mean (SD)	387.72 (527.11)	606.51 (687.41)	<.001
Loneliness, mean (SD)	3.2 (1.98)	2.94 (2.06)	.21
Modeling, mean (SD)	2.55 (1.49)	3.07 (1.14)	<.001
Social support, mean (SD)	1.79 (1.27)	2.33 (1.07)	<.001
Attitude, mean (SD)	3.42 (0.57)	3.49 (0.59)	.28
Self-efficacy, mean (SD)	3.47 (0.98)	3.42 (0.96)	.58
Intention, mean (SD)	6.51 (1.61)	6.99 (1.73)	.01

XSL•FO
RenderX

Table 2. Hierarchical logistic regression to study whether user characteristics predict differences in delivery mode preference.

Determinants	Step 1 (R^2=0.152)[a]				Step 2 (R^2 =0.186)[a]			
	Exp (B)	B	SE	*P* value	Exp (B)	B	SE	*P* value
First block[b]								
Age	0.91	–0.10	0.02	<.001	0.91	–0.10	0.02	<.001
Gender[c]	0.79	–0.24	0.26	.36	0.71	–0.34	0.27	.21
Education[d]	1.43	0.36	0.26	.17	1.32	0.28	0.27	.30
Body mass index	1.01	0.01	0.02	.56	1.02	0.02	0.03	.48
Moderate-to-vigorous physical activity	1.00	0.00	0.00	.40	1.00	0.00	0.00	.49
Loneliness	0.87	–0.14	0.06	.03	0.91	–0.10	0.07	.16
Second block[b]								
Modeling	—[e]	—	—	—	0.95	–0.05	0.13	.71
Social support	—	—	—	—	1.46	0.38	0.14	.01
Attitude	—	—	—	—	0.84	–0.17	0.28	.54
Self-efficacy	—	—	—	—	0.77	–0.27	0.17	.11
Intention	—	—	—	—	1.15	0.14	0.1	.19

[a]Explained variance (Nagelkerke R^2).

[b]Printed coded 0, Web-based coded 1.

[c]Men coded 0, women coded 1.

[d]Low educational attainment coded 0, high educational attainment coded 1.

[e]Not applicable.

Attrition

Attrition differed significantly between the delivery modes, that is, 50% in the printed delivery mode and 71% in the Web-based delivery mode (*P*<.001). Table 3 provides an overview of the predictors of attrition during the intervention.

The assessed demographic variables in step 1 were all nonsignificant. The delivery mode, added in step 2, was a significant predictor of attrition (B=1.34; SE 0.27; Exp (B)=3.81; *P*<.001), with attrition higher among participants in the Web-based delivery mode than in the printed delivery mode. By adding the psychosocial determinants to the analyses in step 3, delivery mode remained significant (B=1.28; SE 0.28; Exp (B)=3.58; *P*<.001) and educational attainment also became significant (B=–0.53; SE 0.28; Exp (B)=0.59; *P*=.049): attrition was higher among participants in the Web-based delivery mode and among participants with a low educational attainment than for those in the printed delivery mode and for participants with a high educational attainment. Explained variance (R^2) in the steps ranged between 0.02 and 0.12. The interactions that were assessed in step 4 (ie, delivery mode x gender [*P*=.11], delivery mode x age [*P*=.22], and delivery mode x education [*P*=.26]) were all nonsignificant.

XSL•FO
RenderX

Table 3. Hierarchical logistic regression to study whether user characteristics, delivery mode, and interaction predict differences in attrition.

Determinants	Step 1 (R^2=0.023)[a]				Step 2 (R^2=0.101)[a]				Step 3 (R^2=0.118)[a]			
	Exp (B)	B	SE	*P* value	Exp (B)	B	SE	*P* value	Exp (B)	B	SE	*P* value
First block[b]												
Gender[c]	0.86	–0.16	0.25	.53	0.91	–0.09	0.26	.73	0.85	–0.16	0.27	.55
Age	0.97	–0.03	0.02	.11	1.00	0.00	0.02	.92	1.00	0.00	0.02	.94
Education[d]	0.70	–0.36	0.24	.14	0.61	–0.49	0.26	.06	0.59	–0.53	0.28	.049
Body mass index	0.97	–0.04	0.02	.13	0.96	–0.04	0.02	.08	0.96	–0.04	0.03	.09
Moderate-to-vigorous physical activity	1.00	0.00	0.00	.29	1.00	0.00	0.00	.18	1.00	0.00	0.00	.23
Loneliness	0.94	–0.07	0.06	.26	0.97	–0.03	0.06	.64	0.98	–0.02	0.07	.73
Second block[b]												
Delivery mode[e]	—[f]	—	—	—	3.81	1.34	0.27	<.001	3.58	1.28	0.28	<.001
Third block[b]												
Self-efficacy	—	—	—	—	—	—	—	—	0.80	–0.22	0.16	.16
Intention	—	—	—	—	—	—	—	—	0.99	0.00	0.10	.99
Modeling	—	—	—	—	—	—	—	—	1.16	0.15	0.13	.23
Social support	—	—	—	—	—	—	—	—	1.08	0.07	0.14	.60
Attitude	—	—	—	—	—	—	—	—	1.11	0.11	0.27	.69

[a]Explained variance (Nagelkerke R^2).

[b]Nonattrition coded 0, attrition coded 1.

[c]Men coded 0, women coded 1.

[d]Low coded 0, high coded 1.

[e]Printed coded 0, Web-based coded 1.

[f]Not applicable.

Discussion

Principal Findings

This study aimed to determine which user characteristics predict the preference for either a Web-based or printed delivery mode of a PA intervention for single older adults with physical impairments. In addition, this study examined which user characteristics and delivery mode predict attrition. This provides insights into which factors should be considered when designing PA interventions for this target population.

Delivery Mode Preference

A total of 41% of the participants chose to start in the Web-based delivery mode. Although this demonstrated a potential interest of single older adults with physical impairments in Web-based delivered interventions, the majority still preferred a printed delivery mode. This is in agreement with data on the existing digital divide, showing that only 50% of older adults regularly use the internet, with only 15% using health applications [11-13]. These findings corroborate previous research and data suggesting that despite the increase in internet use among older adults over the last decade, it may still take many years for internet delivery mode to be the leading preference among all age groups [50-52]. Therefore, presently, intervention developers should not rule out printed delivery modes for this target population, as this could lead to the exclusion of a large segment of the target population.

Age was found to be a significant predictor of delivery mode preference, with older participants preferring the printed delivery mode more often. This finding is consistent with our hypothesis, as well as with previous research [23,24,26]. This finding is also corroborated by the Unified Theory of Acceptance and Use of Technology (UTAUT) [53]. In this model, performance expectancy, effort expectancy, and social influence determine usage intention, and through intention, they influence behavior. According to UTAUT, it may be that as older adults have less experience with the internet [54,55], they may expect the Web-based delivery mode of Active Plus65 to be more difficult and consequently choose the printed delivery mode; this preference may be enhanced by the social influence of peers who have the same expectations.

In contrast to our hypothesis, educational attainment was not found to be a significant predictor of delivery mode preference. It could be that the lower use of internet for health-enhancing interventions among people with a lower educational attainment [52,56-58] is outweighed by a general increase in availability and use of internet by older adults [54]. Another explanation is

XSL•FO
RenderX

that assistance with internet offered when inviting participants for Active Plus65 could have given less educated participants enough confidence to participate in the Web-based delivery mode. In practice, only 10% of participants used this offer, but other assistance may have been received, such as that by the participants' own social network. This explanation is supported by the review of Kampmeijer et al [59] who found support to be essential to give older adults the confidence to experiment with new technologies. Another explanation may lie in the overrepresentation of the female gender (65%) and low educational attainment (63%) at baseline. As educational opportunities have long been to the disadvantage of women, there's a likelihood of educational attainment being less indicative of intelligence or digital literacy for older women; therefore, it may not have a predictive value for delivery mode.

In step 1 of the exploratory analyses into the predictive value of PA and loneliness on delivery mode preference, only loneliness was found to be a significant predictor: a higher degree of loneliness was found among participants in the printed delivery mode. In step 2 of the analyses, loneliness became nonsignificant, and social support for being physically active emerged as a significant predictor: participants with lower social support for being physically active preferred the printed delivery mode. Possibly, participants who receive less social support for being physically active also receive less social support for other aspects of life, such as for digitalization, thus making them less inclined to participate in a Web-based intervention. This is supported by previous studies that argue that those with close social support receive explanation and encouragement to use new technologies, such as internet, making it easier for them to adopt Web-based interventions [50,60,61]. Policy makers who strive to increase internet use within health care (eg, for budgetary reasons) should pay special attention to those who are older and have lower social support. Because these groups showed a preference for a printed delivery mode, steering them strictly toward internet delivery could risk losing them altogether for the intervention. For this subgroup, it may be essential to emphasize that the Web-based intervention is easy to use. In addition, future intermediaries of interventions could consider providing internet training opportunities to stimulate the use of internet-delivered interventions. A pre-enrollment questionnaire that assesses the level of internet literacy could be useful to determine the optimum format of the intervention.

Attrition

Overall, attrition from the intervention was 58%. Although this is considerable, it is not uncommon: other studies on PA interventions for older adults have shown widely varying attrition rates, ranging from 22% to 76% [62-64]. However, several more recent studies have shown relatively low attrition rates (ranging from 0% to 51%, with a mean of 21%) [65,66], and these provide indications that an association between lower attrition and higher age may be present [26,66]. When considering these studies, the attrition from our intervention appears relatively high. Considering the relatively high attrition, for future research, a deeper analysis into the appreciation of the intervention would be useful.

Only delivery mode and educational attainment were found to be significant predictors of attrition: attrition was higher among participants in the Web-based delivery mode and among those with a lower educational attainment. The fact that only 2 determinants were found to predict attrition indicates that computer tailoring in Active Plus65 delivers advice that is equally valued in a broad range of participants. To provide corroboration for the finding that both the Web-based delivery mode and low educational attainment are predictors of attrition, 2 models can be outlined, that is, the Senior Technology Acceptance and Adoption Model (STAM) [67] and the Cycle of Technology Acquirement by Independent Living Seniors Model (C-TAILS) [68]. In STAM, the ease of learning is a crucial determinant for conversion to a new technology. It may be that the lower educated participants who took part in the Web-based delivery mode are unable to succeed in comfortably using the Web-based delivery mode and consequently stop using the intervention. In addition, C-TAILS stipulates that a new technology needs to be aligned with an individual's needs: because Active Plus65 is not acquired on the participants' initiative, it may be that their need for an intervention is lower and for the lower educated participants in particular, initial difficulties with using the Web-based intervention results in attrition. It may have some practical implications that attrition is higher in the Web-based delivery mode and among those with a lower educational attainment. For future intervention development, including targeted retention techniques specifically for Web-based delivery, such as email prompts, or delivering not all advice at once but in stages, could decrease attrition [14,69]. Several studies show that presentation strategies of interventions may need to be tailored considering those with low educational attainment to decrease attrition, for example, using more graphic materials instead of text and using entertaining or interactive elements [70,71].

It was also hypothesized that older age would be a predictor of higher attrition, but this was not established. An explanation may lie in the specific characteristics of our target population: participants who are older, have poor health status, and are unemployed will more often use an intervention as intended and will thus show lower attrition [72]. Those demographic determinants are comparable with the characteristics of the participants of Active Plus65 who are older, have physical impairments, and are mostly retired; although being retired may not be directly comparable with being unemployed, there are obvious similarities. Conversely, another characteristic specific to our target population, being single, may have had an opposite effect on attrition: it has been demonstrated that not having a life partner negatively influences internet use [50], which could contribute to a higher attrition rate.

Our assumption that a stronger presence of the psychosocial determinants associated with behavior change would be related to lower attrition was not confirmed. At 3 months in the intervention, participants already received advice 2 times. It could be that from this advice, participants obtain the anticipated aid they needed from the intervention and decide to discontinue use. That may even be more so in cases where higher levels of variables associated with behavior change are present: Active Plus65 focuses strongly on stimulating the motivation for PA,

XSL•FO
RenderX

and there is a possibility that for participants who already have a higher commitment to behavior change, the additional value of Active Plus65 is less distinct. It has been suggested that especially in Web-based interventions, participants may stop using an intervention once they achieve outcomes they consider adequate [73]. Attrition from this point of view may not even be negative but rather be an affirmation of realizing what participants had expected to gain. This shows that a solid insight into the preintervention characteristics that are predictive of attrition may be useful before enrolling participants in the intervention. In line with that, more insight into the appreciation of the intervention could provide valuable information.

Strengths and Limitations

As far as could be determined, this is the only study assessing the reach and attrition of the Web-based and printed delivery mode of an intervention with identical content among a population of single older adults with a physical impairment. As this population is growing fast, this study provides valuable insights. However, some limitations need to be acknowledged.

First, with 6%, the response rate appears to be quite low. Although a recent review showed that this is consistent with similar interventions [74], low response rates limit the public health impact of such interventions. The relatively low response rates and nonavailability of information on nonparticipants make it impossible to perform predictive analyses on who is interested in such interventions. We can only provide insight into the older adults who actually chose to participate. Second, only the baseline characteristics were included as potential predictors of attrition: other variables, such as digital literacy, engagement, or satisfaction with the intervention, may be related to attrition, but this could not be determined. Third, attrition from the intervention was relatively high, although this is not uncommon in eHealth interventions for older adults and in agreement with comparable studies [75]. Fourth, our study focuses on a specific subpopulation of older adults, that is, those who have a physical impairment and are single. As older age is generally accompanied by the onset of physical impairments, most older adults will meet this particular characteristic of our target population. However, this will not be applicable for the characteristic of being single, which may have implications for the generalizability of our findings. Considering the aim of our intervention, that is, stimulating PA preferably done with others, there's a possibility that our intervention impacts singles and nonsingles differently. It may thus be advisable to repeat our studies in a population of mixed-marital status. Finally, the proportion of variance explained by our analyses appears relatively low (2%-19%), despite the inclusion of a broad range of potential demographic, health, and psychosocial determinants for delivery mode preference or for attrition. Nonetheless, these results are in line with comparable studies [23,26].

Conclusions

The findings of our study outline which delivery modes are likely to be the most advisable for specific target populations, thus increasing the impact that interventions can potentially have on public health. Our results show that participants who are older and have lower levels of social support for PA are more attracted to the printed delivery mode of Active Plus65. Attrition was higher among those with a lower educational attainment, indicating that for these participants, print-delivered interventions would yield higher participation rates than Web-based delivered interventions. Although the Web-based delivery mode showed a higher attrition rate, printed delivery modes in general have the downside of being more expensive. It may therefore be advisable that printed delivery modes and Web-based delivery modes are offered alongside each other. Further research may also provide potential solutions to decrease attrition among those with lower education attainment. Considering the high speed at which changes in internet use occur, a continuous research into delivery mode preference and attrition is needed.

Acknowledgments

This research was supported by a grant from Meer Veerkracht, Langer Thuis, a program of FNO, reference number 100698. FNO is a Dutch foundation committed to increasing opportunities for better health, quality of life, and future prospects of people who are vulnerable because of chronic illness or disability. FNO approved the design of the study, collection, and analysis of data, but it had no role in writing the manuscript.

Authors' Contributions

LL and CB designed and wrote the original proposal for obtaining the funding. DP and BB were also involved in the original proposal. BB, DP, and JB were responsible for conducting interviews and for writing and programming the intervention content. LL and CB critically reviewed and approved the intervention content. BB, DP, and JB were responsible for the recruitment procedure. JB was responsible for drafting the manuscript. All authors contributed to the writing of the manuscript; critically revised the manuscript for important intellectual content; and read and approved the final manuscript. This manuscript is original and is not under consideration or published by any other journal. There are no potentially overlapping publications.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Advice that focuses on being active with a chronic disease.

XSL•FO
RenderX

[PNG File, 3MB - jmir_v21i8e13416_app1.png]

Multimedia Appendix 2
Example of PA exercise done at home from emailed advice.
[PNG File, 528KB - jmir_v21i8e13416_app2.PNG]

References

1. World Health Organization. 2010. Prevalence of Insufficient Physical Activity URL: http://www.who.int/gho/ncd/risk_factors/physical_activity_text/en/ [accessed 2017-01-19] [WebCite Cache ID 6q9zrB5Db]
2. World Health Organization. 2010. Global Recommendations on Physical Activity for Health URL: http://www.who.int/dietphysicalactivity/publications/9789241599979/en/ [WebCite Cache ID 6qA00OgKe]
3. Muellmann S, Forberger S, Möllers T, Bröring E, Zeeb H, Pischke CR. Effectiveness of ehealth interventions for the promotion of physical activity in older adults: a systematic review. Prev Med 2018 Dec;108:93-110. [doi: 10.1016/j.ypmed.2017.12.026] [Medline: 29289643]
4. 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]
5. Golsteijn RH, Peels DA, Evers SM, Bolman C, Mudde AN, de Vries H, et al. Cost-effectiveness and cost-utility of a web-based or print-delivered tailored intervention to promote physical activity among adults aged over fifty: an economic evaluation of the active plus intervention. Int J Behav Nutr Phys Act 2014 Sep 28;11:122 [FREE Full text] [doi: 10.1186/s12966-014-0122-z] [Medline: 25262435]
6. Davies CA, Spence JC, Vandelanotte C, Caperchione CM, Mummery WK. Meta-analysis of internet-delivered interventions to increase physical activity levels. Int J Behav Nutr Phys Act 2012 Apr 30;9:52 [FREE Full text] [doi: 10.1186/1479-5868-9-52] [Medline: 22546283]
7. Donkin L, Christensen H, Naismith SL, Neal B, Hickie IB, Glozier N. A systematic review of the impact of adherence on the effectiveness of e-therapies. J Med Internet Res 2011 Aug 5;13(3):e52 [FREE Full text] [doi: 10.2196/jmir.1772] [Medline: 21821503]
8. 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]
9. 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]
10. Tennant B, Stellefson M, Dodd V, Chaney B, Chaney D, Paige S, et al. eHealth literacy and web 2.0 health information seeking behaviors among baby boomers and older adults. J Med Internet Res 2015 Mar 17;17(3):e70 [FREE Full text] [doi: 10.2196/jmir.3992] [Medline: 25783036]
11. Anderson GO. AARP. 2017. Technology Use and Attitudes Among Mid-Life and Older Americans URL: https://www.aarp.org/content/dam/aarp/research/surveys_statistics/technology/info-2018/atom-nov-2017-tech-module.doi.10.26419%252Fres.00210.001.pdf [accessed 2019-02-15]
12. Negreiro M. European Parliament. 2015. Bridging the Digital Divide in the EU URL: http://www.europarl.europa.eu/RegData/etudes/BRIE/2015/573884/EPRS_BRI(2015)573884_EN.pdf [accessed 2019-02-15]
13. Anderson M, Perrin A. Pew Research Center. 2017. Tech Adoption Climbs Among Older Adults URL: https://www.pewinternet.org/2017/05/17/tech-adoption-climbs-among-older-adults/ [accessed 2019-02-15]
14. Brouwer WB, Kroeze W, Crutzen R, de Nooijer J, de Vries N, 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 Jan 6;13(1):e2 [FREE Full text] [doi: 10.2196/jmir.1639] [Medline: 21212045]
15. Boekhout JM, Peels DA, Berendsen BA, Bolman CA, Lechner L. An ehealth intervention to promote physical activity and social network of single, chronically impaired older adults: adaptation of an existing intervention using intervention mapping. JMIR Res Protoc 2017 Nov 23;6(11):e230 [FREE Full text] [doi: 10.2196/resprot.8093] [Medline: 29170146]
16. Barlow MA, Liu SY, Wrosch C. Chronic illness and loneliness in older adulthood: the role of self-protective control strategies. Health Psychol 2015 Aug;34(8):870-879. [doi: 10.1037/hea0000182] [Medline: 25528177]
17. Hawthorne G. Perceived social isolation in a community sample: its prevalence and correlates with aspects of peoples' lives. Soc Psychiatry Psychiatr Epidemiol 2008 Feb;43(2):140-150. [doi: 10.1007/s00127-007-0279-8] [Medline: 17994175]
18. Vlaming RH, Haveman-Nies A, Groeniger IB, de Groot L, van't Veer P. Determinants of trends in loneliness among Dutch older people over the period 2005-2010. J Aging Health 2014 Apr;26(3):422-440. [doi: 10.1177/0898264313518066] [Medline: 24452851]
19. Jones LW, Courneya KS. Exercise counseling and programming preferences of cancer survivors. Cancer Pract 2002;10(4):208-215. [doi: 10.1046/j.1523-5394.2002.104003.x] [Medline: 12100105]

XSL•FO
RenderX

20. Short CE, Vandelanotte C, Duncan MJ. Individual characteristics associated with physical activity intervention delivery mode preferences among adults. Int J Behav Nutr Phys Act 2014 Feb 25;11(1):25 [FREE Full text] [doi: 10.1186/1479-5868-11-25] [Medline: 24568611]
21. Daley A, Stokes-Lampard H, Wilson S, Rees M, Roalfe A, Macarthur C. What women want? Exercise preferences of menopausal women. Maturitas 2011 Feb;68(2):174-178. [doi: 10.1016/j.maturitas.2010.11.011] [Medline: 21168292]
22. Forbes CC, Plotnikoff RC, Courneya KS, Boulé NG. Physical activity preferences and type 2 diabetes: exploring demographic, cognitive, and behavioral differences. Diabetes Educ 2010;36(5):801-815. [doi: 10.1177/0145721710378538] [Medline: 20736386]
23. Golsteijn RH, Bolman C, Peels DA, Volders E, de Vries H, Lechner L. A web-based and print-based computer-tailored physical activity intervention for prostate and colorectal cancer survivors: a comparison of user characteristics and intervention use. J Med Internet Res 2017 Dec 23;19(8):e298 [FREE Full text] [doi: 10.2196/jmir.7838] [Medline: 28835353]
24. Greaney ML, Puleo E, Bennett GG, Haines J, Viswanath K, Gillman MW, et al. Factors associated with choice of web or print intervention materials in the healthy directions 2 study. Health Educ Behav 2014 Feb;41(1):52-62 [FREE Full text] [doi: 10.1177/1090198113486803] [Medline: 23720532]
25. Plantinga S, Kaal M. Mediawijzer. 2018. How Media-Wise is the Netherlands?: Research Media Literacy URL: https://www.mediawijzer.net/wp-content/uploads/sites/6/2018/09/Rapport-Mediawijsheid-volwassenen-2018.pdf [accessed 2018-11-09] [WebCite Cache ID 73nME4gJo]
26. Peels DA, Bolman C, Golsteijn RH, de Vries H, Mudde AN, van Stralen MM, et al. Differences in reach and attrition between web-based and print-delivered tailored interventions among adults over 50 years of age: clustered randomized trial. J Med Internet Res 2012 Dec 17;14(6):e179 [FREE Full text] [doi: 10.2196/jmir.2229] [Medline: 23246790]
27. Ajzen I. The theory of planned behavior. Organ Behav Hum Perform 1991 Dec;50(2):179-211. [doi: 10.1016/0749-5978(91)90020-T]
28. de Vries H, Mudde A, Leijs I, Charlton A, Vartiainen E, Buijs G, et al. The European Smoking Prevention Framework Approach (EFSA): an example of integral prevention. Health Educ Res 2003 Oct;18(5):611-626. [doi: 10.1093/her/cyg031] [Medline: 14572020]
29. Fishbein M, Ajzen I. Predicting And Changing Behavior: The Reasoned Action Approach. New York: Psychology Press; 2010.
30. McGuire WJ. Attitudes and attitude change. In: Fiske ST, Gilbert DT, Lindzey G, editors. Handbook of Social Psychology. New York: Random House; 1985:233-346.
31. Guertler D, Vandelanotte C, Kirwan M, Duncan MJ. Engagement and nonusage attrition with a free physical activity promotion program: the case of 10,000 steps Australia. J Med Internet Res 2015 Jul 15;17(7):e176 [FREE Full text] [doi: 10.2196/jmir.4339] [Medline: 26180040]
32. Luhmann M, Hawkley LC. Age differences in loneliness from late adolescence to oldest old age. Dev Psychol 2016 Dec;52(6):943-959. [doi: 10.1037/dev0000117] [Medline: 27148782]
33. Victor CR, Yang K. The prevalence of loneliness among adults: a case study of the United Kingdom. J Psychol 2012;146(1-2):85-104. [doi: 10.1080/00223980.2011.613875] [Medline: 22303614]
34. Barnard Y, Bradley MD, Hodgson F, Lloyd AD. Learning to use new technologies by older adults: perceived difficulties, experimentation behaviour and usability. Comput Hum Behav 2013 Jul;29(4):1715-1724 [FREE Full text] [doi: 10.1016/j.chb.2013.02.006]
35. Tsai HS, Shillair R, Cotten SR. Social support and 'playing around': an examination of how older adults acquire digital literacy with tablet computers. J Appl Gerontol 2017 Jan;36(1):29-55 [FREE Full text] [doi: 10.1177/0733464815609440] [Medline: 26491029]
36. Boekhout JM, Berendsen BA, Peels DA, Bolman CA, Lechner L. Evaluation of a computer-tailored healthy ageing intervention to promote physical activity among single older adults with a chronic disease. Int J Environ Res Public Health 2018 Dec 15;15(2):pii: E346 [FREE Full text] [doi: 10.3390/ijerph15020346] [Medline: 29462862]
37. Boekhout JM, Berendsen BA, Peels DA, Bolman C, Lechner L. Physical impairments disrupt the association between physical activity and loneliness: a longitudinal study. J Aging Phys Act 2019 May 6:1-10. [doi: 10.1123/japa.2018-0325] [Medline: 30859889]
38. Lechner L, de Vries H. Starting participation in an employee fitness program: attitudes, social influence, and self-efficacy. Prev Med 1995 Nov;24(6):627-633. [doi: 10.1006/pmed.1995.1098] [Medline: 8610087]
39. Lechner L, de Vries H. Participation in an employee fitness program: determinants of high adherence, low adherence, and dropout. J Occup Environ Med 1995 Apr;37(4):429-436. [doi: 10.1097/00043764-199504000-00014] [Medline: 7670898]
40. Resnick B, Jenkins LS. Testing the reliability and validity of the self-efficacy for exercise scale. Nurs Res 2000;49(3):154-159. [doi: 10.1097/00006199-200005000-00007] [Medline: 10882320]
41. Sallis JF, Grossman RM, Pinski RB, Patterson TL, Nader PR. The development of scales to measure social support for diet and exercise behaviors. Prev Med 1987 Nov;16(6):825-836. [doi: 10.1016/0091-7435(87)90022-3] [Medline: 3432232]
42. Sheeran P, Orbell S. Implementation intentions and repeated behaviour: augmenting the predictive validity of the theory of planned behaviour. Eur J Soc Psychol 1999;29(2/3):349-369 [FREE Full text] [doi: 10.1002/(SICI)1099-0992(199903/05)29:2/33.0.CO;2-Y]

XSL•FO
RenderX

43. Wendel-Vos GC, Schuit AJ, Saris WH, Kromhout D. Reproducibility and relative validity of the short questionnaire to assess health-enhancing physical activity. J Clin Epidemiol 2003 Dec;56(12):1163-1169. [doi: 10.1016/S0895-4356(03)00220-8] [Medline: 14680666]
44. de Hollander EL, Zwart L, de Vries SI, Wendel-Vos W. The SQUASH was a more valid tool than the OBiN for categorizing adults according to the Dutch physical activity and the combined guideline. J Clin Epidemiol 2012 Jan;65(1):73-81. [doi: 10.1016/j.jclinepi.2011.05.005] [Medline: 21840174]
45. Wagenmakers R, van den Akker-Scheek I, Groothoff J, Zijlstra W, Bulstra S, Kootstra J, et al. Reliability and validity of the short questionnaire to assess health-enhancing physical activity (SQUASH) in patients after total hip arthroplasty. BMC Musculoskelet Disord 2008 Oct 17;9:141 [FREE Full text] [doi: 10.1186/1471-2474-9-141] [Medline: 18928545]
46. de Jong GJ, van Tilburg T. The de Jong Gierveld short scales for emotional and social loneliness: tested on data from 7 countries in the UN generations and gender surveys. Eur J Ageing 2010 Jun;7(2):121-130 [FREE Full text] [doi: 10.1007/s10433-010-0144-6] [Medline: 20730083]
47. Krosnick JA, Presser S. Question and questionnaire design. In: Marsden PV, Wright JD, editors. Handbook of Survey Research. West Yorkshire, England: Emereld Group Publishing Limited; 2010:263-312.
48. Remillard ML, Mazor KM, Cutrona SL, Gurwitz JH, Tjia J. Systematic review of the use of online questionnaires of older adults. J Am Geriatr Soc 2014 Apr;62(4):696-705 [FREE Full text] [doi: 10.1111/jgs.12747] [Medline: 24635138]
49. Twisk JW. Applied Multilevel Analysis: A Practical Guide for Medical Researchers. Cambridge: Cambridge University Press; 2006.
50. Friemel TN. The digital divide has grown old: determinants of a digital divide among seniors. New Media Soc 2014 Jun 12;18(2):313-331 [FREE Full text] [doi: 10.1177/1461444814538648]
51. Hall AK, Bernhardt JM, Dodd V, Vollrath MW. The digital health divide: evaluating online health information access and use among older adults. Health Educ Behav 2015 Apr;42(2):202-209 [FREE Full text] [doi: 10.1177/1090198114547815] [Medline: 25156311]
52. 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 Dec 27;19(4):e136 [FREE Full text] [doi: 10.2196/jmir.6731] [Medline: 28450271]
53. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Q 2003;27(3):425. [doi: 10.2307/30036540]
54. Statistics Netherlands. 2016. Eight Percent of Dutch People Never Use the Internet URL: https://www.cbs.nl/nl-nl/nieuws/2016/22/acht-procent-van-de-nederlanders-nooit-op-internet [WebCite Cache ID 75ILUV5bU]
55. Statista. 2018. Adult Internet Usage Penetration in the United States From 2000 to 2018, by Age Group URL: https://www.statista.com/statistics/184389/adult-internet-users-in-the-us-by-age-since-2000/ [accessed 2019-01-09] [WebCite Cache ID 75ILgh5Us]
56. Lustria ML, Noar SM, Cortese J, van Stee K, 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]
57. van Deursen AJ, van Dijk JA, ten Klooster PM. Increasing inequalities in what we do online: a longitudinal cross sectional analysis of internet activities among the Dutch population (2010 to 2013) over gender, age, education, and income. Telemat Inform 2015 May 1;32(2):259-272. [doi: 10.1016/j.tele.2014.09.003]
58. de Veer AJ, Peeters JM, Brabers AE, Schellevis FG, Rademakers JJ, Francke AL. Determinants of the intention to use e-health by community dwelling older people. BMC Health Serv Res 2015 Mar 15;15:103 [FREE Full text] [doi: 10.1186/s12913-015-0765-8] [Medline: 25889884]
59. Kampmeijer R, Pavlova M, Tambor M, Golinowska S, Groot W. The use of e-health and m-health tools in health promotion and primary prevention among older adults: a systematic literature review. BMC Health Serv Res 2016 Dec 5;16(Suppl 5):290 [FREE Full text] [doi: 10.1186/s12913-016-1522-3] [Medline: 27608677]
60. Hill R, Betts LR, Gardner SE. Older adults' experiences and perceptions of digital technology: (dis)empowerment, wellbeing, and inclusion. Comput Hum Behav 2015 Jul;48:415-423. [doi: 10.1016/j.chb.2015.01.062]
61. Vroman KG, Arthanat S, Lysack C. 'Who over 65 is online?' Older adults' dispositions toward information communication technology. Comput Hum Behav 2015 Feb;43:156-166 [FREE Full text] [doi: 10.1016/j.chb.2014.10.018]
62. Jancey J, Lee A, Howat P, Clarke A, Wang K, Shilton T. Reducing attrition in physical activity programs for older adults. J Aging Phys Act 2007 Apr;15(2):152-165. [doi: 10.1123/japa.15.2.152] [Medline: 17556782]
63. Mouton A, Cloes M. Web-based interventions to promote physical activity by older adults: promising perspectives for a public health challenge. Arch Public Health 2013 Jul 2;71(1):16 [FREE Full text] [doi: 10.1186/0778-7367-71-16] [Medline: 23819885]
64. Mullen SP, Wójcicki TR, Mailey EL, Szabo AN, Gothe NP, Olson EA, et al. A profile for predicting attrition from exercise in older adults. Prev Sci 2013 Oct;14(5):489-496 [FREE Full text] [doi: 10.1007/s11121-012-0325-y] [Medline: 23412942]
65. Alley SJ, Kolt GS, Duncan MJ, Caperchione CM, Savage TN, Maeder AJ, et al. The effectiveness of a web 2.0 physical activity intervention in older adults - a randomised controlled trial. Int J Behav Nutr Phys Act 2018 Dec 12;15(1):4 [FREE Full text] [doi: 10.1186/s12966-017-0641-5] [Medline: 29329587]

XSL•FO
RenderX

66. van der Deijl M, Etman A, Kamphuis CB, van Lenthe FJ. Participation levels of physical activity programs for community-dwelling older adults: a systematic review. BMC Public Health 2014 Dec 18;14:1301 [FREE Full text] [doi: 10.1186/1471-2458-14-1301] [Medline: 25523712]
67. Renaud K, van Biljon J. Predicting Technology Acceptance and Adoption by the Elderly: A Qualitative Study. In: Proceedings of the 2008 Annual Research Conference of the South African Institute of Computer Scientists and Information Technologists on IT Research in Developing Countries: Riding the Wave of Technology. 2008 Presented at: SAICSIT'08; October 6-8, 2008; Wilderness, South Africa p. 210-219. [doi: 10.1145/1456659.1456684]
68. Peek ST, Luijkx KG, Vrijhoef HJ, Nieboer ME, Aarts S, van der Voort CS, et al. Origins and consequences of technology acquirement by independent-living seniors: towards an integrative model. BMC Geriatr 2017 Dec 22;17(1):189 [FREE Full text] [doi: 10.1186/s12877-017-0582-5] [Medline: 28830444]
69. Schneider F, de Vries H, Candel M, van de Kar A, van Osch L. Periodic email prompts to re-use an internet-delivered computer-tailored lifestyle program: influence of prompt content and timing. J Med Internet Res 2013 Jan 31;15(1):e23 [FREE Full text] [doi: 10.2196/jmir.2151] [Medline: 23363466]
70. 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]
71. Milward J, Drummond C, Fincham-Campbell S, Deluca P. What makes online substance-use interventions engaging? A systematic review and narrative synthesis. Digit Health 2018;4:2055207617743354 [FREE Full text] [doi: 10.1177/2055207617743354] [Medline: 29942622]
72. Reinwand DA, Schulz DN, Crutzen R, Kremers SP, de Vries H. Who follows ehealth interventions as recommended? A study of participants' personal characteristics from the experimental arm of a randomized controlled trial. J Med Internet Res 2015 May 11;17(5):e115 [FREE Full text] [doi: 10.2196/jmir.3932] [Medline: 25963607]
73. 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]
74. Zubala A, MacGillivray S, Frost H, Kroll T, Skelton DA, Gavine A, et al. Promotion of physical activity interventions for community dwelling older adults: a systematic review of reviews. PLoS One 2017;12(7):e0180902 [FREE Full text] [doi: 10.1371/journal.pone.0180902] [Medline: 28700754]
75. 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]

Abbreviations

BMI: body mass index
C-TAILS: Cycle of Technology Acquirement by Independent Living Seniors Model
MVPA: moderate to vigorous physical activity
PA: physical activity
SQUASH: Short Questionnaire to Assess Health Enhancing Physical Activity
STAM: Senior Technology Acceptance and Adoption Model
UTAUT: Unified Theory of Acceptance and Use of Technology

Edited by G Eysenbach; submitted 16.01.19; peer-reviewed by D Reinwand, D Van Dyck; comments to author 07.02.19; revised version received 03.04.19; accepted 16.04.19; published 28.08.19.

Please cite as:
Boekhout JM, Peels DA, Berendsen BAJ, Bolman C, Lechner L
A Web-Based and Print-Delivered Computer-Tailored Physical Activity Intervention for Older Adults: Pretest-Posttest Intervention Study Comparing Delivery Mode Preference and Attrition
J Med Internet Res 2019;21(8):e13416
URL: http://www.jmir.org/2019/8/e13416/
doi:10.2196/13416
PMID:31464186

©Janet Maria Boekhout, Denise Astrid Peels, Brenda Angela Juliette Berendsen, Catherine Bolman, Lilian Lechner. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 28.08.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of

XSL•FO
RenderX

Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

How Prefrail Older People Living Alone Perceive Information and Communications Technology and What They Would Ask a Robot for: Qualitative Study

Katia Daniele[1,2,3], MSc; Maura Marcucci[2,4], MD; Cesarina Cattaneo[3], MA; Nunzio Alberto Borghese[1], MSc; Lucia Zannini[3], MA, PhD

[1]Department of Computer Science, University of Milan, Milan, Italy

[2]Geriatric Unit, Fondazione IRCCS Ca' Granda Ospedale Maggiore Policlinico, Milan, Italy

[3]Department of Biomedical Sciences for Health, University of Milan, Milan, Italy

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

Corresponding Author:
Lucia Zannini, MA, PhD
Department of Biomedical Sciences for Health
University of Milan
Via Carlo Pascal, 36
Milan, 20133
Italy
Phone: 39 02503 15101
Fax: 39 0250315105
Email: lucia.zannini@unimi.it

Abstract

Background: In the last decade, the family system has changed significantly. Although in the past, older people used to live with their children, nowadays, they cannot always depend on assistance of their relatives. Many older people wish to remain as independent as possible while remaining in their homes, even when living alone. To do so, there are many tasks that they must perform to maintain their independence in everyday life, and above all, their well-being. Information and communications technology (ICT), particularly robotics and domotics, could play a pivotal role in aging, especially in contemporary society, where relatives are not always able to accurately and constantly assist the older person.

Objective: The aim of this study was to understand the needs, preferences, and views on ICT of some prefrail older people who live alone. In particular, we wanted to explore their attitude toward a hypothetical caregiver robot and the functions they would ask for.

Methods: We designed a qualitative study based on an interpretative phenomenological approach. A total of 50 potential participants were purposively recruited in a big town in Northern Italy and were administered the Fried scale (to assess the participants' frailty) and the Mini-Mental State Examination (to evaluate the older person's capacity to comprehend the interview questions). In total, 25 prefrail older people who lived alone participated in an individual semistructured interview, lasting approximately 45 min each. Overall, 3 researchers independently analyzed the interviews transcripts, identifying meaning units, which were later grouped in clustering of themes, and finally in emergent themes. Constant triangulation among researchers and their reflective attitude assured trustiness.

Results: From this study, it emerged that a number of interviewees who were currently using ICT (ie, smartphones) did not own a computer in the past, or did not receive higher education, or were not all young older people (aged 65-74 years). Furthermore, we found that among the older people who described their relationship with ICT as negative, many used it in everyday life. Referring to robotics, the interviewees appeared quite open-minded. In particular, robots were considered suitable for housekeeping, for monitoring older people's health and accidental falls, and for entertainment.

Conclusions: Older people's use and attitudes toward ICT does not always seem to be related to previous experiences with technological devices, higher education, or lower age. Furthermore, many participants in this study were able to use ICT, even if they did not always acknowledge it. Moreover, many interviewees appeared to be open-minded toward technological devices, even toward robots. Therefore, proposing new advanced technology to a group of prefrail people, who are self-sufficient and can live alone at home, seems to be feasible.

XSL•FO
RenderX

(J Med Internet Res 2019;21(8):e13228) doi:10.2196/13228

KEYWORDS

frail elders; independent living; attitude; technology; information technology; robotics; qualitative research; interview

Introduction

Background

Aging is associated with physiological decay, higher risk of multiple acute and chronic diseases, and ultimately, loss of independence and disability [1]. The constant growth of the proportion of populations represented by older people is challenging the civil society and the health and social systems [1-4]. In this scenario, preventive actions aiming at promoting an active and healthy aging, as opposed to treatment, have the largest potential to reduce the societal burden associated with population aging.

From the individual perspective, many older people want to remain as independent as possible and to remain in their own home; aging in place rather than in nursing home, even when they live alone, is an essential part of this wish [1,4-6]. Aging in place is often considered a good alternative to expensive institutional care by policy makers [6]. Furthermore, "[...] research has shown that there may be significant benefits from delaying or avoiding moving to skilled nursing residences" [5]. However, living alone in older age does not necessarily mean autonomy; rather, it might simply reflect the changes in the structure of the contemporary society, especially in cities, in which children more often move definitively from the nuclear family. In fact, older people living alone are often already in an initial state of vulnerability (prefrail condition), while to safely maintain their independence in their home, one should be able to perform at least some instrumental and basic activities of daily living [1]. Moreover, having a sufficient level of functioning does not necessarily imply satisfaction and well-being, which might also depend on the ability and opportunity to enjoy a social life [7].

There is a growing interest in exploring the potential of information and communications technology (ICT) in interventions to assist frail older people and also to promote active and healthy aging [1,3,6-11]. However, it is commonly agreed that to increase the success of ICT-based solutions, these technologies must be designed according to the end users' needs. This is particularly relevant, and also challenging, when the user is an older person [3,7,9,12,13]. Their opinion and needs become crucial to plan technological devices, which need to be considered from the beginning of their development. Pursuing the acceptance and adaptation of the older user to an already designed technology has a higher chance of failure [5]. This is one of the reasons why there has been an increased interest among researchers in exploring older people's perception of technologies. Studies conducted so far have found that different factors might affect older people's aptitude and attitude toward ICT-based solutions, including psychological and physical condition, personality, life history, culture, and socioeconomic status.

Studies conducted so far have often showed age as a limiting factor for the use of ICT devices [14,15]. Other factors have been also identified as predictors or determinants of technology use among older people, such as education, mood, culture, motivation [2,9,10,14,15], previous experience with technology, ease of use, perceived utility [5,8,11,12,15-17], and the social environment [5,6,8,10,14,16,17].

To our knowledge, no study conducted to explore older people's preferences on ICT-based solutions to promote independent living has so far focused on a prefrail older population, who might not apparently require assistance for daily activities but, being at risk of deterioration, might still benefit from a support to an independent living. Moreover, we could not find studies looking at the relationship between older people and technologies, conducted in Italy or with participants belonging to a Latin culture. In this sociocultural reality, *familism* —which denotes the centrality of family in the life of people and expectations of mutual emotional and instrumental support among family members throughout the life span [18]—can generate resistance to ICT-based solutions to assist the older person. It is commonly agreed that the users' cultural background impacts the views and expectations toward technological devices.

Objective

With this background, we conducted a qualitative study aimed at answering the following question: what are the needs, wishes, preferences, and views about technologies, in general and as a support to independent living, in a sample of Italian prefrail older people who live alone?

This study was part of a wider European project, HORIZON 2020 number 732158, named MoveCare (Multiple-actors Virtual Empathic Caregiver for the Elder), which aims at supporting the independent living of the elderly at home.

Methods

Study Design

Assuming that older persons' attitudes and preferences regarding a complex phenomenon such as ICT are mediated by past events and lived experience with technology, a qualitative study based on the interpretative phenomenological approach (IPA) [19] was designed. This method allows exploring people's lived experience about a specific phenomenon.

Data were gathered through individual semistructured interviews which, compared with focus groups or group interviews, we believed would have a higher chance of making the older interviewees at their ease, especially when discussing personal issues. This data collection method was largely used in previous qualitative studies based on IPA and on similar topics [4,8,10,14,17,20].

XSL•FO
RenderX

This study has been approved by the Ethical Committee of *Fondazione Istituto di Ricovero e Cura a Carattere Scientifico Ca' Granda Ospedale Maggiore Policlinico* of Milan, on May 17, 2017. All the participants signed an informed consent containing clear and standardized information about study aims and procedures.

Sampling

This study was conducted in Milan, Italy.

First, we identified a set of 50 potential participants according to a purposeful sampling. We looked for older people who, at first approach, were willing to talk about their habits and opinions. Moreover, candidates had to satisfy the following criteria:

- age ≥65 years
- living alone (or living alone and receiving assistance for a maximum of 1 hour a day)
- speaking Italian fluently.

From February to May 2017, we sampled members and users of an association of older volunteers named *Associazione Nazionale Tutte le Età Attive per la Solidarietà (ANTEAS;* n=32), people identified through the social work of the municipality of Milan (n=13), patients and relatives referred to the geriatric outpatient clinics of the hospital Policlinico of Milan (n=3), and members of the Italian Union of Retired People of Milan (n=2). The 50 potential participants were contacted by phone. Overall, 4 declined because of health problems and 5 gave no explanation, 4 did not reply to the phone call, 1 did not show up at the appointment, and 1 was not fluent in Italian. The remaining 35 persons were then invited for an interview.

Procedure

Before the interview, phenotype frailty assessment [21] and Mini-Mental State Examination (MMSE) [22] were administered. To be included in the study, the potential participants had to meet 1 or 2 Fried criteria (ie, prefrail category) and score ≥26 at the MMSE (ie, excluding people with clinically relevant cognitive impairment).

When a potential participant satisfied those inclusion criteria, the full interview was conducted. We drafted an initial interview grid, which in the first part of the interview included questions on the participants' sociodemographic characteristics, social condition and lifestyle, relationships with relatives and friends, and current health problems and therapies. To explore the experience and attitude toward ICT, the interviewee was asked to imagine a new technology (ie, *like a robot*) available at home and to describe the functions he/she would have asked for it to have. The interviewees were then shown a short video in which an aged woman talked about her experience with Giraff, a previous version of a robot implemented in the MoveCare project [23]. Video was shown from the beginning to 1-min [24,25]. At the end of the video, we again posed the question on the desired functions of the technology, also trying to explore the interviewee's feelings toward the opportunities offered by a service robot at home. Finally, we investigated the participants' level of education and asked the interviewees for the income they thought was necessary to have a satisfactory quality of life in their (urban) context. Even though our interview guide was not theory-driven, we considered some studies [1,4,6,8] when formulating some questions of our grid, which was slightly modified after 2 test interviews.

When a candidate did not satisfy the inclusion criteria, we completed a *courtesy interview* on some neutral topics concerning everyday life.

After analyzing the 25 transcriptions, the researchers agreed that data saturation was reached and no more interviews were scheduled. Richards and Morse describe the process of data saturation as follows: "When data offer no new direction, no new question, then there is no need to sample further – the account satisfies. Often, the first sign is that investigator has a sense of having heard or seen it all" [26].

After 2 test interviews, minor modifications were done to the interview grid. The final list of questions is reported in Multimedia Appendix 1. Each interview was performed by one of the 3 investigators involved, who did not have any previous personal or professional relationship with the participants. To guarantee homogeneity and adherence to standards among the interviewers, the senior investigator conducted the test interviews, with the other 2 researchers attending as observers. After the interview and verbatim transcription, the 3 investigators met for debriefing.

Interviews took place in a quiet room in the Policlinico, or at the participant's home when it was difficult for the older person to travel to the hospital. Each interview lasted about 45 min. At the end of the interview, and later, during its analysis, the researchers wrote memos on what happened during and after the interview and all the reflections that could have been relevant to data analysis [19].

Analysis

Demographics

Sociodemographic and clinical data were extracted from the interviews and processed to convert descriptive/qualitative information into semantic clusters based on the literature and with the help of a geriatrician. For example, a diet was considered *adequate* if the older person declared to have 3 quite balanced meals a day; it was judged *partially adequate* if the interviewee stated to have less than 3 meals a day or 3 meals a day but unbalanced. The diet was considered *inadequate* if the participant declared to have less than 3 meals a day and those meals resulted to be unbalanced. The final classification is shown in Multimedia Appendices 2 and 3. Finally, descriptive statistics were used to report the characteristics of our sample.

Qualitative Analyses

Data analysis began as soon as each interview was completed. A total of 3 researchers (with psychopedagogical and/or care background) independently performed the analysis of the interviews. Each interview, transcribed verbatim, was read several times to grasp its global meaning. The coding was data driven. Indeed, each researcher independently identified the meaning units (named in other qualitative methods codes/labels), which were later grouped into clustering of themes. These clustering of themes were later clustered into emergent themes,

XSL•FO
RenderX

which were named according to their content. As the clustering of themes emerged, the researchers went back to the transcription to verify its adherence to the participant's words. The emergent themes were finally related to each other to create a meaningful network. The researchers then discussed each phase of the process until an agreement was reached and the results were reconciled.

Results

Participants

A total of 26 candidates eventually met the inclusion criteria. After the full interview, one of the participants retired his/her consent to participate in the study for personal reasons and the file of interview was destroyed.

The characteristics of the 25 participants are summarized in Tables 1-3.

Table 1. Participants' demographic data (N=25).

Characteristic	Values
Age (years), mean (SD)	77.5 (6.6)
Age (years), n (%)	
65-69	4 (16)
70-74	5 (20)
75-79	6 (24)
80-84	6 (24)
85-89	3 (12)
90-94	1 (4)
Gender, n (%)	
Male	6 (24)
Female	19 (76)
Nationality, n (%)	
Italian	23 (92)
Egyptian	1 (4)
Austrian	1 (4)
Education, n (%)	
Primary school	4 (16)
Lower secondary school	10 (40)
Upper secondary school (ie, high school)	5 (20)
University	6 (24)

Table 2. Participants' clinical screening data (N=25).

Clinical screening scale	Values
Mini-Mental State Examination, mean (SD)	28 (1.3)
Fried scale, n (%)	
1	13 (52)
2	12 (48)

XSL•FO
RenderX

Table 3. Participants' quality of life data (N=25).

Variable	n (%)
Aids and prostheses	
None	2 (8)
Hearing aid	3 (12)
Cane	4 (16)
Glasses	22 (88)
Autonomy in managing money	
Partially dependent	6 (24)
Independent	17 (68)
Not available	2 (8)
Autonomy in shopping	
Dependent	2 (8)
Partially dependent	10 (40)
Independent	13 (52)
Diet	
Inadequate	1 (4)
Partially adequate	6 (24)
Adequate	18 (72)
Help for housekeeping	
Never	7 (28)
Occasional	6 (24)
Routinely but minimal	5 (20)
Frequent	4 (16)
Very frequent	3 (12)
Medicines	
No	3 (12)
Yes	22 (88)
Physical activity	
None	1 (4)
Weak	10 (40)
Moderate	9 (36)
Good or high	5 (20)
Quality of sleeping	
Bad	8 (32)
Fair	10 (40)
Good or high	6 (24)
Social life—level of engagement in social activities	
Weak	3 (12)
Moderate	8 (32)
Good or high	14 (56)
Supposed monthly income	
≤€1000	4 (16)
€1000 to €2000	18 (72)

XSL•FO
RenderX

Variable	n (%)
≥€2000	3 (12)

Qualitative Results

As part of the qualitative analysis of the interviews, 477 meaningful units were identified and then grouped into 110 clustering of themes. From these, 13 emergent themes resulted. Overall, 9 of these clustering of themes were related to the quality of life in prefrail older people and will be presented elsewhere. Furthermore, 4 emergent themes were concerning the older people's attitudes toward ICT-based home services. These were (1) previous and current experiences with ICT, (2) participants' views of their own relationship with technologies, (3) functions they would ask an imaginary robot for, and (4) sensitivity of the older person's opinion to others' experiences with existing ICT solutions for independent living (see Multimedia Appendix 4).

Previous and Current Experiences With Information and Communications Technology Devices: Between Continuity and Discontinuity

A clear digital divide was identified among participants in relation to the current use of technologies. One group (about one-third of the participants) declared to use an old mobile phone (number of participants=6) or the landline phone and no other technological devices (n=1); some affirmed they only used mobile phones during holidays (n=2).

The other group (about two-thirds) stated to use 1 or more ICT devices, such as a smartphone (n=13), a computer (n=9), and a tablet (n=1). Among smartphone owners, 3 also owned an old mobile phone. Furthermore, 7 participants affirmed they used WhatsApp. Those who had a smartphone also used some functions/apps such as the camera (n=6), the alarm clock, the calculator, and the memo. Moreover, among the smartphone/tablet owners, 1 used Skype and 2 used Facebook. Those who owned a smartphone (or a computer) used it to play Web (eg, Burako, a card game; n=3) or not Web-based (solitaire; n=1). In addition, 5 smartphone owners declared that they had an internet connection but were unable to use it at that time. A total of 3 participants declared that they had some difficulties in using the touch function of their smartphone; however, 2 of them affirmed that they could find some strategies to overcome those difficulties.

Technological devices were mainly used by participants to communicate and interact with relatives or friends. To this purpose, some informants specified that they used the smartphone to make phone calls (n=4), for emails (n=1), and for the app WhatsApp (n=1). The computer (internet) was used for email (n=8), video calls via Skype (n=4), to receive news from relatives (n=3), and to receive photos (n=2).

One interviewee (Interviewee 12) said she had an emergency call device, but she never used it as she was afraid of making unintentional calls that would have alarmed her caregivers unnecessarily.

Technological devices were not only used by the participants for communication purposes; internet was also used to search for recipes (n=1), for home banking (n=3), to search for the meanings of new words, or to look for information (n=7):

> *[...] Google Earth, that map used for searching Geography. [...] Well yes, [I search] any information, from Geography to historical figures, some historical figures, some art works* [Interviewee 9]

Moreover, the participants declared that they used computers for some applications such as Office (n=3; "Of course, [I know] how to write documents using Word, working on tables..."—Interviewee 5), for their personal accountancy (n=1; "I use it to manage the household accounts"—Interviewee 5), to archive photos (n=1), and to watch Digital Versatile Discs (n=1).

Furthermore, from the analysis of the interviews, 3 trends emerged concerning the current use of ICT in relation with the past use of ICT, that is, at a younger age. About one-third of the participants used a computer in the past (both for work and for personal interests) and still used either a computer (n=6) or a smartphone (n=6). Most had an internet connection. An exception in this group was represented by 1 interviewee who used a computer in the past, but at the time of the interview, the only ICT he/she owned was an old mobile phone.

A second trend was represented by those (about one-third of the participants) who did not use a computer in the past and just used an old mobile phone at the time of the study.

Finally, a third group of participants did not use a computer in the past, but declared to currently use 1 or more ICT devices, such as a smartphone (n=7), a computer (n=3), or a tablet (n=1).

With regard to the association with the level of education, about half of the participants who had a computer in the past, received a postsecondary level of education. Conversely, among the 13 smartphone owners, only one-third went through higher education. Surprisingly, we did not find any obvious association between the participants' age and their use of technology.

Older People's Views of Their Relationship With Information and Communications Technology: Thinking Negatively, Acting Positively

Almost one-third of the interviewees (n=11) did not have an opinion about their relationship with technological devices. One-third (n=9) claimed to have a negative relationship with ICT, that is, endorsing a feeling of denial, inadequacy, or a lack of expertise:

> *[...] The only thing...because I'm not good at these technological things* [Interviewee 2]

> *My relationship with technologies is very bad, very bad, very bad. [...] I think I refuse [to learn] it, because I do not think I'm so stupid not to be able to* [Interviewee 6]

XSL•FO
RenderX

> *For goodness sake, It's already so much that I've used this [the phone]! Absolutely! I'm not able to use them [technological devices].* [Interviewee 7]

This type of feeling was not necessarily associated with a refusal to use ICT. In fact, 8 of those who considered their relationship with technologies as negative (n=11) used several devices in their daily life (ie, smartphones and computers).

A total of 8 participants declared that they needed, and someway sought for, some help to be able to use ICT:

> *Yes, but, for the most tremendous things, my brother in law used to come, he was the king of technology.* [Interviewee 6]

> *Then, if a payment via home banking has to be completed, well...it's made by my son or my daughter.* [Interviewee 9]

In addition, 5 of those who could not define their relationship with ICT actually used several devices.

A small group of participants affirmed that they were in favor of new technologies (n=1), or stated that they wanted to improve their relationship with ICT or they wanted to adapt themselves to "advancing technology" (Interviewee 8), for example, by learning how to use their smartphone better (n=2), or by attending an informatics course (n=1). One participant said he/she wanted "to live in the technological era" (Interviewee 11).

Finally, 3 participants clearly stated that ICT was not a fundamental part of their lives, as "you can work the same without having a computer" (Interviewee 7) or was not "essential" (Interviewee 16). One interviewee declared he/she still used a paper telephone book "only because I often lose my cell phone" (Interviewee 15). However, 2 of these participants actually used a computer (Interviewee 15) or a smartphone (Interviewee 16).

Functions That Older People Would Ask a Hypothetical Robot for: Between Housekeeping and Need for Company

Most of the participants (n=18) wished to own a robot that helped them with cleaning the house. According to an interviewee, a robot could be a valid substitute of the housekeeper:

> *[A small robot should] first of all, do the cleaning and then I don't know, nothing else, because I'm only interested in having the house cleaned and tidied up.* [Interviewee 1]

Among these 18 interviewees, only 7 did not have any help with housekeeping activities.

Cooking was another desired task (n=5). One participant imagined that the robot could even be used as an oven.

Some participants (4) asked for a multitasking robot, able to act as a formal caregiver or babysitter or maid. According to 1 participant, it might be a machine with "its own intelligence" (Interviewee 7), able to perform the tasks that a human does not want to do or cannot do. Following this request of a *multitasking robot*, 1 participant imagined that this mechanical device could be able to drive a car. Another participant imagined the robot to go shopping with him. Others (n=2) wanted this hypothetical robot to go shopping for them or to buy their medicines, if they were unable to go by themselves (n=2). One participant asked for help with getting washed and dressed. Finally, 1 participant asked for support with physiotherapy.

Other participants (n=3) believed that a robot should be able to monitor the old person and ask for help in case of emergency:

> *If I fall, I fall. I can't get up and knock at my neighbour's door. And if...if I had the robot instead...very good [...] If the robot is present, it doesn't sleep, it hears that something doesn't work, or it rather has a sensor which tells it directly and it...does what it has to...nips, calls 113, I don't know, 118 [emergency telephone numbers], it does anything [...] well [I would like] it to monitor me...and that it should intervene in case I couldn't make it.* [Interviewee 2]

On the other hand, about one-third of the participants would ask for a robot with more entertaining functions. The imaginary robot could itself be a companion (n=3) to listen to music (n=1), sing together (n=1), play cards (n=1), or share hobbies with (n=1). It was otherwise seen as a provider of general leisure activities (n=1), or to keep abreast of news (n=1), or as a tool to learn writing and reading (n=1). Most of the interviewees in this second group had a *weak* or *fair* social life.

Finally, a small group of participants (n=3) claimed that a robot could not only be useless but also a deterrent for those people who can perform some tasks autonomously (eg, housekeeping).

Sensitivity of the Old Persons' Opinion to Others' Experiences With Existing Information and Communications Technology Solution for Independent Living

After watching the video about Giraff, a robot carer which offers company and some monitoring functions [23], the participants' opinions on ICT undertook a substantial change.

The most requested function from a robot was still related to cleaning the house, but this request was expressed by 9 participants instead of 19.

The second most frequent request, after watching the video, became the robot's ability to ask for help in case of emergency (n=9), falls (n=1), and danger, in general (n=1). In particular, according to 5 participants, the robot could be useful in case of gas leaks, open windows, or an attempted break-in (n=5).

Moreover, only after watching the video, the participants could endorse the usefulness of the robot in monitoring vital signs (n=4). One person even spoke about "being monitored 365 days per year" (Interviewee 6). Giraff was not only perceived by some participants (n=3) as a useful instrument to remind the older person to perform some health-related tasks (eg, measuring their blood pressure and taking pills), but also as a reminder for buying groceries. The function of preparing meals was still considered important by 2 participants.

XSL•FO
RenderX

The robot was also acknowledged, but only minimally, as a communication instrument (n=2); indeed, it was considered by 1 participant as a valid alternative to a computer or a tablet. Moreover, an informant speculated an amplification function for people who suffer from hearing problems.

However, the idea that a robot could be a companion for an older person increased after watching the video (mentioned by 7 participants). One participant sustained that Giraff could be useful to go for walks together, and another imagined it could drive a car and take a person out. The fact that the robot could not "speak, offend, or get angry like a human being" was mentioned as an advantage of this special companion by 1 participant (Interviewee 9).

Some participants considered Giraff as a useful caregiver for those who are in old age (n=4) or have cognitive problems. However, although some informants still endorsed that the robot could perform important assistance functions in the case of people with lower levels of autonomy, overall, after watching the video, the participants tended to scale down the robot's functions compared with the multitasking robot they had imagined before.

Interestingly, Giraff was seen as inadequate if an old person was self-sufficient in performing certain tasks (n=4) or even as a deterrent to keep oneself active and self-sufficient (n=2):

> *I don't know, these things...all those possibilities, it's true, they exist and in my opinion they are very good for those who, unfortunately, can't move and then I can understand, but, as long as a person is self-sufficient and can move freely, I think they are wasted potentialities.* [Interviewee 4]

> *Yes, well it does anything so, then...I don't close the window anymore, I don't switch off the light, it turns the light on, it does anything! I stay there completely still...so, O.K. when I am 102 years old...I agree...But if it does these things now using technology...when it gets dark all the lights go on, there are lots of things around...well, human beings don't do anything anymore! [Technologies] can do all these things, if the door or the window are open, they close them...so I sit on my chair or in an armchair with that thing opposite me and I don't move anymore.* [Interviewee 8]

One participant highlighted that Giraff cannot be useful in increasing social relationships, particularly for self-sufficient people. Another participant did not consider this robot as a valid alternative or as an addition to other devices such as computer and telephone. One participant stressed the fact that Giraff could be annoying for those who live in a small house.

We noticed that the participants' capacity to change their mind (about the possible functions of the robot), after watching the Giraff video, were mainly related to their baseline relationship with ICT, that is, the more positive it was, the more easily they changed their mind and could grasp the potentialities of a robot aimed at assisting an older person who lives alone.

Discussion

Principal Findings

Our study partially challenges the previously reported positive association between older people's past and current use of ICT, as well as the positive association between the current use of ICT and higher education and lower age [4,12,14,15]. From our study, it emerged indeed that a number of interviewees who were currently using ICT (ie, smartphone) did not own a computer in the past, or did not receive higher education, or were not all *youngolder people* (aged 64-75 years). Furthermore, we found that a negative view of ICT (thinking negatively) not always corresponds to its actual rejection or underuse in everyday life.

Referring to robotics, our interviewees appeared quite open-minded. In particular, robots were considered suitable for housekeeping, for monitoring older people's health and accidental falls, and for entertainment. After watching the Giraff video [23], many interviewees acknowledged those functions and even glimpsed other potentialities.

Although ICT is providing a wealth of new devices and possibilities, because of the common belief that older people are not keen to adapt or use new technologies in the same way younger generations would, researchers have been very cautious in developing new platforms, applications, and systems for older people [27].

Indeed, literature reported some correlation between the positive or negative perceptions of ICT and older people's level of education, age, and previous familiarity with technological devices [4,12,14,15], thus supporting the theory of an educationally based digital divide. However, according to our findings, being open-minded to technologies seems to be only loosely related to older people's level of education, contrary to what was reported by other studies [14]. In our study, just half of the participants who had a computer owned a degree, and among smartphone owners, just one-third had received higher education. Furthermore, we did not find correspondence between older people's lower age and greater use of technology, contrary to previous studies [14,15]. In accordance with the existing research [4,12], we found that the actual use of technology seems more correlated with the previous use of technological tools. We also found that those who used a computer in the past seemed to be more open-minded toward new technologies, particularly smartphones. However, having used technology in the past does not seem a necessary condition to accept technology in the older age [15]; even if some participants in our study did not own a computer in the past, they did currently have a smartphone and were able to use it, not only to communicate, but also for other functions offered by the device. This is quite an important finding and it may be explained by the fact that today's technologies are certainly more available and immediate than in the past, and therefore more accessible to older people too, with little adaptation. This process seems to improve older people's digital literacy. This is confirmed by the fact that previous familiarity with technology seems to have a greater impact on the current use of computers and less on

XSL•FO
RenderX

smartphones and tablets use, whose interfacing modality is more intuitive [28].

Concerning the actual use of technological devices, we found that they are used by more *technological* older people for several aims. The most frequent use is to communicate with their children and relatives, but also to search for the meaning of some words, recipes, and to manage their money (home banking).

In accordance with a recent study [6], our results indicate that the greatest impact on the use of technologies is determined by their views: the more positive it was, more participants said they used technologies with satisfaction. However, the relationship between technologies views and their use is not straightforward. In fact, we detected some ambivalence in the participants' words. A conflicting position sometimes emerged from the interviews about the object (the technologies that the participants actually used) and the way older people represent technologies to themselves. We discovered that even among the participants who claimed to have a negative relationship with technologies, some familiarity with them can be found in practice. It seems that negative views of technology do not always affect their actual use; on the contrary, sometimes older people tend to think negatively about technology, but act positively toward it.

In general, the older people in our study had many negative preconceptions about technology. They often declared that they were not able to use it despite the fact that they actually used several devices in their daily life. This feeling of inadequacy or incompetence in using technologies also emerged from other studies [8-10], and it seems to be reflected in our interviewees' need to receive some help when using technologies. Therefore, providing technical support appears to be fundamental when proposing new technologies to older people.

With regard to the functions that the participants would ask a hypothetical robot for, when assisting them at home, we found that they require, above all, cleaning functions. In some cases, they asked for support in preparing meals and in other cases they requested to have their health state monitored or to receive monitoring and aid on request, as already reported in literature [1,7]. From the analysis of the participants' requests, it seems that the desire to maintain autonomy and avoiding the help of a caregiver at home, especially for cleaning, could make them accept a robot in their home, as reported by existing literature [4,17,29]. Those participants who claimed for a multitasking robot seemed to be moved by the same reasons, although simultaneously they made an unrealizable request. It could be interpreted as a rejection of a robot at home or conversely as an unexpressed request for human aid, which can effectively act as a maid. A few participants stated that a robot could be useless or discouraging for those people who could still perform some tasks and therefore have a negative impact on their health and autonomy.

Health status is a common concern of all participants. A vital parameter control function was also considered important by participants, especially for the anxiolytic effect it may have on them. They also requested an option for Web-based shopping (although some have pointed out that the everyday shopping is important for the older people's health) and functions that are similar to tools such as the *Bimby or Thermomix*, a kitchen robot which helps to prepare meals.

The need of sociality also emerged clearly. Some older people (especially those who stated that they had a poor or average social life) would ask a hypothetical robot for entertainment-related functions such as listening to music, singing, playing cards, receiving information or keeping up-to-date, and even sharing hobbies. A robot supporting older people living alone could thus have management and monitoring functions, as well as entertainment and company functions.

After watching the Giraff video, many of the above functions advocated by participants were confirmed, some reinforced, and new possibilities on socialization envisaged. The interviewed participants understood more clearly the usefulness of a robot in monitoring and managing emergencies, particularly the Giraff's function of calling in case of emergency, which was acknowledged by one-third of the participants. One-third of the participants highlighted the perceived usefulness of Giraff in detecting gas leaks or intrusion of strangers into the house. From the socialization point of view, the participants' acknowledgment of a robot's ability to communicate and entertain increased after watching the video. This is an interesting aspect as people may not be automatically disposed to recognize a nonhuman interlocutor (the robot) as a converser in communication processes or as a leisure companion. One participant even emphasized that the robot "could not speak, offend, and get angry like a human being"; so, this mechanical device could be sometimes preferred to the company of a person.

We found that although only a few participants considered that a robot with the abovementioned functions would be useful to those who lost their autonomy, many realized that such a device would help maintain one's own autonomy and be able to keep living alone in their home.

The way participants changed their mind (before and after watching the video) on the possible functions of a robot assisting them at home, seems to be correlated to the views of the technologies that the informants had previously expressed: the more positive it was, more they easily changed their mind and could grasp the potentialities of a robot aimed at assisting the older person who lives alone. Therefore, when proposing a robot to a prefrail elder living alone, the relationship with technologies and their views should be carefully evaluated. Both the older person and the other significant ones around them must perceive the usefulness of technologies, for example, in the contribution that they can give to their safety.

The functions that our participants asked a robot for are perfectly in line with those recommended by others belonging to different cultures, such as house cleaning, help in finding information, detecting falls or domestic accidents, monitoring vital parameters, and even in walking around [1,7]. In addition to these functions, more related to house management and to health monitoring, some participants expressed their desire for a company or entertainment function that may counteract a condition of loneliness referred by many older people as indeed a cause of frailty.

XSL•FO
RenderX

In conclusion, the results of our study indicate that a technological device such as Giraff could be well received by prefrail older people who live alone at home, even when they belong to a Latin culture (Mediterranean countries and Latin America), where the elder's assistance is commonly delegated to family members, usually women. In some cases, the robot was considered useful in communication processes or even as a leisure companion. This means that some participants could consider such a device as a remedy to their loneliness. This does not mean considering a technology such as Giraff as a substitute of a human person, but as a tool to maintain and broaden one's relationships and interests.

Limitations

The qualitative design of this study entails both strengths and limitations. Though we were able to examine data not easily accessible to quantitative research, these results are not transferable. An additional limitation stems from our recruitment mostly coming from ANTEAS, in which participants volunteered or participated in recreational activities. Therefore, even if purposively sampled, the majority of our participants could be defined as *socially active older people*. Furthermore, our participants lived in a big city. Different perspectives on technologies could be gathered from prefrail older people who live alone in a countryside.

Conclusions

This study provides insights on the perceptions of new technologies in a group of prefrail older people. This is a target of potential users that has received little attention by the literature. Our study suggests that many participants were able to use technologies, even if they did not always recognize it. Moreover, many seemed to be open-minded toward technological devices, even toward robots.

Moreover, some older people, even from Latin culture where *familism* —a strong commitment to the family as a system of support, learning, socialization, and assistance—can be a core cultural value, could recognize among the functions that a robot should have, some not trivial ones, such as company and entertainment.

Therefore, proposing new advanced technology to a group of prefrail older people, who are self-sufficient and live alone at home, is both attainable and desirable. However, before proposing robots to this target group, it is highly recommended to conduct a pilot study on the development of such new technology to make it more suitable for the end users, as we began to do in this study.

Acknowledgments

This research was supported by European project, HORIZON 2020 number 732158, named MoveCare (Multiple-actors Virtual Empathic Caregiver for the Elder). Thanks to Sarah Damanti, for her contribution in translating the initial research report, and to Giulia Dolci for her suggestions in the analysis of quantitative data. The authors are very grateful to all the older people who participated in this study.

Authors' Contributions

MM and LZ designed the study. KD, CC, and LZ administered tests and conducted the interviews. KD, CC, LZ, and MM analyzed the data. MM, KD, AB, and LZ wrote the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Interview grid.

[PDF File (Adobe PDF File), 57KB - jmir_v21i8e13228_app1.pdf]

Multimedia Appendix 2

Classification of the participant data on function and lifestyle—1.

[PDF File (Adobe PDF File), 68KB - jmir_v21i8e13228_app2.pdf]

Multimedia Appendix 3

Classification of the participant data on function and lifestyle—2.

[PDF File (Adobe PDF File), 76KB - jmir_v21i8e13228_app3.pdf]

Multimedia Appendix 4

The emergent themes resulting from the older people's interviews on information and communications technology.

XSL•FO
RenderX

[PDF File (Adobe PDF File), 24KB - jmir_v21i8e13228_app4.pdf]

References

1. Smarr C, Mitzner TL, Beer JM, Prakash A, Chen TL, Kemp CC, et al. Domestic robots for older adults: attitudes, preferences, and potential. Int J Soc Robot 2014 Apr 1;6(2):229-247 [FREE Full text] [doi: 10.1007/s12369-013-0220-0] [Medline: 25152779]
2. Wilkowsk W, Ziefle M, Alagöz F. How user diversity and country of origin impact the readiness to adopt E-health technologies: an intercultural comparison. Work 2012;41 Suppl 1:2072-2080. [doi: 10.3233/WOR-2012-0433-2072] [Medline: 22317022]
3. Mitzner TL, Chen TL, Kemp CC, Rogers WA. Identifying the potential for robotics to assist older adults in different living environments. Int J Soc Robot 2014 Apr 1;6(2):213-227 [FREE Full text] [doi: 10.1007/s12369-013-0218-7] [Medline: 24729800]
4. Sallinen M, Hentonen O, Kärki A. Technology and active agency of older adults living in service house environment. Disabil Rehabil Assist Technol 2015 Jan;10(1):27-31. [doi: 10.3109/17483107.2013.836685] [Medline: 24044649]
5. Demiris G, Rantz MJ, Skubic M, Aud MA, Tyrer HW. Home-based assistive technologies for elderly: attitudes and perceptions. AMIA Annu Symp Proc 2005:935 [FREE Full text] [Medline: 16779222]
6. Peek STM, Luijkx KG, Rijnaard MD, Nieboer ME, van der Voort CS, Aarts S, et al. Older adults' reasons for using technology while aging in place. Gerontology 2015 Jun 5;62(2):226-237 [FREE Full text] [doi: 10.1159/000430949] [Medline: 26044243]
7. Smarr C, Prakash A, Beer JM, Mitzner TL, Kemp CC, Rogers WA. Older adults' preferences for and acceptance of robot assistance for everyday living tasks. Proc Hum Factors Ergon Soc Annu Meet 2012 Sep;56(1):153-157 [FREE Full text] [doi: 10.1177/1071181312561009] [Medline: 25284971]
8. Chen K, Chan AH. Use or non-use of gerontechnology--a qualitative study. Int J Environ Res Public Health 2013 Oct;10(10):4645-4666 [FREE Full text] [doi: 10.3390/ijerph10104645] [Medline: 24084674]
9. Blažun H, Vošner J, Kokol P, Saranto K, Rissanen S. Elderly people's interaction with advanced technology. Stud Health Technol Inform 2014;201:1-10. [Medline: 24943518]
10. Gjevjon ER, Oderud T, Wensaas GH, Moen A. Toward a typology of technology users: how older people experience technology's potential for active aging. Stud Health Technol Inform 2014;201:25-31. [Medline: 24943521]
11. Peek ST, Wouters EJ, Luijkx KG, Vrijhoef HJ. What it takes to successfully implement technology for aging in place: focus groups with stakeholders. J Med Internet Res 2016 May 3;18(5):e98 [FREE Full text] [doi: 10.2196/jmir.5253] [Medline: 27143097]
12. Ezer N, Fisk AD, Rogers WA. Attitudinal and intentional acceptance of domestic robots by younger and older adults. Univers Access Hum Comput Interact (2009) 2009;5615:39-48 [FREE Full text] [doi: 10.1007/978-3-642-02710-9_5] [Medline: 25584365]
13. Broadbent E, Tamagawa R, Patience A, Knock B, Kerse N, Day K, et al. Attitudes towards health-care robots in a retirement village. Australas J Ageing 2012 Jun;31(2):115-120. [doi: 10.1111/j.1741-6612.2011.00551.x] [Medline: 22676171]
14. Chen K, Chan AH. Gerontechnology acceptance by elderly Hong Kong Chinese: a senior technology acceptance model (STAM). Ergonomics 2014;57(5):635-652. [doi: 10.1080/00140139.2014.895855] [Medline: 24655221]
15. Rodeschini G. Gerotechnology: a new kind of care for aging? An analysis of the relationship between older people and technology. Nurs Health Sci 2011 Dec;13(4):521-528. [doi: 10.1111/j.1442-2018.2011.00634.x] [Medline: 21929580]
16. Künemund H, Tanschus NM. The technology acceptance puzzle. Results of a representative survey in Lower Saxony. Z Gerontol Geriatr 2014 Dec;47(8):641-647. [doi: 10.1007/s00391-014-0830-7] [Medline: 25412802]
17. Courtney KL, Demiris G, Rantz M, Skubic M. Needing smart home technologies: the perspectives of older adults in continuing care retirement communities. Inform Prim Care 2008;16(3):195-201 [FREE Full text] [Medline: 19094406]
18. Sabogal F, Marín G, Otero-Sabogal R, Marín B, Perez-Stable EJ. Hispanic familism and acculturation: what changes and what doesn't? Hisp J Behav Sci 2016 Jul 25;9(4):397-412. [doi: 10.1177/07399863870094003]
19. Smith JA, Flowers P, Michael L. Interpretative Phenomenological Analysis: Theory, Method And Research. London: Sage Publications; 2009.
20. Wang RH, Korotchenko A, Clarke LH, Mortenson WB, Mihailidis A. Power mobility with collision avoidance for older adults: user, caregiver, and prescriber perspectives. J Rehabil Res Dev 2013;50(9):1287-1300 [FREE Full text] [doi: 10.1682/JRRD.2012.10.0181] [Medline: 24458968]
21. Fried LP, Tangen CM, Walston J, Newman AB, Hirsch C, Gottdiener J, Cardiovascular Health Study Collaborative Research Group. Frailty in older adults: evidence for a phenotype. J Gerontol A Biol Sci Med Sci 2001 Mar;56(3):M146-M156. [Medline: 11253156]
22. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975 Nov;12(3):189-198. [Medline: 1202204]
23. Digital Single Market. YouTube. Nonna Lea and a robot called Mr. Robin URL: https://www.youtube.com/watch?v=9pTPrA9nH6E [accessed 2018-10-22] [WebCite Cache ID 73MXiCRn3]

XSL•FO
RenderX

24. Cesta A, Cortellessa G, Orlandini A, Tiberio L. ong-term evaluation of a telepresence robot for the elderly: methodology and ecological case study. Int J of Soc Robotics 2016 Jan 20;8(3):421-441. [doi: 10.1007/s12369-016-0337-z]
25. Camanio Care. Giraff URL: https://www.camanio.com/us/products/giraff/ [WebCite Cache ID 73MUiN2Zo]
26. Richards L, Morse JM. Readme First For A User's Guide To Qualitative Methods. Los Angeles, CA: Sage Publications; 2013.
27. Mostaghel R. Innovation and technology for the elderly: systematic literature review. J Bus Res 2016 Nov;69(11):4896-4900. [doi: 10.1016/S2213-8587(19)30109-3] [Medline: 30872122]
28. Isaías P, Blashki K. Human-Computer Interfaces and Interactivitymergent Research and Applications. Pennsylvania, United States: IGI Global; 2014.
29. Damholdt MF, Nørskov M, Yamazaki R, Hakli R, Hansen CV, Vestergaard C, et al. Attitudinal change in elderly citizens toward social robots: the role of personality traits and beliefs about robot functionality. Front Psychol 2015 Nov;20(6):1701. [Medline: 26635646]

Abbreviations

ANTEAS: Associazione Nazionale Tutte le Età Attive per la Solidarietà
ICT: information and communications technology
IPA: interpretative phenomenological approach
MMSE: Mini-Mental State Examination

Edited by G Eysenbach; submitted 22.12.18; peer-reviewed by A Manca, S Shah, L Sequeira; comments to author 14.02.19; revised version received 03.04.19; accepted 20.04.19; published 06.08.19.

Please cite as:
Daniele K, Marcucci M, Cattaneo C, Borghese NA, Zannini L
How Prefrail Older People Living Alone Perceive Information and Communications Technology and What They Would Ask a Robot for: Qualitative Study
J Med Internet Res 2019;21(8):e13228
URL: https://www.jmir.org/2019/8/e13228/
doi:10.2196/13228
PMID:31389341

©Katia Daniele, Maura Marcucci, Cesarina Cattaneo, Nunzio Alberto Borghese, Lucia Zannini. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 06.08.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Perspectives From Authors and Editors in the Biomedical Disciplines on Predatory Journals: Survey Study

Andrew J Cohen[1], MD; German Patino[1], MD; Puneet Kamal[1], BS; Medina Ndoye[1], MD; Anas Tresh[1], BS; Jorge Mena[1], BS; Christi Butler[1], MD; Samuel Washington[1], MD; Benjamin N Breyer[1], MD

Department of Urology, University of California-San Francisco, San Francisco, CA, United States

Corresponding Author:
Benjamin N Breyer, MD
Department of Urology
University of California-San Francisco
400 Parnassus Avenue, Suite A610
San Francisco, CA, 94110
United States
Phone: 1 415 476 3372
Email: Benjamin.Breyer@ucsf.edu

Abstract

Background: Predatory journals fail to fulfill the tenets of biomedical publication: peer review, circulation, and access in perpetuity. Despite increasing attention in the lay and scientific press, no studies have directly assessed the perceptions of the authors or editors involved.

Objective: Our objective was to understand the motivation of authors in sending their work to potentially predatory journals. Moreover, we aimed to understand the perspective of journal editors at journals cited as potentially predatory.

Methods: Potential online predatory journals were randomly selected among 350 publishers and their 2204 biomedical journals. Author and editor email information was valid for 2227 total potential participants. A survey for authors and editors was created in an iterative fashion and distributed. Surveys assessed attitudes and knowledge about predatory publishing. Narrative comments were invited.

Results: A total of 249 complete survey responses were analyzed. A total of 40% of editors (17/43) surveyed were not aware that they were listed as an editor for the particular journal in question. A total of 21.8% of authors (45/206) confirmed a lack of peer review. Whereas 77% (33/43) of all surveyed editors were at least somewhat familiar with predatory journals, only 33.0% of authors (68/206) were somewhat familiar with them (P<.001). Only 26.2% of authors (54/206) were aware of Beall's list of predatory journals versus 49% (21/43) of editors (P<.001). A total of 30.1% of authors (62/206) believed their publication was published in a predatory journal. After defining predatory publishing, 87.9% of authors (181/206) surveyed would not publish in the same journal in the future.

Conclusions: Authors publishing in suspected predatory journals are alarmingly uninformed in terms of predatory journal quality and practices. Editors' increased familiarity with predatory publishing did little to prevent their unwitting listing as editors. Some suspected predatory journals did provide services akin to open access publication. Education, research mentorship, and a realignment of research incentives may decrease the impact of predatory publishing.

(J Med Internet Res 2019;21(8):e13769) doi:10.2196/13769

KEYWORDS

predatory journals; open access publication; global; citation; literature

Introduction

Increased access to the Internet has allowed for open access publishing to flourish. Traditional modes of scholarly publication involve the transfer of copyright from authors to publishers, with journal fees collected to provide access to articles. Traditional publishing, whether due to cost or perceived prejudicial peer review, is not embraced by all [1,2]. In contrast, in an open access model, authors typically retain rights to their work, it is immediately available to all readers, and funding is provided by authors themselves in the form of publication fees. Almost exclusively, publication and circulation of open access

XSL•FO
RenderX

articles occurs through the Internet. Both models involve peer review, editing, and some degree of article promotion.

So called "predatory" journals take advantage of the open access publication model to prey on unsuspecting authors [3]. Such journals advertise the same services as open access journals but may fail to provide adequate peer review, licensing, quality control, and content preservation. Predatory journals are more likely to solicit authors via email for papers in exchange for publishing fees. Given that monetary incentives comes from author submissions, they are not beholden to the same motivation of high, perceived journal quality to entice library subscription as in traditional journals. By design, early-career physicians may not suspect such journals as illegitimate and are strongly incentivized to publish to advance their careers [4]. Predatory journals' articles are not typically indexed on accepted forums and are objectively poor venues for dissemination. Predatory journals go to great lengths to appear legitimate; the fact that 400,000 items are published online per year under their banner speaks to their success [5].

Defining what is and is not a predatory journal has been a point of contention among researchers [6-8]. Beall's list, a list of publishers thought to be possibly predatory based on a single researcher's criteria, has been controversial [8]. It is now only available as an online resource. Recently, other groups have independently identified characteristics that suggest a journal is predatory [9]. Lists that purposefully provide certification of legitimate open access publishing, such as the Directory of Open Access Journals (DOAJ) [10] and the Open Access Scholarly Publishers Association (OASPA) [11], offer an alternative way to ensure high-quality open access publishing. Up-and-coming open access journals could appear to share some of the characteristics of a predatory journal and have been critical of blacklists [7,12].

What remains unknown amid this controversy is the perspective of both editors who manage predatory journals and those choosing to publish in them. While the lay press and literature has increasingly drawn attention to this issue, these journals continue to exist [3,5,13]. No literature to date addresses whether authors are unknowing victims or complacent coconspirators in the predatory publishing scheme. Our objective was to understand the motivation of authors in sending their work to potentially predatory journals. Moreover, we aimed to understand the perspective of journal editors at journals publicly cited as potentially predatory. We hypothesize that authors and editors are unaware that journals in which they are involved are possibly predatory.

Methods

Beall's list [8] was accessed on August 1, 2018, at which time 2567 publishers were noted. Beall's list was created using 48 criteria for a publisher-at-large or an individual journal [8]. The criteria assess varied factors such as grammatical or spelling errors on official communications, availability of complete editorial contact information, and the presence of false claims of indexing by services such as PubMed [14]. At present, this list is exclusively available online [8].

A total of 350 publishers were randomly selected from this list, using a random-number generator. Eight reviewers (AC, GP, PK, MN, AT, JM, CB, and SW) evaluated publishers using the following criteria. First, we confirmed an active link to the publisher website. Second, we reviewed the titles of journals within the portfolio of a publisher's work to assess if any were biomedical journals in scope. Specifically, we applied the MEDLINE criteria: "[Journals] predominantly devoted to reporting original investigations in the biomedical and health sciences, including research in the basic sciences; clinical trials of therapeutic agents; effectiveness of diagnostic or therapeutic techniques; or studies relating to the behavioral, epidemiological, or educational aspects of medicine" [15]. Journals meeting this criteria were noted. Third, we cross-referenced listed journals with the DOAJ [10], the OASPA [11], and the US National Library of Medicine [14]. We excluded journals listed on any of these sites. Fourth, we assessed whether links to specific journals were live and whether the journals had published articles. Fifth, the reviewers selected the most recent, original online article from the journal and recorded the name, country of origin, and email of both author and editor from the website. The selection criteria and ultimate composition of the cohort are summarized in Figure 1. All data were stored in REDCap [16].

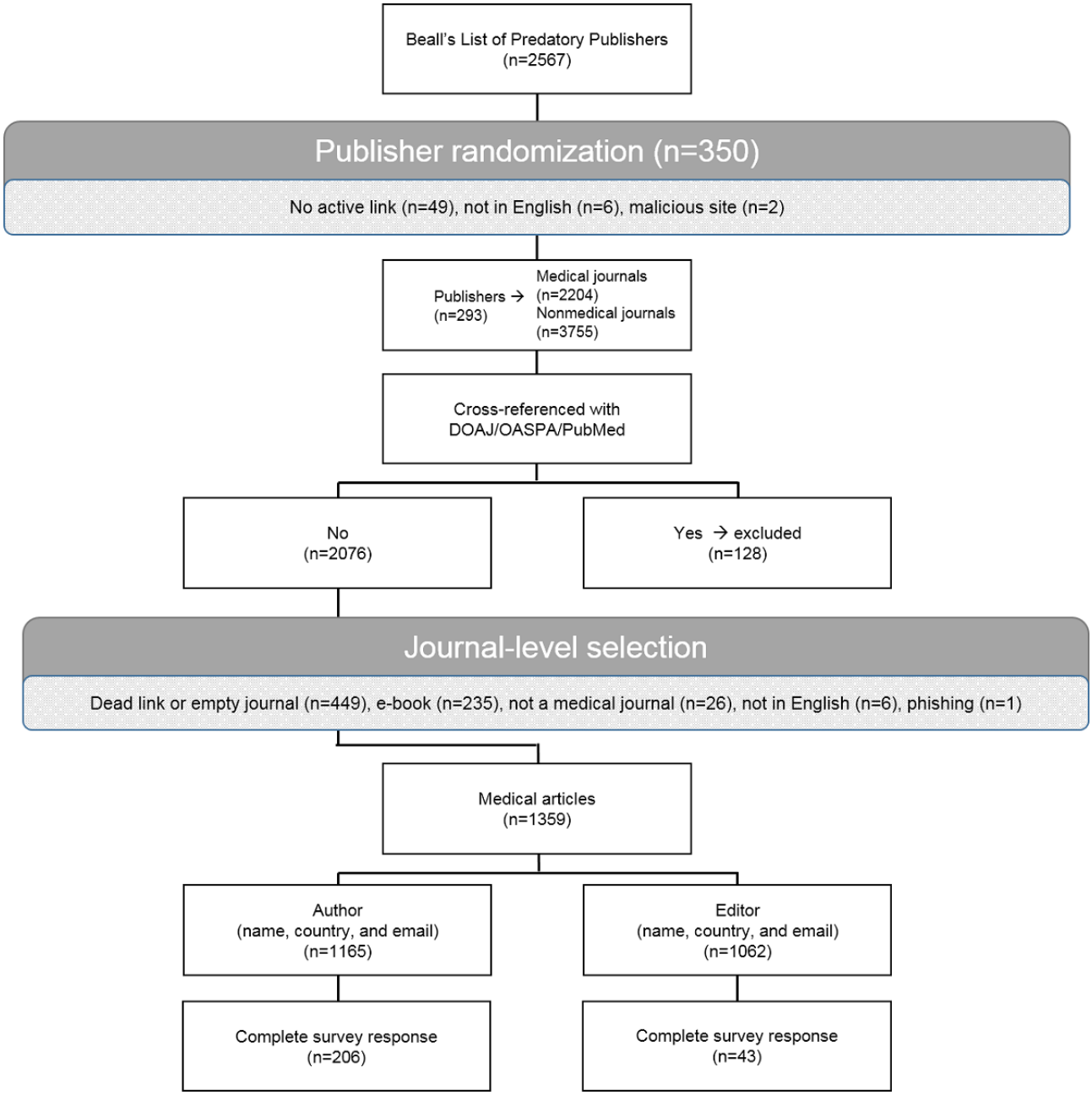
Surveys for both authors and editors were created during a round-table discussion among study authors. The surveys underwent iterative testing for clarity, content, and length. Next, we invited editors and authors to participate via email. Surveys were administered using REDCap, with one reminder sent to nonresponders. We simultaneously performed a nested randomized controlled trial on incentives to encourage survey response, which is published separately [17]. Invitees were randomized into a control group, a group eligible for a cash prize (US $100), and a group whose response lead to monetary donation to charity (US $2.50 to Rotary International per respondent). Rotary International is a nondenominational international charity with 35,000 worldwide clubs that have been instrumental in multiple projects, including the fight to eradicate Polio [18]. To maximize recruitment, survey invitations were personalized to include the article and journal name for authors and the journal name for editors.

Surveys assessed authors' basic demographics, publication history, their recollection of the editorial process for the article in question, and their knowledge regarding predatory journals (see Multimedia Appendix 1). Editors were queried about basic demographics, the editorial process for the publication in the journal in question, cost of publication, and their knowledge regarding predatory journals (see Multimedia Appendix 2). Of note, the mention of predatory journals was not made until a separate page of the survey, so as to not prejudice responses. For the purposes of this study, we defined a predatory journal as "an exploitative open-access academic publishing business model that involves charging publication fees to authors without providing the editorial and publishing services typically associated with legitimate journals." Citations of articles as reported by authors were confirmed using a search in Google Scholar [19]. Partial survey responses were excluded, but those

XSL•FO
RenderX

answering *Prefer not to answer* were not considered incomplete responses.

Figure 1. The CONsolidated Standards Of Reporting Trials (CONSORT) flow diagram. DOAJ: Directory of Open Access Journals; OASPA: Open Access Scholarly Publishers Association.



Summary statistics were used to describe the cohort. Means and standard deviations as well medians and interquartile ranges (IQRs) were used for continuous variables. Frequency tables were used for categorical variables. The chi-square statistic was used to compare frequencies between groups. Developed-nation status was based on the World Bank listing for high-income countries derived from gross national income higher than US $12,056 per capita [20]. Statistics were calculated using Stata 15 (StataCorp). The datasets used and/or analyzed during this study are available from the corresponding author upon reasonable request.

Consent for publication was granted by participants when they responded to the survey in accordance with Institutional Review Board (IRB) approval. Ethical clearance was granted by the IRB of the University of California, San Francisco, CA (approval number: 18-25351).

Results

Overview

Journals meeting selection criteria came from 181 distinct publishers. There was a substantial range in the number of journals per publisher represented in our cohort. A total of 58.0% of publishers (105/181) were singular entities, having only a single journal in their portfolio. However, several larger publishers were also included; indeed, 52.00% of (1146/2204) journals came from just 3 out of 181 publishing companies (1.7%). Of the articles selected, 59.68% (811/1359) were

XSL•FO
RenderX

published in 2018, 17.73% (241/1359) were published in 2017, and none were published before 2014. Authors and editors represented a global academic community (see Figure 2). Of all articles selected, 40.03% (544/1359) had corresponding authors from high-income countries, as defined by the World Bank [20]. The overall survey response rate was 13.0%, with editor response rates significantly lower than those for authors (6.5% vs 18.9%, respectively, *P*<.01). Several respondents provided incomplete responses (n=40). Complete responses were collated from 206 authors and 43 editors, out of a potential 1165 and 1062, respectively, for a final response rate of 11.18% (249/2227).

Authors

Responding authors had a median age of 43 years (IQR 33-54), with 47.0% (95/202) reporting that they had been in practice for over 15 years (see Table 1). A total of 7.4% (15/202) of authors self-identified as still in a training program. Among authors, 80.1% (165/206) reported that publication of articles is part of their academic promotion process. Authors published a median of 3 (IQR 2-6) manuscripts in the year prior to survey completion; lifetime median articles published was 15 (IQR 5-45). A total of 56.1% (110/196) of funding for research projects was from personal funds. Authors reported a median cost of US $190 (IQR 0-520) for publication. A total of 78.2% of authors (158/202) recalled a peer review, with 38.6% (76/197) reporting 16-30 days between submission and acceptance. A total of 68.0% of authors (140/206) had to submit revisions and 67.0% of authors (138/206) did not submit their article elsewhere before selecting this particular journal. Of 206 studies, 22.3% (n=46) were observational, 11.2% (n=23) were basic science studies, 10.7% (n=22) were case series, and 9.7% (n=20) were systematic reviews.

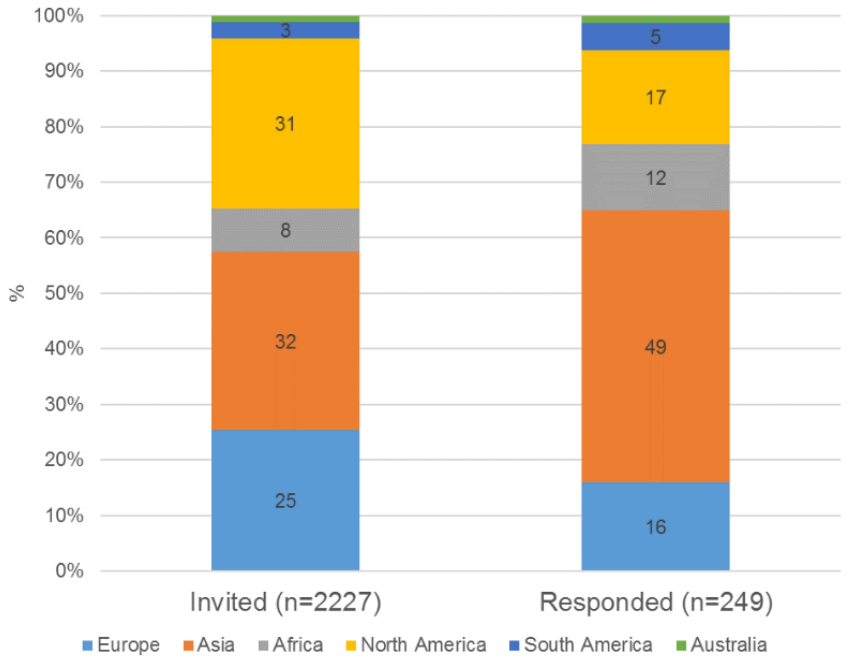
A total of 12.1% of authors (25/206) felt publication in their chosen journal was both prestigious and had a positive impact on their career. In addition, 19.6% of authors (39/199) noted a positive career impact if they reported their paper was cited versus 10.6% (21/199) that reported a positive impact when their paper was not cited (*P*=.04), with 27.2% (56/206) of total articles reportedly cited. Authors' perceptions of the journals' prestige was not impacted by whether their paper was cited. Google Scholar searching confirmed that only 16% (9/56) of those articles purported to be cited had a recorded citation. A total of 40.8% of authors (84/206) felt this particular publication was neutral in terms of career advancement and 14.1% (29/206) had published in the same journal prior to this particular publication. Most common reasons for selecting this particular journal for publication were open access (74/206, 35.9%), solicitation (52/206, 25.2%), and affordability (31/206, 15.0%).

Authors that paid fees in the top quartile (>US $519) for publication had no difference in perception of journal prestige, impact on career, shorter publication times, or fewer revision requests than authors paying lower fees (see Table 2). They were significantly more likely to use private, nongovernment funds (*P*<.01). These authors reporting higher expenses were not necessarily more likely from high-income countries (80/206, 38.8% vs 60/206, 29.1%; *P*=.08). Authors came from 54 countries, spanning all continents except Antarctica; 23.3% (48/206) were from India and 13.6% (28/206) were from the United States.

Editors

Editors' median age was 41 years (IQR 33-53) (see Table 1). Among survey responders, 40% of editors (16/40) were from India, 20% (8/40) were from the United States, and the residual editors were from varied locations. Editors reported performing editorial duties for a median of 2 journals (IQR 2-4), and for 79% of the editors (34/43), such involvement promotes their academic advancement. A total of 42% of editors (18/43) work primarily in an academic center and 79% (34/43) noted that being an editor positively impacts their career.

Figure 2. Geographic distribution of survey invitees and responses.



XSL•FO
RenderX

Table 1. Authors' and editors' basic demographics.

Demographics	Authors (N=206)	Editors (N=43)	*P* value
Age in years, median (IQR[a])	43 (33-54)	41 (33-53)	.72
Gender (male; N=205 authors), n (%)	147 (71.7)	29 (67)	.43
From high-income country, n (%)	71 (34.5)	13 (42)	.40
Years in practice (N=202 authors; N=42 editors), n (%)			
1-5	37 (18.3)	8 (19)	.77
6-10	34 (16.8)	9 (21)	
11-15	21 (10.4)	4 (9)	
>15	95 (47.0)	20 (48)	
In training	15 (7.4)	1 (2)	
Estimated publication cost in US $, median (IQR)	190 (0-520)	634 (75-1360)	.02
Estimated days from submission to acceptance (N=197 authors; N=25 editors), n (%)			
0-15	30 (15.2)	4 (16)	.01
16-30	76 (38.6)	4 (16)	
31-45	45 (22.8)	14 (56)	
46-60	25 (12.7)	2 (8)	
>60	21 (10.7)	1 (4)	
Articles did not undergo peer review (N=25 editors), n (%)	45 (21.8)	2 (8)	N/A[b]
Type of study (author), n (%)			
Observational	46 (22.3)	N/A	N/A
Other[c]	32 (15.5)	N/A	
Basic science	23 (11.2)	N/A	
Case series	22 (10.7)	N/A	
Survey research	20 (9.7)	N/A	
Systematic review	20 (9.7)	N/A	
Qualitative research	15 (7.3)	N/A	
Cross-sectional	15 (7.3)	N/A	
Editorial or letter to the editor	13 (6.3)	N/A	
Reasons for publishing in this particular journal (author), n (%)			
Open access for dissemination	74 (35.9)	N/A	N/A
Other[d]	52 (25.2)	N/A	
Solicited by editor	52 (25.2)	N/A	
Affordability	31 (15.0)	N/A	
Influenced by online advertising	22 (10.7)	N/A	
Recommendation from peer	22 (10.7)	N/A	

[a]IQR: interquartile range.

[b]Not applicable.

[c]Other types of studies include meta-analyses, randomized controlled trials, case controls, and critical reviews.

[d]Other reasons include impact factors, recommendations from supervisor, print ads, and perceived journal prestige.

XSL•FO
RenderX

Table 2. Publication costs and authors' perceptions.

Perceptions and other factors	Bottom 75th percentile of cost	Top 25th percentile of cost	*P* value
Perception of journal prestige (N=126 bottom 75th; N=71 top 75th), n (%)			
Not prestigious	19 (15.1)	9 (13)	.81
Little prestige	35 (27.8)	15 (21)	
Moderate prestige	52 (41.3)	34 (48)	
Very prestigious	15 (11.9)	10 (14)	
Most prestigious	5 (4.0)	3 (4)	
Impact on career (N=129 bottom 75th; N=72 top 75th), n (%)			
Large negative impact	2 (1.6)	1 (1)	.49
Small negative impact	3 (2.3)	1 (1)	
Neutral	54 (41.9)	28 (39)	
Small positive impact	51 (39.5)	24 (33)	
Large positive impact	19 (14.7)	18 (25)	
Estimated days from submission to acceptance (N=127 bottom 75th; N=70 top 75th), n (%)			
0-15	16 (12.6)	14 (20)	.69
16-30	50 (39.7)	26 (37)	
31-45	31 (24.4)	14 (20)	
46-60	17 (13.4)	8 (11)	
>60	13 (10.2)	8 (11)	
Revisions required (N=127 bottom 75th; N=72 top 75th), n (%)	85 (66.9)	51 (71)	.57
Funding source (N=130 bottom 75th; N=66 top 75th), n (%)			
Personal	82 (63.1)	27 (41)	<.001
Department	21 (16.2)	13 (20)	
Private	5 (3.9)	15 (23)	
Public	22 (16.9)	11 (17)	
Article subsequently cited (N=129 bottom 75th; N=71 top 75th), n (%)	33 (25.6)	21 (30)	.43

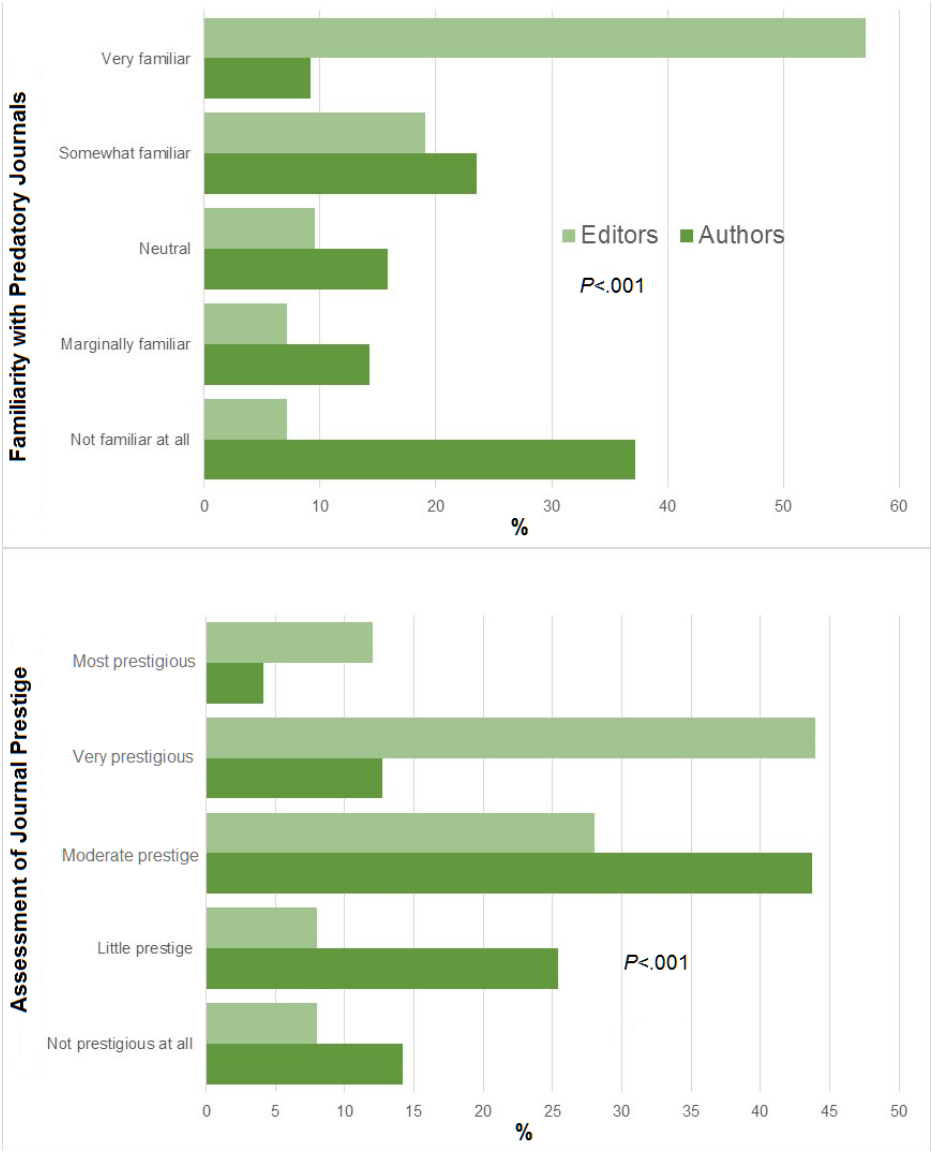
Editors' and Authors' Views on Predatory Journals

A total of 40% of editors (17/43) who responded were not aware that they were listed as an editor for the particular journal in question. These editors were excluded from subsequent analysis. Among remaining editors, 96% (25/26) felt their journal was not a predatory journal and 50% (13/26) were already aware their journal was listed on Beall's list. A total of 92% (24/26) of editors reported that their journal requires peer review. Editors reported that the mean cost to publish in their journals was US $634 (IQR 75-1360). Editors' estimated, on average, that 50% of manuscripts required revisions and that 35% were rejected. A total of 54% (14/26) of editors reported that 31-45 days elapsed between submission and acceptance.

Whereas 77% of all surveyed editors (33/43) were at least somewhat familiar with predatory journals, only 33.0% (68/206) of authors were somewhat familiar with them ($P<.001$) (see Figure 3). Similarly, only 26.2% of authors (54/206) were aware of Beall's list of predatory journals versus 49% (21/43) of editors ($P<.001$). A total of 30.1% of authors (62/206) believed that their publication was published in a journal that could be defined as predatory, once given the definition. Given knowledge that a journal was definitely predatory, 87.9% of authors (181/206) surveyed would not publish in the same journal in the future.

XSL•FO
RenderX

Figure 3. Discrepancy in author and editor opinions on predatory journals.



Narrative comments were also collected via the survey. These provided surprising and conflicting insight into the attitudes and opinions of both authors and editors. Additionally, many survey respondents felt passionate enough about this topic to contact the research staff via email to share additional long-form opinions. Selected comments are presented in an anonymized fashion in Multimedia Appendix 3. Some editors expressed surprise and outrage about their listing as editors, whereas others were defensive of their journals, described continued editorial improvements, and in rare cases disparaged Beall's list. Authors expressed regret about their publication choice, whereas others felt they had no other options or provided negative commentary about traditional publishing models. The comments emphasize the controversial and divergent opinions surrounding predatory publishing. These issues will not be solved overnight; as stated by one author, "In [country name], one must have two or three publications in [a] journal with impact, so you do not say what or why; simply you must do it."

Discussion

Principal Findings

We found editors were at least somewhat familiar with predatory journals, but the vast majority of authors (67%) were not. Per authors and editors, in limited cases, predatory journals seemingly provide some editorial support via revisions, rejections, and circulation that was enough to drive citations. Predatory journal authorship is a global phenomenon in our study, with higher penetrance in India and the United States. Alarmingly, 39% of editors reported not even being aware of being a journal editor for the journal in question. Several editors sent us comments stating they had previously asked that their name and contact information be removed from the journal websites, given no purposeful affiliation. After alerting authors that their recent publication was in a potentially predatory journal, 88% would avoid the same journal, demonstrating that via dissemination of knowledge regarding predatory practices, predatory publishing may decrease.

XSL•FO
RenderX

Services Provided by Predatory Journals

On the surface, these predatory journals are providing at least some service to authors. In our cohort, authors frequently recalled a peer review and a need to submit revisions. Editors similarly stated that a peer review was performed and approximately 35% of articles were rejected. A potential marker of journal prestige and circulation is ultimately citation, reported by 27% of surveyed authors [21]. Of note, we could only verify 16% of author-reported citations. While assessing the quality of dissemination or peer review is beyond the scope of this work, from provided comments, there is clearly a stigma with being associated with a predatory journal for both authors and editors. There is little debate that any journal continuing to list editors that renounce that affiliation are illegitimate. Nevertheless, this work urges caution in branding a journal predatory without insight provided by editors and authors; ongoing work defining predatory journals should incorporate their perspectives.

Further Analysis

We found that 40% of authors were from countries designated by the World Bank as high income; prior work suggests that authors who published in predatory journals were primarily from developing nations [22]. A predominance of Indian (34.7%), African (16.4%), and Asian (25.6%) authorship was shown previously, using a complex stratified selection criteria of publishers; in our cohort, 23% of authors were from India [5]. Recent growth in research enterprise in nations such as India has created a need for forums of publication. Publication in the mainstream scientific press may be limited by access, bias against international research work, lack of shared scientific interest, or a true marker of low-quality work [12]. India may be a case study for this situation and this could explain why it is overrepresented in authors and editors in predatory journals, as typical journals do not and cannot accept the quantity of work produced [23].

Prior work proposed that young naïve authors are the target of predatory journals, but 47% of responding authors in this study have been in practice for more than 15 years [4]. Similarly, in economics, a surprising 11% of 1284 articles in predatory journals were published by authors registered as part of a prestigious international authorship group [24]. By way of strategy, publishers target emerging markets rather than new researchers specifically [25]. By creating multiple publishing sites, each with 60-100 unrelated journals tailored to a particular culture, it can be difficult to complete an online journal search and not link to one of these journals. Further analysis suggested that such journals essentially had the same editorial team, despite glaringly different topic areas and very few, if any, publications [25].

A high proportion of responding authors in our cohort reported that publication is an important facet of academic advancement. There is a growing concern that the motivation to publish may supersede desires for research quality [26]. Both authors and editors are incentivized to publish and edit to advance their careers. While the scholarly advancement of knowledge is a noble pursuit in and of itself, a growing number of scientific articles are never cited, suggesting inherently that they are flawed, do not contribute to knowledge, or are in poorly indexed journals rendering their research essentially invisible [27]. It is a fact that due to limited resources and human capital, not all academic centers are set up to succeed as a research enterprise. This may limit the potential impact of research from such sources. Mixed with incentives to "publish or perish," authors may be tempted to lower their standards for publication and provide a market for continued predatory publishing.

Potential solutions to predatory publishing include demanding promotion committees judge faculty on the quality of publications in lieu of sheer number [27]. Authors should receive training regarding the existence of predatory publishers. The onus remains on authors to independently verify credentials of journals and to confirm journal indexing claims. New initiatives such as *Think. Check. Submit.* [28], an international collaboration providing practical resources to educate researchers, promote integrity, and build credible research publications, should be publicized [29]. Research mentors should assist colleagues and steer them toward legitimate journals. White lists or other affirming criteria should be publicized among authors seeking publication to motivate higher-quality outlets for their work. Authors in our study were most motivated by cost and the open access model, so high-quality journals meeting that criteria should be lauded. They were also motivated by solicitations, so education surrounding such email or print invitations should be circulated [30]. Accountability for publishers should be demanded, particularly for blatant false claims involving indexing, editors, and cost. Regulatory oversight in the country of origin of predatory publishers should be strengthened—in some cases, no such oversight exists—and criminal charges pursued, if relevant.

Some authors shared already-published concerns regarding blacklists, such as Beall's stifling of innovation in open access publishing, which is why we cross-referenced our findings with well-publicized white lists to generate a cohort of authors and editors [6-8]. Prior work on predatory journals has not made this effort [5-8]. White lists provide an independently verified listing of journals with ethical and quality publication standards. We elected to use the US National Library of Medicine [14] as one white-list source in lieu of Embase or other publisher-operated sites. Given that Embase and similar sites are publishing-company owned, their ability to assess the suitability of journals for inclusion may be biased, given their own financial incentives. In contrast, the US National Library of Medicine has transparent requirements for inclusion. Nonetheless, different white lists could provide a slightly different cohort of authors or editors to study and, hence, caution is advised in drawing generalizations.

Limitations

A major limitation to this study is our poor response rates from both authors and editors, but particularly editors. Given that 39% of responding editors were not even aware that they were editors at the journal in question, it is not surprising that editors-at-large may have faced confusion with our survey invitation. Moreover, unlike author email addresses, the editorial email addresses often took a generic form, such as "editor@journalname.org," and may have encountered staff

XSL•FO
RenderX

that did not forward the survey to the intended recipient. Our data comes from an international cohort, but we randomized the selection of publishers and not the countries of publishers. As such, some countries only contributed via a single survey response and the responses should not be seen as representative of all authors of that region. All authors published their papers in English journals. Nonetheless, English may not be a first language for respondents, limiting responsiveness or leading to response error. Due to monetary and time constraints, we did not perform survey adaptation or cross-cultural validation for global participants [31,32]. Survey response may have been low due to survey fatigue, disinterest, or prejudice against the topic area, which may have increased response bias [31,32]. Ultimately, this was a volunteer sample, which selected for respondents over nonrespondents. Given that our demographic data exclusively came from voluntary respondents, we cannot compare to the nonresponder group, reducing the generalizability of our work.

Conclusions

Predatory journal authorship is a global phenomenon not unique to early-career researchers. The majority of studied authors were not familiar with predatory publishing practices, despite being published in a suspected predatory journal. Alarmingly, 39% of editors reported that they were not even aware of being an editor for the journal in question, clearly confirming the unethical practices of such journals. Education, research mentorship, and a realignment of research incentives may decrease the impact of predatory publishing.

Acknowledgments

Kelly Johnson provided invaluable assistance in supporting this research. Anthony Enriquez is recognized for his tireless efforts in assisting with IRB approval. We acknowledge the generous philanthropic support of the Alafi Fund, which was awarded to BNB and supported this work.

Authors' Contributions

BNB and AC conceived and conceptualized the study, coordinated the data collection activity, carried out the statistical analysis, and wrote the initial draft of the manuscript. SW, CB, PK, GP, AT, JM, MN, and AC collected the raw data, including contact information for potential survey respondents. They also iteratively developed the survey tool and participated in the design of the study, as well as edited the manuscript. PK specifically assisted with methodology development. GP was instrumental in editing the initial drafts. GP also assisted with graphic design and overall scope of the work. AC and GP coordinated the REDCap database housing the data and created the survey invitations. BNB provided mentorship and leadership for the project. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Author online survey.
[PDF File (Adobe PDF File), 43KB - jmir_v21i8e13769_app1.pdf]

Multimedia Appendix 2
Editor online survey.
[PDF File (Adobe PDF File), 44KB - jmir_v21i8e13769_app2.pdf]

Multimedia Appendix 3
Free-form comments, edited for length, grammar, and spelling.
[PDF File (Adobe PDF File), 297KB - jmir_v21i8e13769_app3.pdf]

References

1. Teixeira da Silva JA, Dobránszki J. Problems with traditional science publishing and finding a wider niche for post-publication peer review. Account Res 2015;22(1):22-40. [doi: 10.1080/08989621.2014.899909] [Medline: 25275622]
2. Eriksson S, Helgesson G. The false academy: Predatory publishing in science and bioethics. Med Health Care Philos 2017 Jun;20(2):163-170 [FREE Full text] [doi: 10.1007/s11019-016-9740-3] [Medline: 27718131]
3. Moher D, Shamseer L, Cobey K, Lalu M, Galipeau J, Avey M, et al. Stop this waste of people, animals and money. Nature 2017 Sep 06;549(7670):23-25. [doi: 10.1038/549023a] [Medline: 28880300]
4. Cariappa M, Dalal S, Chatterjee K. To publish and perish: A Faustian bargain or a Hobson's choice. Med J Armed Forces India 2016 Apr;72(2):168-171 [FREE Full text] [doi: 10.1016/j.mjafi.2016.03.005] [Medline: 27274612]
5. Shen C, Björk BC. 'Predatory' open access: A longitudinal study of article volumes and market characteristics. BMC Med 2015 Oct 01;13:230 [FREE Full text] [doi: 10.1186/s12916-015-0469-2] [Medline: 26423063]

XSL•FO
RenderX

6. Strielkowski W. Predatory journals: Beall's List is missed. Nature 2017 Apr 26;544(7651):416. [doi: 10.1038/544416b] [Medline: 28447641]
7. Teixeira da Silva JA. Caution with the continued use of Jeffrey Beall's "predatory" open access publishing lists. AME Med J 2017 Jul;2:97. [doi: 10.21037/amj.2017.06.14]
8. Beall's List of Predatory Journals and Publishers. URL: https://beallslist.weebly.com/
9. Shamseer L, Moher D, Maduekwe O, Turner L, Barbour V, Burch R, et al. Potential predatory and legitimate biomedical journals: Can you tell the difference? A cross-sectional comparison. BMC Med 2017 Mar 16;15(1):28 [FREE Full text] [doi: 10.1186/s12916-017-0785-9] [Medline: 28298236]
10. DOAJ (Directory of Open Access Journals). URL: https://doaj.org/
11. Open Access Scholarly Publishers Association (OASPA). URL: https://oaspa.org/
12. Stojanovski J, Marušić A. Does small equal predatory? Analysis of publication charges and transparency of editorial policies in Croatian open access journals. Biochem Med (Zagreb) 2017 Jun 15;27(2):292-299 [FREE Full text] [doi: 10.11613/BM.2017.032] [Medline: 28694721]
13. Neuroskeptic. Discover Magazine. 2017 Jul 22. Predatory journals hit by 'Star Wars' sting URL: http://blogs.discovermagazine.com/neuroskeptic/2017/07/22/predatory-journals-star-wars-sting/#.WsQntejwZnJ
14. National Center for Biotechnology Information (NCBI). PubMed URL: https://www.ncbi.nlm.nih.gov/pubmed/ [accessed 2019-08-01]
15. US National Library of Medicine (NLM). Fact sheet: MEDLINE® journal selection URL: https://www.nlm.nih.gov/pubs/factsheets/jsel.html [accessed 2018-08-01]
16. Harris P, Taylor R, Thielke R, Payne J, Gonzalez N, Conde J. Research electronic data capture (REDCap): A metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
17. Cohen A, Washington S, Butler C, Kamal P, Patino G, Tresh A, et al. Altruistic donation to improve survey responses: A global randomized trial. BMC Res Notes 2019 Feb 28;12(1):113 [FREE Full text] [doi: 10.1186/s13104-019-4146-y] [Medline: 30819217]
18. Rotary International. URL: https://www.rotary.org/en [accessed 2019-06-19]
19. Google Scholar. URL: https://scholar.google.com/ [accessed 2019-06-20]
20. World Bank Data Help Desk. World Bank country and lending groups URL: https://datahelpdesk.worldbank.org/knowledgebase/articles/906519#High_income
21. Leydesdorff L, Bornmann L, Comins JA, Milojevi S. Citations: Indicators of quality? The impact fallacy. Front Res Metr Anal 2016 Aug 02;1:1-15 [FREE Full text] [doi: 10.3389/frma.2016.00001]
22. Xia J, Harmon J, Connolly K, Donnelly R, Anderson M, Howard H. Who publishes in predatory journals? J Assoc Inf Sci Technol 2015 Jul;66(7):1406-1417. [doi: 10.1002/asi.23265]
23. Satyanarayana K. Time for 'Publish in India' movement. Indian J Med Res 2004 May;119(5):vii-vix. [Medline: 15218977]
24. Wallace F, Perry T. Munich Personal RePEc Archive (MPRA). 2016 Jun 01. Economists behaving badly: Publications in predatory journals URL: https://mpra.ub.uni-muenchen.de/73075/1/MPRA_paper_73075.pdf [accessed 2019-08-01]
25. Bolshete P. Analysis of thirteen predatory publishers: A trap for eager-to-publish researchers. Curr Med Res Opin 2018 Jan;34(1):157-162. [doi: 10.1080/03007995.2017.1358160] [Medline: 28722493]
26. Rawat S, Meena S. Publish or perish: Where are we heading? J Res Med Sci 2014 Feb;19(2):87-89 [FREE Full text] [Medline: 24778659]
27. Dahlia R. London School of Economics (LSE) Impact Blog. 2014 Apr 23. Are 90% of academic papers really never cited? Reviewing the literature on academic citations URL: http://blogs.lse.ac.uk/impactofsocialsciences/2014/04/23/academic-papers-citation-rates-remler/
28. Think. Check. Submit. URL: https://thinkchecksubmit.org/ [accessed 2019-06-25]
29. Beall J. Best practices for scholarly authors in the age of predatory journals. Ann R Coll Surg Engl 2016 Feb;98(2):77-79 [FREE Full text] [doi: 10.1308/rcsann.2016.0056] [Medline: 26829665]
30. Clemons M, de Costa E Silva M, Joy A, Cobey K, Mazzarello S, Stober C, et al. Predatory invitations from journals: More than just a nuisance? Oncologist 2017 Feb;22(2):236-240 [FREE Full text] [doi: 10.1634/theoncologist.2016-0371] [Medline: 28188258]
31. Epstein J, Santo R, Guillemin F. A review of guidelines for cross-cultural adaptation of questionnaires could not bring out a consensus. J Clin Epidemiol 2015 Apr;68(4):435-441. [doi: 10.1016/j.jclinepi.2014.11.021] [Medline: 25698408]
32. Lee J, Jung D. Measurement issues across different cultures. Taehan Kanho Hakhoe Chi 2006 Dec;36(8):1295-1300. [Medline: 17215601]

Abbreviations

CONSORT: CONsolidated Standards Of Reporting Trials
DOAJ: Directory of Open Access Journals
IQR: interquartile range

XSL•FO
RenderX

IRB: Institutional Review Board
OASPA: Open Access Scholarly Publishers Association

Edited by G Eysenbach; submitted 21.02.19; peer-reviewed by O Beiki, S Nimbalkar; comments to author 11.04.19; revised version received 24.04.19; accepted 07.05.19; published 30.08.19.

Please cite as:
Cohen AJ, Patino G, Kamal P, Ndoye M, Tresh A, Mena J, Butler C, Washington S, Breyer BN
Perspectives From Authors and Editors in the Biomedical Disciplines on Predatory Journals: Survey Study
J Med Internet Res 2019;21(8):e13769
URL: http://www.jmir.org/2019/8/e13769/
doi:10.2196/13769
PMID:31471960

©Andrew J Cohen, German Patino, Puneet Kamal, Medina Ndoye, Anas Tresh, Jorge Mena, Christi Butler, Samuel Washington, Benjamin N Breyer. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 30.08.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Privacy-Preserving Methods for Feature Engineering Using Blockchain: Review, Evaluation, and Proof of Concept

Michael Jones[1,2], MS; Matthew Johnson[1,2], MS; Mark Shervey[1,2], BA; Joel T Dudley[1,2], PhD; Noah Zimmerman[1,2], PhD

[1]Center for Biomedical Blockchain Research, Icahn School of Medicine at Mount Sinai, Redwood City, CA, United States

[2]Institute for Next Generation Healthcare, Icahn School of Medicine at Mount Sinai, New York City, NY, United States

Corresponding Author:
Noah Zimmerman, PhD
Center for Biomedical Blockchain Research
Icahn School of Medicine at Mount Sinai
10-234 Marshall St
Redwood City, CA, 94063
United States
Phone: 1 650 352 3879
Email: noah.zimmerman@mssm.edu

Abstract

Background: The protection of private data is a key responsibility for research studies that collect identifiable information from study participants. Limiting the scope of data collection and preventing secondary use of the data are effective strategies for managing these risks. An ideal framework for data collection would incorporate feature engineering, a process where secondary features are derived from sensitive raw data in a secure environment without a trusted third party.

Objective: This study aimed to compare current approaches based on how they maintain data privacy and the practicality of their implementations. These approaches include traditional approaches that rely on trusted third parties, and cryptographic, secure hardware, and blockchain-based techniques.

Methods: A set of properties were defined for evaluating each approach. A qualitative comparison was presented based on these properties. The evaluation of each approach was framed with a use case of sharing geolocation data for biomedical research.

Results: We found that approaches that rely on a trusted third party for preserving participant privacy do not provide sufficiently strong guarantees that sensitive data will not be exposed in modern data ecosystems. Cryptographic techniques incorporate strong privacy-preserving paradigms but are appropriate only for select use cases or are currently limited because of computational complexity. Blockchain smart contracts alone are insufficient to provide data privacy because transactional data are public. Trusted execution environments (TEEs) may have hardware vulnerabilities and lack visibility into how data are processed. Hybrid approaches combining blockchain and cryptographic techniques or blockchain and TEEs provide promising frameworks for privacy preservation. For reference, we provide a software implementation where users can privately share features of their geolocation data using the hybrid approach combining blockchain with TEEs as a supplement.

Conclusions: Blockchain technology and smart contracts enable the development of new privacy-preserving feature engineering methods by obviating dependence on trusted parties and providing immutable, auditable data processing workflows. The overlap between blockchain and cryptographic techniques or blockchain and secure hardware technologies are promising fields for addressing important data privacy needs. Hybrid blockchain and TEE frameworks currently provide practical tools for implementing experimental privacy-preserving applications.

(J Med Internet Res 2019;21(8):e13600) doi:10.2196/13600

KEYWORDS

privacy; machine learning; confidentiality; data collection; mobile health; feature engineering; geolocation; blockchain; smart contract; cryptography; trusted execution environment

XSL•FO
RenderX

Introduction

Background

Data Privacy Issues With New Technologies

The emergence of social networks, smartphones, wearable devices, and internet of things (IoT) devices introduces unprecedented avenues for the mass collection of personal data about behaviors, biology, and health. The ubiquity of these technologies presents novel challenges when considering how to protect the privacy of individuals, and the potential to reveal sensitive and identifiable information intentionally or unintentionally has grown.

A recent Pew Research Center report found that physical location data represent one of the most sensitive data types [1]; yet more than 1000 popular smartphone apps track precise location data, some of which sell that data to third parties for targeted ads or analytics [2]. Prompts that grant an app permission to collect location data rarely reflect how the data will be used, with specifics buried in an app's privacy policy. Although location companies claim that the data collected are used to analyze aggregate patterns, not individual identities, employees and clients still have access to raw data and could identify users without their consent. Major telecommunications carriers sell user location data, and reporters have shown that data can be resold to a long chain of downstream companies. The lack of regulation in this data ecosystem has resulted in a black market for the sale of user location data [3].

Once a third party collects user data, it is difficult to guarantee that the data are not misused or mishandled. Between 2013 and 2014, Cambridge Analytica collected social media data from Facebook users for academic research, but later repurposed the data for political advertising [4]. In the past decade, major data breaches have exposed billions of user accounts [5]. There are also several instances of malicious apps that directly expose private information without user consent [6]. Regulatory efforts, including the "right to be forgotten" directive under the General Data Protection Regulation, aim to curb this trend in an effort to protect user privacy [7].

These issues present difficulties for biomedical researchers conducting studies that would otherwise benefit from convenient, passive, and longitudinal methods of data collection to identify novel biomarkers and develop digital therapeutics. There is a need for an open and trusted method for sharing data with untrusted third parties that ensures (1) posterior privacy, where personal data are not shared beyond the study for which the individual has consented and (2) that the data are only used for the intended purpose of the study.

In this paper, we reviewed the current state of privacy-preserving techniques for personal data, motivated by a location-sharing use case with applications in health care. We compared privacy-preserving techniques along several axes, including the level of trust required in the research team, the generalizability of the technique, and the availability of open source tool support. It is our intention to provide a pragmatic road map to help researchers make informed decisions about the utilization and processing of sensitive personal data. We provide a reference software implementation for the location-sharing example use case, using one of the examined techniques for privacy preservation.

Predictive Modeling in Health Care Using Biomedical and Location Data

Smartphone phone usage, and geolocation data in particular, is consequential for several health care applications. Location data have already been used in a variety of applications in health, for example, to monitor behavioral and environmental risk factors [8,9], to improve disease management and treatment delivery [10], and to inform public health policy in substance abuse [11]. In a representative example, researchers found that features extracted from global positioning system (GPS; movement and locations) and phone usage (social connectedness) strongly related to symptom severity in depression. The availability of smartphone tools provides a vector for continuous, passive assessments that could one day augment current data collection methods in clinical psychopharmacology [12]. However, it is important to stress that although geolocation data can be valuable for health care research, it is also one of the most fundamentally sensitive pieces of personal information.

Feature Engineering

Feature engineering is the process of transforming raw data into a representation that is amenable to machine learning algorithms. For example, say you are building a system to forecast driving time between two locations in a major metropolitan area. You are given data that contain the date, time of day, and driving time between the two locations for the previous year. The raw date data (YYYY-MM-DD) are unlikely to be useful for predicting drive time, but knowing whether the day is a weekday or weekend may be very useful. A machine learning scientist might write code that returns true if the date is a weekday and false if it is a weekend. The newly engineered Boolean feature, *weekday*, encodes important domain knowledge—that traffic patterns are different on weekdays compared with weekends—and may improve the accuracy of the predictions from the machine learning model.

Historically, feature engineering has been a manual process, based on the experience and domain expertise of the machine learning scientist [13]. More recently, automated systems that learn feature representations automatically from the data, such as sparse coding and auto encoders, have demonstrated good performance as the basis for deep learning models. Here, we describe a framework for feature engineering that preserves the privacy of identifiable data and is applicable to either manual or automated feature engineering procedures.

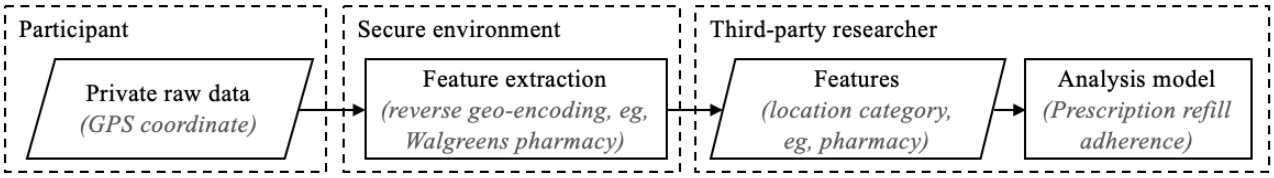
Minimal Exposure Feature Engineering

Our approach is based on the premise of minimal exposure; that participants should only reveal the minimal data required for the study and researchers should only collect the data required for the study. The feature engineering step of an analysis pipeline offers an opportunity to limit exposure by transforming identifiable, sensitive, or otherwise private data into deidentified or anonymized features. This minimal exposure approach to feature engineering creates a framework that benefits both

XSL•FO
RenderX

participants and researchers. By making it openly difficult for researchers to obtain raw personal data, participants may feel more willing to share their data and contribute to research studies. At the same time, removing researcher data access may simplify and expedite research studies by reducing the resources diverted toward maintaining secured data servers and limiting exposure to personally identifying information. In Figure 1, we illustrate the approach whereby raw data and feature extraction are encapsulated in a secure environment, removed from the researchers who are primarily interested in the underlying features.

Figure 1. A minimal exposure approach to feature engineering, where sensitive raw data are not exposed to a third party. As an example, reverse geo-encoding is performed in a secure environment to extract a location category, which could be used to determine population models on prescription refill adherence.



Interest in Blockchain Technology for Data Privacy

Over the course of 6 months in 2018, the landscape map of blockchain projects within the health care sector tripled in size, with nearly 150 projects that raised more than US $660 million [14]. The most common function of health care and biomedical blockchains is the management of data and digital assets (38%), which includes identity management, patient data, health systems operations data, and more [14]. This suggests that one of the more popular applications of blockchain technology centers around the idea that individuals may desire control of their data as a way of feeling that their privacy and data are kept more secure.

A blockchain consists of a distributed network of unaffiliated computers (nodes) that maintain an immutable record of transactions that are verified using a cryptographic protocol. Blockchain networks are further characterized as public, private, or consortium networks depending on who can participate in the network, and how transactions are verified. In public blockchains, transactions are verified and a global state of *truth* (distributed ledger) is maintained by a *trustless* network. A *trustless* network refers to a decentralized network with a consensus protocol. The consensus protocol incorporates sender authenticity via public key cryptography, game theory and cryptoeconomic (digital currency) incentives, and computational complexity to ensure that honest nodes are rewarded, and dishonest nodes are penalized to maintain the canonical truth. By making each transaction auditable and permissionless, public blockchains ensure data integrity, trust, and verifiability.

Advances in blockchain technology have enabled the deployment of rule-based, self-executing software code called smart contracts. Smart contracts remove the need for intermediaries by acting as predefined arbiters. In addition, smart contracts are immutable and publicly verifiable when the contract code is made public. The combination of smart contracts with a *trustless* environment is what eliminates the need for trusted third parties that are responsible for managing private data. These features make smart contracts particularly relevant to this study.

Aim of This Study

The aim of this study was to examine and compare current privacy-preserving methods based on their ability to maintain the privacy of personal shared data. Methods were compared based on the level of trust required of a third party, and the practicality of implementing these techniques framed as a feature engineering step. This study also aimed to identify the more promising techniques that researchers and software developers can use when building applications concerned with preserving data privacy.

The examination is set against a practical use case of collecting location data from individual participants, from which interesting features related to health can be extracted. To make this example as accessible to researchers as possible, we provide an open-sourced software project that implements one of the examined techniques for the location sharing use case.

Methods

Primary Outcomes

The primary outcomes of this study are as follows:

1. Define a set of properties on which to evaluate the privacy-preserving properties of each approach.
2. A qualitative comparison, grounded in a geolocation feature engineering use case, of the privacy-preservation properties of each approach.
3. A proof-of-concept software implementation for extracting the category of a location from GPS coordinate data while maintaining privacy using one of the more practical blockchain techniques.

Literature Review

We conducted a review of literature, health care–related blockchain use cases, and applied blockchain projects on the Web. These techniques were identified using keyword searches in electronic databases (Google Scholar and PubMed) and search engine (Google) results. The keywords were *privacy blockchain*, *deidentification*, and *privacy feature engineering*. The results at the time of the search (January 2019) consisted of methodologies described in a variety of formats, including 4 academic papers in peer-reviewed journals, 6 academic papers in conference proceedings, 2 literature and product surveys, 1

XSL•FO
RenderX

doctoral dissertation, 7 scientific journal preprints, 11 product specifications, and 1 academic lecture materials.

The techniques were divided into the following categories: (1) methods that rely on a trusted third party, (2) cryptographic methods, (3) trusted execution environments (TEE), and (4) methods incorporating blockchain. Examples about existing implementations of each of these technologies are included in Multimedia Appendix 1 [15-43].

Evaluation Properties

Data privacy laws [44-46] offer a regulatory perspective on the several dimensions in which data privacy can be compromised. Table 1 summarizes some of the key regulatory principles.

These regulatory guidelines make it clear that data privacy is highly dependent on the responsibilities of trusted organizations, and the capabilities of the technologies they implement. We predict that future data-sharing systems will be informed by these privacy guidelines and that a framework for evaluating privacy-preserving technologies should map to these guidelines. In this paper, each privacy-preserving approach is evaluated based on the following properties:

1. The level of trust required in a third party because of the following:
   - Third-party access to raw data
   - Participant visibility of data use
   - Third-party ability to reuse data
   - Centralization and single points of trust
   - Potential for security vulnerabilities
2. The generalizability and implementation practicality of the technique:
   - Computational or communication complexity
   - Implementation complexity
   - Availability of developer tools
   - Availability of open source tool support

Table 1. Data privacy laws in the European Union and the United States.

Source and guideline	Summarized text
General Data Protection Regulation Article 5	
“data minimisation”	Personal data collection is limited to what is necessary
“lawfulness, fairness and transparency”	Personal data are processed in a transparent manner
“purpose limitation”	Personal data are collected with an explicit purpose, and further processing adheres to the initial purpose
“accountability”	Third parties are responsible for adhering to privacy laws
“integrity and confidentiality”	Personal data are securely processed and there are protections against unauthorized use
Health Insurance Portability and Accountability Act Privacy Rule	
Limits who can view and share an individual’s health information	Health information cannot be used for purposes not directly related to providing health treatment without an individual’s consent (with exceptions)
Health Information Technology for Economic and Clinical Health Act Subtitle D	
Data security of digital health information	Electronic medical records must be secured, and data breaches must be reported

Description of Geolocation Use Case

Like most complex data types, GPS data are typically transformed before being used in an analysis through feature engineering. There are 2 broad classes of geolocation features that underlie most of the current health care research applications of geolocation data.

Statistical Descriptors

They compute summary statistics from the raw GPS data. For example, total distance traveled in a day, the variance in number of locations visited, and the travel radius.

Semantic Descriptors

They combine the GPS data with a third-party geospatial information system to determine location types, such as library, gym, or house of worship or broad location themes (eg, neighborhoods with high rates of crime defined by census data).

A few examples of application use cases that would use geolocation features include replacing active monitoring tasks [8-10,47-49], triggering just-in-time interventions [10,49,50], and accessibility to health services [11,51]. Geospatial applications that incorporate blockchain include the management of IoT devices, crowd-sourced data collection, and emergency response [52].

Reference to *geolocation feature extraction* will be made in the Results and Discussion sections to ground the investigation in a practical use case while evaluating different approaches for preserving privacy.

Results

Trusted Third-Party Methods

In a traditional biomedical research setting, the protection of human subjects is managed by an institutional review board (IRB) at the research institution. The role of the IRB is to certify

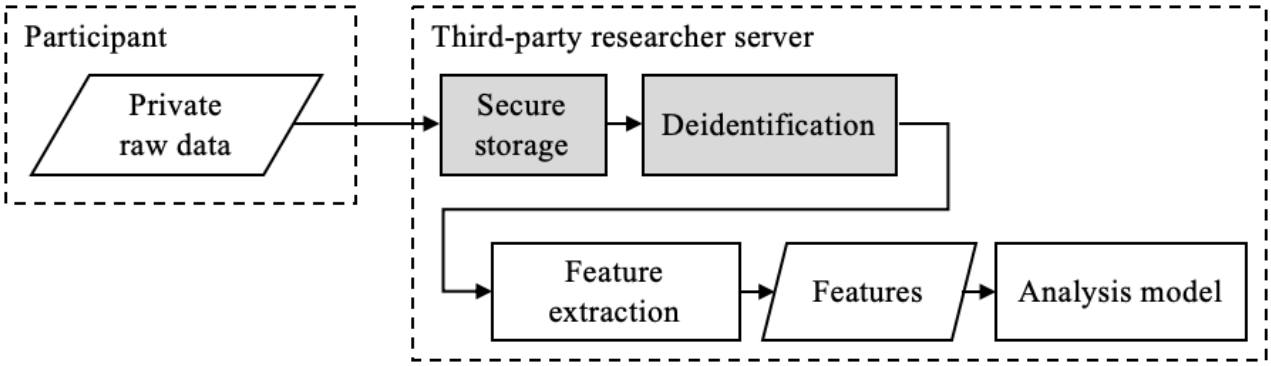
XSL•FO
RenderX

that research subjects are informed of the risks of participating in research, data security guidelines are followed, and risks and safeguards are clearly outlined and mitigated. In this model, the research institution operates as a trusted third party with a responsibility to protect patient data privacy. However, new forms of research enabled by smartphone technology, biosensors, and the routine collection of large datasets are changing the nature of research and straining the traditional process whereby a single institution can act as the trusted third party [53]. In the following sections, we cover 2 conventional approaches to privacy preservation that rely on a trusted third party.

Server-Side Deidentification

A typical server-side data collection pipeline for a research study will ingest raw participant data from a client-side application and incorporate encryption, access control, deidentification procedures, or some other method to ensure that the raw data are not irresponsibly exposed. This approach is straightforward to implement and can afford a research team strict control over the feature engineering pipeline. Software updates can be made server-side without the need to force users to update client-side. However, a high degree of trust must be placed in the research team, as private data are exposed at several stages in the pipeline (Figure 2).

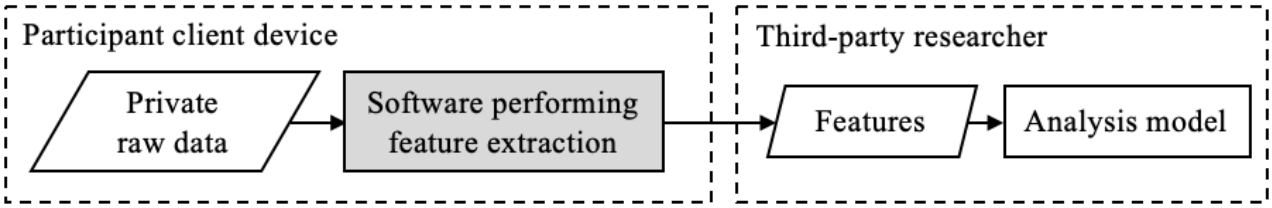
Figure 2. Server-side deidentification: (1) raw data may be exposed during feature extraction and analysis [54], (2) data access control is centrally controlled and mutable, so there is no strict enforcement of how data will be used, and (3) deidentification procedures are often single-use custom software implementations and are unlikely to be open sourced or certified to be thorough and secure. In the case of global positioning system location data, the raw data itself can sometimes be combined with external sources of data (eg, social media) to identify an individual [55-57]. Secure storage and deidentification are highlighted in gray to indicate steps in data pipeline that require trust with handling private data.



Client-Side Feature Extraction

Deployment of software that maintains raw data on device and performs feature extraction locally is another viable method for privacy preservation and increasingly a gold standard when it comes to data collected on smartphones. Localizing data on a device so that only the participant has access to it eliminates the risk that third parties can expose, misuse, or repurpose the raw data, but it relies on the integrity of the installed software (Figure 3).

Figure 3. Client-side feature extraction, where the installed software is highlighted in gray to indicate that some level of trust is needed that the software is secure and honest: (1) participants must maintain an updated version of the software so that the feature engineering is appropriate and secure and (2) participants must trust software developers or software validators that the installed application is performing as intended. Open-source software can increase visibility and provide stronger assurances of data privacy but practically requires additional security verification.



Cryptographic Techniques

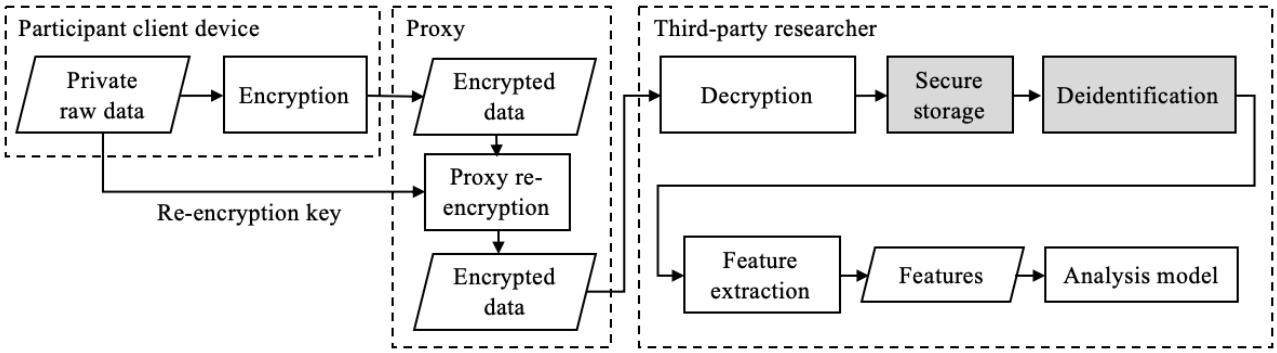
Proxy Re-Encryption

Proxy re-encryption (PRE) is a technique in public-key encryption that allows a proxy to delegate decryption access to encrypted data from one party to another (Figure 4). An important characteristic of PRE is that the proxy does not learn any information about the contents of the encrypted data. As such, it is a powerful data access control technique.

However, PRE is limited in its utility for privacy-preserving feature engineering because data access control does not provide a mechanism for posterior privacy. This still requires trust that the research team will manage the decrypted data securely and honestly.

XSL•FO
RenderX

Figure 4. A participant can provide decryption rights to a researcher through proxy re-encryption. The researcher must still be responsible for secure storage and deidentification of the data, along with posterior privacy.



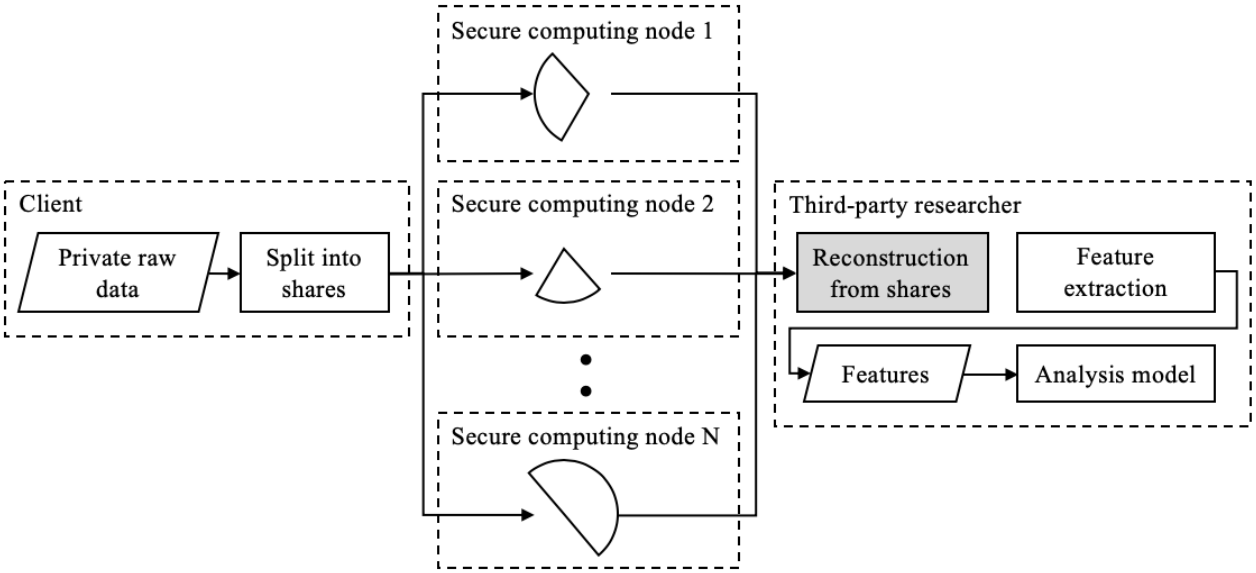
Secure Multiparty Computation

Multiparty computation (MPC) is a class of cryptographic protocols that emulate computations evaluated by a trusted party but instead distributes trust among several parties (Figure 5). MPC is best suited to feature engineering problems that compute a result on aggregate data rather than a single participant's data. In addition, MPC suffers from being exponential in communication (between parties), which limits performance.

Some implementations tackle private data sharing using MPC protocols and attempt to mitigate security risk further by supporting distributed storage architectures. However, these distributed systems are typically limited in the number of nodes and managed by a single organization (ie, single point of trust). In theory, a single point of trust could have an agenda and exert a degree of control over every node in the distributed network. However, if all the nodes are controlled by a single organization, it is feasible for the organization to access private user data.

The principle behind using a distributed network of computing parties is that there is no single point of trust and is an important one that will be revisited when examining blockchain methods.

Figure 5. Distributing trust among multiple parties with multiparty computation. In one type of multiparty computation, private data can be decomposed into secret shares and stored on several computing nodes; reconstruction of the private data is only possible with all (or a majority) of the secret shares. As shown in the figure, this alleviates the issue with secure storage of the data but does not secure the data once it is reconstructed. Trust is still required when reconstructing a single participant's private data and is highlighted in gray.



Homomorphic Encryption

Homomorphic encryption (HE) is a form of encryption where a computation on encrypted data will produce the same result as performing the computation on the unencrypted data before encryption. This can be formalized as illustrated in Figure 6. There are different schemes of HE, including partial and fully HE; partial homomorphic encryption (PHE) indicates one or more operations can be run on encrypted data and preserve homomorphic properties, while any arbitrary computation is possible when dealing with fully homomorphic encryption (FHE). HE is strong from a privacy-preserving standpoint because a participant can maintain exclusive ownership rights over data but potentially limits the feasibility of feature extraction (Figure 7).

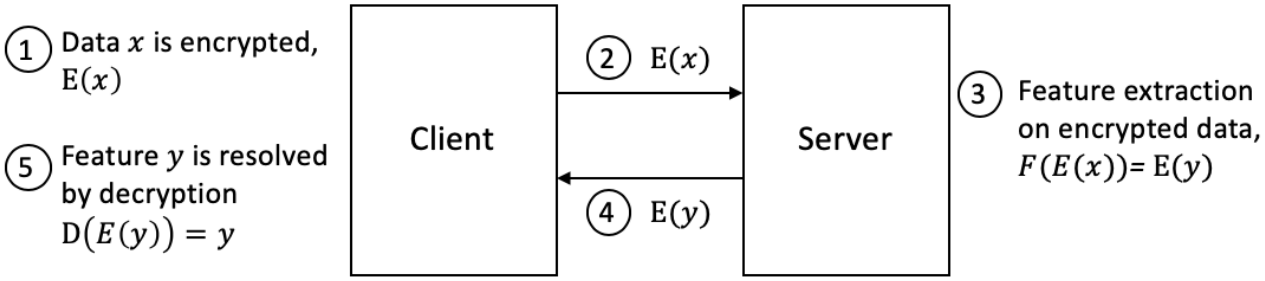
XSL•FO
RenderX

Figure 6. Equation illustrating property of homomorphic encryption, where E(x) represents the encryption of data x.

$$E(x \otimes y) = E(x) \oplus E(y)$$

$\otimes$ and $\oplus$ are some choice of operator

Figure 7. Typical client-server homomorphic encryption (HE) pipeline. Raw data ownership is maintained client-side but inaccessible to other parties. HE is limited for the feature engineering use case because it requires decryption as a last step. If researchers are provided the ability to decrypt the feature data from step 5, then they would also have the ability to decrypt the raw, sensitive data from step 1.



In a data-sharing context, HE may be suitable in specific use cases for feature extraction when the encrypted data vector is an interesting feature itself. There is also a broader applicability for privacy-preserving feature extraction when features from individual participants are not required, and data can be aggregated in the encrypted domain. However, the applicability of HE is not suitable for general purpose feature engineering in its current state and depends on the nature of the data and of the feature extraction. For example, resolving a location type from a GPS coordinate is not computable but rather the result of a lookup function. HE would not serve this kind of scheme.

Another major hurdle in the adoption of HE is the increased computational complexity of processing encrypted data, resulting in extremely long processing times. Sophisticated computations cannot be achieved with PHE, and FHE suffers from very low computational performance [17]. This makes all but the simplest operations impractical using HE.

Zero-Knowledge Proof

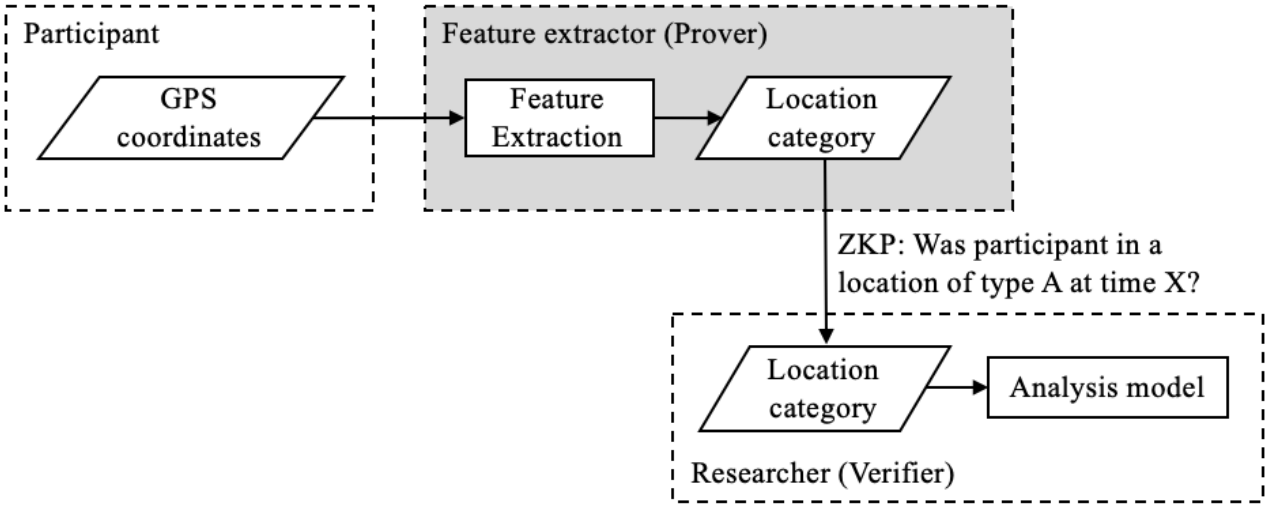
A zero-knowledge proof (ZKP) is a cryptographic method whereby a *Prover* convinces a *Verifier* whether some mathematical statement is true or false, without revealing any underlying data. Privacy preservation is built into ZKPs, making it a strong approach for handling private data.

ZKPs can be further classified based on 2 kinds of statements that should be proved in *zero knowledge*: statements about facts (eg, that a participant's GPS coordinate corresponds to a *hospital*) and statements about knowledge (eg, that a participant's GPS coordinate is known) [19]. The latter kind is a ZKP of knowledge and is the most common application of ZKPs, which is identification and authentication (eg, password authentication). However, in the context of feature engineering, it is the former type of problem that extracts some metadata that is relevant.

The ZKP concept closely mirrors the minimal exposure framework described in Figure 1. To address the geolocation feature extraction use case, a data pipeline as shown in Figure 8 could be implemented.

XSL•FO
RenderX

Figure 8. Hiding global positioning system (GPS) coordinates behind a simplistic zero-knowledge proof (ZKP) implementation. Practically, this looks like a simplistic black box whereby the ZKP is a subroutine that performs feature extraction without revealing the raw GPS coordinate to the researcher. For example, this subroutine could consist of implementing a lookup table that maps GPS coordinates to category of location. The challenge then lies in careful implementation and permeability of this subroutine to ensure security of the data (highlighted in gray). A malicious party should not be able to identify a participant's GPS coordinates through a process of elimination by trying several inputs. The analogue would be trying to guess a password via brute force.



As ZKPs more broadly correspond to a variety of techniques, it is difficult to recommend it summarily for general feature extraction problems, and they should instead be evaluated on a case-by-case basis. In addition, some common challenges include implementation and computational complexity. Sometimes, they still require a trusted third party to prove a statement [58]. It is our view that ZKPs hold potential as building blocks for privacy-preserving protocols but is currently an active area of research [59] rather than a practical and accessible tool for implementation.

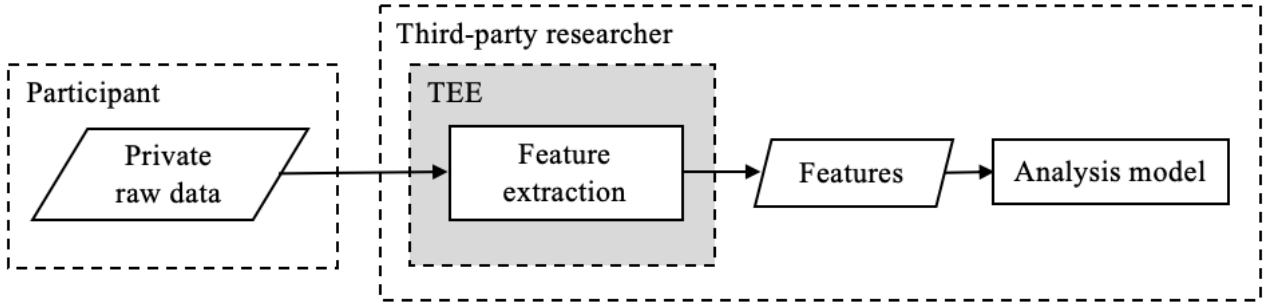
Trusted Execution Environments

A TEE (also referred to as secure hardware enclave) is a chip-level hardware design in modern processors that enables isolated execution over confidential data. Figure 9 illustrates how TEEs could encapsulate private raw data without exposing to the researcher. A major benefit of TEEs is that they have little performance overhead over native computations, making them practical for a wide range of applications [25], while providing guarantees that malicious applications cannot tamper with the computations running on secure enclaves.

An important consideration when using TEEs is the very real risk of hardware vulnerabilities that can be exploited. In early 2018, hardware vulnerabilities in modern commercial processors were reported that can expose private data to rogue processes (Meltdown) or attacks on processors that perform branch prediction (Spectre) [61]. Another vulnerability (Foreshadow) explicitly affected Intel SGX processors [62], bringing to question how trustworthy a secure enclave can be. In an effort to address these concerns, emerging open-source TEE projects argue that security by obscurity in commercial designs is insufficient and that community-driven, open security will lead to more reliable designs [25].

Another criticism of TEEs regarding data privacy is that practical applications might use a single or handful of TEEs, which centralizes the management of data. The thought is that it is still a "very strong assumption to require all participants to globally trust a single or handful of (TEE) processors" [63]. To address this limitation, some projects have emerged that incorporate blockchain with TEEs to decentralize the network of computing nodes. This approach will be examined further in the next section on blockchain methods.

Figure 9. A trusted execution environment (TEE; highlighted in gray) provides data privacy through encapsulation. Security features include memory isolation, memory encryption, isolated architecture, and secure key provisioning. A remote attestation process follows to verify correct execution of a program and provide a proof of origin [60]. A level of trust is still required in TEEs (highlighted in gray) because of the risk of hardware exploits.



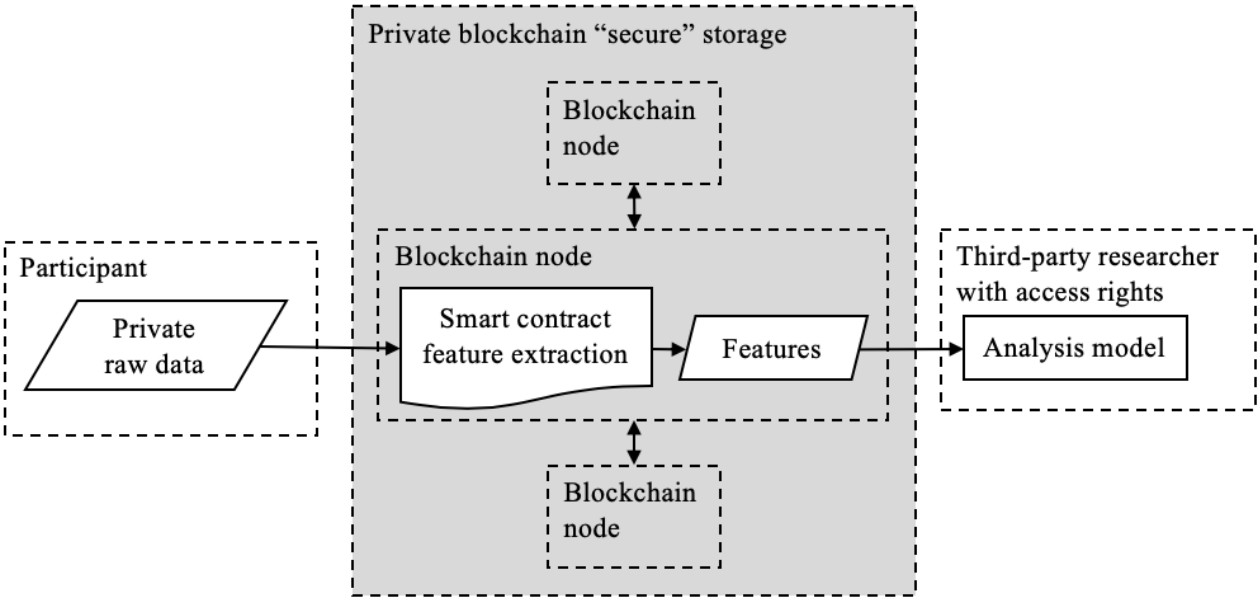
XSL•FO
RenderX

Blockchain Methods

Private and Consortium Blockchains

Private and consortium blockchain networks limit who can participate in the network, usually by creating access controls that are managed by highly trusted entities. Often, additional rules are included to create a permissions system, control which nodes can verify transactions, and make transaction data private to the parties involved. The last of the reasons is particularly appealing in a privacy-centric data–sharing context, but it comes at the price of placing trust in the maintainers of the private network. However, the very same feature that makes public blockchains so appealing—a trustless, decentralized network state secured by cryptoeconomic incentives—is missing from private blockchains and is illustrated in Figure 10.

Figure 10. Private blockchain where it represents a secure data environment but requires similar trust as other centralized techniques (participant consent to release data, trusted parties). For this reason, the entire private blockchain network is highlighted in gray.



Public Blockchain Smart Contract

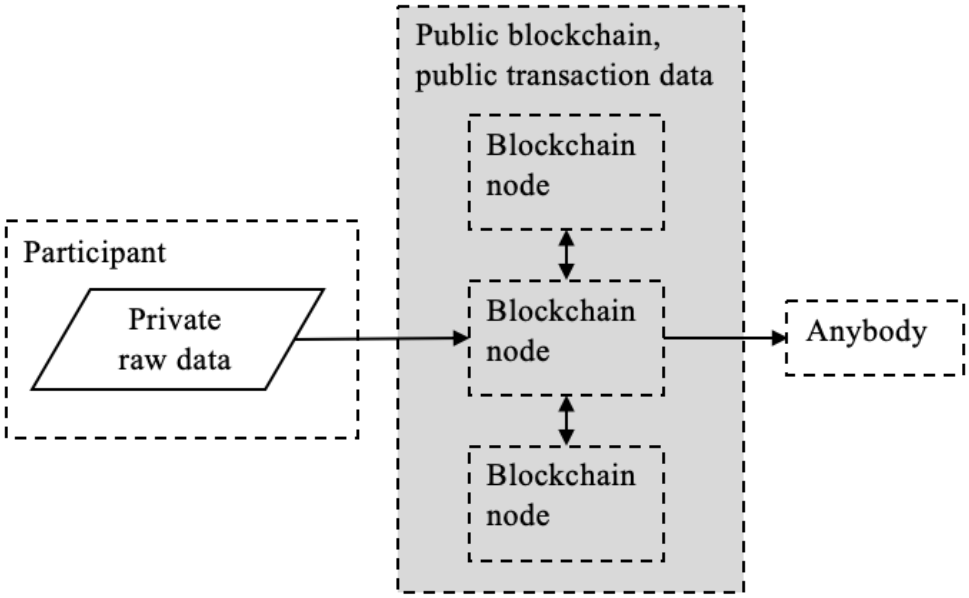
Smart contracts on a public blockchain are small, modular pieces of software that cannot be changed once they are deployed on the network. This is an advantageous quality for privacy-preserving software, because a user of the smart contract is guaranteed that their data will always be processed the same way. The functionality of deployed smart contracts is verifiable when the smart contract code is made public.

However, the first important acknowledgment when considering traditional public blockchains (eg, Ethereum) for data sharing should be that input data uploaded to a public blockchain is publicly visible and permanently recorded. This allows for all fund senders and recipients, all transaction data, and the state of every contract variable to be visible to any observer, as illustrated in Figure 11.

One argument is that blockchains offer privacy because the originator and recipient of data transactions are described only by a randomly generated account address. Therefore, pseudonymity is possible if participants generate new addresses for each transaction. However, Web trackers have shown it is possible to deanonymize users by analyzing transactions [64], and in the case of certain sensitive data such as GPS coordinates, it is possible to reidentify a large fraction of users by comparison with other structured data available (eg, from social media) [33,55-57].

XSL•FO
RenderX

Figure 11. Public blockchain implementation, where transaction data are public. The entire network is highlighted in gray to indicate data exposure. The standard workaround is to ensure that any sensitive data recorded are encrypted. As a result, it is generally impractical to run feature extraction on a smart contract and would require some off-chain computation on a centralized server that requires trust.



Privacy-Preserving Blockchains Incorporating Zero-Knowledge Proofs

The allure behind incorporating ZKPs with a public blockchain is to enable data privacy while maintaining the benefits of blockchain: no single point of trust and immutability of transactions. One implementation of this technique that captured public attention was the ability to hide the origin, destination, and amount in a cryptocurrency transaction [18,34,35].

The idea of extending ZKPs on blockchain to include smart contract logic would be a powerful catalyst for privacy-preserving, trustless applications. A blockchain proposal called Hawk [33] uses ZKPs to verify transactions and executes private smart contracts off-chain. Unfortunately, Hawk cannot guarantee posterior privacy because it relies on a minimally trusted manager, which is disincentivized from revealing sensitive data during transactions but provides no guarantees after the transactions are complete. In addition, the Hawk paper has yet to materialize into a usable software release. Similarly, there are also no details that describe if and when private smart contracts on Ethereum will be available in the near future.

The limitation of ZKP computational complexity is heightened in the context of blockchain, which requires the technology to be deployed at the distributed scale. ZKPs on blockchain is an active research area, making private data sharing and feature engineering in this context inaccessible for the time being. An application for sharing GPS coordinate data using ZKPs with blockchain smart contracts will have to wait for the technology to develop.

Privacy-Preserving Blockchains Incorporating Trusted Execution Environments

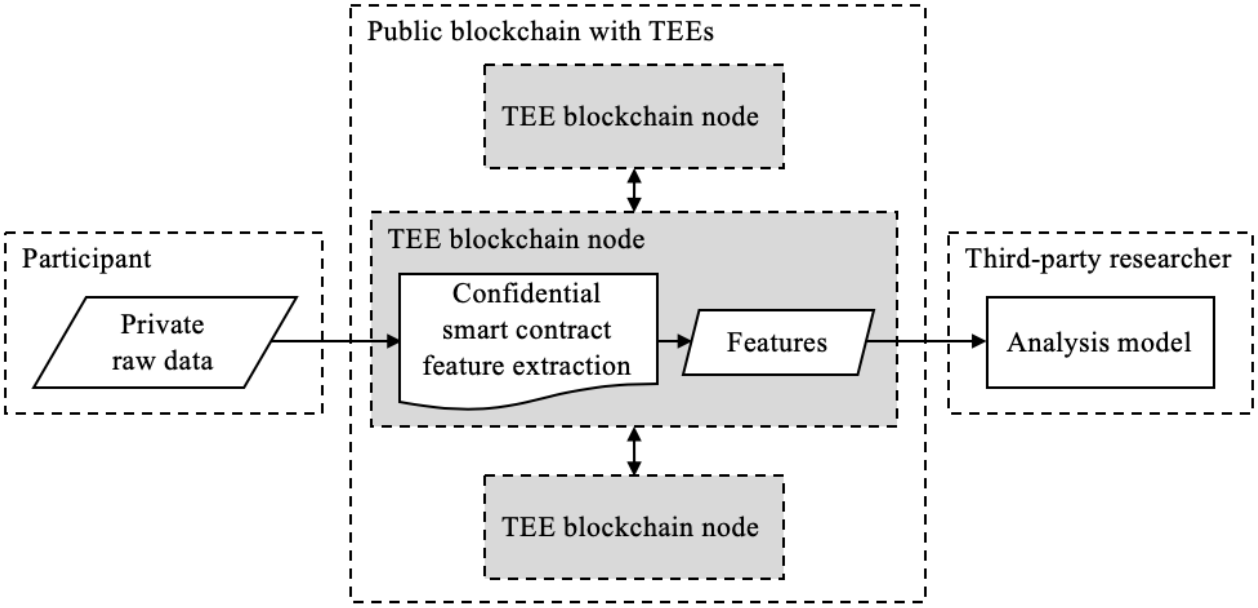
In an effort to encourage less data siloing, platforms are emerging that hybridize TEEs and blockchain smart contracts. This approach boasts the strengths of modular, immutable software with isolated computation environments so that the data pipeline is transparent and secure. This technique is illustrated in Figure 12.

Nevertheless, this approach still hinges on the security of the underlying TEE hardware and its vulnerabilities. Platforms that incorporate blockchain and TEEs are new technologies, and experimental by nature, so underlying security threats remain to be uncovered.

This approach also provides benefits in terms of implementation practicality and accessibility. The Oasis [32] and Enigma [36] projects are developer friendly, releasing documentation, tutorials, and testnets on which to develop and deploy applications.

XSL•FO
RenderX

Figure 12. Blockchain that incorporates confidential smart contracts on trusted execution environments (TEEs) for feature extraction. Similar to single TEEs, the risk of hardware exploits should not be ignored.



Software Implementation of Location-Sharing Use Case

In this paper, generalizability and practicality of implementation were two of the evaluation criteria used when comparing privacy-preserving techniques. To further evaluate our findings that the hybrid blockchain-TEE technique provides a practical platform for developing privacy-preserving software, we implemented a proof-of-concept software application that addresses the location-sharing use case that frames this paper. This use case is based on the scenario where a research study participant shares useful features about their location data with a third-party research team, without revealing their raw GPS coordinates.

The implementation consists of the following:

1. A smart contract deployed on the Oasis Devnet.
2. A smartphone (iOS) app with a graphical user interface for participants and third parties to interact with the smart contract.

Confidential smart contracts on the Oasis Devnet enable private transaction data and private smart contract state (Table 2), which are used to maintain participant confidentiality and can be used to conceal a participant's raw geolocation data. The Oasis Devnet manages per-session encryption keys that are used to encrypt communication between a client and smart contract instance, such that nobody else can view the unencrypted transaction data [65].

Figure 13 illustrates user interactions with the deployed smart contract. The contract provides a publicly accessible method for participants to post their timestamped location data. Participant identity is kept confidential by maintaining a private mapping in the smart contract state between participant wallet address and a participant identifier.

The smart contract also provides a publicly accessible method for third parties to register a geocoordinate with a predetermined category of location (ie, *hospital*, *gym*, or *pharmacy*). Third parties can query the smart contract to view participant visits to categories of location. This data could be used to build a model of participant visits to registered pharmacies, for example.

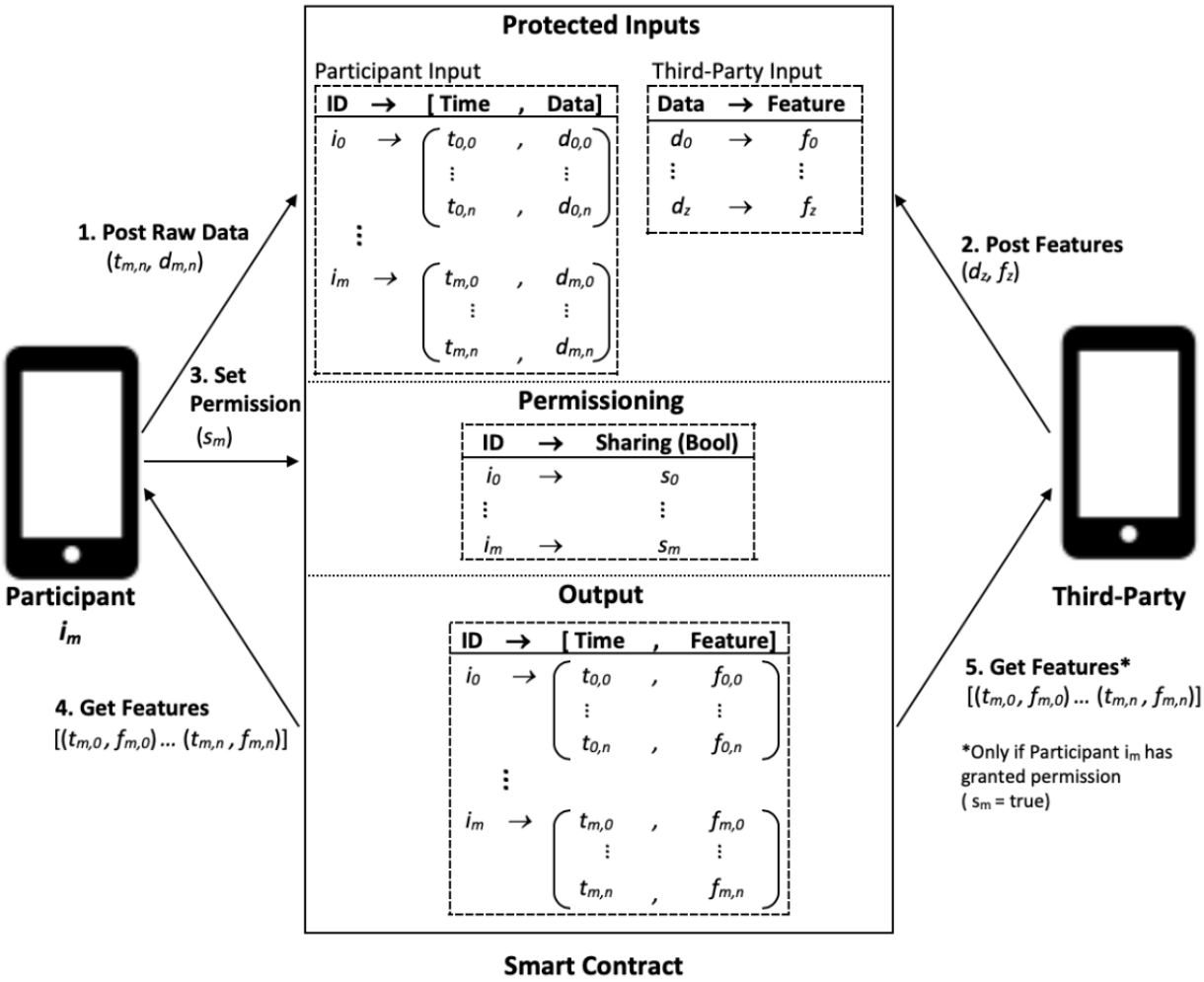
Table 2. Information visibility on Oasis Devnet.

Visibility	Information
Public	• Transaction sender address (ie, participant wallet address and third-party wallet address) • Transaction recipient address (ie, smart contract address) • Transaction value transaction (ie, amount in *DEV*, the Oasis Devnet token used to fund transactions)
Private	• Transaction argument data (ie, raw GPS[a] coordinate data) • Transaction result data (ie, returned feature data) • Method name called by a transaction (ie, "postParticipantLocation") • Smart contract state (ie, mapping of participant wallet addresses to participant ID) • Event data (not used for this prototype; events (logs) can be emitted in response to transactions)

[a]GPS: global positioning system.

XSL•FO
RenderX

Figure 13. Implementation of the proof-of-concept software application that would address a location sharing use case. A participant, identified by an identifier im, would (1) post their raw identifying location data, dm,n, along with its respective timestamp, tm,n using their iOS device. A third party, from another iOS device, is able to (2) post a raw location data, dz, with a respective feature, fz. For example, a feature such as the category of a location like the string Hospital or Pharmacy. If the location data, dm,n, posted by the participant matches the location data, dz, posted by the third party then the participant's respective timestamp, tm,n, is mapped to that respective feature, fz. The participant is, at any time, able to (3) set the sharing permission, sm, of all of their previously posted timestamps, tm,0…tm,n, and associated features, fm,0…fm,n, to third parties. The participant is also able to (4) get all of their previously posted timestamps, tm,0…tm,n, and associated features, fm,0…fm,n. The third party is also able to (5) get these same timestamps, tm,0…tm,n, and associated features, fm,0…fm,n if and only if the participant has granted permission, sm=true.



The smart contract is currently deployed on the Oasis Devnet as a traditional smart contract, not as a *confidential* smart contract. Only *confidential* smart contracts run on TEEs and maintain the privacy of smart contract state values and transactional data. However, the library (web3swift) for deploying smart contracts from an iOS smartphone app does not yet support *confidential* smart contract deployment to the Oasis Devnet at the time of this writing. We plan to implement support for this in the near future when web3swift library support is available.

In addition, we hope this proof-of-concept software can serve as a starting ground for future research interested in privacy preservation for feature engineering use cases. Additional details about the software design, development stack, implementation, and trade-offs are described in a tutorial manuscript currently under review [66]. The software source code is publicly available on GitHub at HD2i/GeolocationSmartContract [67] and HD2i/Geolocation-iOS [68]. Full details on the usage are included with the software.

Discussion

Comparison of Privacy-Preserving Methods

We have found that conventional methods that rely on a trusted third party for securing participant data generally fall short of providing full guarantees that sensitive data cannot be accessed for unintended purposes. Participants must provide a high degree of trust in researchers that use server-side deidentification procedures and maintain data warehouses themselves. Numerous data breaches on centralized servers in the past decade have illustrated that the responsible question to ask oneself is when, rather than if, private information will be exposed. To combat this, researchers should try to limit the exposure of raw data. A valid approach is to instead perform client-side feature extraction on a personal device under the control of each participant, such as a personal smartphone or private data server. The main drawback to this approach is a high burden for researchers to develop secure, validated software; meanwhile, participants still need to trust that the software is only collecting the intended

XSL•FO
RenderX

data and that there are no other data collection routines present that are only described in the fine print of a privacy policy.

Several sophisticated classes of cryptographic techniques exist that provide privacy-preserving methods, and this is an active area of research. PRE can be used for access control to encrypted data, but it falls short of guaranteeing posterior privacy. Secure MPC removes the need for a trusted third party when collecting encrypted data from 2 or more parties and computing aggregate results. However, practical implementations concerned with performance rely on a small number of computing nodes that are managed by a single party, which requires trust in the operators, software, and the security of the computing nodes. HE is considered a holy grail of privacy-preserving methods, under which computations can be performed on private data in the encrypted space. However, there are few use cases where encrypted data can be used as a feature, or applications are limited by computational performance. ZKPs are a broad class of cryptographic techniques that provide powerful guarantees of data privacy but need to be evaluated for a particular application. At a superficial level, they can be used for authentication, but other applications usually incur implementation and computational complexity.

Advances in computing hardware have enabled privacy-preservation through the design of TEEs. These chip-level designs create an isolated memory space where computations can be performed on sensitive data. However, as with software data breaches, it is difficult to guarantee that no hardware vulnerabilities can be exploited. In addition, a participant must still trust that the software running on the TEE computing node is the one advertised.

The previous 3 approaches all share a common element: the feature extraction runs on a centralized server or computing node that is managed by a trusted third party. Although some methods provide higher levels of data security, there is still a shortcoming in terms of visibility. Participants must trust that a third party is doing what it says it is doing, and nothing else. This is where blockchain technology provides a unique benefit.

Blockchains provide a trustless environment that features visibility and immutability by running on a decentralized network that is secure from tampering through cryptoeconomic incentives. In addition, they offer the unique benefit of being immutable pieces of software (smart contracts) that can be verified to do exactly what is promised if the contract code is made public. Of course, this is only guaranteed in public blockchains, where private or consortium blockchains tend to incorporate trusted parties and are more centralized. Unfortunately, blockchains were designed for security and data integrity but not for data privacy.

The combination of cryptographic techniques and TEEs with blockchain addresses the single point of trust weakness and holds the highest potential for trustworthy privacy-preserving platforms. Two promising hybrids are blockchains incorporating ZKPs and blockchains incorporating TEEs. Blockchain with ZKPs has been successful in providing transactional privacy with cryptocurrency, but it has not developed to the same level for smart contract data privacy. Blockchain with TEEs is a developing technology, but it has reached a level of maturity where developers can begin to develop and deploy real applications on these platforms. Naturally, potential hardware vulnerabilities on these platforms do not make this approach ideal; however, until cryptographic methods like FHE can be widely applied at scale, blockchains with TEEs seem to be the best approach available currently. In addition, we found that robust developer documentation and tools were available, which makes this approach accessible for product implementations as well. Our software implementation for the geolocation use case was in part encouraged by the practical direction and developer support provided by the Oasis platform and illustrates that privacy-preserving methods are realizable today on nonproduction developer networks. However, it is important to stress that blockchains with TEE developer networks are experimental at the time of this submission, and a conservative approach should be taken.

Table 3 summarizes our findings and attempts to qualitatively compare each approach in terms of how much trust must be placed in other entities. In addition, a rough indication of each method's practicality (based on implementation and computational complexity) is defined. Examples of real-world implementations or developing projects for each method are also identified; more detail is provided in Multimedia Appendix 1. Finally, the major limitations of each method are summarized.

XSL•FO
RenderX

Table 3. Comparison of trusted third party, cryptographic, and blockchain approaches to data privacy for research studies.

Method	Level of trust	Practicality	Limitations	Examples
Server-side deidentification	High	Medium	Centralized; vulnerable to data reuse and data breaches; no visibility	Strava GPS[a] devices
Client-side feature extraction	Medium	Medium	No visibility	Apple device predictive keyboard; Open PDS/SafeAnswers
Proxy re-encryption	High	Low	Only for data access control; vulnerable to data reuse and data breaches	NuCypher pyUmbral
Multiparty computation	Low	Medium	Specific use cases; centralized; no visibility; communication complexity	Jana, Sharemind, Partisia, Sepior
Homomorphic encryption	Low	Low	Limited operations or extremely low performance	NuCypher nuFHE
Zero-knowledge proof (ZKP)	Low	Low	Specific use cases; centralized; no visibility; implementation and computational complexity	zk-SNARK
Trusted execution environment (TEE)	Low	Medium	Potential for hardware vulnerabilities; no visibility	Intel SGX, ARM TrustZone, Keystone Project
Private or consortium blockchain	Medium	Medium	Pseudocentralized; depends on design	Hyperledger Fabric
Public blockchain smart contract	High	Medium	Only for data access control	Ethereum
Public blockchain with ZKPs	High	Low	Proof of concept, no software release available	ZCash, Hawk
Public blockchain with TEE	Lowest	Medium	Potential for hardware or other vulnerabilities; nonproduction stages	Enigma, Oasis

[a]GPS: global positioning system.

Limitations

In the introductory section on minimal exposure feature engineering, we identified that the feature engineering step in an analysis pipeline offers an opportunity to limit exposure and remove identifiable features. Although we promote this framework for minimizing the exposure of private data where possible, we recognize that not all feature engineering problems are amenable to deidentification. In these scenarios, we recommend that data security safeguards be in place, including encryption and secured servers.

Transactions on a public blockchain network have an inherent cost, which the parties involved in the transaction must pay for in cryptocurrency. How large this financial cost is varies based on the value of the cryptocurrency and on the congestion of the network at any given time, so no quantifiable amount is provided here. The cost may be a significant consideration for practical implementations, but our focus was on identifying methods that maintain privacy.

When posting transactions to any internet-connected network, including public blockchains, a reasonable concern is that a participant would reveal their internet protocol (IP) address, which itself is a piece of identifying information. One workaround to this concern would be to implement an internet request proxy (eg, a thin-client of the Tor software), which can relay internet traffic to conceal a user's location and usage [69]. However, the authors have not implemented this feature, and it is left to future work.

Conclusions

We believe that the boundaries of data privacy are being pushed forward with blockchain technology. The fundamental limitation of privacy-preserving protocols that run on a single point of trust with centralized servers is addressed by immutable smart contracts. As different cryptographic and software techniques overlap with blockchain, stronger guarantees of privacy become possible. In particular, we think the combination of blockchain with TEEs seems like a practical and forward-thinking approach to privacy-preserving feature engineering. This conclusion is supported by our development and deployment of a proof-of-concept private geolocation data-sharing software on a hybrid blockchain-TEE developer platform. However, no system is free from all vulnerabilities and should be thoroughly tested when interacting with highly sensitive, private data such as GPS coordinates or other biomedical data.

Acknowledgments

This work was supported by the Institute for Next Generation Healthcare at the Icahn School of Medicine at Mount Sinai and a gift from the Harris Family Charitable Foundation (to JTD).

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1
Examples of existing privacy-preserving technologies.
[PDF File (Adobe PDF File), 275KB - jmir_v21i8e13600_app1.pdf]

References

1. Madden M. Pew Research Center. 2014. Social Security Numbers, Health Info and Phone Conversations Among the Most Sensitive Data URL: http://www.pewinternet.org/2014/11/12/public-privacy-perceptions/pi_2014-11-12_privacy-perceptions_16/ [accessed 2019-02-01] [WebCite Cache ID 75qnMH4B6]
2. Valentino-DeVries J, Singer N, Keller MH, Krolik A. The New York Times. 2018. Your Apps Know Where You Were Last Night, and They're Not Keeping It Secret URL: https://www.nytimes.com/interactive/2018/12/10/business/location-data-privacy-apps.html [accessed 2019-01-31] [WebCite Cache ID 75qZ7Fvc3]
3. Cox J. Vice Media. 2019. I Gave a Bounty Hunter $300. Then He Located Our Phone URL: https://motherboard.vice.com/en_us/article/nepxbz/i-gave-a-bounty-hunter-300-dollars-located-phone-microbilt-zumigo-tmobile [accessed 2019-01-31] [WebCite Cache ID 75qZQsGJ4]
4. Cadwalladr C, Graham-Harrison E. The Guardian. 2018. Revealed: 50 Million Facebook Profiles Harvested for Cambridge Analytica in Major Data Breach URL: https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election [accessed 2019-01-31] [WebCite Cache ID 75qYuklWy]
5. Armerding T. CSO. 2018. The 18 Biggest Data Breaches of the 21st Century URL: https://www.csoonline.com/article/2130877/data-breach/the-biggest-data-breaches-of-the-21st-century.html [accessed 2019-01-31] [WebCite Cache ID 75qZiZ8Jz]
6. Zhou Y, Zhang X, Jiang X, Freeh VW. Taming Information-Stealing Smartphone Applications (on Android). In: Proceedings of the 4th International Conference on Trust and Trustworthy Computing. 2011 Presented at: TRUST'11; June 22-24, 2011; Pittsburgh, PA p. 93-107 URL: https://dl.acm.org/citation.cfm?id=2022255
7. European Commission. 2016. Data Protection in the EU: Fundamental Rights URL: https://ec.europa.eu/info/law/law-topic/data-protection/data-protection-eu_en [accessed 2019-01-31] [WebCite Cache ID 75qak0hGe]
8. Saeb S, Zhang M, Karr CJ, Schueller SM, Corden ME, Kording KP, et al. Mobile phone sensor correlates of depressive symptom severity in daily-life behavior: an exploratory study. J Med Internet Res 2015 Jul 15;17(7):e175 [FREE Full text] [doi: 10.2196/jmir.4273] [Medline: 26180009]
9. Sano A, Picard RW. Stress Recognition Using Wearable Sensors and Mobile Phones. In: Proceedings of the 2013 Humaine Association Conference on Affective Computing and Intelligent Interaction. 2013 Presented at: ACII'13; September 2-5, 2013; Geneva, Switzerland p. 671-676. [doi: 10.1109/ACII.2013.117]
10. Müller AM, Blandford A, Yardley L. The conceptualization of a just-in-time adaptive intervention (JITAI) for the reduction of sedentary behavior in older adults. Mhealth 2017;3:37 [FREE Full text] [doi: 10.21037/mhealth.2017.08.05] [Medline: 29184889]
11. Bertz JW, Epstein DH, Preston KL. Combining ecological momentary assessment with objective, ambulatory measures of behavior and physiology in substance-use research. Addict Behav 2018 Dec;83:5-17. [doi: 10.1016/j.addbeh.2017.11.027] [Medline: 29174666]
12. Moskowitz DS, Young SN. Ecological momentary assessment: what it is and why it is a method of the future in clinical psychopharmacology. J Psychiatry Neurosci 2006 Jan;31(1):13-20 [FREE Full text] [doi: 10.1146/annurev.clinpsy.3.022806.091415] [Medline: 16496031]
13. Anderson M, Antenucci D, Bittorf V, Burgess M, Cafarella M, Kumar A, et al. Brainwash: A Data System for Feature Engineering. In: Proceedings of the Conference on Innovative Data Systems Research. 2013 Presented at: CIDR'13; January 6-9, 2013; Asilomar, CA, USA.
14. Coravos A, Zimmerman N. Stat - Reporting From the Frontiers of Health and Medicine. 2018. Blockchains for Biomedicine and Health Care Are Coming. Buyer: Be Informed URL: https://www.statnews.com/2018/07/25/blockchains-biomedicine-health-care-buyer-be-informed/ [accessed 2019-01-31] [WebCite Cache ID 75qS7Rgor]
15. Gentry C. Applied Cryptography Group | Stanford University. 2009. A Fully Homomorphic Encryption Scheme URL: https://crypto.stanford.edu/craig/craig-thesis.pdf
16. Streit S, Streit B, Suffian S. arXiv. 2017. Privacy-Enabled Biometric Search URL: https://arxiv.org/abs/1708.04726
17. Hesamifard E, Takabi H, Ghasemi M. arXiv. 2017. CryptoDL: Deep Neural Networks over Encrypted Data URL: https://arxiv.org/abs/1711.05189
18. Sasson EB, Chiesa A, Garman C, Green M, Miers I, Tromer E, et al. Zerocash: Decentralized Anonymous Payments from Bitcoin. In: Proceedings of the IEEE Symposium on Security and Privacy. 2014 Presented at: IEEE'14; May 18-21, 2014; San Jose, CA, USA p. 459-474. [doi: 10.1109/SP.2014.36]

XSL•FO
RenderX

19. Green M. Cryptography Engineering Blog. 2017. Zero Knowledge Proofs: An Illustrated Primer, Part 2 URL: https://blog.cryptographyengineering.com/2017/01/21/zero-knowledge-proofs-an-illustrated-primer-part-2/ [accessed 2019-05-01] [WebCite Cache ID 75qba3Pzn]
20. NuFHE. 2019. NuFHE, A GPU-Powered Torus FHE Implementation URL: https://nufhe.readthedocs.io/en/latest/ [accessed 2019-01-31] [WebCite Cache ID 75qckinV8]
21. Archer DW, Bogdanov D, Kamm L, Lindell Y, Nielsen K, Pagter JI, et al. Cryptology ePrint Archive. 2018. From Keys to Databases – Real-World Applications of Secure Multi-Party Computation URL: https://eprint.iacr.org/2018/450.pdf [accessed 2019-01-31] [WebCite Cache ID 75qdCoBMK]
22. Ateniese G, Fu K, Green M, Hohenberger S. Improved proxy re-encryption schemes with applications to secure distributed storage. Acm T Inform Syst Se 2006 Feb 1;9(1):1-30. [doi: 10.1145/1127345.1127346]
23. Myers S, Shull A. Cryptology ePrint Archive. 2017. Efficient Hybrid Proxy Re-Encryption for Practical Revocation and Key Rotation URL: https://eprint.iacr.org/2017/833.pdf [accessed 2019-01-31] [WebCite Cache ID 75qdwcWMd]
24. Nuñez D. GitHub. 2018. Umbral: A Threshold Proxy Re-Encryption Scheme URL: https://raw.githubusercontent.com/nucypher/umbral-doc/master/umbral-doc.pdf [accessed 2019-02-01] [WebCite Cache ID 75qeCUGd1]
25. Keystone Enclave. URL: https://keystone-enclave.org/ [accessed 2019-02-01] [WebCite Cache ID 75qf37EZw]
26. Intel® Software. 2017. Introduction to Intel Software Guard Extensions (Intel SGX) URL: https://software.intel.com/sites/default/files/managed/81/61/intel-sgx-webinar.pdf [accessed 2019-02-01] [WebCite Cache ID 75qggTMQr]
27. Arm Developer. Introducing Arm TrustZone URL: https://developer.arm.com/technologies/trustzone [accessed 2019-02-01] [WebCite Cache ID 75qgqRtLo]
28. Chen F, Wang C, Dai W, Jiang X, Mohammed N, Al Aziz MM, et al. PRESAGE: pRivacy-preserving gEnetic testing via SoftwAre Guard Extension. BMC Med Genomics 2017 Dec 26;10(Suppl 2):48 [FREE Full text] [doi: 10.1186/s12920-017-0281-2] [Medline: 28786365]
29. Dai W, Deng J, Wang Q, Cui C, Zou D, Jin H. SBLWT: a secure blockchain lightweight wallet based on trustzone. IEEE Access 2018 Jul 17;6:40638-40648. [doi: 10.1109/ACCESS.2018.2856864]
30. Zyskind G, Nathan O, Pentland A. arXiv. 2015. Enigma: Decentralized Computation Platform with Guaranteed Privacy URL: https://arxiv.org/abs/1506.03471
31. Cheng R, Zhang F, Kos J, He W, Hynes N, Johnson N, et al. arXiv. 2018. Ekiden: A Platform for Confidentiality-Preserving, Trustworthy, and Performant Smart Contract Execution URL: https://arxiv.org/abs/1804.05141v5
32. Oasis Documentation. Oasis Overview URL: https://docs.oasiscloud.io/en/latest/overview/ [WebCite Cache ID 75qiGaqlD]
33. Khazbak Y, Cao G. Deanonymizing Mobility Traces With Co-Location Information. In: Proceedings of the Conference on Communications and Network Security. 2017 Presented at: CNS'17; October 9-11, 2017; Las Vegas, Nevada, USA. [doi: 10.1109/CNS.2017.8228621]
34. Ethereum Team. Ethereum. 2017. Byzantium HF Announcement URL: https://blog.ethereum.org/2017/10/12/byzantium-hf-announcement/ [accessed 2019-02-01] [WebCite Cache ID 75qm3P5Hu]
35. O'Leary RR. CoinDesk. 2017. Ethereum's Byzantium Testnet Just Verified A Private Transaction URL: https://www.coindesk.com/ethereums-byzantium-testnet-just-verified-private-transaction [accessed 2019-02-01] [WebCite Cache ID 75qmEPz2u]
36. Enigma. Enigma Protocol: Overview URL: https://enigma.co/protocol/ [accessed 2019-02-01] [WebCite Cache ID 75qmORtk3]
37. Goldwasser S, Micali S, Rackoff C. The Knowledge Complexity of Interactive Proof-Systems. In: Proceedings of the Seventeenth Annual ACM Symposium on Theory of Computing. 1985 Presented at: STOC'85; May 6-8, 1985; Providence, Rhode Island, USA p. 291-304. [doi: 10.1145/22145.22178]
38. Strava Developers. Strava API v3: API and SDK Reference URL: https://developers.strava.com/docs/reference/ [accessed 2019-02-01] [WebCite Cache ID 75qtKIrI1]
39. Manot V. Ask Different - Stack Exchange. 2014. Where Does Apple's Predictive Keyboard Get Its 'Contextual' Content From? URL: https://apple.stackexchange.com/questions/145237/where-does-apples-predictive-keyboard-get-its-contextual-content-from [WebCite Cache ID 75qtT4z8v]
40. Damgård I, Pastro V, Smart N, Zakarias S. Cryptology ePrint Archive. 2011. Multiparty Computation from Somewhat Homomorphic Encryption URL: https://eprint.iacr.org/2011/535.pdf [accessed 2019-02-01] [WebCite Cache ID 75qto7UXm]
41. Maslove DM, Klein J, Brohman K, Martin P. Using blockchain technology to manage clinical trials data: a proof-of-concept study. JMIR Med Inform 2018 Dec 21;6(4):e11949 [FREE Full text] [doi: 10.2196/11949] [Medline: 30578196]
42. Ichikawa D, Kashiyama M, Ueno T. Tamper-resistant mobile health using blockchain technology. JMIR Mhealth Uhealth 2017 Jul 26;5(7):e111 [FREE Full text] [doi: 10.2196/mhealth.7938] [Medline: 28747296]
43. Blum M, Feldman P, Micali S. Non-Interactive Zero-Knowledge and Its Applications. In: Proceedings of the Twentieth Annual ACM Symposium on Theory of Computing. 1988 Presented at: STOC'88; May 2-4, 1988; Chicago, Illinois, USA p. 103-112. [doi: 10.1145/62212.62222]
44. General Data Protection Regulation. 2018. Article 5: Principles Relating to Processing of Personal Data URL: https://gdpr-info.eu/art-5-gdpr/ [accessed 2019-05-10] [WebCite Cache ID 78Fiddsoy]

XSL•FO
RenderX

45. US Department of Health and Human Services. HIPAA Administrative Simplification URL: https://www.hhs.gov/sites/default/files/ocr/privacy/hipaa/administrative/combined/hipaa-simplification-201303.pdf [WebCite Cache ID 78FixOyHQ]
46. US Department of Health and Human Services. Title XIII - Health Information Technology for Economic and Clinical Health Act URL: https://www.hhs.gov/sites/default/files/ocr/privacy/hipaa/understanding/coveredentities/hitechact.pdf [WebCite Cache ID 78FjHK0PW]
47. Jaques N, Taylor S, Azaria A, Ghandeharioun A, Sano A, Picard R. Predicting Students' Happiness From Physiology, Phone, Mobility, and Behavioral Data. In: Proceedings of the 2015 International Conference on Affective Computing and Intelligent Interaction. 2015 Presented at: ACII'15; September 21-24, 2015; Xi'an, China p. 222-228 URL: http://europepmc.org/abstract/MED/28515966 [doi: 10.1109/ACII.2015.7344575]
48. Pratap A, Atkins DC, Renn BN, Tanana MJ, Mooney SD, Anguera JA, et al. The accuracy of passive phone sensors in predicting daily mood. Depress Anxiety 2019 Dec;36(1):72-81. [doi: 10.1002/da.22822] [Medline: 30129691]
49. Onnela JP, Rauch SL. Harnessing smartphone-based digital phenotyping to enhance behavioral and mental health. Neuropsychopharmacology 2016 Dec;41(7):1691-1696 [FREE Full text] [doi: 10.1038/npp.2016.7] [Medline: 26818126]
50. 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]
51. Kirby RS, Delmelle E, Eberth JM. Advances in spatial epidemiology and geographic information systems. Ann Epidemiol 2017 Dec;27(1):1-9. [doi: 10.1016/j.annepidem.2016.12.001] [Medline: 28081893]
52. Kamel BM, Wilson JT, Clauson KA. Geospatial blockchain: promises, challenges, and scenarios in health and healthcare. Int J Health Geogr 2018 Dec 5;17(1):25 [FREE Full text] [doi: 10.1186/s12942-018-0144-x] [Medline: 29973196]
53. Bloss C, Nebeker C, Bietz M, Bae D, Bigby B, Devereaux M, et al. Reimagining human research protections for 21st century science. J Med Internet Res 2016 Dec 22;18(12):e329 [FREE Full text] [doi: 10.2196/jmir.6634] [Medline: 28007687]
54. Wikipedia. List of Datasets for Machine-Learning Research URL: https://en.wikipedia.org/wiki/List_of_datasets_for_machine_learning_research [accessed 2019-02-01] [WebCite Cache ID 75qu0dJow]
55. Ji S, Li W, Srivatsa M, Beyah R. Structural Data De-Anonymization: Quantification, Practice, and Implications. In: Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security. 2014 Presented at: CCS'14; November 3-7, 2014; Scottsdale, Arizona, USA p. 1040-1053. [doi: 10.1145/2660267.2660278]
56. Gambs S, Killijian MO, del Prado CM. De-anonymization attack on geolocated data. J Comput Syst Sci 2013 Jul;80(8):789-797. [doi: 10.1109/TrustCom.2013.96]
57. Wang H, Gao C, Li Y, Wang G, Jin D, Sun J. De-Anonymization of Mobility Trajectories: Dissecting the Gaps between Theory and Practice. In: Proceedings of the Conference on Network and Distributed System Security Symposium. 2018 Presented at: NDSS'18; February 18-21, 2018; San Diego, CA, USA. [doi: 10.14722/ndss.2018.23222]
58. Buchanan B. A Security Site. Zero-Knowledge Proof: Proving Age With Hash Chains URL: https://asecuritysite.com/encryption/age [accessed 2019-01-31] [WebCite Cache ID 75qcQ8X1d]
59. Zhu X, Badr Y. Identity management systems for the internet of things: a survey towards blockchain solutions. Sensors (Basel) 2018 Dec 1;18(12):pii: E4215 [FREE Full text] [doi: 10.3390/s18124215] [Medline: 30513733]
60. Vill H. Institute of Computer Science: Courses. 2017. SGX Attestation Process: Research Seminar in Cryptograpy URL: https://courses.cs.ut.ee/MTAT.07.022/2017_spring/uploads/Main/hiie-report-s16-17.pdf [accessed 2019-02-01] [WebCite Cache ID 75qfE7ggw]
61. Meltdown and Spectre. 2018. URL: https://spectreattack.com/ [accessed 2019-02-01] [WebCite Cache ID 75qibUYhu]
62. van Bulck J, Minkin M, Weisse O, Genkin D, Kasicki B, Piessens F, et al. Foreshadow: Extracting the Keys to the Intel SQX Kingdom With Transient Out-of-Order Execution. In: Proceedings of the 27th USENIX Conference on Security Symposium. 2018 Presented at: SEC'18; August 15-17, 2018; Baltimore, Maryland, USA p. 991-1008.
63. Kosba A, Miller A, Shi E, Wen Z, Papamanthou C. Hawk: The Blockchain Model of Cryptography and Privacy-Preserving Smart Contracts. In: Proceedings of the Symposium on Security and Privacy. 2016 Presented at: SP'16; May 22-26, 2016; San Jose, CA, USA. [doi: 10.1109/SP.2016.55]
64. Goldfeder S, Kalonder H, Reisman D, Narayanan A. When the cookie meets the blockchain: privacy risks of web payments via cryptocurrencies. Proc Privacy Enhanc Technol 2018;2018(4):179-199. [doi: 10.1515/popets-2018-0038]
65. Oasis Documentation. Confidential Smart Contracts URL: https://docs.oasiscloud.io/en/latest/confidentiality-overview/
66. Johnson M, Jones M, Shervey M, Dudley JT, Zimmerman N. Building a secure biomedical data sharing DApp (forthcoming). J Med Internet Res 2019:- [FREE Full text] [doi: 10.2196/preprints.13601]
67. GitHub. 2019. GeolocationSmartContract URL: https://github.com/HD2i/GeolocationSmartContract [accessed 2019-02-01] [WebCite Cache ID 75s4vJTip]
68. GitHub. 2019. Geolocation-iOS URL: https://github.com/HD2i/Geolocation-iOS [accessed 2019-02-01] [WebCite Cache ID 75s534uIq]
69. Tor Project | Anonymity Online. Tor Project: Overview URL: https://www.torproject.org/about/overview.html.en [WebCite Cache ID 75qmY5rTk]

XSL•FO
RenderX

Abbreviations

FHE: fully homomorphic encryption
GPS: global positioning system
HE: homomorphic encryption
IoT: internet of things
MPC: multiparty computation
PHE: partial homomorphic encryption
PRE: proxy re-encryption
TEE: trusted execution environment
ZKP: zero-knowledge proof

Edited by P Zhang, K Clauson; submitted 01.02.19; peer-reviewed by D Maslove, M Hölbl, JT te Gussinklo; comments to author 01.04.19; revised version received 13.05.19; accepted 01.07.19; published 14.08.19.

Please cite as:
Jones M, Johnson M, Shervey M, Dudley JT, Zimmerman N
Privacy-Preserving Methods for Feature Engineering Using Blockchain: Review, Evaluation, and Proof of Concept
J Med Internet Res 2019;21(8):e13600
URL: http://www.jmir.org/2019/8/e13600/
doi:10.2196/13600
PMID:31414666

©Michael Jones, Matthew Johnson, Mark Shervey, Joel T Dudley, Noah Zimmerman. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 14.08.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

A Blockchain Framework for Patient-Centered Health Records and Exchange (HealthChain): Evaluation and Proof-of-Concept Study

Ray Hales Hylock[1], PhD; Xiaoming Zeng[1], MD, PhD

Department of Health Services and Information Management, College of Allied Health Sciences, East Carolina University, Greenville, NC, United States

Corresponding Author:
Ray Hales Hylock, PhD
Department of Health Services and Information Management
College of Allied Health Sciences
East Carolina University
4340 Health Sciences Building, Mail Stop 668
Greenville, NC, 27858
United States
Phone: 1 252 744 6184
Email: hylockr@ecu.edu

Abstract

Background: Blockchain has the potential to disrupt the current modes of patient data access, accumulation, contribution, exchange, and control. Using interoperability standards, smart contracts, and cryptographic identities, patients can securely exchange data with providers and regulate access. The resulting comprehensive, longitudinal medical records can significantly improve the cost and quality of patient care for individuals and populations alike.

Objective: This work presents HealthChain, a novel patient-centered blockchain framework. The intent is to bolster patient engagement, data curation, and regulated dissemination of accumulated information in a secure, interoperable environment. A mixed-block blockchain is proposed to support immutable logging and redactable patient blocks. Patient data are generated and exchanged through Health Level-7 Fast Healthcare Interoperability Resources, allowing seamless transfer with compliant systems. In addition, patients receive cryptographic identities in the form of public and private key pairs. Public keys are stored in the blockchain and are suitable for securing and verifying transactions. Furthermore, the envisaged system uses proxy re-encryption (PRE) to share information through revocable, smart contracts, ensuring the preservation of privacy and confidentiality. Finally, several PRE improvements are offered to enhance performance and security.

Methods: The framework was formulated to address key barriers to blockchain adoption in health care, namely, information security, interoperability, data integrity, identity validation, and scalability. It supports 16 configurations through the manipulation of 4 modes. An open-source, proof-of-concept tool was developed to evaluate the performance of the novel patient block components and system configurations. To demonstrate the utility of the proposed framework and evaluate resource consumption, extensive testing was performed on each of the 16 configurations over a variety of scenarios involving a variable number of existing and imported records.

Results: The results indicate several clear high-performing, low-bandwidth configurations, although they are not the strongest cryptographically. Of the strongest models, one's anticipated cumulative record size is shown to influence the selection. Although the most efficient algorithm is ultimately user specific, Advanced Encryption Standard–encrypted data with static keys, incremental server storage, and no additional server-side encryption are the fastest and least bandwidth intensive, whereas proxy re-encrypted data with dynamic keys, incremental server storage, and additional server-side encryption are the best performing of the strongest configurations.

Conclusions: Blockchain is a potent and viable technology for patient-centered access to and exchange of health information. By integrating a structured, interoperable design with patient-accumulated and generated data shared through smart contracts into a universally accessible blockchain, HealthChain presents patients and providers with access to consistent and comprehensive medical records. Challenges addressed include data security, interoperability, block storage, and patient-administered data access, with several configurations emerging for further consideration regarding speed and security.

XSL•FO
RenderX

(J Med Internet Res 2019;21(8):e13592) doi:10.2196/13592

KEYWORDS

blockchain; chameleon hashing; health information exchange; health information management; HL7 FHIR; patient-centered health; medical records; proxy re-encryption; redactable blockchain; smart contracts; digital health; electronic health records

Introduction

Overview

Health care is a data-intensive domain with vast amounts of information generated, accessed, and disseminated daily. Unfortunately, patient records are typically isolated in institution-centric *electronic health records* (EHRs), resulting in fragmentation with consequences ranging from inefficient care coordination to lack of critical information during emergencies [1-3]. Interoperability requirements were instituted as a remedy, but a system supporting comprehensive patient record integration remains elusive. Furthermore, the *Office of the National Coordinator for Health Information Technology* (ONC), in a 2015 Congressional report, detailed how technology vendors and providers limit patient access, by what has since been codified as *information blocking* [4,5]. Thus, not only do patient data remain disjointed but also barriers erected by data holders dissuade patient engagement and information exchange, culminating in the loss of agency. Blockchain technology coupled with nationally recognized interoperability standards (eg, Fast Healthcare Interoperability Resources, FHIR) has been presented as a viable solution to the said concerns [1-3,6-24].

Traditional health information exchange follows 1 of the following 3 models: *push* (sending information from 1 location to another), *pull* (extracting information from a source), or *view* (peering into a system). Although these practices technically achieve health information exchange, they are not sustainable, en masse solutions to patient-centered care. Thus, Halamka et al [23] proposed blockchains are a fourth model, with a stated goal of restoring patient agency [3,24].

Blockchains are distributed and decentralized repositories of information secured by various cryptographic primitives. Ideally, participants (eg, patients, providers, and payers) upload data to the chain in a secure, authenticated fashion. The result is a comprehensive medical record accessible by those with patient permission as enforced by smart contracts. As participants only need to communicate with the blockchain using recognized interoperability standards (eg, FHIR), once trust is established, all information is securely exchanged. That is, instead of having multiple points of connection, document formats, and exchange protocols to follow (each a security risk and potentially costly to address), a universally accessible blockchain minimizes the overall risk to the participating entity while simultaneously enriching information exchange and patient engagement.

More specifically, deploying blockchain in health care is suggested to break down information exchange barriers inherent in disparate, siloed EHR systems; empower patients through data consolidation and access controls and enabling (eg, secure and verifiable authorizations, form completion, discharge instruction review, and patient-generated data contribution); improve quality of care while reducing costs and fraud; promote data integrity, validation, and provenance; track medical devices and pharmaceuticals; facilitate clinical trial accountability and auditability; and support research through access to large-scale, longitudinal, aggregated patient records [3,6-22]. Virtually, all previous works are proof of concepts or pilots, endeavoring to address the information security, interoperability, data integrity, identity validation, and scalability concerns hindering adoption [3,6-19].

Herein, we submit technical solutions to address these concerns, culminating in a detailed framework and open-source proof-of-concept tool—*HealthChain* —for a patient-centered blockchain. In the Methods section, HealthChain's components are defined and, when appropriate, compared with an immutable blockchain design. Contributions include a mixed-block blockchain, redactable patient blocks, amendable smart contracts, adoption of *proxy re-encryption* (PRE) for granting and revoking data sharing rights, formulation of a *2-party PRE decryption* (2PD) scheme to facilitate mediated exchange, 4 configurable modes of operation, and a comprehensive set of experiments.

Blockchain Applications in Health Care

Heralded as a disruptive technology, blockchain research has intensified in recent years. Researchers and developers in health care have proposed, conceptualized, and implemented blockchain-based platforms to transform patient data sharing and information interoperability.

OmniPHR is a patient-centered blockchain emphasizing the distributed and interoperable principles of *personal health records* (PHRs). Patient data are recorded in encrypted, hierarchical blocks signed by the inserting entity (eg, provider, patient, or medical device). As data may be stored off-chain, OmniPHR maintains location pointers [25]. Another well-known system—MedRec—was conceived by researchers at Harvard and MIT. This Ethereum-based system links global patient identities to records held by providers. MedRec authenticates participants and stores provider pointers and record hashes (for data integrity). Patients interface with providers through MedRec to view data through smart contracts. Furthermore, patients manage third party access through the creation of smart contracts [3,23,24]. FHIRChain, developed by Zhang et al [2], is a blockchain-based architecture faceted in accordance with the secure and scalable sharing requirements of the ONC's Shared Nationwide Interoperability Roadmap [26] and leverages the emerging FHIR standard [27]. As with MedRec, data are stored off-chain and accessible through pointers and smart contract–controlled access tokens. In addition to interoperability, Kuo et al's [28] ModelChain performs privacy-preserving machine learning in the blockchain [28,29].

Beyond patient-centric applications, other blockchain solutions have been presented in a health care setting including supply chain management [1,12,17,30-32], clinical research and data

XSL•FO
RenderX

monetization [14,16,33-36], medical and research fraud detection [34,35,37], public health surveillance [13,18,20,38], and managing internet of (healthy) things [12,39-43].

Background and Terminology

Consensus and Hashing

A *consensus* algorithm is a protocol followed by the blockchain when determining the truthfulness and timeliness of blocks under consideration. On reaching consensus, blocks are accepted or rejected. There are many consensus algorithms from which to choose proof of work [44], proof of stake [45,46], proof of elapsed time [47,48], or Kafka [47,49]—a discussion of each is outside the scope of this work, and we refer the reader to the supplied references for further exploration.

Each block is identified by a *hash*, which is essentially a unique and verifiable fingerprint (Figure 1). A *1-way cryptographic hashing function* produces said hash, representing the content of a *message*. In blockchain, the message consists of block data and the previous block's hash; the inclusion of the latter creates an unbreakable bond (ie, chain; Figure 2). Hashing functions satisfy 2 key principles: (1) each input has a distinct output (ie, uniqueness), and (2) a given input has the same output (ie, verifiability).

Point 1 seeks to prevent *collisions* —a phenomenon where 2 distinct messages have the same hash. Attackers can theoretically exploit collisions by *forging* blocks with desirable modifications (eg, a financial transfer), replacing authentic blocks as the forged hashes are verifiable (point 2)—this is an oversimplification but suitable for illustrative purposes. Thus, it is imperative to use hash functions without known vulnerabilities or collisions.

Figure 1. Block schematic with sample financial data and hashes.

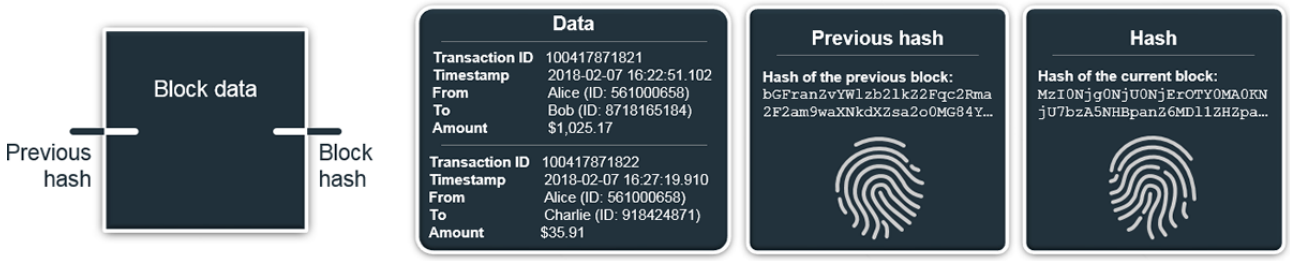
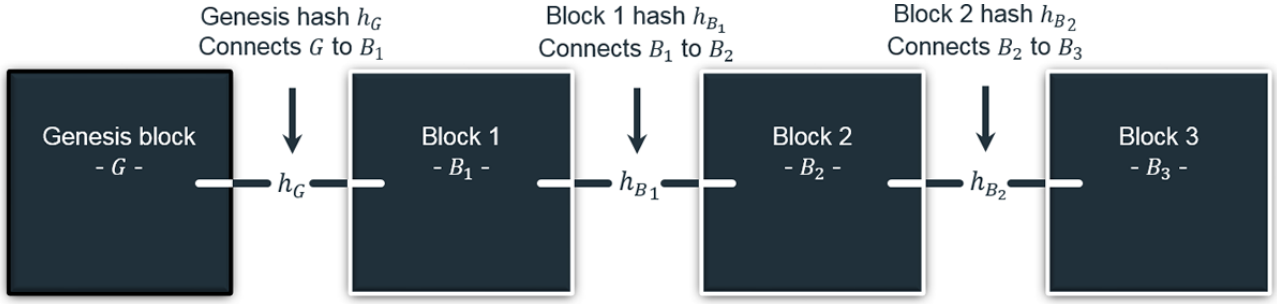


Figure 2. Blockchain diagram with several blocks, including the foundational genesis block, and noted hash connections.



Chameleon Hashing

Before Ateniese et al's *redactable blockchain* [50], a block was axiomatically held as immutable. The authors, however, made several keen observations as to when mutability was desirous if not mandatory to comply with the legal, regulatory, and general usability requirements. These include legal violations (eg, illicit content or intellectual property rights infringements), amending changes to information (eg, avoiding data fragmentation [see section on Patient Blocks]), patching executable code (eg, debugging smart contracts or removing an embedded virus), expunging data (eg, right to be forgotten, General Data Protection Regulation, or privacy breach), and allowing for blockchain consolidation (eg, a merger) [7,8,15,50-52].

Chameleon hashing was posited as a viable alternative to traditional functions. As proposed by Krawczyk and Rabin [53-55], chameleon hashing satisfies the 2 hashing principles while introducing a *trapdoor*, allowing for efficient generation of collisions by the possessor of the trapdoor (ie, private) key. However, without the said key, collisions are just as unlikely as nontrapdoor functions [50,53-55].

Smart Contracts

Smart contracts are autonomous transactions executed when stipulated terms of an agreement are met [3,22,51,52,56-59]. The creation of such contracts arose from the need to engender trust in an inherently untrustworthy environment. For instance, if some condition is met, how can one guarantee each party will comply with the agreement? Once executed, it will carry out the specified terms without fail.

Smart contracts, popularized by cryptocurrencies such as Bitcoin and Ethereum, have many practical applications in health care. Patients, for instance, can provide authorization through smart contracts to participate in studies or share information. They can codify rules leading to patient notification, for example, data accessed or communication received. They can also be used as a form of context-based access control, stipulating access rights to covered entities, business associates, and subcontractors [3,21,22].

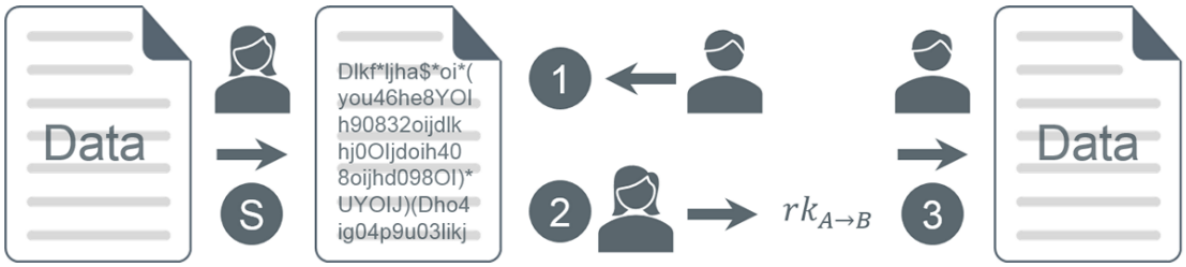
XSL•FO
RenderX

Proxy Re-Encryption

PRE enables the *delegation* of decryption rights by a *delegator* to a *delegatee* through an intervening *proxy* [60-63]. The notion is quite intuitive and shown in Figure 3. A user, for example, Alice, encrypts a message forming a ciphertext. Bob requests a *re-encryption key* to facilitate decryption of Alice's ciphertext without exposing either party's secret information. PRE has been generally deployed in the cloud [63-70] for network storage [60], distributed file systems [71], email forwarding [65,71], and information exchange [72]. In health care, it has been suggested to safeguard patient data and identities in cloud-based systems [73,74], secure mobile health monitoring and telehealth [66-68,75], and control disclosure of information in PHRs [65] and health information exchanges [72].

Most PRE schemes use *elliptic curve cryptography* (ECC) [76-80], an asymmetric cryptosystem (*Federal Information Processing Standard* [FIPS] 186-4 [81]). Advancements in quantum computing [82-85] and modern attacks [86-88], however, foreshadow its demise. One promising replacement is quantum-resistant *lattices* [71,89-95]. Kirshanova [71] and Kim and Jeong [93] have recently published frameworks for implementing PRE using lattices. Thus, as we enter a postquantum age, so too will PRE, ensuring its longevity.

Figure 3. Proxy re-encryption process overview.



Process overview

Setup (S): Alice (A, delegator) encrypts her data using proxy re-encryption and stores it.

1. Bob (B, delegatee) requests access to Alice's data (directly (ie, here) or through a proxy).
2. Alice generates a re-encryption key ($rk_{A\to B}$) using Bob's public key and her private key, sending the re-encryption key and encrypted data to Bob (or provides him with the data's location).
3. Bob uses a combination of his private key and the re-encryption key to decrypt Alice's data.

Methods

Proposed HealthChain Framework

Herein, an overview of the HealthChain framework is presented in Textbox 1. Specific details are provided in the proceeding subsections.

Patient Centered

As a patient-centric framework, HealthChain presents patients with a holistic view of their medical record, restoring agency through interaction. It encourages the accumulation, modification, generation, and review of information; ensures data integrity; authenticates identities; promotes unambiguous exchange; and executes user-granted access rights through smart contracts. Thus, HealthChain does not suffer from the common ailments plaguing PHR adoption such as data security and validity concerns, interoperability challenges, trust, and technological barriers to adoption [96-98]. Of note, HealthChain is not intended to replace EHRs but to serve as an interface between patients and third parties (eg, providers or payers).

Permissioned Blockchain

HealthChain is defined as a *permissioned* blockchain; only trusted parties (eg, hospitals, research institutions, universities, and government agencies) have the authority to manipulate the blockchain within a *private* network. These parties form a consortium (eg, a Regional HealthChain Organization) to manage the HealthChain, ensuring compliance with, for example, relevant statutes (eg, Health Insurance Portability and Accountability Act of 1996, HIPAA) and interoperability standards. Permissioning is further extended to patients, who are validated by a consortium member before account creation; this process can imitate those for patient portals. Although permissioning is a HealthChain requisite, a specific implementation is not. Thus, adopters may incorporate any system of choice.

Textbox 1. The 6 components of the HealthChain framework.

1. Patient centered
2. Permissioned blockchain: nodes and users
3. Interoperability: nationally recognized interoperability standards
4. Mixed-block blockchain: log and patient blocks
5. Smart contracts: permissioned interoperability
6. Health Insurance Portability and Accountability Act of 1996 and HealthChain: legal requirements and supporting components

XSL•FO
RenderX

Interoperability

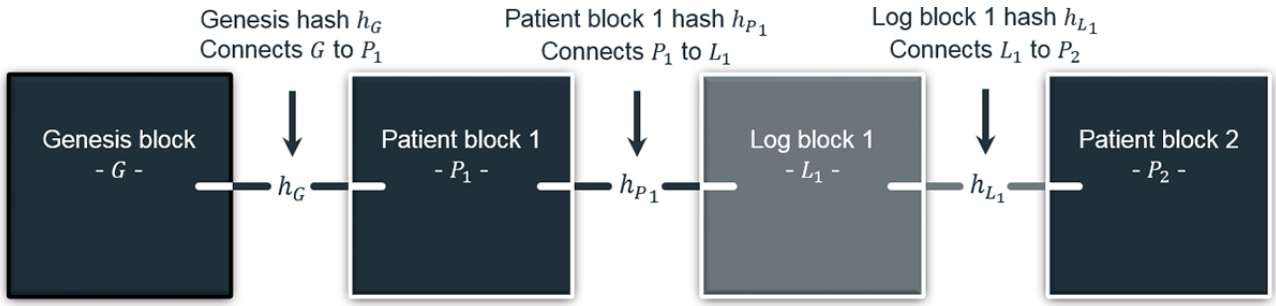
Patient-managed health information systems must conform to nationally recognized interoperability standards to be successful [99,100]. Using proprietary or lesser known standards (if at all) erects interoperability barriers, diminishing utility and adoption. Although any standard(s) is acceptable in this framework, HL7 FHIR [27] is recommended.

Although not yet federally required, HL7 FHIR is, according to the 2019 Interoperability Standards Advisory by ONC, under consideration for 26 interoperability needs [100]. Furthermore, ONC's Interoperability Proving Ground (an open community platform for sharing interoperability projects) indicates 21.3% (96/450) of projects use FHIR. Notable participants include Allscripts, the American Medical Association, Cisco, NewYork-Presbyterian Hospital, OpenHIE, the Sequoia project, and the US Department of Veterans Affairs [101]. Apple has also invested in FHIR for their PHR app [102]. Clearly, the anticipation is standard acceptance, thus our selection.

Mixed-Block Blockchain

The proposed blockchain integrates 2 semantically distinct block types: *log* and *patient*, each detailed in subsequent sections. The distinction lies in block *redaction* or editability. Patient blocks are proposed to be redactable, whereas log blocks are not. Figure 4 augments Figure 2 to account for the 2 types. As detailed, the architecture is unaffected. A simple flag in each block distinguishes the types.

Figure 4. Mixed-block blockchain adaptation of Figure 2.



Log Blocks

Log blocks are an immutable, historical account of operations on the blockchain, such as added patients and blocks, patient block modification metadata, and the issuance and execution of smart contracts. Hence, traditional hashing algorithms (eg, Secure Hash Algorithm [SHA]-256) suffice. In addition, as the data are not sensitive (ie, contain no identifying information), encryption is unnecessary. Thus, log blocks are added to the blockchain following contemporary means (eg, through consensus).

Patient Blocks

Structurally, patient blocks consist of plaintext metadata (eg, unique identifier, type flag, patient's anonymous identifier, timestamp, hash, and the issuing miner's identifier and signature), encrypted patient data, and smart contracts. Principally, they adhere to Ateniese et al's redaction scheme [50] and are established and updated as follows.

When a patient requests an account, an authorized node prepares a block and submits it through a selected consensus method for inclusion. This represents the only instance of consensus in the patient block process (Figure 5, account flow). Once accepted, the patient assumes control. Block alterations are hashed, signed, and broadcasted to the network. Nodes validate the transactions and apply the addendums (Figure 5, redact flow). Multimedia data require special handling, as their large size disproportionately (likely prohibitively) consumes bandwidth and computational resources. We adopt the methodology in MedRec [3], where data remain at the source with a location reference stored in the block for ad-hoc retrieval. Throughout this process, transactions in the form of logs are collected. As stated in the previous section, log blocks are added following a traditional consensus methodology and the chain amended (Figure 5, log flow).

Redaction addresses 3 shortcomings of immutable blocks and patient transaction isolation. The first is data fragmentation. Immutable blockchains insert new transactions in contemporaneous blocks, splintering medical records, which are collections of temporal events, rather than isolated, for example, financial, transactions (Figure 6, initial and subsequent flow). Retrieval, consequently, necessitates (1) blockchain (or a potentially corruptible, off-chain index) scanning by nodes to recover and transmit encrypted fragments and (2) decrypting and reconstituting said fragments by requesters (Figure 7). The greater the fragmentation, the more resource intensive the process. Redaction, as implemented in HealthChain, colocates patient information (Figure 5, redact flow), minimizing overall system effort (Figure 7). Second is immutability itself. Legitimate modifications (eg, adding encounter notes, modifying medical histories, or removing incorrect user-generated data) are simulated in immutable blockchains through new transactions. Requesters must apply these transactions in the proper temporal order (eg, overwrite older data with newer) to reconstruct consistent records. Resources (ie, time, space, and bandwidth) are thus depleted with each modification. Redaction modifies data in place, nullifying this effect. The third, and final, shortcoming is consensus. Modifications produce new blocks necessitating consensus. Redactions to established blocks avoid this costly process, conserving time, effort, bandwidth, and storage for users and nodes.

XSL•FO
RenderX

Figure 5. Patient account (ie, block) establishment, redaction, and logging processes in HealthChain.

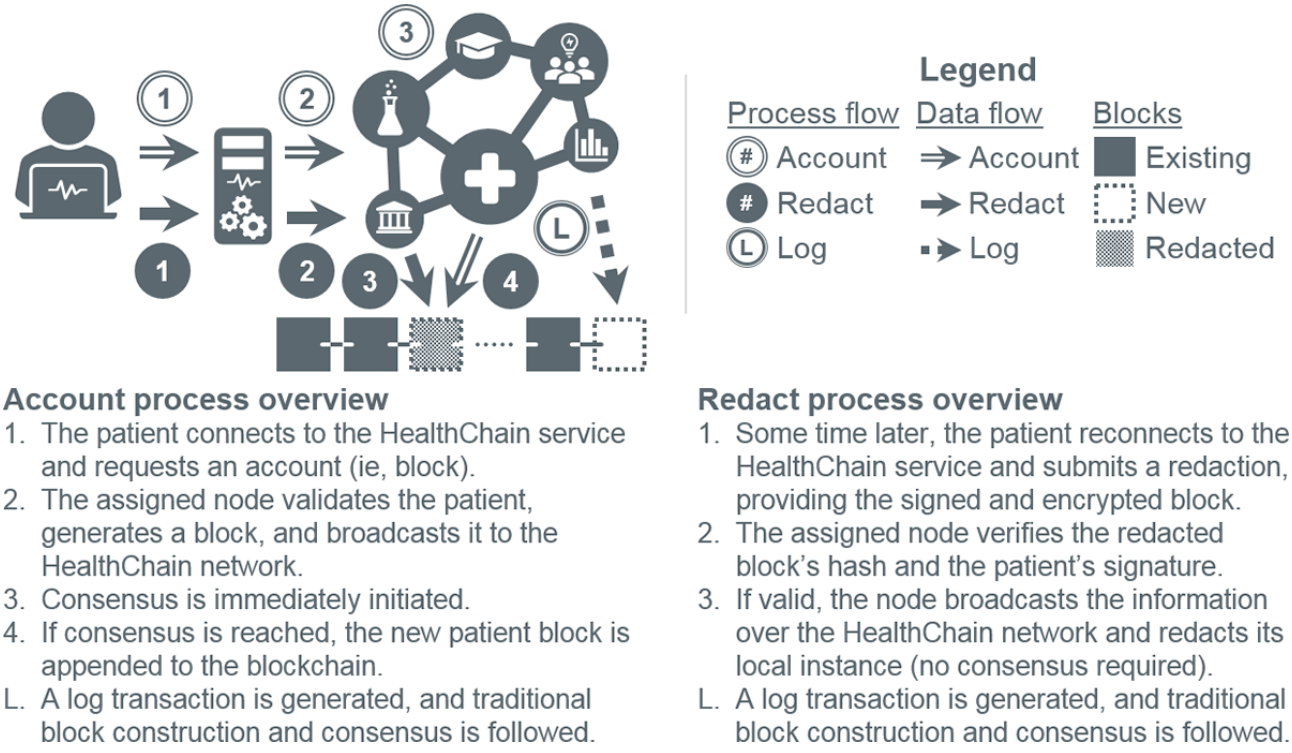
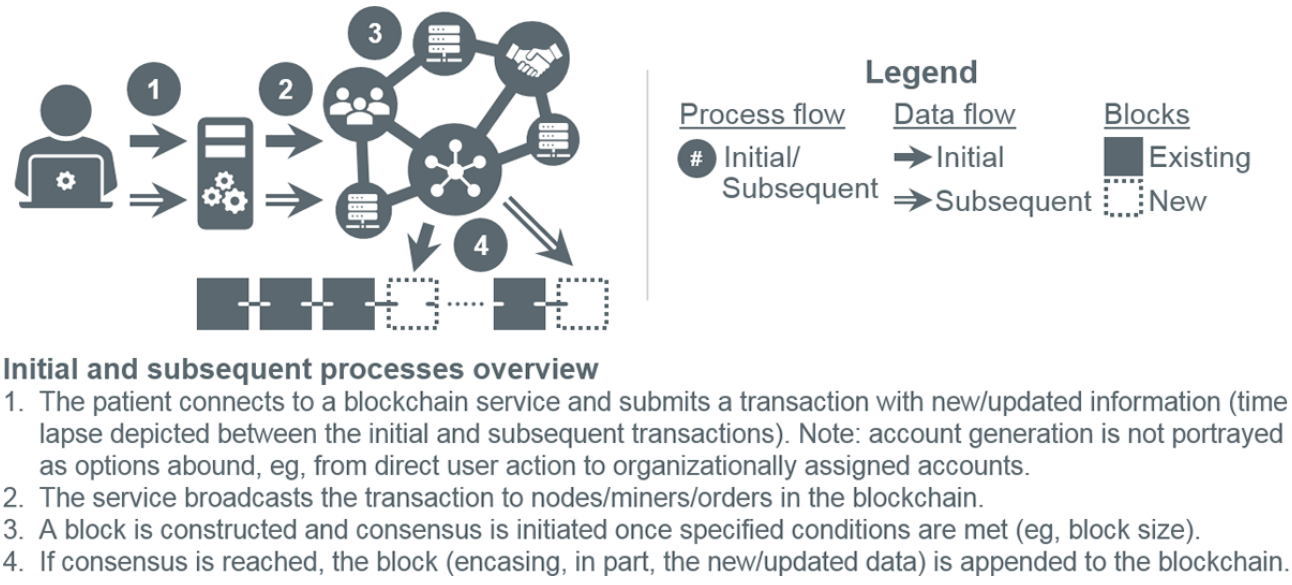
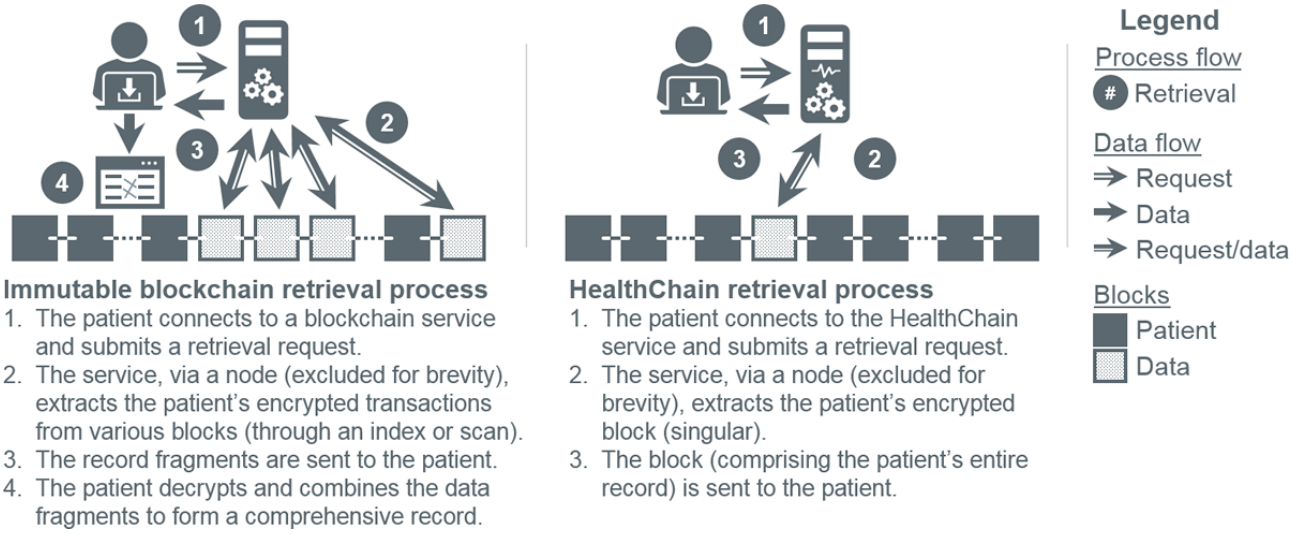


Figure 6. Initial and subsequent patient data entry in a traditional blockchain.



XSL•FO
RenderX

Figure 7. Information retrieval in an immutable blockchain versus HealthChain.



Hashing Redactable Blocks

Once appended to the blockchain, a block's hash cannot change; otherwise, the chain breaks. Redaction, therefore, precludes traditional hash functions (eg, SHA-256) as, by definition, different inputs produce distinct hashes. Hence, the deployment of chameleon hashing.

Ateniese et al [50] identified several chameleon hash formulations suitable for blockchains. Here, *public-coin chameleon hashing* (PCCH) is applied. PCCH uses public-key cryptography for verification (using public keys) and redaction (by generating collisions with private keys). In terms of security, PCCH hashes are as hard to forge as nontrapdoor functions. Hence, PCCH maintains a secure and valid state in redactable blockchains [50].

Block Encryption

Classified as *electronic protected health information* under the HIPAA, patient data must be encrypted (45 Code of Federal Regulations [CFR] section 164.304) per the US Department of Health and Human Services–issued guidelines (Health Information Technology for Economic and Clinical Health 13402(h)(2)). These guidelines reference National Institute of Standards and Technology Special Publication 800-111, which recommends the Advanced Encryption Standard (AES), although any FIPS-approved cryptosystem (eg, ECC) is acceptable [103]. Any selection must be mindful of the proposed framework, which includes information exchange.

Traditional symmetric (eg, AES) and asymmetric (eg, ECC) primitives are insufficient as they compel 1 of the 3 insecure or infeasible information exchange solutions [69]. The first is exchanging secret information. This jeopardizes data integrity (through altering, corrupting, or re-encrypting data) and digital identities (secured by private keys). Second, originators (eg, patients) can re-encrypt data under requester public keys. Although secure, originators must be omnipresent and dedicate considerable personal resources to the process; otherwise, sharing ceases. Finally, third parties can represent originators for re-encryption. However, this exposes plaintext and secret information to the said third parties. As a potential solution to these challenges, we endorse PRE.

PRE facilitates information exchange through dedicated third parties without exposing sensitive information. Beyond proposing PRE, we identified 4 operations critical to securing and validating stored and exchanged data in the proposed framework: (1) encryption and decryption, (2) re-encryption, (3) sign and verify (eg, digital signatures), and (4) encrypt sign and decrypt-verify (eg, encrypted sign and verify). *AFGH* (so named for the authors' last names) [60] is the chosen PRE scheme as it fulfills operations 1 and 2 and encourages the formulation of 3 and 4. In addition, its re-encryption keys are *unidirectional* (decrypt only), *noninteractive* (no secret information exchanged), and *nontransitive* (cannot combine keys to forge privileges) [60]; Multimedia Appendix 1 provides a more thorough introduction.

Operations 1 and 2 are instrumental in securing and sharing sensitive information and are the foundation of any PRE scheme. Operation 3 facilitates message authentication and data integrity. A message *signed* (ie, encrypted) by a sender using its private key can be *verified* (ie, decrypted) by anyone with its public key. As only the sender has its private key, verification proves authenticity and integrity. Multimedia Appendix 1 offers our posited *sign and verify* AFGH modification. Operation 4 unites 1 and 3 to protect sensitive, signed messages. Signed messages are encrypted with the recipient's public key, ensuring only it can decrypt before verification. Multimedia Appendix 1 presents the conceived *encrypt-sign* and *decrypt-verify* AFGH extension.

Moreover, some PRE-derived ciphertexts (eg, *ElGamal* based [104] in AFGH) are *malleable*. This is advantageous when optimizing, for instance, data rekeying, which amounts to multiplication in ElGamal, instead of decrypt-encrypt in AES. In addition, it is compulsory for *homomorphic encryption*, that is, computation on ciphertexts. Although this is an active area of research, *fully* homomorphic schemes (ie, those that add *and* multiply) are presently impractical [94,95,105-111]. The research, however, is progressing, with lattices emerging as a promising area [94,95,105,106]. Lattices, therefore, have the

XSL•FO
RenderX

potential to bring postquantum, fully homomorphic encryption to this framework.

PRE is traditionally deployed for key exchange [60,76]. Data are encrypted using, for example, AES, and the key encrypted and exchanged through PRE. An alternative, as proposed herein, is PRE-encrypted data. Both have merits and were extensively tested. Table 1 compares these approaches by 4 properties. We focus on PRE implementations over elliptic curves (ECs) because of their prominence [76-80].

EC PRE is slower and larger than AES because of key and *bilinear map* sizes. EC keys are twice AES for the same cryptographic strength—for example, EC-256 equates to AES-128 [112]. Bilinear maps (see Multimedia Appendix 1) are even larger—for example, up to 3072 bits for EC-256 [112-115]. Consequently, the system uses more space and time, hindering performance. Furthermore, AFGH eliminates re-encryption data integrity concerns through decrypt-only keys [60], a property unsupported by AES. Finally, the malleable AFGH cipher supports dynamic rekeying (ie, altering encryption keys) through multiplication, instead of the delegator having to decrypt, encrypt, and retransmit all data as under AES.

Table 1. A comparison of Advanced Encryption Standard encrypted blocks with proxy re-encryption encrypted keys to proxy re-encryption encrypted blocks.

Property	AES[a] block encryption	PRE[b] block encryption
Encrypt and decrypt speed	Faster	Slower
Size of ciphertext	Smaller	Larger
Key operations	Decrypt and encrypt	Decrypt only
Rekeying ciphers	Decrypt then encrypt	Multiplication

[a]AES: Advanced Encryption Standard.

[b]PRE: proxy re-encryption.

Smart Contracts

Smart contracts herein enable conditional information exchange. Their mutability (from patient block storage) permits modification and revocation without duplicate, conflicting, or vulnerable contracts remaining on the blockchain. During instantiation, a template is automatically populated (once authorization is furnished); no programming is necessary. On execution, an engine hardcoded into the server platform applies a series of instructions, given a contract's parameters. There are several reasons for this approach. The proposed smart contracts are structurally uniform (eg, identifiers and signatures of involved parties, terms, and re-encryption keys), eliminating the need for arbitrary code support. Furthermore, logical errors [56-58] and exploitable vulnerabilities [51,52,58,59] in programmed contracts can compromise patient data. An automated, templatized design with a parameterized, hardcoded engine eliminates these vulnerabilities.

Figure 8 demonstrates the smart contract initiation process, in which a re-encryption key is created and stored in a signed contract on the delegator's block. Figure 9 shows the execution of the contract generated in step 1 by the delegatee in step 2, facilitating data decryption. In this example (and proof of concept), data are stored as FHIR messages for direct interoperability with capable systems. This represents a basic application of PRE to smart contracts. It is, however, insecure.

Although granting access with PRE is simple, revoking it is not. Consider the following, 2 parties enter into a 1-week smart contract. If the delegatee notes the re-encryption key during valid execution, nothing explicitly prevents it from decrypting the delegator's block after contract termination.

A naïve refinement to close this vulnerability is implementing PRE as originally defined—proxies decrypt delegator data, then encrypt for delegatees [61]. However, as numerous studies have concluded, proxies cannot be trusted with delegator private keys and plaintext [65,70,90,116]. Another approach is subkeying by, for instance, time [60,117-119]. Each period has a unique, random variable to which all ciphertexts and re-encryption keys are bound. It is argued that this ensures delegatees cannot access *new* information. However, if a contract is terminated within a period, new information *will* be available as access is not revoked, only confined. In addition, one must manage many keys, data are fragmented over time, and interperiod interoperability is cumbersome.

To address revocation, we submit 2PD (Multimedia Appendix 1), a variant of the original formulation [61] whereby data decryption requires 2 parties with complementary re-encryption keys. Figures 10 and 11 adapt Figures 8 and 9, respectively, to 2PD. In premise, the *intermediary* (eg, node; an augmented proxy, thus the distinction) applies its re-encryption key to the delegator's data (as it is malleable), producing a ciphertext discernable to the delegatee alone. In terms of security, intermediaries no longer require private keys, and its re-encryption key does not expose plaintext. Furthermore, delegatee re-encryption keys cannot decrypt data on the blockchain, thus realizing revocation.

XSL•FO
RenderX

Figure 8. Smart contract initiation using standard proxy re-encryption.

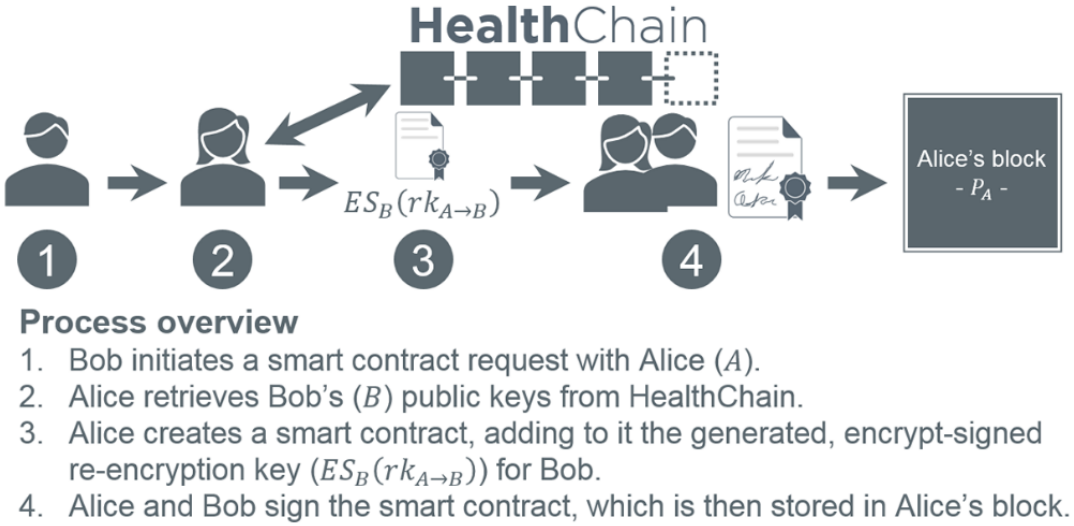


Figure 9. Smart contract execution using standard proxy re-encryption. FHIR: Fast Healthcare Interoperability Resources.

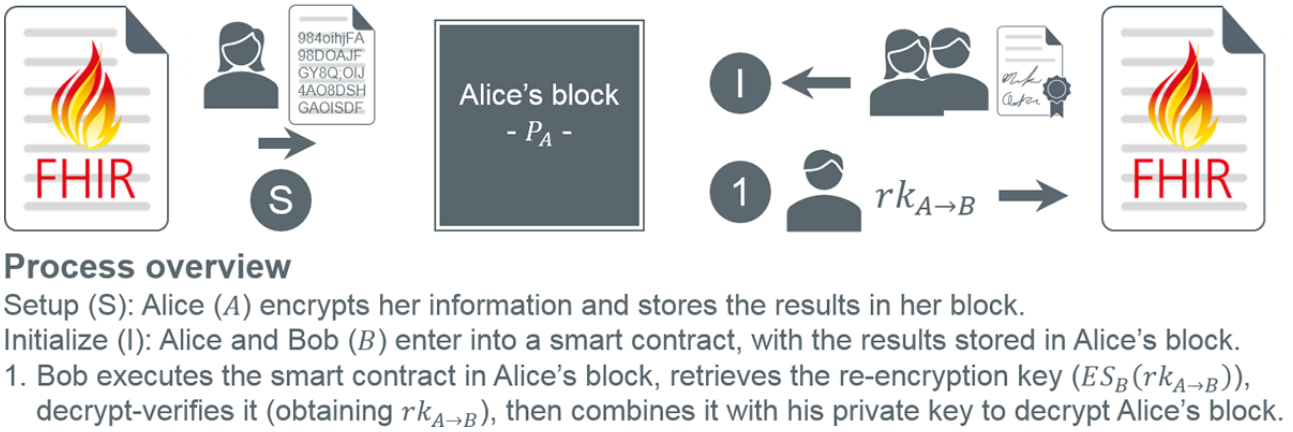
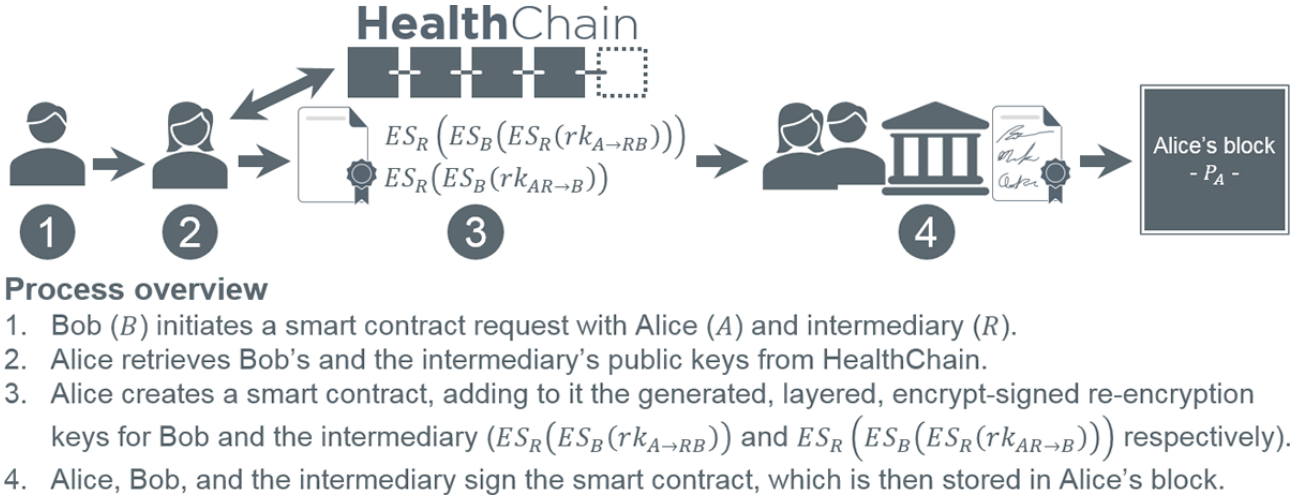
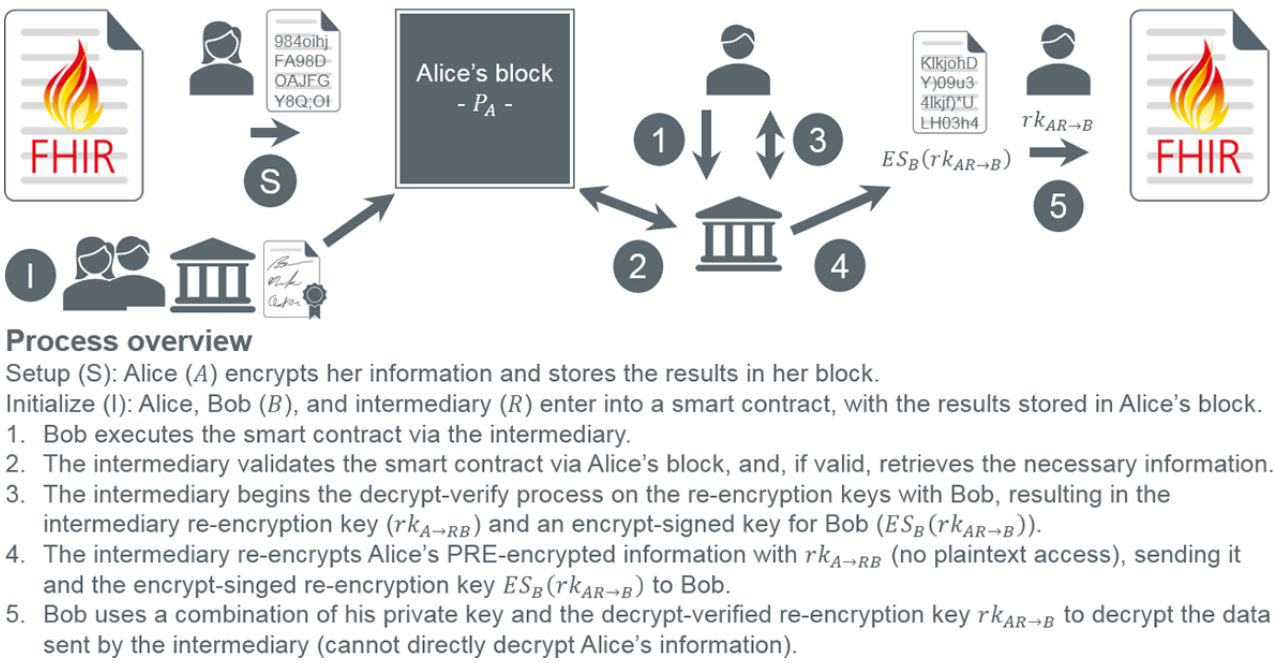


Figure 10. Smart contract initiation using 2-party proxy re-encryption decryption.



XSL•FO
RenderX

Figure 11. Smart contract execution using 2-party proxy re-encryption decryption. FHIR: Fast Healthcare Interoperability Resources; PRE: proxy re-encryption.



There are several additional 2PD factors to consider. First, it demands an intermediary be *semitrusted* in that it will not maliciously alter the hardcoded smart contract engine or distribute keys to unauthorized entities. As our framework implements a permissioned blockchain, intermediaries (ie, nodes) are inherently trustworthy. That said, private keys and plaintext are withheld to deter collusion and improper data use. Second, re-encryption keys must mathematically prohibit manipulation resulting in forged privileges. This is imparted by AFGH's nontransitive property [60]. Finally, re-encryption key retrieval mandates intermediary and delegatee consent without revealing said keys. Encrypt-sign *layering*, as documented in Figures 10 and 11, provides such support, as 1 layer must be decrypt verified by the opposing entity (ie, consent) and the final layer by the intended recipient, simultaneously averting unilateral access and key exposure.

Health Insurance Portability and Accountability Act of 1996 and HealthChain

Administrative rule 45 CFR section 164.524 grants patients the right to request copies of their records, which are to be delivered *in the form and format requested by the individual, if they are readily producible in such as format* (45 CFR section 164.524(c)(2)(i)). With the growing acceptance of HL7 FHIR (section Interoperability), one can envisage a future with it being *readily producible*. Thus, the proposed framework can leverage patient access rights through FHIR to seamlessly communicate with health information systems, eliminating the burden of ad-hoc data extraction, manual data entry, and data transformation placed on users and providers.

Systems in this space must be HIPAA compliant. Table 2 maps facets of the posited framework to pertinent HIPAA administrative rules, suggesting compliance. A thorough assessment, however, must be conducted before deployment.

XSL•FO
RenderX

Table 2. Health Insurance Portability and Accountability Act of 1996 administrative rule specifications (privacy rule and security rule) and submitted HealthChain components supporting compliance.

Specification	Rule: 45 CFR[a] section 164	HealthChain
Authorization and revocation (PR[b])	508, 510	Smart contracts, sign and verify, and encrypt-sign and decrypt-verify: confidential communications and verifiable requests and authorizations
Restriction requests (PR)	522(a)(1)	Smart contracts, sign and verify, and encrypt-sign and decrypt-verify: confidential communications and verifiable requests and authorizations
Amendments (PR)	526	Smart contracts, sign and verify, and encrypt-sign and decrypt-verify: confidential communications and verifiable requests and authorizations
Confidential communications (PR)	522(b)(2)	Encrypt-sign and decrypt-verify: message integrity, verifiable identity, and encryption
Unique user authentication (SR[c])	312(a)(2)(i)	Unique encryption and hashing key pairs, sign and verify, and encrypt-sign and decrypt-verify: verifiable identity (key possession and signing) and patient block hashes and patient block encryption
Encryption and decryption (SR)	312(a)(2)(iv)	Unique encryption and hashing key pairs, sign and verify, and encrypt-sign and decrypt-verify: verifiable identity (key possession and signing) and patient block hashes and patient block encryption
Integrity (SR)	312(c)(1)	Unique encryption and hashing key pairs, sign and verify, and encrypt-sign and decrypt-verify: verifiable identity (key possession and signing) and patient block hashes and patient block encryption
Audit controls (SR)	312(b)	Log blocks
Person or entity authentication (SR)	312(d)	Sign and verify, encrypt-sign and decrypt-verify, re-encryption key layering, and delegatee re-encryption: verifiable identity (verification algorithms and construction of the delegatee re-encryption process)
Transmission security—integrity controls and encryption (SR)	312(e)(1), (2)(i), and (2)(ii)	Patient block encryption, intermediary re-encryption, sign and verify, and encrypt-sign and decrypt-verify (layering): verifiable identity and transfer of encrypted data only by design

[a]CFR: 45: Code of Federal Regulations.

[b]PR: privacy rule.

[c]SR: security rule.

Experimental Design

The experimental design facilitates the examination of HealthChain's novel components. Additional services such as permissioning and consensus along with a comprehensive distributed network were not evaluated for the following reasons. First, although essential to the framework's practical implementation (whereas here we are exploring new functionality), no improvements to those areas were proposed in this work; hence, testing is unwarranted. Second, ancillary services and an arbitrarily sized experimental network inject considerable overhead (potentially orders of magnitude above the measured item) into the process, rending the subject of analysis indistinguishable from noise. As such, we intentionally limited the components implemented in the proof of concept to only those necessary to successfully evaluate the processes defined in the subsequent sections.

Regarding experimental block operations, recall that as patient blocks are generated during account initialization, all subsequent actions necessitate redactions to said block section Patient Blocks). Therefore, all experiments conducted herein are redactions.

Configurable Modes

HealthChain operations are dictated by 4 configurable modes (see Multimedia Appendix 2 for a detailed summary). The first is the 2-option *block encryption mode*: (1) AES-encrypted data with PRE-encrypted key (denoted as A) and (2) PRE-encrypted data (signified by P), as defined and justified in section Block Encryption.

The second is the *storage mode* with 2 possibilities: *full block* (F) and *incremental* (I). Full block incorporates all data into a single block. Incremental, in all but 1 case (see next mode), transmits only new and modified entries or removal instructions. Table 3 compares the 2 options over 6 properties. Full block transmits 1 block (ie, transaction), which is more efficient than multiple blocks. Cipher padding in full block mode is negligible, whereas potentially considerable in incremental mode (cumulative padding). Record isolation is trivial in incremental mode as each is its own entry. Full block masks transactions in an encrypted block, making isolation and metadata attacks difficult. Regarding transmission speed and size, incremental mode may be smaller and thus faster if only transmitting minimal changes relative to the medical record. Finally, a single block is bound in size by the storage mechanism (eg, a database byte array attribute). Larger medical records may require multiple blocks, fragmenting the information and increasing the complexity of management. Incremental does not suffer from this limitation.

XSL•FO
RenderX

Table 3. Comparison of full block and incremental storage mode options.

Property	Full block	Incremental
Transactions	1	1 or more
Cipher padding	Negligible	Potentially considerable
Record isolation	No	Yes
Speed of transmission	Slow	Potentially fast
Size of transmission	Large	Potentially small

Third is *encryption key mode*, with *static* (S) and *dynamic* (D) choices. Static mode encrypts a patient block using the same key in perpetuity, whereas dynamic mode re-encrypts data under a new one with each update. Dynamic mode enhances security, as compromised information is of limited use, but consumes more resources. AES-encrypted data in incremental block storage (AI) mode requires the patient to re-encrypt and transmit all data. Being malleable, PRE-encrypted data in incremental storage (PI) mode encrypts updates using a new key, sending them and a scalar to a node for appending and rekeying, respectively.

Server-side encryption mode is the last mode: simply on (Y) or off (N). If enabled, the server encrypts (by block encryption mode) user data using an ephemeral key for each entry, renewed under the dynamic key policy (does not impact PRE operations). This protects against improper access by dynamically rekeying accessed entries.

Tables 4 and 5 present listings of configurable mode abbreviations and descriptions used in our experiments by AES and PRE encryption respectively. Refer to Multimedia Appendix 2 for more details.

Table 4. Advanced Encryption Standard (AES) configurable experimental mode abbreviations and descriptions.

Mode	Description
AF	AES-encrypted data, full block storage
AI	AES-encrypted data, incremental storage
ADF	AES-encrypted data, dynamic encryption key full block storage
ADI	AES-encrypted data, dynamic encryption key incremental storage
ASF	AES-encrypted data, static encryption key full block storage
ASI	AES-encrypted data, static encryption key incremental storage
AFN	AES-encrypted data, full block storage no server-side encryption
AFY	AES-encrypted data, full block storage server-side encryption
AIN	AES-encrypted data, incremental storage no server-side encryption
AIY	AES-encrypted data, incremental storage server-side encryption
ADFN	AES-encrypted data, dynamic encryption key, full block storage, no server-side encryption
ADFY	AES-encrypted data, dynamic encryption key, full block storage, server-side encryption
ADIN	AES-encrypted data, dynamic encryption key, incremental storage, no server-side encryption
ADIY	AES-encrypted data, dynamic encryption key, incremental storage, server-side encryption
ASFN	AES-encrypted data, static encryption key, full block storage, no server-side encryption
ASFY	AES-encrypted data, static encryption key, full block storage, server-side encryption
ASIN	AES-encrypted data, static encryption key, incremental storage, no server-side encryption
ASIY	AES-encrypted data, static encryption key, incremental storage, server-side encryption

XSL•FO
RenderX

Table 5. Proxy re-encryption (PRE) configurable experimental mode abbreviations and descriptions.

Mode	Description
PF	PRE-encrypted data, full block storage
PI	PRE-encrypted data, incremental storage
PDF	PRE-encrypted data, dynamic encryption key full block storage
PDI	PRE-encrypted data dynamic encryption key incremental storage
PSF	PRE-encrypted data, static encryption key full block storage
PSI	PRE-encrypted data, static encryption key incremental storage
PFN	PRE-encrypted data, full block storage no server-side encryption
PFY	PRE-encrypted data, full block storage server-side encryption
PIN	PRE-encrypted data, incremental storage no server-side encryption
PIY	PRE-encrypted data, incremental storage server-side encryption
PDFN	PRE-encrypted data, dynamic encryption key, full block storage, no server-side encryption
PDFY	PRE-encrypted data, dynamic encryption key, full block storage, server-side encryption
PDIN	PRE-encrypted data, dynamic encryption key, incremental storage, no server-side encryption
PDIY	PRE-encrypted data, dynamic encryption key, incremental storage, server-side encryption
PSFN	PRE-encrypted data, static encryption key, full block storage, no server-side encryption
PSFY	PRE-encrypted data, static encryption key, full block storage, server-side encryption
PSIN	PRE-encrypted data, static encryption key, incremental storage, no server-side encryption
PSIY	PRE-encrypted data, static encryption key, incremental storage, server-side encryption

Experiments

A total of 5 system dimensions are measured over 16 mode combinations for AES-128 and EC-256: transmission size, network latency, client processing time, server processing time, and smart contract execution.

Transmission size refers to the number of bytes sent from client to server. Typical consumer internet connections have low upload rates as households consume more content than contribute; thus, upload bandwidth is a concern. Scalability for users with metered connection is also of interest, as they may incur costs associated with overages or plan adjustments.

Network latency assesses the effect internet-based transmissions have on the proposed framework. The results are analyzed independently and integrated into client processing time and smart contract execution.

Client processing time represents the time devoted by clients to, for example, insert records, generate synchronization instructions, regenerate smart contracts (if dynamic keying), compute block hashes, and broadcast the previous to the blockchain.

Server processing time is the time incurred by servers during, for example, AES PRE key renewing (dynamic AES), instruction application (all), PRE scalar multiplication (PDIN/Y), and smart contract updating (all dynamic).

Smart contract execution measures the time required to run a smart contract. The process uses 2PD and writes the output to a zipped file on the delegatee's machine.

Datasets

Each of the 16 mode combinations was evaluated by insertion and scaling costs. Insertion costs are determined by adding records en masse to a clean system instance (only contains the account request profile). Insights are garnered on cost amortization and limits associated with bulk and single record processing without existing record influences (eg, re-encryption and retransmission), which may direct policy on block synchronization. Moreover, 4 datasets of observations (1 per day) were synthetically generated for testing (Table 6). Each was statically sized at 400 bytes for record-level evaluation. As each record is of identical size, cryptographic processing time, cipher length, and bytes transmitted are comparable by record across the various configurations.

Scaling is scrutinized by gauging the effect existing records have on insertions. As this system accumulates records, these experiments facilitate the examination of existing medical records on overall performance. For insertion, care was taken to avoid interaction between new and existing data. Here, the reverse is of interest on how existing data affect record insertions. The results inform decisions on configuration selection and synchronization strategy. Testing began by instantiating a system with 1 of the 3 datasets (Table 6)—representative of small, medium, and large patient records from a deidentified instructional medical database [120]. Then, the 4 insertion datasets were applied, with the average per record reported as an indicator of general performance.

XSL•FO
RenderX

Table 6. Number of records and byte range per record by experimental dataset (all records were formatted as Health Level-7 Fast Healthcare Interoperability Resources [FHIR] JSON messages using HAPI FHIR).

Experiment	Records	Bytes/record (average)	Notes
Insertion	1	400	1 day
Insertion	30	400	30 days
Insertion	365	400	1 year
Insertion	1461	400	4 years
Scaling	334	395-761 (581)	26 encounters, 33 conditions, 145 medication requests, and 130 observations
Scaling	945	395-768 (675)	27 encounters, 159 conditions, 624 medication requests, and 135 observations
Scaling	2361	394-770 (732)	109 encounters, 119 conditions, 2029 medication requests, and 104 observations

Testing Environment

The testing environment aligns with the minimal requirements defined in the Experimental Design section. In its simplest form, HealthChain is a medium of information exchange between an entity (eg, patient) and a server (eg, node). Every process in HealthChain can be reduced to a series of entity-server interactions; therefore, our testing environment emulates this 2-machine structure.

The first machine is a Lenovo T540p running Windows 7 Enterprise with 16 GB of memory, an Intel i7-4800 MQ processor, and a wired, consumer internet connection. The second is a Dell Optiplex 9010 running Linux Mint 17.1 with 8 GB of memory, an Intel i7-3770 processor, and a wired, business internet connection. Communication rates (download/upload in megabits per second [Mbps]) are as follows: 32.1/5.9 and 955.4/176.3 Mbps for each respective machine [121].

To facilitate direct processing time comparisons between client and server, 1 machine (the Lenovo) assumed both roles. This colocation, however, failed to address networking concerns. Thus, experiments incorporating transmission costs (ie, client processing time and smart contract execution) were duplicated using both machines, which are situated 5 miles apart. These times replaced those in the 2 identified measures for a more realistic outcome while still affording client and server relative performance comparisons.

Proof-of-Concept Implementation

The proof of concept facilitates the examination of the novel HealthChain elements as specified in section Experimental Design; it is not a production-ready blockchain system. The realization of the proposed framework requires the blending of the components defined herein with an existing blockchain technology such as Hyperledger Fabric or Ethereum. The proof of concept consists of 2 systems and 2 libraries written primarily in Java 8 and JSP and uses HAPI FHIR for document formatting [122] (Multimedia Appendix 3).

The first system, *HealthChainServer*, instantiates a single node that, for instance, establishes the blockchain, processes patient account requests, validates and manages patient identities and block transactions, supports 2PD, and transmits updates to patients (eg, the latest encounter). A multinode system (and therefore broadcasting capabilities) is unnecessary for testing. In addition, log blocks are not examined as they are a trivial extension to existing blockchain technologies.

The second system, *HealthChainWebClient*, is a simple JSP-enabled version of the Gentelella Alela template [123]. Through the Web portal, patients can, for example, request accounts, create and manage smart contracts, manually add records, import and export FHIR messages, and receive updates (eg, from a participating hospital). It was through the file upload feature that experimental data were added, which were then transmitted to the server over a socket connection. Although sufficient for testing, a robust, security-focused, application programming interface–driven approach (such as the one developed using SMART on FHIR [124]) should be implemented before use in production.

Regarding libraries, the first provides chameleon hash support by way of PCCH [50] as outlined in the section on Hashing Redactable Blocks. The second realizes PRE through AFGH [60] (using, as a foundation, the Java Pairing-Based Cryptography library [jPBC] [113]) as justified in the section on Block Encryption.

Results

Transmission Size

Here, transmission size is analyzed by insertion (Figure 12) and scaling (Figure 13). Server-side encryption does not impact transmission, hence its exclusion.

In Figure 12, PDI and PSI are nearly double the others (in overall and per record transmission size). Both have 789−400=389 bytes of padding (ie, wasted space) per record inserted, thus the disparity with all AES configurations. In addition, PDF and PSF are in line with AES (about 2% larger) rather than PDI and PSI. This is attributable to the full block generation process, as all records are fused into one, then divided into 789-byte ciphers, markedly reducing padding (see the section on Dataset Effects on Cipher Size). Furthermore, by 365, all ciphers near saturation (ie, negligible padding and amortized overhead).

Regarding overhead, dynamic options include 9137 bytes per smart contract (one here), ADI and ADF incorporate a new 397-byte AES PRE key, PDI uses a 384-byte scalar, and all send a 96-byte block hash.

XSL•FO
RenderX

To scaling (Figure 13), ASI, PDI, and PSI are unaffected by the number of existing records (only transmit modifications), requiring 416, 789, and 810 bytes on average of encrypted data per added record, respectively.

Figure 12. Transmission size in kilobytes and bytes per record by the number of records inserted. ADF: AES-encrypted data, dynamic keys, full block storage; ADI: AES-encrypted data, dynamic keys, incremental storage; AES: Advanced Encryption Standard; ASF: AES-encrypted data, static keys, full block storage; ASI: AES-encrypted data, static keys, incremental storage; PDF: PRE-encrypted data, dynamic keys, full block storage; PDI: PRE-encrypted data, dynamic keys, incremental storage; PRE: proxy re-encryption; PSF: PRE-encrypted data, static keys, full block storage; PSI: PRE-encrypted data, static keys, incremental storage.

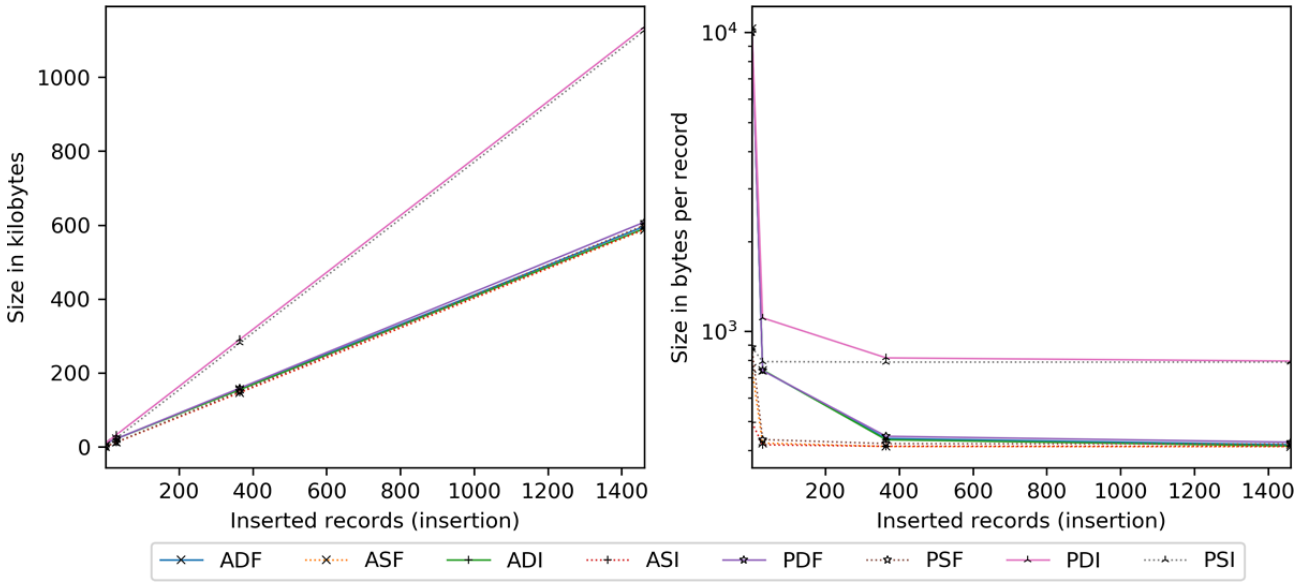
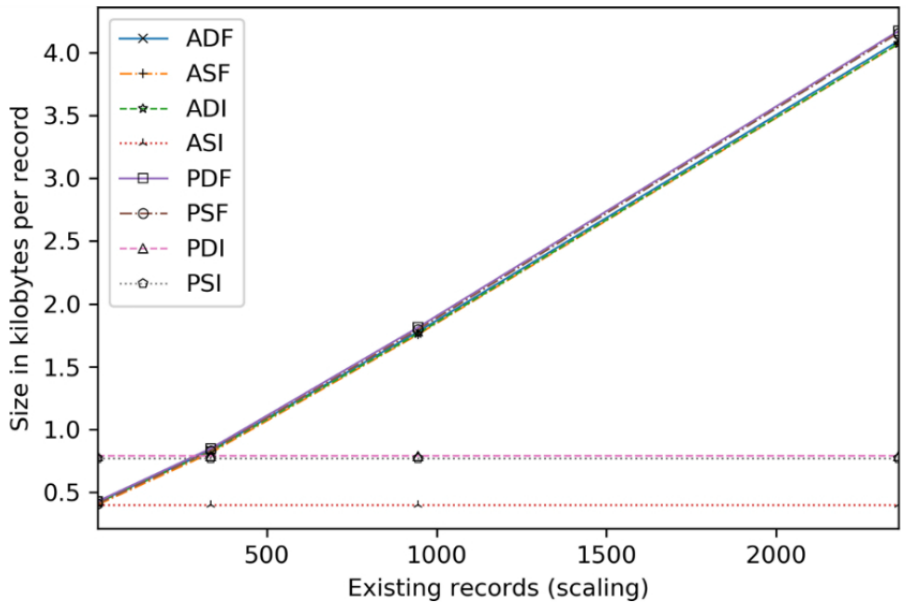


Figure 13. Transmission size in kilobytes per record added given an existing record set. ADF: AES-encrypted data, dynamic keys, full block storage; ADI: AES-encrypted data, dynamic keys, incremental storage; AES: Advanced Encryption Standard; ASF: AES-encrypted data, static keys, full block storage; ASI: AES-encrypted data, static keys, incremental storage; PDF: PRE-encrypted data, dynamic keys, full block storage; PDI: PRE-encrypted data, dynamic keys, incremental storage; PRE: proxy re-encryption; PSF: PRE-encrypted data, static keys, full block storage; PSI: PRE-encrypted data, static keys, incremental storage.



Dataset Effects on Cipher Size

AES and PRE cipher sizes are dissected in Tables 7 and 8 for incremental and full block encryption modes, respectively. Comparisons are drawn at the record level; thus, overhead bytes were removed (see the section on Transmission Size and 231 bytes for the patient record generated during account activation). AES is slightly larger on average than the underlying data (0.7%-4.1%) regardless of mode. PRE incremental is extensively padded (26.5%-97.3%), whereas minimal in full block mode (3.1%-5.2% beyond 1 insertion). Thus, PRE is subject to extreme variability, relative to AES, on input file size.

XSL•FO
RenderX

Table 7. Incremental storage: byte range per file and average Advanced Encryption Standard- and proxy re-encryption-encrypted cipher sizes by dataset (insertion and scaling).

Dataset	Bytes/file range (average)	AES[a] average (difference, %)	PRE[b] average (difference, %)
130,365, and 1461	400	416 (4.0)	789 (97.3)
334	395-761 (581)	585 (0.6)	802 (38.0)
945	395-768 (675)	679 (0.7)	840 (24.3)
2361	394-770 (732)	737 (0.7)	926 (26.5)

[a]AES: Advanced Encryption Standard.

[b]PRE: proxy re-encryption.

Table 8. Full block storage: total bytes per dataset and average Advanced Encryption Standard- and proxy re-encryption-encrypted cipher sizes by dataset (insertion and scaling).

Dataset	Total bytes	AES[a] average (difference, %)	PRE[b] average (difference, %)
1	400	416 (4.0)	789 (97.3)
30	12,000	12,368 (3.1)	12,624 (5.2)
365	146,000	150,112 (2.8)	153,066 (4.8)
1465	584,400	600,848 (2.8)	613,053 (4.9)
334	193,815	196,304 (1.3)	200,406 (3.4)
945	637,934	644,688 (1.1)	658,026 (3.1)
2361	1,727,714	1,744,032 (0.9)	1,779,981 (3.0)

[a]AES: Advanced Encryption Standard.

[b]PRE: proxy re-encryption.

Network Latency

Network latency is analyzed in isolation to understand client-to-server (Figures 14 and 15) and server-to-client (Figure 16) effects. For insertion costs, Figure 14, incremental PRE grows at about twice the pace of others, proportional to cipher size (Tables 7 and 8). In addition, by 365 in Figure 15, the Mbps transmitted saturate the connection, whereupon ADI/ASI and PDI/PSI stabilize at 0.6 and 1.1 milliseconds per fragment (ie, a single record or full block), respectively. Regarding scaling (Figure 14), ASI, PDI, and PSI are constant, whereas full block and ADI grow as they reprocess all entries per update.

Figure 16 examines server-to-client transmissions as anticipated during record downloads from participating entities (eg, clinics) and intermediate smart contract results (if a delegatee). Data include the insertion sets as well as existing with 1461 additions to demonstrate scale. Transmission time depends on the block encryption (cipher size) and storage (padding effects) modes. AF, AI, and PF are indistinguishable from one another, whereas PI is on average 55% slower because of larger, excessively padded ciphers. Regarding bandwidth, a limit at approximately 18.4 Mbps (3.4 times the update limit in Figure 15), first experienced by the larger PI, is revealed. This corresponds to increased trajectories in time.

XSL•FO
RenderX

Figure 14. Client-to-server network latency in seconds per inserted record (ie, insertion) and seconds per record added given an existing record set (ie, scaling) — includes connection establishment, termination, and transmission time. ADF: AES-encrypted data, dynamic keys, full block storage; ADI: AES-encrypted data, dynamic keys, incremental storage; AES: Advanced Encryption Standard; ASF: AES-encrypted data, static keys, full block storage; ASI: AES-encrypted data, static keys, incremental storage; PDF: PRE-encrypted data, dynamic keys, full block storage; PDI: PRE-encrypted data, dynamic keys, incremental storage; PRE: proxy re-encryption; PSF: PRE-encrypted data, static keys, full block storage; PSI: PRE-encrypted data, static keys, incremental storage.

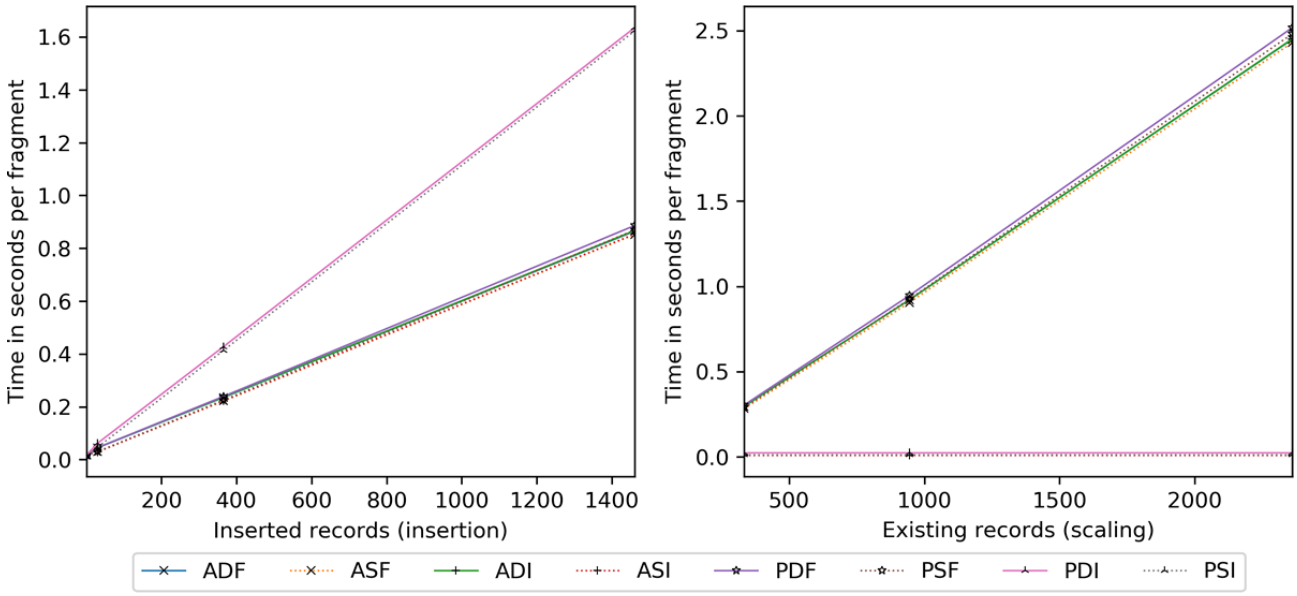
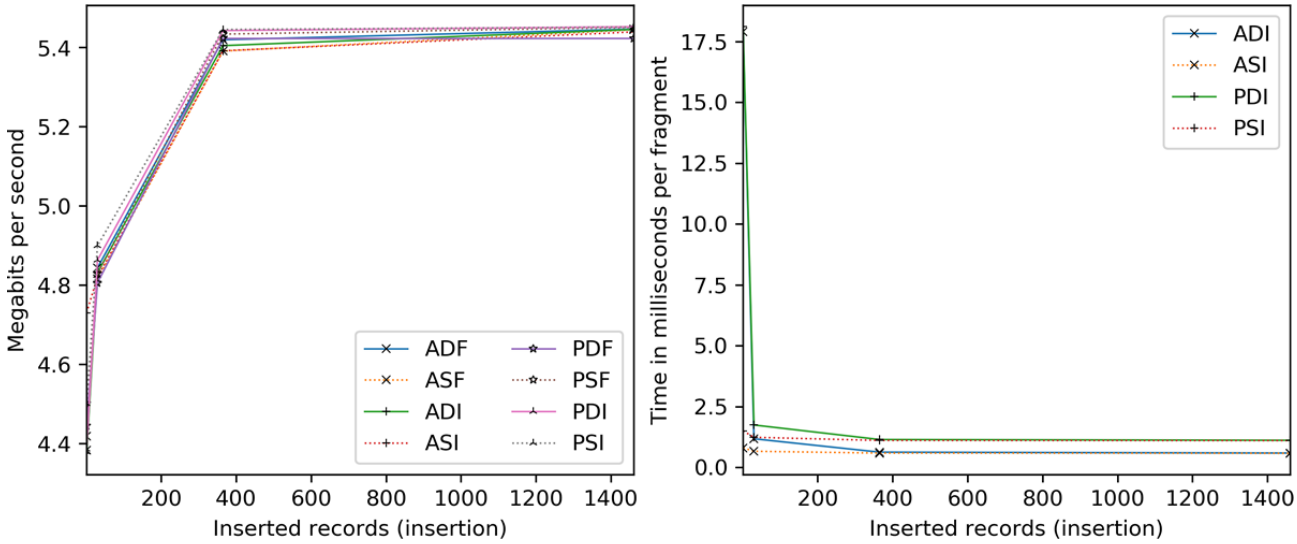
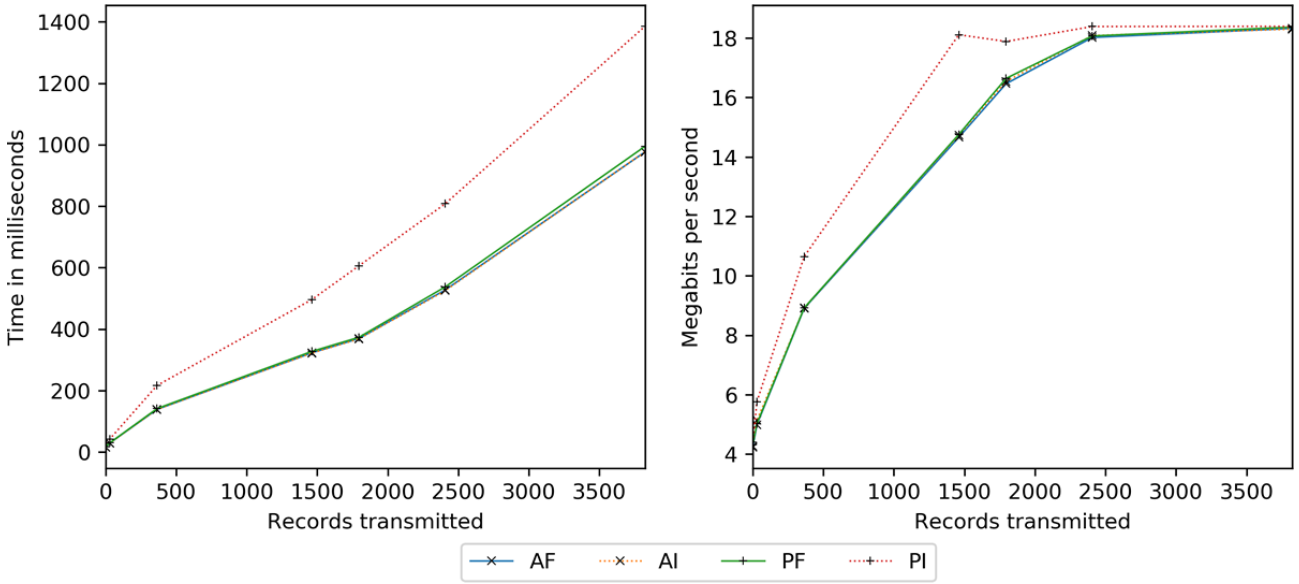


Figure 15. Client-to-server network latency (transmission only) measured in megabits per second and milliseconds per fragment by the number of records inserted. ADF: AES-encrypted data, dynamic keys, full block storage; ADI: AES-encrypted data, dynamic keys, incremental storage; AES: Advanced Encryption Standard; ASF: AES-encrypted data, static keys, full block storage; ASI: AES-encrypted data, static keys, incremental storage; PDF: PRE-encrypted data, dynamic keys, full block storage; PDI: PRE-encrypted data, dynamic keys, incremental storage; PRE: proxy re-encryption; PSF: PRE-encrypted data, static keys, full block storage; PSI: PRE-encrypted data, static keys, incremental storage.



XSL•FO
RenderX

Figure 16. Server-to-client network latency (transmission only) measured in milliseconds and megabits per second by the number of records transmitted to the client. AES: Advanced Encryption Standard; AF: AES-encrypted data, full block storage; AI: AES-encrypted data, incremental block storage; PF: PRE-encrypted data, full block storage; PI: PRE-encrypted data, incremental storage; PRE: proxy re-encryption.



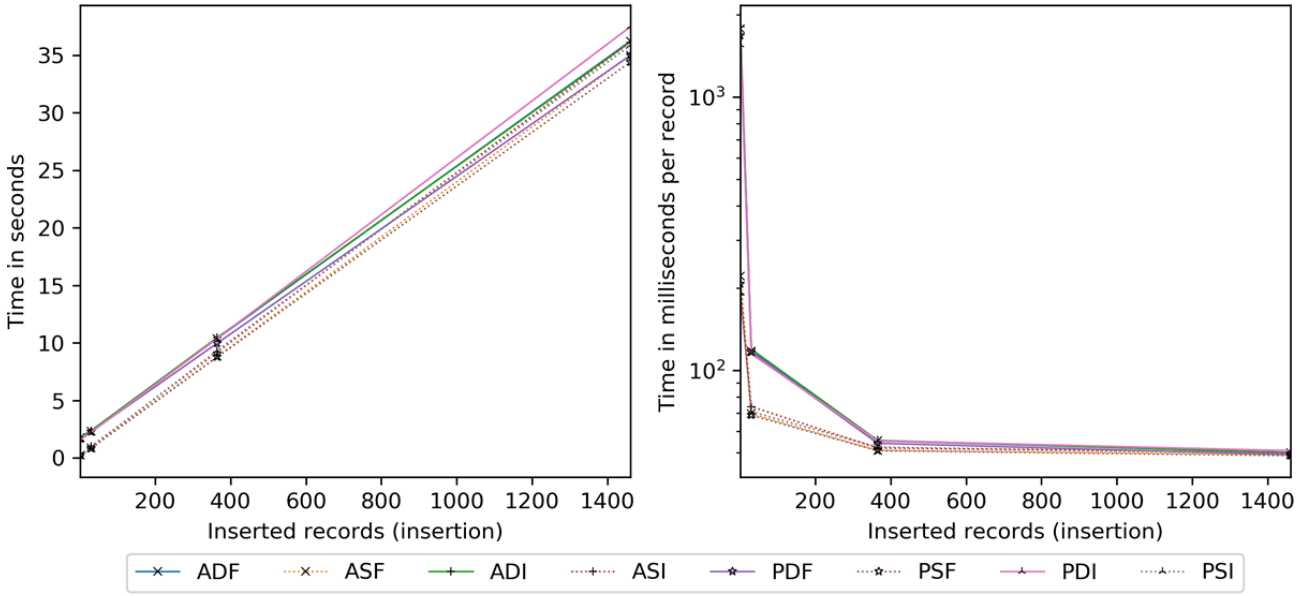
Client Processing Time

Client processing time is explored for insertions (Figure 17) and scaling (Figure 18). Server-side encryption does not impact client performance, hence its exclusion. For these tests, the client and server are the same machine (for relative comparison), whereas the transmission time is taken from the network latency experiments.

All configurations require a similar amount of time per Figure 17. The quickest are the static full block approaches, followed by dynamic full block (2% slower), static incremental (8% slower), and dynamic incremental (12% slower) approaches. Per record, by 365, all are within 5 milliseconds and narrowing.

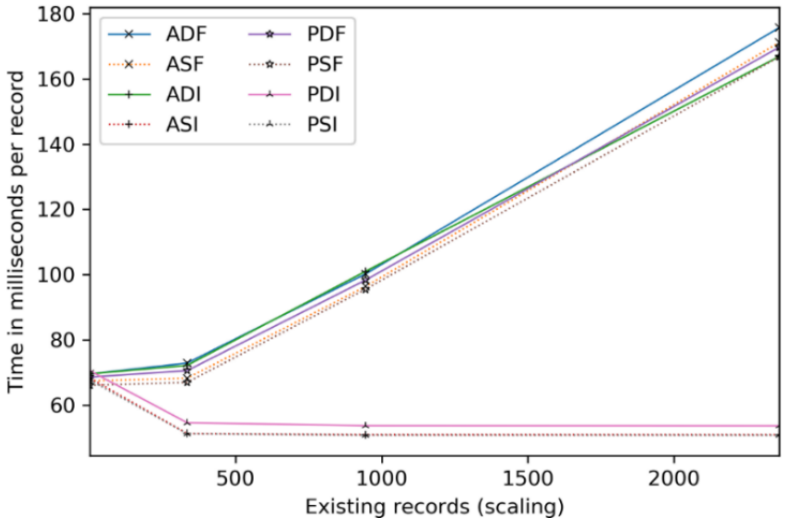
Concerning scaling (Figure 18), network latency accounts for 24% to 27% of the overall cost at 1 existing record, dropping precipitously to 1% to 4% by 365. Beyond 365, ASI, PDI, and PSI are constant time as only new and modified information are processed. Although the others are 4 to 9 milliseconds apart, ASF and PSF tend to be slightly faster. However, by 2362, ASI, PDI, and PSI are 68% to 71% more efficient than the other methods.

Figure 17. Client processing time in seconds and milliseconds per record by the number of records inserted. ADF: AES-encrypted data, dynamic keys, full block storage; ADI: AES-encrypted data, dynamic keys, incremental storage; AES: Advanced Encryption Standard; ASF: AES-encrypted data, static keys, full block storage; ASI: AES-encrypted data, static keys, incremental storage; PDF: PRE-encrypted data, dynamic keys, full block storage; PDI: PRE-encrypted data, dynamic keys, incremental storage; PRE: proxy re-encryption; PSF: PRE-encrypted data, static keys, full block storage; PSI: PRE-encrypted data, static keys, incremental storage.



XSL•FO
RenderX

Figure 18. Client processing time in milliseconds per record added given an existing record set. ADF: AES-encrypted data, dynamic keys, full block storage; ADI: AES-encrypted data, dynamic keys, incremental storage; AES: Advanced Encryption Standard; ASF: AES-encrypted data, static keys, full block storage; ASI: AES-encrypted data, static keys, incremental storage; PDF: PRE-encrypted data, dynamic keys, full block storage; PDI: PRE-encrypted data, dynamic keys, incremental storage; PRE: proxy re-encryption; PSF: PRE-encrypted data, static keys, full block storage; PSI: PRE-encrypted data, static keys, incremental storage.



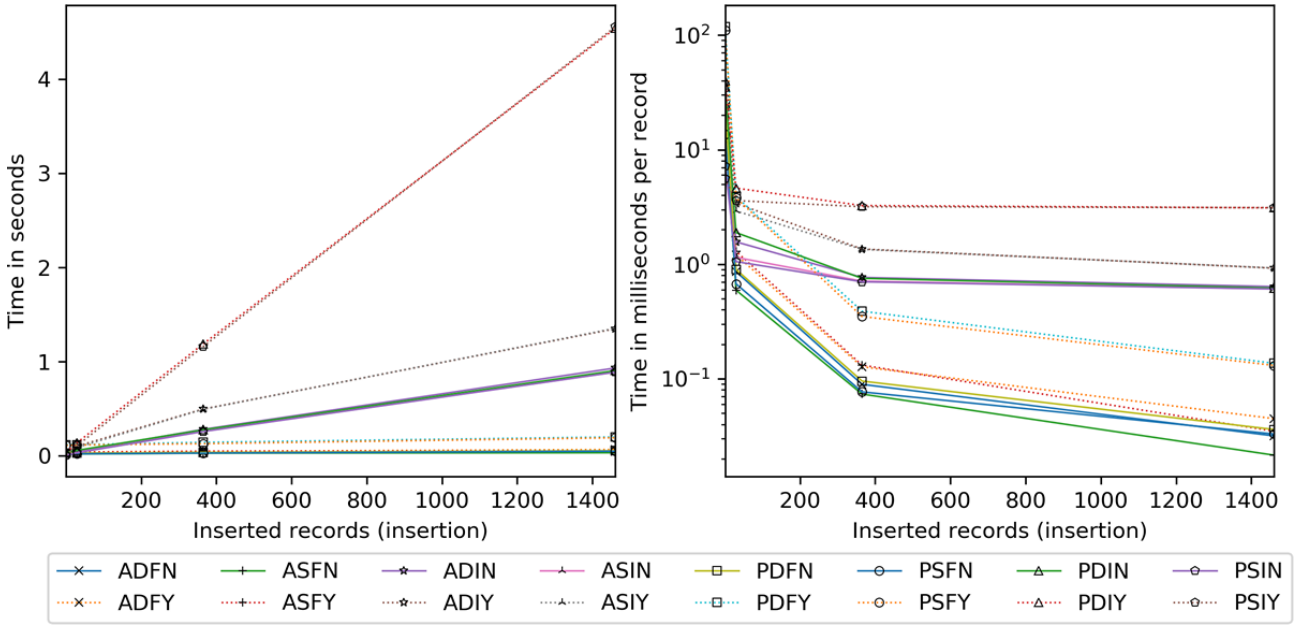
Server Processing Time

Server insertion costs are presented in Figure 19. Full block approaches process insertions the fastest. The order is in tens of milliseconds for AES (19-65 milliseconds), and PRE without server-side encryption (14-53 milliseconds). PRE with server-side encryption is measured in the 111 to 199 milliseconds range. ASIN and PSIN extend nearly uniformly from 7 to 889 milliseconds over the sets, with a slight dynamic keying penalty of 12 to 42 milliseconds. ADIY and ASIY are roughly 40% to 70% costlier than ADIN and ASIN. PDIY and PSIY are the slowest, reaching 4.5 seconds at 1461 insertions. Per record, incremental methods level off at 365, whereas full block approaches continue to decline at a rate greater than 57% at 1461.

Regarding scaling (Figure 20), ASIN, ASIY, PSIN and PSIY are constant time. ASIN and PSIN take 0.6, ASIY 1.1, and PSIY 3.1 milliseconds per record. The rest are affected in various ways by existing records. The quickest configurations are ADFN, ADFY, ASFN, ASFY, PDFN, and PSFN. By 2362, they are just shy of ASIN and PSIN. PDFY and PSFY are minimally affected by existing records, with times ranging from 0.3 to 1.1 milliseconds. PDIN is next at roughly 0.7 to 3.6 milliseconds. ADIN and ADIY increase sharply from 0.7 to 6.9 milliseconds and 1.1 to 9.1 milliseconds, respectively. PDIY, at 3.5 to 6.1 milliseconds, is cheaper than ADIY at 945 and ADIN around 1800.

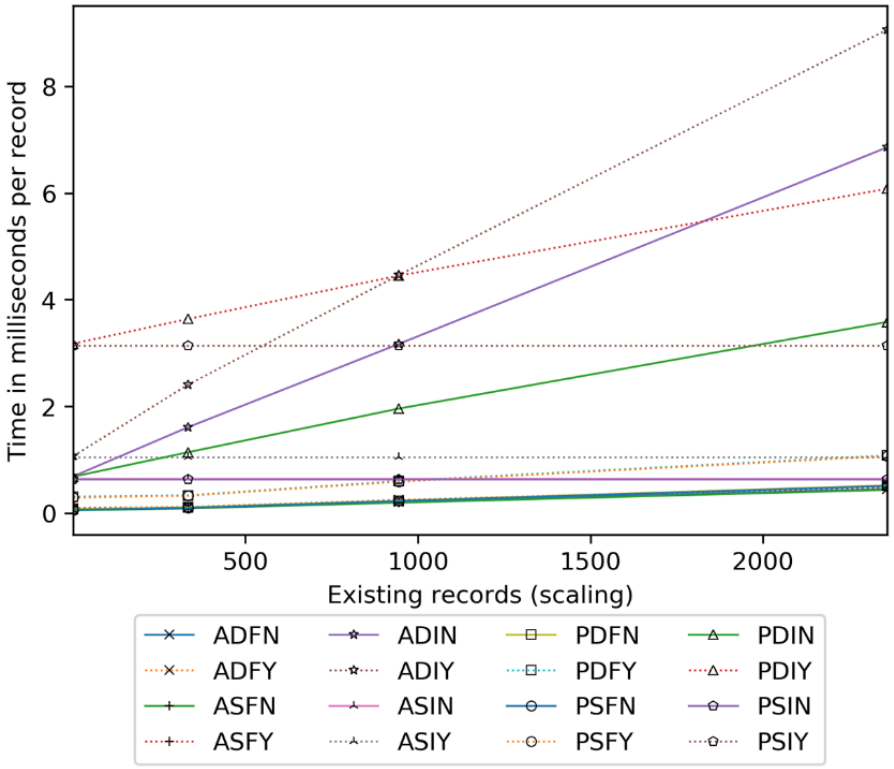
XSL•FO
RenderX

Figure 19. Server processing time in seconds and milliseconds per record by the number of records inserted. ADFN: AES-encrypted data, dynamic keys, full block storage, no server-side encryption; ADFY: AES-encrypted data, dynamic keys, full block storage, server-side encryption; ADIN: AES-encrypted data, dynamic keys, incremental storage, no server-side encryption; ADIY: AES-encrypted data, dynamic keys, incremental storage, server-side encryption; AES: Advanced Encryption Standard; ASFN: AES-encrypted data, static keys, full block storage, no server-side encryption; ASFY: AES-encrypted data, static keys, full block storage, server-side encryption; ASIN: AES-encrypted data, static keys, incremental storage, no server-side encryption; ASIY: AES-encrypted data, static keys, incremental storage, server-side encryption; PDFN: PRE-encrypted data, dynamic keys, full block storage, no server-side encryption; PDFY: PRE-encrypted data, dynamic keys, full block storage, server-side encryption; PDIN: PRE-encrypted data, dynamic keys, incremental storage, no server-side encryption; PDIY: PRE-encrypted data, dynamic keys, incremental storage, server-side encryption; PRE: proxy re-encryption; PSFN: PRE-encrypted data, static keys, full block storage, no server-side encryption; PSFY: PRE-encrypted data, static keys, full block storage, server-side encryption; PSIN: PRE-encrypted data, static keys, incremental storage, no server-side encryption; PSIY: PRE-encrypted data, static keys, incremental storage, server-side encryption.



XSL•FO
RenderX

Figure 20. Server processing time in milliseconds per record added given an existing record set. ADFN: AES-encrypted data, dynamic keys, full block storage, no server-side encryption; ADFY: AES-encrypted data, dynamic keys, full block storage, server-side encryption; ADIN: AES-encrypted data, dynamic keys, incremental storage, no server-side encryption; ADIY: AES-encrypted data, dynamic keys, incremental storage, server-side encryption; AES: Advanced Encryption Standard; ASFN: AES-encrypted data, static keys, full block storage, no server-side encryption; ASFY: AES-encrypted data, static keys, full block storage, server-side encryption; ASIN: AES-encrypted data, static keys, incremental storage, no server-side encryption; ASIY: AES-encrypted data, static keys, incremental storage, server-side encryption; PDFN: PRE-encrypted data, dynamic keys, full block storage, no server-side encryption; PDFY: PRE-encrypted data, dynamic keys, full block storage, server-side encryption; PDIN: PRE-encrypted data, dynamic keys, incremental storage, no server-side encryption; PDIY: PRE-encrypted data, dynamic keys, incremental storage, server-side encryption; PRE: proxy re-encryption; PSFN: PRE-encrypted data, static keys, full block storage, no server-side encryption; PSFY: PRE-encrypted data, static keys, full block storage, server-side encryption; PSIN: PRE-encrypted data, static keys, incremental storage, no server-side encryption; PSIY: PRE-encrypted data, static keys, incremental storage, server-side encryption.



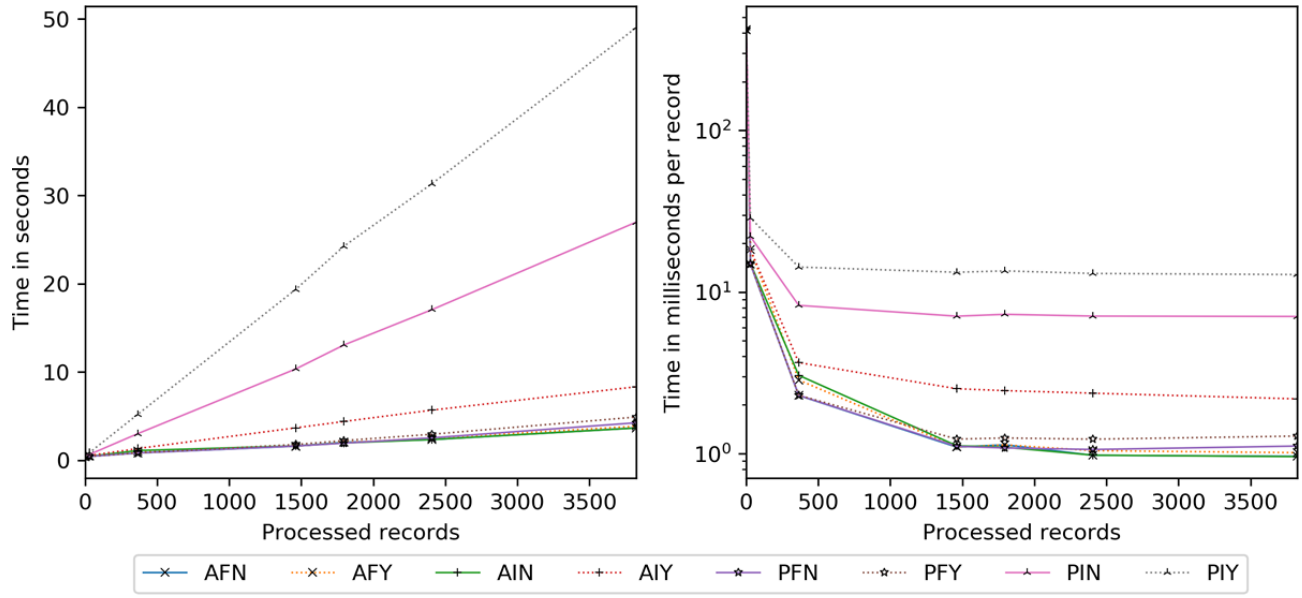
Smart Contract Execution

Smart contract execution by time and per record is assessed in this section, with results conveyed in Figure 21. Encryption key mode is not a factor in smart contract execution, as there is only 1 re-encryption key for all data; thus, it is not reported. As with the server-to-client network latency experiments, data include insertion and existing with 1461 additions sets.

In absolute terms, AFN, AFY, and AIN are the fastest, followed by PFN and PFY (progressing from 1% to 20% slower), AIY (23%-183%), PIN (46%-854%), and PIY (94%-1679%). Incremental server-side encryption is considerably expensive, doubling PRE and tripling AES times. Network latency accounts for around 9% of incremental and 19% of full block time. Per record, AIY, PIN, and PIY noticeably level-off by 365, and AFY, PFN, and PFY by 1461. AFN and AIN, however, descend beyond 3822 (2361+1461) at 2%.

XSL•FO
RenderX

Figure 21. Smart contract execution time in seconds and milliseconds per record by the number of records processed. AES: Advanced Encryption Standard; AFY: AES-encrypted data, full block storage, server-side encryption; AFN: AES-encrypted data, full block storage, no server-side encryption; AIN: AES-encrypted data, incremental block storage, no server-side encryption; AIY: AES-encrypted data, incremental storage, server-side encryption; PFN: PRE-encrypted data, full block storage, no server-side encryption; PFY: PRE-encrypted data, full block storage, server-side encryption; PIN: PRE-encrypted data, incremental block storage, no server-side encryption; PIY: PRE-encrypted data, incremental storage, server-side encryption; PRE: proxy re-encryption.



Discussion

Principal Findings

Figure 22 integrates pertinent results into a single visual for high-level performance analysis of client and server insertion (per record insertion cost given *n* loaded records) and scaling (per record insertion cost given *n* existing records) operations. Proceeding from top to bottom is client time in milliseconds (ie, section on Client Processing Time), server time in milliseconds (ie, section on Server Processing Time), and size in kilobytes (ie, section on Transmission Size).

The impracticality of full block approaches is apparent from the patient's vantage point. Whether inserting or scaling, block formation and transmission are resource intensive. For those with basic computers or mobile devices, or those operating on metered or low-bandwidth networks, these options should be avoided.

Catering to constrained environments are ASIN, PSIY, PDIN, PDIY, PSIN, and PSIY. ASIN, ASIY, PSIN, and PSIY are constant in bytes and server processing time, with decreasing client processing time because of network latency amortization over an otherwise constant process. PDIN and PDIY are constant in bytes and amortize latency as do the previous but irregular for the server (ie, server-side rekeying). Overall, ASIN and ASIY are the top performers. They require minimal time and bandwidth for record insertion, hold constant when scaling, and quickly execute smart contracts (refer to the section on Smart Contract Execution for details). The compromise is security. Static approaches, although fast, are susceptible to attacks (see the section on Smart Contracts and Multimedia Appendix 1).

Dynamic selections are more secure as data are continually rekeyed, preventing decryption by old or compromised keys. ADIN and ADIY are arguably the worst and operationally infeasible, as rekeying requires (1) the client to decrypt, encrypt, and transmit all information with each update and (2) the server to replace the existing block with the new, server-side–encrypted data.

PRE incremental methods are byte intensive because of excessive padding of the small experimental files and slow during smart contract execution. This is mostly mitigated through full block approaches. Unlike AES, the variability of PRE cipher size is vast. It has the potential to be compact and efficient or bloated and wasteful. PSIN and PSIY are subject to the same static key vulnerabilities as ASIN and ASIY and are more expensive. One must decide if cipher malleability justifies increased resource expenditure.

PDIN and PDIY are the only viable dynamic options. Server processing is insignificant for record insertions (a few milliseconds) but rises with scale. From 1 to 2361, PDIN is 1% to 7% and PDIY 4% to 11%, the magnitude of the client. Server processing is projected to eclipse client by around 40,000 and 37,000 records for PDIN and PDIY respectively. However, with appropriate hardware and in-memory databases, this cost can be reduced. Compared with ASIN and ASIY, both are marginally slower on the client, but roughly twice in bytes and latency.

Ultimately, several candidates emerge. ASIN and ASIY for speed, PSIN and PSIY for malleability, and PDIN and PDIY for malleability and security. Refer to Table 9 for a detailed comparison.

XSL•FO
RenderX

Figure 22. Relative comparison of client and server processing time in milliseconds and transmission size in kilobytes per record by insertion and scaling. ADFN: AES-encrypted data, dynamic keys, full block storage, no server-side encryption; ADFY: AES-encrypted data, dynamic keys, full block storage, server-side encryption; ADIN: AES-encrypted data, dynamic keys, incremental storage, no server-side encryption; ADIY: AES-encrypted data, dynamic keys, incremental storage, server-side encryption; AES: Advanced Encryption Standard; ASFN: AES-encrypted data, static keys, full block storage, no server-side encryption; ASFY: AES-encrypted data, static keys, full block storage, server-side encryption; ASIN: AES-encrypted data, static keys, incremental storage, no server-side encryption; ASIY: AES-encrypted data, static keys, incremental storage, server-side encryption; PDFN: PRE-encrypted data, dynamic keys, full block storage, no server-side encryption; PDFY: PRE-encrypted data, dynamic keys, full block storage, server-side encryption; PDIN: PRE-encrypted data, dynamic keys, incremental storage, no server-side encryption; PDIY: PRE-encrypted data, dynamic keys, incremental storage, server-side encryption; PRE: proxy re-encryption; PSFN: PRE-encrypted data, static keys, full block storage, no server-side encryption; PSFY: PRE-encrypted data, static keys, full block storage, server-side encryption; PSIN: PRE-encrypted data, static keys, incremental storage, no server-side encryption; PSIY: PRE-encrypted data, static keys, incremental storage, server-side encryption.

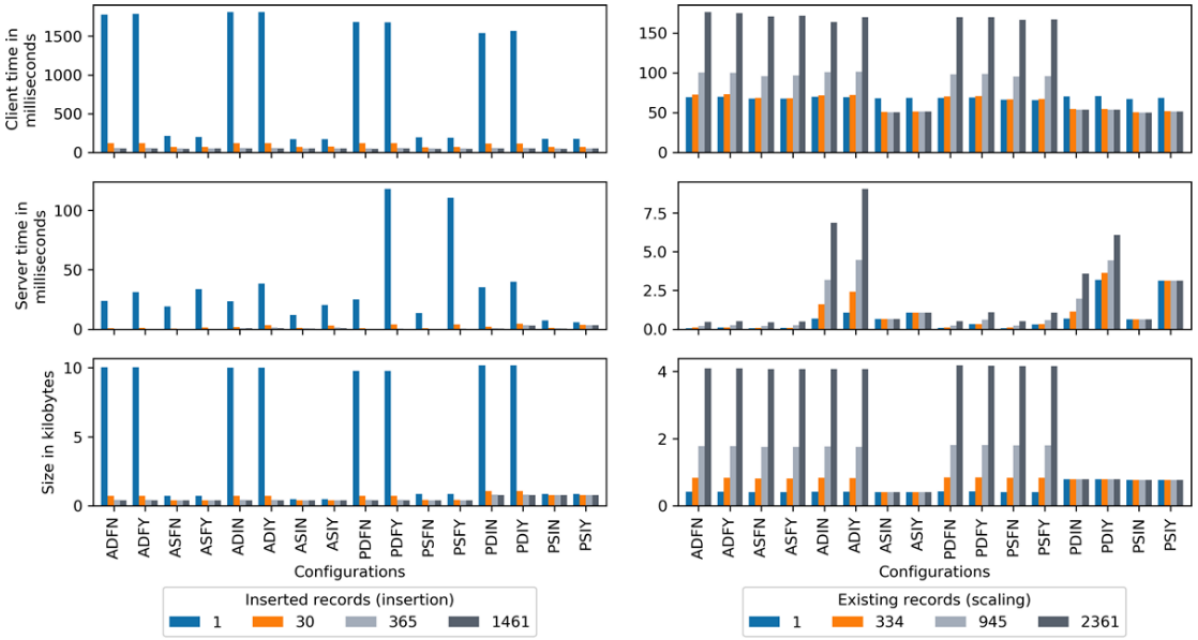


Table 9. Comparison of practical configurations by cipher malleability; security; insertion, scaling, and smart contract execution time; and transmitted bytes.

Property	ASIN[a]	ASIY[b]	PSIN[c]	PSIY[d]	PDIN[e]	PDIY[f]
Cipher malleability	No	No	Yes	Yes	Yes	Yes
Security—dynamic keying	No	No	No	No	Yes	Yes
Security—server-side encryption	No	Yes	No	Yes	No	Yes
Client insertions time	Fastest	Fastest	Fast	Fast	Fast if >30	Fast if >30
Client insertions bytes	Smallest	Smallest	Largest	Largest	Largest	Largest
Server insertions time	Fast	Slow	Slowest	Slow	Fast	Slowest
Client scaling time	Fastest, constant[g]	Fastest, constant[g]	Fastest, constant[g]	Fastest, constant[g]	Fast, constant[g]	Fast, constant[g]
Client scaling bytes	Smallest, constant	Smallest, constant	Small, constant	Small, constant	Small, constant	Small, constant
Server scaling time	Fast, constant	Slow, constant	Fast, constant	Slower, constant	Slower	Very slow
Smart contract execution	Fastest	Fastest	Very slow	Very slow	Slowest	Slowest

[a]ASIN: Advanced Encryption Standard–encrypted data, static keys, incremental storage, no server-side encryption.

[b]ASIY: Advanced Encryption Standard–encrypted data, static keys, incremental storage, server-side encryption.

[c]PSIN: proxy re-encryption–encrypted data, static keys, incremental storage, no server-side encryption.

[d]PSIY: proxy re-encryption–encrypted data, static keys, incremental storage, server-side encryption.

[e]PDIN: proxy re-encryption–encrypted data, dynamic keys, incremental storage, no server-side encryption.

[f]PDIY: proxy re-encryption–encrypted data, dynamic keys, incremental storage, server-side encryption.

[g]Constant time if latency, which is amortized over records, is not factored.

XSL•FO
RenderX

Limitations

Our study has the following limitations. First, the results are consistent with AES-128 and EC-256 alone. It is impossible to extrapolate the effects a change may have. Second, the small experimental files resulted in excess PRE cipher padding. Although the records were legitimate, EHR data may produce different results. Third, only 1 smart contract, which dynamic options regenerate during an update, was present for testing. With many contracts and few existing records, overall performance may diminish. In addition, smart contract regeneration was not optimized as the entire contract was reproduced and transmitted instead of just the re-encryption keys. This modification has the potential to decrease the size by 41%. Finally, server-side encryption only operates in dynamic mode. A static or periodic (eg, daily or after *x* number of transactions) option would reduce server-side processing at the expense of security. This will especially benefit PRE schemes, as they suffer a tremendous penalty under the weight of rekeying data after each read.

Conclusions and Future Work

In this study, a proof-of-concept patient-centered blockchain—HealthChain—was presented. The posited framework promotes patient engagement and facilitates secure, mediated information exchange between patients and providers. Redactable patient blocks, by way of chameleon hashing, were introduced to minimize data fragmentation, allow for in-place editing, and reduce resource consumption. PRE, smart contracts, and HL7 FHIR form the foundation of our proposed information exchange model, along with our 2PD PRE scheme and signature methods. A total of 16 experimental configurations were examined over 5 system dimensions by the cost of record insertion and scaling. Results indicate ASIN was the fastest and least bandwidth intensive, whereas PDIY was the best cryptographically, although the ultimate configuration rests with implementers and their desired level of speed and security.

Furthermore, 3 areas are targeted for future work. First, Barreto-Lynn-Scott [125] and Kachisa-Schaefer-Scott [126] EC families will be explored as potential replacements for the outdated jPBC curves. Second, as the proof-of-concept client tool is a deployed Web service on a client's machine (sufficient for testing), practical application necessitates an architectural redesign, wherein clients access HealthChain through a hosted, browser-based system. Hence, the cryptographic services will be ported to JavaScript for client-side execution, ensuring plaintext and generated secrets remain unexposed to nodes. Finally, we integrate our solution into Hyperledger Fabric to make use of their consensus, permissioning, and communications infrastructure.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Proxy re-encryption mathematical foundation, proposed extensions, and proofs.
[PDF File (Adobe PDF File), 514KB - jmir_v21i8e13592_app1.pdf]

Multimedia Appendix 2
Experimental configurations.
[PDF File (Adobe PDF File), 248KB - jmir_v21i8e13592_app2.pdf]

Multimedia Appendix 3
Source code files.
[ZIP File (Zip Archive), 33MB - jmir_v21i8e13592_app3.zip]

References

1. Yaeger K, Martini M, Rasouli J, Costa A. Emerging blockchain technology solutions for modern healthcare infrastructure. J Sci Innov Med 2019;2(1):1. [doi: 10.29024/jsim.7]
2. Zhang P, White J, Schmidt DC, Lenz G, Rosenbloom ST. FHIRChain: applying blockchain to securely and scalably share clinical data. Comput Struct Biotechnol J 2018;16:267-278 [FREE Full text] [doi: 10.1016/j.csbj.2018.07.004] [Medline: 30108685]
3. Azaria A, Ekblaw A, Vieira T, Lippman A. MedRec: Using Blockchain for Medical Data Access and Permission Management. In: Proceedings of the 2nd International Conference on Open and Big Data. 2016 Presented at: OBD'16; August 22-24, 2016; Vienna, Austria. [doi: 10.1109/OBD.2016.11]
4. Library of Congress. 2016. Public Law 114–255—Dec 13, 2016 URL: https://www.congress.gov/114/plaws/publ255/PLAW-114publ255.pdf [accessed 2019-02-01] [WebCite Cache ID 75rG0BbUq]
5. HealthIT. 2015. Report to Congress: Report on Health Information Blocking URL: https://www.healthit.gov/sites/default/files/reports/info_blocking_040915.pdf [WebCite Cache ID 75rGDwumh]
6. Bell L, Buchanan WJ, Cameron J, Lo O. Applications of blockchain within healthcare. Blockchain in Healthcare Today 2018 Jul 9;1:1-7. [doi: 10.30953/bhty.v1.8]

XSL•FO
RenderX

7. Cyran MA. Blockchain as a foundation for sharing healthcare data. Blockchain in Healthcare Today 2018 Mar 23;1:1-6. [doi: 10.30953/bhty.v1.13]
8. Ribitzky R, St Clair J, Houlding DI, McFarlane CT, Ahier B, Gould M, et al. Pragmatic, interdisciplinary perspectives on blockchain and distributed ledger technology: paving the future for healthcare. Blockchain in Healthcare Today 2018 Mar 23;1:1-15. [doi: 10.30953/bhty.v1.24]
9. Yue X, Wang H, Jin D, Li M, Jiang W. Healthcare data gateways: found healthcare intelligence on blockchain with novel privacy risk control. J Med Syst 2016 Oct;40(10):218. [doi: 10.1007/s10916-016-0574-6] [Medline: 27565509]
10. Rabah K. Challenges & opportunities for blockchain powered healthcare systems: a review. Mara Res J Med Health Sci 2017 Oct 16;1(1):45-52 [FREE Full text]
11. Xia Q, Sifah EB, Asamoah KO, Gao J, Du X, Guizani M. MeDShare: trust-less medical data sharing among cloud service providers via blockchain. IEEE Access 2017 Jul 24;5:14757-14767. [doi: 10.1109/ACCESS.2017.2730843]
12. Clauson KA, Breeden EA, Davidson C, Mackey TK. Leveraging blockchain technology to enhance supply chain management in healthcare:. Blockchain in Healthcare Today 2018 Mar 23;1:1-12. [doi: 10.30953/bhty.v1.20]
13. Mettler M. Blockchain Technology in Healthcare: The Revolution Starts Here. In: Proceedings of the 18th International Conference on e-Health Networking, Applications and Services. 2016 Presented at: Healthcom'16; September 14-16, 2016; Munich, Germany. [doi: 10.1109/HealthCom.2016.7749510]
14. Choudhury O, Sarker H, Rudolph N, Foreman M, Fay N, Dhuliawala M, et al. Enforcing human subject regulations using blockchain and smart contracts. Blockchain in Healthcare Today 2018 Mar 23;1:1-14. [doi: 10.30953/bhty.v1.10]
15. Esposito C, de Santis A, Tortora G, Chang H, Choo KR. Blockchain: a panacea for healthcare cloud-based data security and privacy? IEEE Cloud Comput 2018 Jan;5(1):31-37. [doi: 10.1109/MCC.2018.011791712]
16. Maslove DM, Klein J, Brohman K, Martin P. Using blockchain technology to manage clinical trials data: a proof-of-concept study. JMIR Med Inform 2018 Dec 21;6(4):e11949 [FREE Full text] [doi: 10.2196/11949] [Medline: 30578196]
17. Sylim P, Liu F, Marcelo A, Fontelo P. Blockchain technology for detecting falsified and substandard drugs in distribution: pharmaceutical supply chain intervention. JMIR Res Protoc 2018 Sep 13;7(9):e10163 [FREE Full text] [doi: 10.2196/10163] [Medline: 30213780]
18. Balsari S, Fortenko A, Blaya JA, Gropper A, Jayaram M, Matthan R, et al. Reimagining health data exchange: an application programming interface-enabled roadmap for India. J Med Internet Res 2018 Jul 13;20(7):e10725 [FREE Full text] [doi: 10.2196/10725] [Medline: 30006325]
19. Ichikawa D, Kashiyama M, Ueno T. Tamper-resistant mobile health using blockchain technology. JMIR Mhealth Uhealth 2017 Jul 26;5(7):e111 [FREE Full text] [doi: 10.2196/mhealth.7938] [Medline: 28747296]
20. Cisneros JL, Aarestrup FM, Lund O. Public health surveillance using decentralized technologies. Blockchain in Healthcare Today 2018 Mar 23;1:1-14. [doi: 10.30953/bhty.v1.17]
21. Dubovitskaya A, Xu Z, Ryu S, Schumacher M, Wang F. Secure and trustable electronic medical records sharing using blockchain. AMIA Annu Symp Proc 2017;2017:650-659 [FREE Full text] [Medline: 29854130]
22. Griggs KN, Ossipova O, Kohlios CP, Baccarini AN, Howson EA, Hayajneh T. Healthcare blockchain system using smart contracts for secure automated remote patient monitoring. J Med Syst 2018 Jun 6;42(7):130. [doi: 10.1007/s10916-018-0982-x] [Medline: 29876661]
23. Halamka JD, Lippman A, Ekblaw A. Harvard Business Review. 2017. The Potential for Blockchain to Transform Electronic Health Records URL: https://hbr.org/2017/03/the-potential-for-blockchain-to-transform-electronic-health-records
24. Ekblaw A, Azaria A, Halamka JD, Lippman A. A Case Study for Blockchain in Healthcare: 'Medrec' Prototype for Electronic Health Records and Medical Research Data. In: Proceedings of the 2nd International Conference on Open & Big Data. 2016 Presented at: OBD'16; August 22-24, 2016; Vienna, Austria.
25. Roehrs A, da Costa CA, da Rosa RR. OmniPHR: a distributed architecture model to integrate personal health records. J Biomed Inform 2017 Jul;71:70-81 [FREE Full text] [doi: 10.1016/j.jbi.2017.05.012] [Medline: 28545835]
26. DeSalvo KB, Galvez D. HealthIT.gov. 2015. Connecting Health and Care for the Nation: A Shared Nationwide Interoperability Roadmap URL: https://www.healthit.gov/sites/default/files/comments_upload/tascet_comments_to_onc_interoperability_roadmap_2015_04_01.pdf [accessed 2019-02-01] [WebCite Cache ID 75rGOiVgd]
27. HL7 International. 2018. Fast Healthcare Interoperability Resources (FHIR) URL: https://www.hl7.org/fhir [accessed 2019-02-01] [WebCite Cache ID 75r90KGMd]
28. Kuo TT, Ohno-Machado L. arXiv. 2018. ModelChain: Decentralized Privacy-Preserving Healthcare Predictive Modeling Framework on Private Blockchain Networks URL: https://arxiv.org/abs/1802.01746
29. Kuo TT, Kim HE, Ohno-Machado L. Blockchain distributed ledger technologies for biomedical and health care applications. J Am Med Inform Assoc 2017 Nov 1;24(6):1211-1220 [FREE Full text] [doi: 10.1093/jamia/ocx068] [Medline: 29016974]
30. Tseng JH, Liao YC, Chong B, Liao SW. Governance on the drug supply chain via Gcoin blockchain. Int J Environ Res Public Health 2018 May 23;15(6):pii: E1055 [FREE Full text] [doi: 10.3390/ijerph15061055] [Medline: 29882861]
31. Radanović I, Likić R. Opportunities for use of blockchain technology in medicine. Appl Health Econ Health Policy 2018 Oct;16(5):583-590. [doi: 10.1007/s40258-018-0412-8] [Medline: 30022440]
32. Mackey TK, Nayyar G. A review of existing and emerging digital technologies to combat the global trade in fake medicines. Expert Opin Drug Saf 2017 May;16(5):587-602. [doi: 10.1080/14740338.2017.1313227] [Medline: 28349715]

XSL•FO
RenderX

33. Roman-Belmonte JM, de la Corte-Rodriguez H, Rodriguez-Merchan EC. How blockchain technology can change medicine. Postgrad Med 2018 May;130(4):420-427. [doi: 10.1080/00325481.2018.1472996] [Medline: 29727247]
34. Benchoufi M, Ravaud P. Blockchain technology for improving clinical research quality. Trials 2017 Jul 19;18(1):335 [FREE Full text] [doi: 10.1186/s13063-017-2035-z] [Medline: 28724395]
35. Benchoufi M, Porcher R, Ravaud P. Blockchain protocols in clinical trials: transparency and traceability of consent. F1000Res 2017;6:66 [FREE Full text] [doi: 10.12688/f1000research.10531.5] [Medline: 29167732]
36. Cichosz SL, Stausholm MN, Kronborg T, Vestergaard P, Hejlesen O. How to use blockchain for diabetes health care data and access management: an operational concept. J Diabetes Sci Technol 2019 Mar;13(2):248-253 [FREE Full text] [doi: 10.1177/1932296818790281] [Medline: 30047789]
37. Till BM, Peters AW, Afshar S, Meara JG, Meara J. From blockchain technology to global health equity: can cryptocurrencies finance universal health coverage? BMJ Glob Health 2017;2(4):e000570 [FREE Full text] [doi: 10.1136/bmjgh-2017-000570] [Medline: 29177101]
38. Boulos MN, Wilson JT, Clauson KA. Geospatial blockchain: promises, challenges, and scenarios in health and healthcare. Int J Health Geogr 2018 Jul 5;17(1):25 [FREE Full text] [doi: 10.1186/s12942-018-0144-x] [Medline: 29973196]
39. Qu C, Tao M, Yuan R. A hypergraph-based blockchain model and application in internet of things-enabled smart homes. Sensors (Basel) 2018 Aug 24;18(9):pii: E2784 [FREE Full text] [doi: 10.3390/s18092784] [Medline: 30149523]
40. Zhou L, Wang L, Ai T, Sun Y. BeeKeeper 2.0: confidential blockchain-enabled IoT system with fully homomorphic computation. Sensors (Basel) 2018 Nov 05;18(11):pii: E3785 [FREE Full text] [doi: 10.3390/s18113785] [Medline: 30400673]
41. Zhu X, Badr Y. Identity management systems for the internet of things: a survey towards blockchain solutions. Sensors (Basel) 2018 Dec 1;18(12):pii: E4215 [FREE Full text] [doi: 10.3390/s18124215] [Medline: 30513733]
42. Park JS, Youn TY, Kim HB, Rhee KH, Shin SU. Smart contract-based review system for an IoT data marketplace. Sensors (Basel) 2018 Oct 22;18(10):pii: E3577 [FREE Full text] [doi: 10.3390/s18103577] [Medline: 30360413]
43. Dwivedi AD, Srivastava G, Dhar S, Singh R. A decentralized privacy-preserving healthcare blockchain for IoT. Sensors (Basel) 2019 Jan 15;19(2):pii: E326 [FREE Full text] [doi: 10.3390/s19020326] [Medline: 30650612]
44. Nakamoto S. Bitcoin.org. 2008. Bitcoin: A Peer-to-Peer Electronic Cash System URL: https://bitcoin.org/bitcoin.pdf [WebCite Cache ID 75m0ZCttp]
45. Bitcoin Forum. 2011. Proof of Stake Instead of Proof of Work URL: https://bitcointalk.org/index.php?topic=27787.0 [accessed 2019-02-01] [WebCite Cache ID 75r9zR885]
46. Buterin V, Griffith V. arXiv. 2017. Casper the Friendly Finality Gadget URL: https://arxiv.org/abs/1710.09437 [accessed 2019-02-01] [WebCite Cache ID 75rA6am4h]
47. Hyperledger – Open Source Blockchain Technologies. 2017. Introduction to Hyperledger Business Blockchain Design Philosophy and Consensus URL: https://www.hyperledger.org/wp-content/uploads/2017/08/Hyperledger_Arch_WG_Paper_1_Consensus.pdf [accessed 2019-02-01] [WebCite Cache ID 75rA9MVgf]
48. Hyperledger Sawtooth. 2019. Introduction URL: https://sawtooth.hyperledger.org/docs/core/releases/latest/introduction.html [accessed 2019-02-01] [WebCite Cache ID 75rACSbao]
49. Hyperledger Fabric Docs. 2019. A Blockchain Platform for the Enterprise URL: https://hyperledger-fabric.readthedocs.io/en/latest/ [accessed 2019-02-01] [WebCite Cache ID 75rAGGyjw]
50. Ateniese G, Magri B, Venturi D, Andrade E. Redactable Blockchain - or - Rewriting History in Bitcoin and Friends. In: Proceedings of the European Symposium on Security and Privacy. 2017 Presented at: EuroS&P'17; April 26-28, 2017; Paris, France. [doi: 10.1109/EuroSP.2017.37]
51. Destefanis G, Marchesi M, Ortu M, Tonelli R, Bracciali A, Hierons R. Smart Contracts Vulnerabilities: A Call for Blockchain Software Engineering? In: Proceedings of the International Workshop on Blockchain Oriented Software Engineering. 2018 Presented at: IWBOSE'18; March 20, 2018; Campobasso, Italy. [doi: 10.1109/IWBOSE.2018.8327567]
52. Atzei N, Bartoletti M, Cimoli T. A Survey of Attacks on Ethereum Smart Contracts (SoK). In: Proceedings of the International Conference on Principles of Security and Trust. 2017 Presented at: POST'17; April 22-29, 2017; Uppsala, Sweden p. 164-186. [doi: 10.1007/978-3-662-54455-6_8]
53. Krawczyk H, Rabin T. CiteSeerX. 1997. Chameleon Hashing and Signatures URL: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.50.3262 [accessed 2019-02-01] [WebCite Cache ID 75rALtvwJ]
54. Krawczyk H, Rabin T. Chameleon Signatures. In: Proceedings of the Network and Distributed System Security Symposium. 2000 Presented at: NDSS'00; February 3-4, 2000; San Diego, California p. 143-154.
55. Krawczyk HM, Rabin TD. Google Patents. 2000. Chameleon Hashing and Signatures URL: https://patents.google.com/patent/US6108783A/en
56. Luu L, Chu DH, Olickel H, Saxena P, Hobor A. Making Smart Contracts Smarter. In: Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security. 2016 Presented at: CCS'16; October 24-28, 2016; Vienna, Austria p. 254-269. [doi: 10.1145/2976749.2978309]
57. Delmolino K, Arnett M, Kosba A, Miller A, Shi E. Step by Step Towards Creating a Safe Smart Contract: Lessons and Insights from a Cryptocurrency Lab. In: Proceedings of the International Conference on Financial Cryptography and Data

XSL•FO
RenderX

Security. 2016 Presented at: FC'16; February 22-26, 2016; Christ Church, Barbados p. 79-94. [doi: 10.1007/978-3-662-53357-4_6]
58. Kalra S, Goel S, Dhawan M, Sharma S. ZEUS: Analyzing Safety of Smart Contracts. In: Proceedings of the Network and Distributed Systems Security Symposium. 2018 Presented at: NDSS'18; February 18-21, 2018; San Diego, CA p. 18-21.
59. Tsankov P, Dan A, Drachsler-Cohen D, Gervais A, Bünzli F, Vechev M. Securify: Practical Security Analysis of Smart Contracts. In: Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security. 2018 Presented at: CCS'18; October 15-19, 2018; Toronto, Canada p. 67-82. [doi: 10.1145/3243734.3243780]
60. Ateniese G, Fu K, Green M, Hohenberger S. Improved proxy re-encryption schemes with applications to secure distributed storage. ACM T Inform Syst Se 2006 Feb 1;9(1):1-30. [doi: 10.1145/1127345.1127346]
61. Mambo M, Okamoto E. Proxy cryptosystems: delegation of the power to decrypt ciphertexts. IEICE T Fund Electr Commun Comput Sci 1997;80(1):54-63.
62. Blaze M, Bleumer G, Strauss M. Divertible Protocols and Atomic Proxy Cryptography. In: Proceedings of the International Conference on the Theory and Applications of Cryptographic Techniques. 1998 Presented at: EUROCRYPT'98; May 31-June 4, 1998; Espoo, Finland p. 127-144. [doi: 10.1007/BFb0054122]
63. Nuñez D, Agudo I, Lopez J. Proxy re-encryption: analysis of constructions and its application to secure access delegation. J Netw Comput Appl 2017 Jun;87:193-209. [doi: 10.1016/j.jnca.2017.03.005]
64. Xu L, Wu X, Zhang X. CL-PRE: A Certificateless Proxy Re-Encryption Scheme for Secure Data Sharing with Public Cloud. In: Proceedings of the 7th ACM Symposium on Information, Computer and Communications Security. 2012 Presented at: ASIACCS'12; May 2-4, 2012; Seoul, Korea p. 87-88. [doi: 10.1145/2414456.2414507]
65. Ibraimi L, Tang Q, Hartel P, Jonker W. A Type-and-Identity-Based Proxy Re-encryption Scheme and Its Application in Healthcare. In: Proceedings of the Workshop on Secure Data Management. 2008 Presented at: SDM'08; August 24, 2008; Auckland, New Zealand p. 185-198. [doi: 10.1007/978-3-540-85259-9_12]
66. Lin H, Shao J, Zhang C, Fang Y. CAM: cloud-assisted privacy preserving mobile health monitoring. IEEE T Inf Foren Sec 2013 Jun 29;8(6):985-997. [doi: 10.1109/TIFS.2013.2255593]
67. Jia W, Zhu J, Cao Z, Wei L, Lin X. SDSM: A Secure Data Service Mechanism in Mobile Cloud Computing. In: Proceedings of the Conference on Computer Communications Workshops. 2011 Presented at: INFOCOM WKSHPS'11; April 10-15, 2011; Shanghai, China p. 1060-1065. [doi: 10.1109/INFCOMW.2011.5928784]
68. Thilakanathan D, Chen S, Nepal S, Calvo R, Alem L. A platform for secure monitoring and sharing of generic health data in the Cloud. Future Gener Comp Sy 2014 Jun;35:102-113. [doi: 10.1016/j.future.2013.09.011]
69. Qin Z, Xiong H, Wu S, Batamuliza J. A survey of proxy re-encryption for secure data sharing in Cloud computing. IEEE T Serv Comput 2017 Sep;13(9):1. [doi: 10.1109/TSC.2016.2551238]
70. Roy R, Mathai PP. Proxy re-encryption schemes for secure cloud data and applications: a survey. Int J Comput Appl 2017 Apr 17;164(5):1-6. [doi: 10.5120/ijca2017913642]
71. Kirshanova E. Proxy Re-Encryption from Lattices. In: Proceedings of the International Workshop on Public Key Cryptography. 2014 Presented at: PKC'14; March 26-28, 2014; Aires, Argentina p. 77-94. [doi: 10.1007/978-3-642-54631-0_5]
72. Zhang A, Bacchus A, Lin X. Consent-based access control for secure and privacy-preserving health information exchange. Secur Commun Netw 2016 Jul 7;9(16):3496-3508. [doi: 10.1002/sec.1556]
73. Au MH, Yuen TH, Liu JK, Susilo W, Huang X, Xiang Y, et al. A general framework for secure sharing of personal health records in cloud system. J Comput Syst Sci 2017 Dec;90:46-62. [doi: 10.1016/j.jcss.2017.03.002]
74. Wang XA, Ma J, Xhafa F, Zhang M, Luo X. Cost-effective secure e-health cloud system using identity based cryptographic techniques. Future Gener Comp Sy 2017 Feb;67:242-254. [doi: 10.1016/j.future.2016.08.008]
75. Huang Q, Wang L, Yang Y. Secure and privacy-preserving data sharing and collaboration in mobile healthcare social networks of smart cities. Secur Commun Netw 2017 Aug;2017:1-12. [doi: 10.1155/2017/6426495]
76. Syalim A, Nishide T, Sakurai K. Realizing Proxy Re-Encryption in the Symmetric World. In: Proceedings of the International Conference on Informatics Engineering and Information Science. 2011 Presented at: ICIEIS'11; November 14-16, 2011; Kuala Lumpur, Malaysia p. 259-274. [doi: 10.1007/978-3-642-25327-0_23]
77. Veeraragavan N, Rajesh MS. Proxy re-encryption using hybrid encryption scheme secure against chosen ciphertext attack. Int J Comput Sci Mobile Appl 2014 Jan;2(1):94-100 [FREE Full text]
78. Tysowski PK, Hasan MA. Hybrid attribute- and re-encryption-based key management for secure and scalable mobile applications in clouds. IEEE T Cloud Comput 2013 Jul;1(2):172-186. [doi: 10.1109/TCC.2013.11]
79. Nuñez D, Agudo I, Lopez J. Integrating OpenID With Proxy Re-Encryption to Enhance Privacy in Cloud-Based Identity Services. In: Proceedings of the International Conference on Cloud Computing Technology and Science. 2012 Presented at: ICCCS'12; December 3-6, 2012; Taipei, Taiwan. [doi: 10.1109/CloudCom.2012.6427551]
80. NuCypher. 2019. URL: https://www.nucypher.com/ [accessed 2019-02-01] [WebCite Cache ID 75rAoxjek]
81. National Institute of Standards and Technology. Federal Register. 2015. Federal Information Processing Standard (FIPS) 186-4, Digital Signature Standard; Request for Comments on the NIST-Recommended Elliptic Curves URL: https://www.federalregister.gov/documents/2015/10/20/2015-26539/

XSL•FO
RenderX

federal-information-processing-standard-fips-186-4-digital-signature-standard-request-for-comments [accessed 2019-02-01] [WebCite Cache ID 75rIZ8DIs]
82. Committee on National Security Systems. 2019. CNSS Policies URL: https://www.cnss.gov/CNSS/issuances/Policies.cfm [accessed 2019-02-01] [WebCite Cache ID 75rAuufT0]
83. Cryptome. 2015. Advisory Memorandum: Use of Public Standards for The Secure Sharing of Information Among National Security Systems URL: https://cryptome.org/2015/08/CNSS_Advisory_Memo_02-15.pdf [accessed 2019-02-01] [WebCite Cache ID 75rAySL9Y]
84. National Security Agency - Applications. 2018. CNSA Suite and Quantum Computing FAQ URL: https://apps.nsa.gov/iaarchive/library/ia-guidance/ia-solutions-for-classified/algorithm-guidance/cnsa-suite-and-quantum-computing-faq.cfm [accessed 2019-02-01] [WebCite Cache ID 75rB1hWh7]
85. National Security Agency - Applications. Commercial National Security Algorithm Suite URL: https://apps.nsa.gov/iaarchive/ [accessed 2019-02-01] [WebCite Cache ID 75rB8qqQW]
86. Joux A, Vitse V. Elliptic curve discrete logarithm problem over small degree extension fields. J Cryptol 2011 Nov 30;26(1):119-143. [doi: 10.1007/s00145-011-9116-z]
87. Galbraith SD, Gaudry P. Recent progress on the elliptic curve discrete logarithm problem. Design Code Cryptogr 2015 Nov 23;78(1):51-72. [doi: 10.1007/s10623-015-0146-7]
88. Cao Z, Liu L. Cryptology ePrint Archive. 2015. On the Disadvantages of Pairing-Based Cryptography URL: https://eprint.iacr.org/2015/084.pdf
89. Nuñez D, Agudo I, Lopez J. NTRUReEncrypt: An Efficient Proxy Re-Encryption Scheme Based on NTRU. In: Proceedings of the 10th ACM Symposium on Information, Computer and Communications Security. 2015 Presented at: ASIA CCS'15; April 14-17, 2015; Singapore, Republic of Singapore p. 179-189. [doi: 10.1145/2714576.2714585]
90. Aono Y, Boyen X, Phong LT, Wang L. Key-Private Proxy Re-Encryption under LWE. In: Proceedings of the International Conference on Cryptology in India. 2013 Presented at: INDOCRYPT'13; December 7-10, 2013; Mumbai, India p. 1-18. [doi: 10.1007/978-3-319-03515-4_1]
91. Xagawa K. Keita Xagawa. 2010. Cryptography with Lattices URL: http://xagawa.net/pdf/2010Thesis.pdf [accessed 2019-02-01] [WebCite Cache ID 75rBFweUG]
92. Xagawa K, Tanaka K. Kyoto University Research Information Repository. 2010. Proxy Re-Encryption Based on Learning With Errors URL: https://repository.kulib.kyoto-u.ac.jp/dspace/bitstream/2433/141576/1/1691-05.pdf
93. Kim KS, Jeong IR. Collusion-resistant unidirectional proxy re-encryption scheme from lattices. J Commun Netw 2016;18(1):1-7. [doi: 10.1109/JCN.2016.000003]
94. Lauter K. Postquantum opportunities: lattices, homomorphic encryption, and supersingular isogeny graphs. IEEE Security Privacy 2017;15(4):22-27. [doi: 10.1109/MSP.2017.3151338]
95. Martins P, Sousa L, Mariano A. A survey on fully homomorphic encryption: an engineering perspective. ACM Comput Surv 2018 Jan 12;50(6):1-33. [doi: 10.1145/3124441]
96. 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]
97. Agarwal R, Anderson C, Zarate J, Ward C. If we offer it, will they accept? Factors affecting patient use intentions of personal health records and secure messaging. J Med Internet Res 2013 Feb 26;15(2):e43 [FREE Full text] [doi: 10.2196/jmir.2243] [Medline: 23470453]
98. Al-Sahan A, Saddik B. Perceived challenges for adopting the personal health record (PHR) at Ministry of National Guard Health Affairs (MNGHA)- Riyadh. Online J Public Health Inform 2016;8(3):e205 [FREE Full text] [doi: 10.5210/ojphi.v8i3.6845] [Medline: 28210426]
99. HIMSS - Healthcare Information and Management Systems Society. 2008. Defining Key Health Information Technology Terms URL: https://www.himss.org/defining-key-health-information-technology-terms-onc-nahit [WebCite Cache ID 75rDNMvNq]
100. HealthIT. 2019. Interoperability Standards: Advisory Office of the National Coordinator for Health IT URL: https://www.healthit.gov/isa/sites/isa/files/inline-files/2019ISAReferenceEdition.pdf [accessed 2019-02-01] [WebCite Cache ID 75rDJ5owA]
101. HealthIT. 2019. Interoperability Proving Ground URL: https://www.healthit.gov/techlab/ipg/ [WebCite Cache ID 75rDPo3Go]
102. Apple. 2019. Empower Your Patients With Health Records on iPhone URL: https://www.apple.com/healthcare/health-records/ [accessed 2019-02-01] [WebCite Cache ID 75rDXPhy1]
103. Scarfone K, Souppaya M, Sexton M. NIST Computer Security Resource Center. 2007. Guide to Storage Encryption Technologies for End User Devices URL: https://csrc.nist.gov/publications/detail/sp/800-111/final [WebCite Cache ID 75rI4U6Zr]
104. Elgamal T. A public key cryptosystem and a signature scheme based on discrete logarithms. IEEE Trans Inf Theory 1985 Jul;31(4):469-472. [doi: 10.1109/TIT.1985.1057074]
105. Gentry C. Applied Cryptography Group | Stanford University. 2009. A Fully Homomorphic Encryption Scheme URL: https://crypto.stanford.edu/craig/craig-thesis.pdf

XSL•FO
RenderX

106. Kogos KG, Filippova KS, Epishkina AV. Fully Homomorphic Encryption Schemes: The State of the Art. In: Proceedings of the Conference of Russian Young Researchers in Electrical and Electronic Engineering. 2017 Presented at: EIConRus'17; February 1-3, 2017; St Petersburg, Russia. [doi: 10.1109/EIConRus.2017.7910591]
107. El-Yahyaoui A, El Kettani MD. Evaluating AND Gates Over Encrypted Data in Cloud Computing. In: Proceedings of the International Conference on Advanced Communication Technologies and Networking. 2018 Presented at: CommNet'18; April 2-4, 2018; Marrakech, Morocco. [doi: 10.1109/COMMNET.2018.8360255]
108. EL-Yahyaoui A, El Kettani MD. A noise-free homomorphic evaluation of the AES circuits to optimize secure big data storage in cloud computing. In: Mohamed BA, Mohamed AA, editors. Innovations in Smart Cities and Applications: Proceedings of the 2nd Mediterranean Symposium on Smart City Applications. New York City: Springer; 2017:420-431.
109. Alves PG, Aranha DF. A framework for searching encrypted databases. J Internet Serv Appl 2018 Jan 3;9(1):1-18. [doi: 10.1186/s13174-017-0073-0]
110. Silva LV, Barbosa P, Marinho R, Brito A. Security and privacy aware data aggregation on cloud computing. J Internet Serv Appl 2018 Apr 2;9(1):1-13. [doi: 10.1186/s13174-018-0078-3]
111. Gentry CB. Google Patents. 2015. Fully Homomorphic Encryption URL: https://patents.google.com/patent/US9083526B2/en
112. Barker E. NIST Computer Security Resource Center. 2016. Recommendation for Key Management, Part 1: General URL: https://csrc.nist.gov/publications/detail/sp/800-57-part-1/rev-4/final [accessed 2019-02-01] [WebCite Cache ID 75rIBSia5]
113. de Caro A, Iovino V. jPBC: Java Pairing Based Cryptography. In: Proceedings of the Symposium on Computers and Communications. 2011 Presented at: ISCC'11; June 28-July 1, 2011; Kerkyra, Greece. [doi: 10.1109/ISCC.2011.5983948]
114. El Mrabet N, Joye M, editors. Guide to Pairing-Based Cryptography. First Edition. Boca Raton, Florida: Chapman and Hall/CRC; 2016.
115. Lynn B. Applied Cryptography Group | Stanford University. 2007. On the Implementation of Pairing-Based Cryptosystems URL: https://crypto.stanford.edu/pbc/thesis.pdf
116. Ateniese G, Benson K, Hohenberger S. Key-Private Proxy Re-Encryption. In: Proceedings of the Cryptographers' Track at the RSA Conference. 2009 Presented at: CT-RSA'09; April 20-24, 2009; San Francisco, CA p. 279-294. [doi: 10.1007/978-3-642-00862-7_19]
117. Liu Q, Tan CC, Wu J, Wang G. Reliable Re-Encryption in Unreliable Clouds. In: Proceedings of the Global Telecommunications Conference. 2011 Presented at: GLOBECOM'11; December 5-9, 2011; Kathmandu, Nepal. [doi: 10.1109/GLOCOM.2011.6133609]
118. Sanka S, Hota C, Rajarajan M. Secure Data Access in Cloud Computing. In: Proceedings of the 4th International Conference on Internet Multimedia Services Architecture and Application. 2010 Presented at: IMSAA'10; December 15-17, 2010; Bangalore, India. [doi: 10.1109/IMSAA.2010.5729397]
119. Hota C, Sanka S, Rajarajan M, Nair SK. Capability-based cryptographic data access control in cloud computing. Int J Adv Netw Appl 2011;3(3):1152-1161 [FREE Full text]
120. Hylock R, Harris ST. Semantic Scholar. 2017. Healthcare Database Management for Health Informatics and Information Management Students: Challenges and Instruction Strategies—Part 2 URL: https://pdfs.semanticscholar.org/0622/7deb64113cbc81fda84889201f4eceb71e63.pdf?_ga=2.176924453.1425319861.1564981021-1326934932.1560835791
121. Speedtest by Ookla. 2019. URL: https://www.speedtest.net/ [accessed 2019-02-01] [WebCite Cache ID 75rE1UoI9]
122. HAPI FHIR. 2019. URL: http://hapifhir.io/ [accessed 2019-02-01] [WebCite Cache ID 75rDxeSTT]
123. Colorlib. 2019. Network Activities URL: https://colorlib.com/polygon/gentelella [accessed 2019-02-01] [WebCite Cache ID 75rDz4CP3]
124. Mandel JC, Kreda DA, Mandl KD, Kohane IS, Ramoni RB. SMART on FHIR: a standards-based, interoperable apps platform for electronic health records. J Am Med Inform Assoc 2016 Sep;23(5):899-908 [FREE Full text] [doi: 10.1093/jamia/ocv189] [Medline: 26911829]
125. Barreto PS, Lynn B, Scott M. Constructing Elliptic Curves with Prescribed Embedding Degrees. In: Proceedings of the International Conference on Security in Communication Networks. 2002 Presented at: SCN'02; September 11-13, 2002; Amalfi, Italy p. 257-267. [doi: 10.1007/3-540-36413-7_19]
126. Kachisa EJ, Schaefer EF, Scott M. Constructing Brezing-Weng Pairing-Friendly Elliptic Curves Using Elements in the Cyclotomic Field. In: Proceedings of the International Conference on Pairing-Based Cryptography. 2008 Presented at: Pairing'08; September 1-3, 2008; Egham, United Kingdom p. 126-135. [doi: 10.1007/978-3-540-85538-5_9]

Abbreviations

2PD: 2-party proxy re-encryption decryption
AES: Advanced Encryption Standard
ADFN: AES-encrypted data, dynamic keys, full block storage, no server-side encryption
ADI: AES-encrypted data, dynamic keys, incremental storage
ADIN: AES-encrypted data, dynamic keys, incremental storage, no server-side encryption
ADIY: AES-encrypted data, dynamic keys, incremental storage, server-side encryption

XSL•FO
RenderX

AF: AES-encrypted data, full block storage
AFN: AES-encrypted data, full block storage, no server-side encryption
AIN: AES-encrypted data, incremental block storage, no server-side encryption
AIY: AES-encrypted data, incremental storage, server-side encryption
ASI: AES-encrypted data, static keys, incremental storage
ASIN: AES-encrypted data, static keys, incremental storage, no server-side encryption
ASIY: AES-encrypted data, static keys, incremental storage, server-side encryption
CFR: Code of Federal Regulations
EC: elliptic curves
ECC: elliptic curve cryptography
EHR: electronic health record
FHIR: Fast Healthcare Interoperability Resources
FIPS: Federal Information Processing Standard
HIPAA: Health Insurance Portability and Accountability Act of 1996
jPBC: Java Pairing-Based Cryptography library
Mbps: megabits per second
ONC: The Office of the National Coordinator for Health Information Technology
PCCH: Public-coin chameleon hash
PDF: PRE-encrypted data, dynamic encryption key full block storage
PDI: PRE-encrypted data, dynamic keys, incremental storage
PDIN: PRE-encrypted data, dynamic keys, incremental storage, no server-side encryption
PDIY: PRE-encrypted data, dynamic keys, incremental storage, server-side encryption
PF: PRE-encrypted data, full block storage
PFN: PRE-encrypted data, full block storage, no server-side encryption
PHR: personal health record
PI: PRE-encrypted data, incremental storage
PIN: PRE-encrypted data, incremental block storage, no server-side encryption
PIY: PRE-encrypted data, incremental storage, server-side encryption
PRE: proxy re-encryption
PSF: PRE-encrypted data, static keys, full block storage
PSI: PRE-encrypted data, static keys, incremental storage
PSIN: PRE-encrypted data, static keys, incremental storage, no server-side encryption
PSIY: PRE-encrypted data, static keys, incremental storage, server-side encryption
SHA: Secure Hash Algorithm

Edited by P Zhang, K Clauson; submitted 01.02.19; peer-reviewed by TT Kuo, T Ueno; comments to author 27.04.19; revised version received 22.06.19; accepted 19.07.19; published 31.08.19.

Please cite as:
Hylock RH, Zeng X
A Blockchain Framework for Patient-Centered Health Records and Exchange (HealthChain): Evaluation and Proof-of-Concept Study
J Med Internet Res 2019;21(8):e13592
URL: http://www.jmir.org/2019/8/e13592/
doi:10.2196/13592
PMID:31471959

©Ray Hales Hylock, Xiaoming Zeng. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 31.08.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Mapping the Access of Future Doctors to Health Information Technologies Training in the European Union: Cross-Sectional Descriptive Study

Guido Giunti[1], MD, PhD; Estefania Guisado-Fernandez[2,3], MD, PhD; Hrvoje Belani[4], MSc; Juan R Lacalle-Remigio[5], MD, PhD

[1]University of Oulu, Oulu, Finland

[2]University College Dublin School of Public Health, Physiotherapy and Sports Science, Dublin, Ireland

[3]Insight Centre for Data Analytics, Dublin, Ireland

[4]Ministry of Health, Zagreb, Croatia

[5]Universidad de Sevilla, Seville, Spain

Corresponding Author:
Guido Giunti, MD, PhD
University of Oulu
Pentti Kaiteran katu 1
Oulu, 90570
Finland
Phone: 358 294 48 0000
Fax: 358 294 48 0000
Email: drguidogiunti@gmail.com

Abstract

Background: Health information technologies (HITs) such as electronic health records (EHR) and telemedicine services are currently used to assist clinicians provide care to patients. There are many barriers to HIT adoption, including mismatches between investments and benefits, disruptions in the workflow, and concerns about privacy and confidentiality. The lack of HIT training of health professionals as a workforce is an increasingly recognized and understudied barrier.

Objective: The purpose of this study is to describe what courses on HIT topics are available at the graduate level for future health professionals in the European Union (EU) and to explore possible determining factors for their exposure to these courses.

Methods: A cross-sectional descriptive study of EU medical schools was performed to explore the prevalence of HIT courses. The curricula of all identified higher learning institutions that offer a medical degree were manually explored to identify graduate-level courses that offer specific training on HIT topics. HIT topics were defined as courses or subjects that provided knowledge on the design, development, use, and implementation of HIT. Associations among potential factors such as population, yearly medical graduates, total number of physicians, EHR presence, and gross domestic product (GDP) were explored.

Results: A total of 302 medical schools from the 28 member states of the EU were explored. Only about one-third (90/302, 29.80%) of all medical degree curricula offered any kind of HIT course at the graduate level; in the medical schools that offered HIT courses, the courses were often mandatory (58/90, 64.44%). In most EU countries, HIT courses are offered in less than half of the medical schools, regardless of the country's GDP per capita. Countries with the highest percentages of HIT course presence have the lowest GDP per capita. There seems to be a weak inverse correlation (–0.49) between the two variables (GDP per capita and HIT course presence). There is a trend between the availability of medical human resources and an increase in the presence of HIT courses, with Romania, Croatia, and Greece as outliers in this respect.

Conclusions: The current state of medical training in the EU leaves much room for improvement. Further studies are required for in-depth analysis of the content and manner of instruction that would fit present and future needs of HIT.

(J Med Internet Res 2019;21(8):e14086) doi:10.2196/14086

KEYWORDS

medical informatics; health information technologies; medical education; European Union

XSL•FO
RenderX

Introduction

Health information technologies (HITs) [1] are used in health care institutions to assist clinicians provide care with tools such as electronic health records (EHR) and telemedicine services [2]. The increasingly vast amounts of health data available provide health professionals access to new ways to collect, analyze, and use that information [3,4]. The use of EHRs for documentation allows data to be organized around diseases or quality indicators, machine learning and artificial intelligence enable population health analytics to identify predictive characteristics and factors for diseases, and remote tools facilitate disease management in point-of-care and home settings [5].

There are many barriers to HIT adoption, including mismatches between investments and benefits, disruptions in the workflow, and concerns about privacy and confidentiality [6-8]. However, while many health professionals and students see the potential benefit of these technologies in health care, many are also frustrated [9,10], struggling to adapt, without knowing the underlying science of information in these new tools. The lack of HIT training of health professionals as a workforce is an increasingly recognized and understudied barrier [4,11,12].

Focused and concerted educational training in health informatics is essential for health professionals to realize the full benefit of the data and tools that are already part of the practice of medicine and to help develop new and improved tools of the future. There are some initiatives worldwide to provide specialization training on medical informatics, such as the American Medical Informatics Association in the United States, with nearly 1700 board certified professionals [13], or England's efforts with the Topol review [14]; many other countries have not, or at least not sufficiently, established such opportunities until now [4]. A recent publication in the British Medical Journal proposes that literacy in informatics should be a formal requirement of all medical education, biomedical research, and public health training [15].

The current state of how future doctors are trained in HIT contents is largely unknown in many European Union (EU) countries. Further, although there may be opportunities for obtaining education in this field, most are targeted at the postgraduate level, leaving HIT training to be pursued as a professional interest and not a core skill. The purpose of this study is to describe what courses on HIT topics are available at the graduate level for future health professionals in the EU and to explore possible determining factors for their exposure to these courses.

Methods

Study Design

A cross-sectional descriptive study of the state of EU medical education was performed to explore the prevalence of specific HIT courses offered to future doctors. Cross-sectional studies are carried out at one time point or over a short period to estimate the prevalence of the outcome of interest for a given population [16].

The curricula of all identified higher learning institutions such as faculties of medicine, schools of medicine, or universities that offer the medical degree, hereon referred to as "medical schools," were explored to identify graduate-level courses that may offer would-be physicians training on any HIT topics. Potential factors such as population, yearly medical graduates, total number of physicians, EHR presence, and GDP were explored for associations.

Settings

The EU is a political and economic union of 28 member states, with an estimated population of over 513 million and almost 2 million practicing physicians [17]. During the first quarter of 2018, the medical degree programs from all EU member states were systematically explored and classified.

Data Sources

No official list of medical schools in EU exists; therefore, a preliminary list was obtained and refined from gray literature [18]. Multimedia Appendix 1 shows the definitive list of medical schools used.

Data on EU population statistics, GDP per capita, EHR presence, and medical graduates and doctors were extracted from their respective official sources [17,19,20]. Latest available information was used, and cases where no information was available were marked.

Selection Criteria

For the purpose of this study, and given the many different terms in use in the field, the terms medical informatics, health information and communications technology (ICT), eHealth, mHealth, biomedical and health informatics, consumer health informatics, digital health, and other variations were considered to be encapsulated by the term "HIT."

HIT topics were defined as courses or subjects that provided knowledge on the design, development, use, and implementation of HIT. The different course names and programs (where available) were revised for mentions of HIT contents. Courses that used ICT tools to enhance medical education on other subjects (eg, anatomy lessons or clinical case virtual simulations), or who only taught ICT content as a means to an end (eg, programming in R for biostatistics) were not considered to be focused on HIT training and were therefore not included.

A small random sample (10%) was independently reviewed to assess clarity of the selection criteria, and interrater reliability was calculated using the Fleiss-Cohen Coefficient. Once this was established, the remaining data were explored. Disagreements were resolved by consensus involving a third reviewer, when necessary.

Data Extraction and Classification

Medical education data were manually extracted from each medical school's website to obtain the latest publicly available curricula and program. A multilingual and multicultural team (GG, EG, and HB) conducted the data extraction, and any language barrier issue was resolved using the Google Translate feature. The team independently reviewed and classified the

XSL•FO
RenderX

extracted information using structured forms. In Multimedia Appendix 2, the raw data are presented.

Statistical Methods

Categorical variables are presented as absolute and relative frequencies. Quantitative variables are presented as mean and SD or median and interquartile range, depending on the distribution. The Landis and Koch standards for the Fleiss-Cohen coefficient were used [21]. Statistical analysis was performed using R (Vienna, Austria, R Foundation for Statistical Computing; 2013).

Results

Interrater reliability was determined using the Cohen kappa statistic by extracting a random sample (n=32) and independently reviewing the selection criteria. Kappa was found to be more than acceptable at 0.91 (SE 0.06, 95% CI 0.79-1.0).

Table 1. Summary of the presence of health information technology courses in each European Union member state.

Country	Total medical schools (N=302), n (%)	Medical schools with HIT[a] courses (n=90), n (%)	Medical schools with HIT courses, where the course is mandatory (n=58), n (%)	EHR[b] availability
Austria	6 (2.0)	1 (16.67)	1 (100)	Nationwide project in progress
Belgium	10 (3.3)	1 (10.00)	0 (0)	Nationwide project in progress
Bulgaria	6 (2.0)	2 (33.33)	1 (50)	Yes
Croatia	4 (1.3)	3 (75.00)	3 (100)	Nationwide project in progress
Cyprus	4 (1.3)	1 (25.00)	1 (100)	Nationwide project in progress
Czech Republic	9 (2.9)	3 (33.33)	2 (66.67)	Nationwide project in progress
Denmark	4 (1.3)	0 (0.00)	N/A[c]	Yes
Estonia	1 (0.3)	0 (0.00)	N/A	Yes
Finland	5 (1.7)	1 (20.00)	1 (100)	Yes
France	34 (11.26)	10 (29.41)	7 (70)	Nationwide project in progress
Germany	38 (12.6)	16 (42.11)	8 (50)	Nationwide project in progress
Greece	7 (2.3)	6 (85.71)	5 (83.33)	Nationwide project in progress
Hungary	4 (1.3)	1 (25.00)	1 (100)	Yes
Ireland	6 (2.0)	2 (33.33)	1 (50)	Nationwide project in progress
Italy	41 (13.6)	7 (17.07)	7 (100)	Nationwide project in progress
Latvia	2 (0.7)	0 (0)	N/A	Nationwide project in progress
Lithuania	2 (0.7)	0 (0)	N/A	Yes
Luxembourg	1 (0.3)	0 (0)	N/A	Nationwide project in progress
Malta	1 (0.3)	0 (0)	N/A	Yes
Netherlands	9 (3.0)	3 (33.33)	0 (0)	Nationwide project in progress
Poland	19 (6.3)	7 (36.84)	7 (100)	Nationwide project in progress
Portugal	7 (2.3)	2 (28.57)	1 (50)	Nationwide project in progress
Romania[d]	13 (4.3)	10 (76.92)	5 (50)	Nationwide project in progress
Slovakia[d]	3 (1.0)	0 (0)	0 (0)	Nationwide project in progress
Slovenia	2 (0.7)	2 (100)	1 (50)	Nationwide project in progress
Spain	25 (8.3)	8 (32)	4 (50)	Nationwide project in progress
Sweden	7 (2.3)	0 (0)	N/A	Yes
United Kingdom	32 (10.6)	2 (6.25)	2 (100)	Yes

[a]HIT: health information technology.

[b]EHR: electronic health record.

[c]N/A: not available.

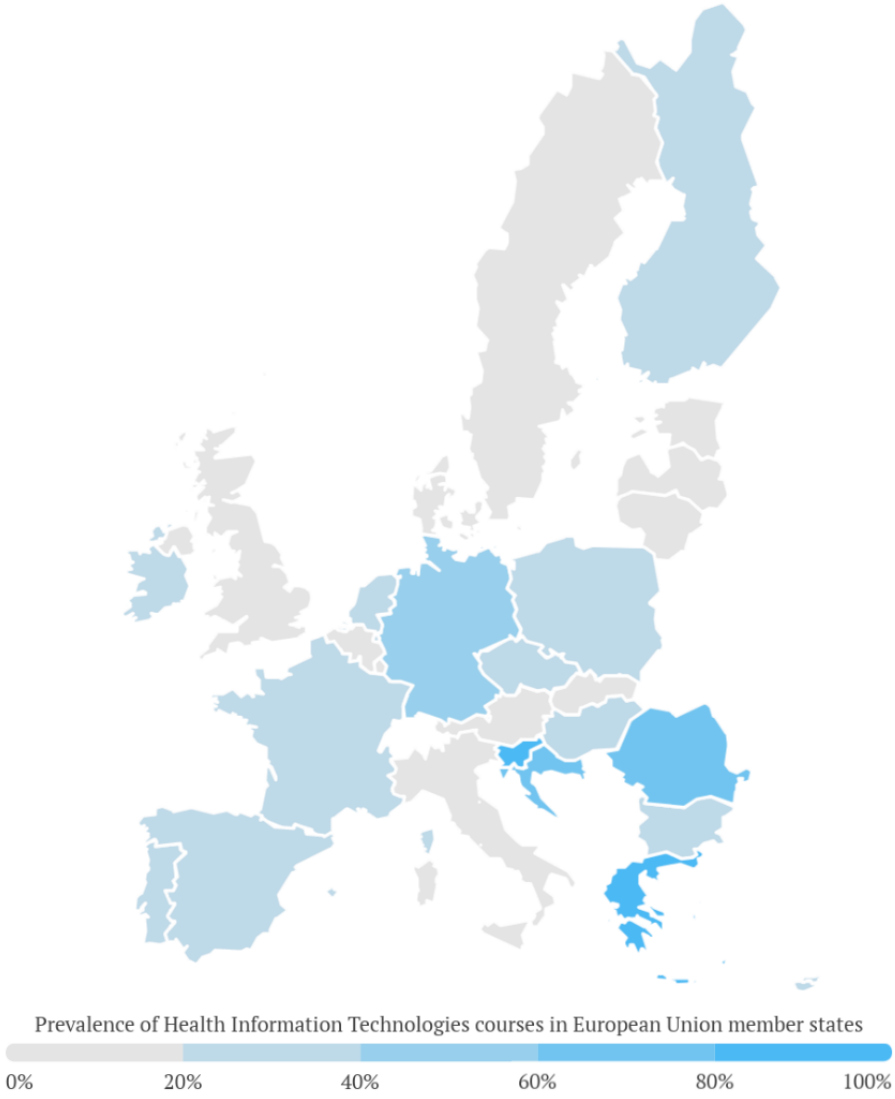
[d]Information on the medical school curricula in one school each in Romania and Slovakia was not available.

XSL•FO
RenderX

A total of 302 medical schools from the 28 member states of the EU were explored. Only one-third (90/302, 29.80%) of all medical degree curricula offered any kind of HIT course on the graduate level; in medical schools that offered HIT courses, the courses were often mandatory (58/90, 64.44%). The prevalence of HIT courses for medical degree students in EU member states is very low. As an indicator of system-wide informatization, we also compared our findings with reports and literature on each EU country's process of EHR implementation. The state of EHR implantation in the EU is a work in progress in most cases (only one-third have nationwide systems). Table 1 presents the information according to each country.

HIT courses offered in most medical schools were titled along the lines of "Medical informatics," "Telemedicine," "e-Health," or "Health informatics" and usually paired with statistics content. Some universities also provided courses on relatively advanced topics such as "AI in Medicine" (Medical University of Lublin, Poland), "Robotics programming with LEGO" (University of Duisburg-Essen, Germany), or workshops on "Application scenarios of Virtual Reality" (University of Ulm, Germany). Entrepreneurship was related to HIT in some cases, for example, in the Galway School of Medicine (National University of Ireland, Ireland), courses on "Becoming a Medical Innovator" are offered. Figure 1 shows a map with EU member states and the presence of HIT courses. A list of all the course names and their frequency can be found in Table 2.

Figure 1. Presence of health information technology courses in the European Union member states.



XSL•FO
RenderX

Table 2. List of all the names of health information technology courses and their frequencies.

Course name	Frequency
Medical informatics	22
Epidemiology, medical biometry and medical informatics	10
Medical informatics and biostatistics	9
Introduction to Medical Informatics	7
Information Technologies	5
Research and New Technologies	3
Biomedical Information Systems	2
Evaluation of methods of analysis applied to life and health sciences	2
Medical IT[a]	2
Advanced Medical Technology	1
Basics of informatics in the health sector	1
Basics of Medical Informatics	1
Becoming a Medical Innovator	1
Bioinformatics	1
Biomedical Research and New Technologies	1
Biometrics and Epidemiology	1
Clinical Informatics and Biostatistics	1
eHealth[b]	1
eHealth & Medical Informatics	1
Health informatics	1
Health Technology Assessment	1
Healthcare Imaging and Information Systems,	1
ICT[c] for Medicine	1
Informatics and Applications of Medical Informatics	1
IT Resources; Telemedicine	1
Lecture Epidemiology and Medical Informatics	1
Legal and Organisational Aspects of Medicine - includes Health IT	1
Medical and Scientific Methodology (includes Bioengineering and Medical Informatics)	1
Medical Applied Informatics	1
Medical Computer Science	1
Medical Informatics and Internet Computer Certificate	1
Medical Informatics, Biomedical Statistics, and scientific English	1
Medical informatics, e-Health[b] and medical statistics	1
Medicine and Technology	1
Modern Informatics in Biomedicine	1
New technologies in biomedicine	1
Robotics and Programming with Lego - An introductory Course to Robotics and Programming for Medical Students	1
Statistics and Bioinformatics	1
Tele-Health & Health Information Technologies In Public Health	1
Telemedicine and eHealth	1
Telemedicine: internet technologies for health	1

XSL•FO
RenderX

Course name	Frequency
Telemedicine: Possibilities and Limitations	1
The Industry Perspectives on Innovative Medicine Intensive Summer Course	1

[a]IT: information technology.

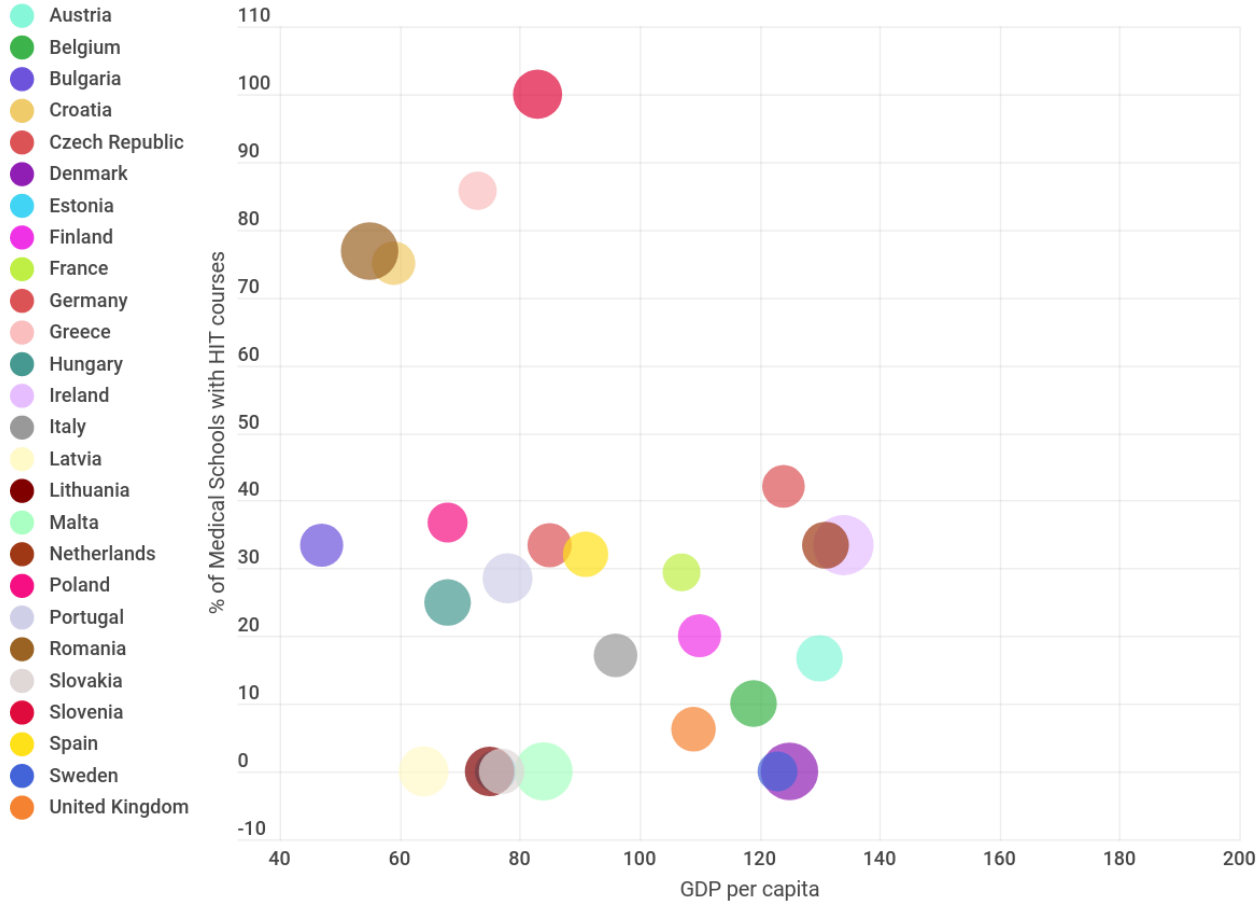
[b]eHealth/e-Health: electronic health.

[c]ICT: information and communication technology.

We also explored possible associations between the presence of HIT courses and EU population statistics, GDP per capita, EHR presence, and yearly new medical graduates per country. Most countries seem to offer HIT courses in less than half of their medical schools, regardless of their GDP per capita. Countries with the highest percentages of HIT course offerings are among the ones with the lowest GDP per capita in EU. There seems to be a weak inverse correlation (–0.49) between the two variables (GDP per capita and HIT course offer). In Figure 2, we show a scatter plot with the relationship between GDP per capita and percentage of medical schools with HIT courses. The exploration of population size, the main demographic variable, resulted in no relation between the variables.

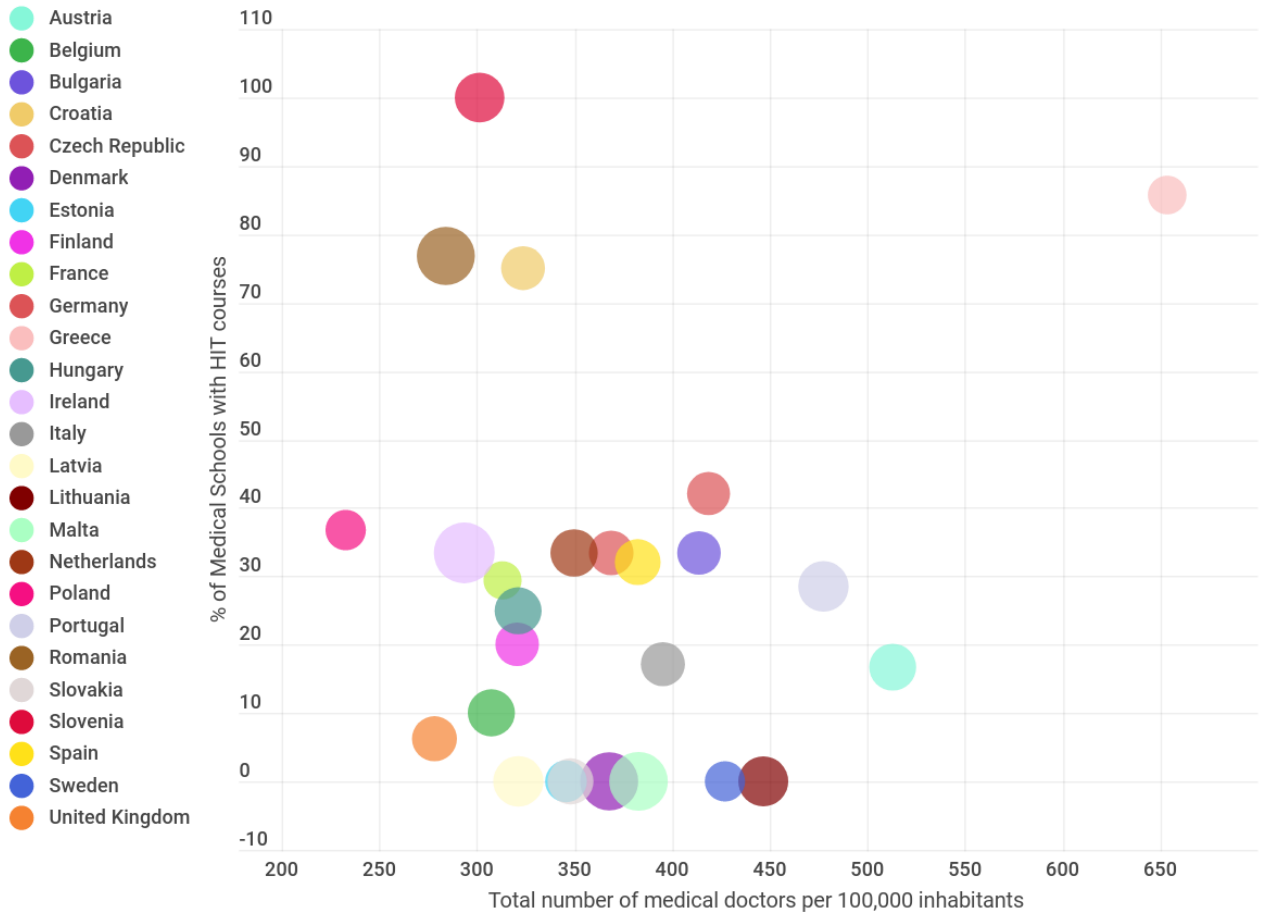
We explored the association of a country's population over the number of physicians and the presence of HIT course offerings. There seems to be a trend between the availability of medical human resources and an increase in the presence of HIT courses, with Romania, Croatia, and Greece as outliers (Figure 3).

Figure 2. Relationship between European Union member states' GDP per capita and presence of HIT courses. The surface represents the number of medical graduates per year over 100,000 inhabitants. Countries where no information on yearly medical graduates was available (Cyprus and Luxembourg) are not shown. GDP: gross domestic product; HIT: health information technology.



XSL•FO
RenderX

Figure 3. Relationship between the total number of physicians per European Union member state and the presence of HIT courses. Countries where information was missing (Cyprus and Luxembourg) are not shown. HIT: health information technology.



Discussion

Principal Findings

This work provides an overview of the state of HIT courses available for graduate-level medical students in the EU. It offers unique insight into the prevalence of HIT courses for each country and explores possible associated factors. Medical students in the EU seem to find it difficult to gain exposure to HIT courses, with only one-third of all medical schools offering any courses on these topics at all. The uneven distribution of training opportunities for medical students on HIT contents between the different countries highlights the need of further work on improving the medical curricula.

Comparison With Prior Work

Despite their promise, HITs have proven difficult to implement [22], and there seems to be a failure to achieve widespread understanding of the benefits of technologies such as electronic record keeping and information exchange [23]. Although technology is increasingly being used in health care settings, health care professionals remain reticent to adopting HIT at different levels [24].

The need to improve the HIT competences of the health care workforce has been frequently emphasized by policymakers at an international levels [25]. One of the projects aiming to identify health care workforce IT skill needs is the CAMEI (Coordination Actions in the scientific era of Medical Education Informatics) project, which is a collaboration between the United States and Europe [26]. Further, the importance of training for the health care workforce in the use of new technologies was also acknowledged in many studies [25,26]. Several authors have pointed out that the training and competences that health care professionals receive regarding ICTs, in general, and as HIT end users, in particular, act as key factors in HIT adoption [27]. The limited exposure that medical students in the EU receive on HIT contents can easily explain the resistance that HIT implementations face later in health services settings. Recommendations from the International Medical Informatics Association [4] suggest that education on HIT contents should be available to all types of health care professionals, regardless of the types of specialization and levels of education as a way of emphasizing the formal penetration of the discipline and its concepts.

A significant factor for the slow uptake of HIT for health care professionals also seems to be the lack of customization that software vendors offer for their solutions [28,29]. The concept of user-centered design places the needs and characteristics of end users at the center of software design [30-32]. However, involving stakeholders in the design and development of technologies is not yet common practice for HIT solutions.

XSL•FO
RenderX

The introduction of HIT should not be viewed as a problem in technology exclusively, but rather as a problem in organizational change [33]. The implementation of HIT in the different EU member states is very diverse, particularly regarding the development of EHRs [27], a central component of an integrated HIT [34]. This heterogeneity seems to also apply to how health professionals are trained at the graduate level with regard to HIT concepts, with countries ranging from no training at all offered to all medical schools requiring the approval of HIT courses to obtain an medical degree.

In addition, it is known that health care spending per capita is positively correlated with GDP per capita [35] and that the use of health information systems may be a method to control increasing health spending [35]. Although more data would be needed to conclude a definitive casual relation, an interesting finding of our study was a weak inverse relationship between the presence of HIT courses and a country's GDP. It would be interesting to ascertain what other factors and how the design of medical program curricula can impact a country's health care expenditure.

Many factors that determine how a degree's curriculum is structured, and while there is a need for change, political factors that prevent much needed modifications may exist. It is important to remember that wide-scale implementation of EHRs and other HIT systems has often been the result of state or national legislations [19]. Initiatives like these exist in many EU countries, but they are relatively recent and the change to curricula takes time. For example, digital competence has been acknowledged as one of the eight key competences for lifelong learning by the EU [36]. Under this framework, governments should encourage universities, particularly medical schools, to introduce formal teaching activities within undergraduate medical studies.

Limitations

The findings of this study should be interpreted in the context of its limitations. The main limitation of this study is that it focuses on courses that specifically teach HIT contents and does not account for how ICTs can be used in different teaching methods such as problem-based learning, or other subjects where HIT concepts may be taught indirectly. Additionally, this study did not perform an in-depth analysis of the medical degree curricula; therefore, it is possible that HIT lessons are included or integrated in other courses. However, given that medical education is usually divided in preclinical studies (eg, basic sciences such as anatomy, physiology, and biochemistry) and clinical studies (eg, various areas of clinical medicine such as internal medicine, pediatrics, obstetrics, and gynecology), it is unlikely that foundational HIT contents would be found within these modules.

Analysis and interpretation was performed using the retrieved information. Newer data could be available through internal channels and publications of each institution or country. Nevertheless, our study used publicly available data, which would also be available to other researchers.

In recent years, new bachelor and master degrees of an interdisciplinary nature have been developed (eg, biomedical engineering). Universities that hold these degrees may have contents or courses available for medical degree students to access; however, as these courses would not appear in the medical degree curricula, these were not explored or included. Further, this study only focuses on medical degree students and does not cover other health care professionals (eg, nurses or physiotherapists) or disciplines that could have some overlap (eg, biomedical engineering).

Finally, an interplay of varying complex factors determines how a degree program and curriculum are designed, both within and outside academic institutions such as local, national, and regional needs; legislations and ruling legal frameworks; and political leadership and visions. These were not included, as they would be too many and exceed the scope of this study.

Conclusions

Technology is transforming health care by slowly changing the way medicine is practiced. As HITs become more prevalent, expectations for health professionals to be fluent users of health technologies will continue to increase. The current state of medical training in the EU in this regard leaves much room for improvement. Further studies are required for in-depth analysis on the type of contents, proper placement in the medical curricula, and manner of instruction that would fit present and future needs of HIT.

Taking into consideration that the use of HIT is likely to continue to increase even further in most European health care systems, the need of future physicians to be literate in both medical sciences and HIT will soon no longer be optional. Empowering the following generations of doctors to allow them to take advantage of the full benefits that technologies have to offer is ever more important. A thorough look at the way we design our medical school curricula is needed.

Acknowledgments

GG and EG-F gratefully acknowledge the grant (number 676201) for the Connected Health Early-stage researcher Support System (CHESS ITN) from the Horizon 2020 Framework Programme of the European Commission. This article is based upon work from COST Action ENJECT TD 1405, supported by COST (European Cooperation in Science and Technology; www.cost.eu).

We would like to thank Luis Fernandez-Luque, PhD, Prof Minna Isomursu, Prof Brian Caulfield, Tim Sangster, Joaquin Chacon-Galvez, Analia Baum, and Diego Giunta for their cooperation and support.

XSL•FO
RenderX

Conflicts of Interest

None declared.

Multimedia Appendix 1
List of medical schools in the EU.
[XLSX File (Microsoft Excel File), 47KB - jmir_v21i8e14086_app1.xlsx]

Multimedia Appendix 2
Data on medical schools.
[XLSX File (Microsoft Excel File), 22KB - jmir_v21i8e14086_app2.xlsx]

References

1. Lorenzi NM. E-Health Strategies Worldwide. Yearb Med Inform 2018 Mar 06;14(01):157-164. [doi: 10.1055/s-0038-1638441]
2. 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 Mar;13(2):121-126 [FREE Full text] [doi: 10.1197/jamia.M2025] [Medline: 16357345]
3. Greenes RA. Medical Informatics. JAMA 1990 Feb 23;263(8):1114. [doi: 10.1001/jama.1990.03440080092030]
4. Mantas J, Ammenwerth E, Demiris G, Hasman A, Haux R, Hersh W, IMIA Recommendations on Education Task Force. Recommendations of the International Medical Informatics Association (IMIA) on Education in Biomedical and Health Informatics. First Revision. Methods Inf Med 2010 Jan 07;49(2):105-120. [doi: 10.3414/ME5119] [Medline: 20054502]
5. Adler-Milstein J, Cohen G. Implementing the IT infrastructure for health reform: adoption of health IT among patient-centered medical home practices. AMIA Annu Symp Proc 2013;2013:11-16 [FREE Full text] [Medline: 24551318]
6. Ball M, Silva J, Bierstock S, Douglas JV, Norcio AF, Chakraborty J, et al. Failure to provide clinicians useful IT systems: opportunities to leapfrog current technologies. Methods Inf Med 2008;47(1):4-7. [Medline: 18213422]
7. Hersh W. Health Care Information Technology. JAMA 2004 Nov 10;292(18):2273. [doi: 10.1001/jama.292.18.2273]
8. Poon EG, Blumenthal D, Jaggi T, Honour MM, Bates DW, Kaushal R. Overcoming barriers to adopting and implementing computerized physician order entry systems in U.S. hospitals. Health Aff (Millwood) 2004;23(4):184-190 [FREE Full text] [Medline: 15318579]
9. Payne T, Corley S, Cullen T, Gandhi TK, Harrington L, Kuperman GJ, et al. Report of the AMIA EHR-2020 Task Force on the status and future direction of EHRs. J Am Med Inform Assoc 2015 Sep;22(5):1102-1110 [FREE Full text] [doi: 10.1093/jamia/ocv066] [Medline: 26024883]
10. Kern J, Fister K. Medical Informatics Meets Medical Education: the Croatian Experience. Eur J Biomed Informatics 2011;7(2):58-63 [FREE Full text] [doi: 10.24105/ejbi.2011.07.2.8]
11. Hersh W. Who are the informaticians? What we know and should know. J Am Med Inform Assoc 2006;13(2):166-170 [FREE Full text] [doi: 10.1197/jamia.M1912] [Medline: 16357354]
12. Hersh W. Health and biomedical informatics: opportunities and challenges for a twenty-first century profession and its education. Yearb Med Inform 2008:157-164. [Medline: 18660890]
13. Detmer D, Lumpkin J, Williamson J. Defining the medical subspecialty of clinical informatics. J Am Med Inform Assoc 2009;16(2):167-168 [FREE Full text] [doi: 10.1197/jamia.M3094] [Medline: 19074293]
14. NHS. 2018. Preparing the healthcare workforce to deliver the digital future URL: https://topol.hee.nhs.uk/wp-content/uploads/HEE-Topol-Review-2019.pdf [accessed 2019-08-01]
15. Fridsma D. Health informatics: a required skill for 21st century clinicians. BMJ 2018 Jul 12;362:k3043. [doi: 10.1136/bmj.k3043] [Medline: 30002063]
16. Thelle D, Laake P. Research in medical and biological sciences. Cambridge, MA: Academic Press; 2015:275-320.
17. European Commission. Eurostat - statistics explained. Healthcare Personnel Statistics - Physicians URL: http://ec.europa.eu/eurostat/statistics-explained/index.php/Healthcare_personnel_statistics_-_physicians#Healthcare_personnel [accessed 2019-08-01]
18. Wikipedia: The Free Encyclopedia. List of medical schools in Europe URL: https://en.wikipedia.org/w/index.php?title=List_of_medical_schools_in_Europe&oldid=868758984 [accessed 2019-08-01]
19. European Comission. 2013. Overview of the National Laws on Electronic Health Records in the EU Member States and Their Interaction with the Provision of Cross-Border EHealth Services URL: https://ec.europa.eu/health/ehealth/projects/nationallaws_electronichealthrecords_en [accessed 2019-08-01]
20. World Health Organization. 2014 Dec. From Innovation to Implementation: eHealth in the WHO European Region URL: http://www.euro.who.int/__data/assets/pdf_file/0012/302331/From-Innovation-to-Implementation-eHealth-Report-EU.pdf [accessed 2019-08-01]
21. Landis JR, Koch GG. The Measurement of Observer Agreement for Categorical Data. Biometrics 1977 Mar;33(1):159. [doi: 10.2307/2529310]

XSL•FO
RenderX

22. Taylor R, Bower A, Girosi F, Bigelow J, Fonkych K, Hillestad R. Promoting Health Information Technology: Is There A Case For More-Aggressive Government Action? Health Affairs 2005 Sep;24(5):1234-1245. [doi: 10.1377/hlthaff.24.5.1234]
23. Achieving Efficiency Improvements in the Health Sector through the Implementation of Information and Communication Technologies. 2010. EHealthNews.eu URL: https://tinyurl.com/y4zybwda [accessed 2019-08-01]
24. Gagnon M, Légaré F, Labrecque M, Frémont P, Pluye P, Gagnon J, et al. Interventions for promoting information and communication technologies adoption in healthcare professionals. Cochrane Database Syst Rev 2009 Jan 21(1):1-100 [FREE Full text] [doi: 10.1002/14651858.CD006093.pub2] [Medline: 19160265]
25. Li S, Bamidis PD, Konstantinidis ST, Traver V, Car J, Zary N. Setting priorities for EU healthcare workforce IT skills competence improvement. Health Informatics J 2017 Apr 25;25(1):174-185. [doi: 10.1177/1460458217704257]
26. Traver V, Konstantinidis S, Bamidis P, Zary N. Analysis of EU-USA cooperation opportunities on IT skills for healthcare workforce. Stud Health Technol Inform 2015;210:561-563. [Medline: 25991210]
27. Lluch M. Healthcare professionals' organisational barriers to health information technologies-a literature review. Int J Med Inform 2011 Dec;80(12):849-862. [doi: 10.1016/j.ijmedinf.2011.09.005] [Medline: 22000677]
28. Westbrook J, Braithwaite J, Iedema R, Coiera E. Evaluating the impact of information communication technologies on complex organizational systems: a multi-disciplinary, multi-method framework. Stud Health Technol Inform 2004;107(Pt 2):1323-1327. [Medline: 15361029]
29. Ford EW, Menachemi N, Phillips MT. Predicting the Adoption of Electronic Health Records by Physicians: When Will Health Care be Paperless? Journal of the American Medical Informatics Association 2006 Jan 01;13(1):106-112. [doi: 10.1197/jamia.M1913]
30. International Organization for Standardization. 2010. Ergonomics of human-system interaction -- Part 210: Human-centred design for interactive systems URL: https://www.iso.org/standard/52075.html [accessed 2019-08-01]
31. De Vito Dabbs A, Myers BA, Mc Curry KR, Dunbar-Jacob J, Hawkins RP, Begey A, et al. User-centered design and interactive health technologies for patients. Comput Inform Nurs 2009;27(3):175-183 [FREE Full text] [doi: 10.1097/NCN.0b013e31819f7c7c] [Medline: 19411947]
32. Pruitt J, Adlin T. The Persona Lifecycle: Keeping People In Mind Throughout Product Design (interactive Technologies). NYC: Morgan Kaufmann; 2019.
33. Wears RL, Berg M. Computer Technology and Clinical Work. JAMA 2005 Mar 09;293(10):1261. [doi: 10.1001/jama.293.10.1261]
34. Black AD, Car J, Pagliari C, Anandan C, Cresswell K, Bokun T, et al. The impact of eHealth on the quality and safety of health care: a systematic overview. PLoS Med 2011 Jan 18;8(1):e1000387 [FREE Full text] [doi: 10.1371/journal.pmed.1000387] [Medline: 21267058]
35. Anderson GF, Frogner BK, Johns RA, Reinhardt UE. Health Care Spending And Use Of Information Technology In OECD Countries. Health Affairs 2006 May;25(3):819-831. [doi: 10.1377/hlthaff.25.3.819]
36. Carretero S, Vuorikari R, Punie Y. European Commission. 2017. DigComp 2.1: The Digital Competence Framework for Citizens with eight proficiency levels and examples of use URL: http://publications.jrc.ec.europa.eu/repository/bitstream/JRC106281/web-digcomp2.1pdf_(online).pdf [accessed 2019-08-01]

Abbreviations

CAMEI: Coordination Actions in the scientific era of Medical Education Informatics
eHealth: electronic health
EHR: electronic health record
EU: European Union
GDP: gross domestic product
HIT: health information technology
ICT: information and communication technology
IT: information technology

Edited by B Caulfield; submitted 21.03.19; peer-reviewed by V Traver Salcedo, WH Ho, Y Acharya, C Perrin; comments to author 16.04.19; revised version received 05.06.19; accepted 19.06.19; published 12.08.19.

Please cite as:
Giunti G, Guisado-Fernandez E, Belani H, Lacalle-Remigio JR
Mapping the Access of Future Doctors to Health Information Technologies Training in the European Union: Cross-Sectional Descriptive Study
J Med Internet Res 2019;21(8):e14086
URL: http://www.jmir.org/2019/8/e14086/
doi:10.2196/14086
PMID:31407668

XSL•FO
RenderX

©Guido Giunti, Estefania Guisado-Fernandez, Hrvoje Belani, Juan R Lacalle-Remigio. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 12.08.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

The Attitudes of Therapists and Physicians on the Use of Sex Robots in Sexual Therapy: Online Survey and Interview Study

Christiane Eichenberg[1], PhD; Marwa Khamis[1], BSc; Lisa Hübner[1], MSc

Faculty of Medicine, Sigmund Freud PrivatUniversität, Vienna, Austria

Corresponding Author:
Christiane Eichenberg, PhD
Faculty of Medicine
Sigmund Freud PrivatUniversität
Freudplatz 3A
Vienna, 1020
Austria
Phone: 43 1 90 500 70
Email: christiane@rz-online.de

Abstract

Background: Various types of robots have already been successfully used in medical care, and the use of new technologies is also playing an increasing role in the area of sexuality. Sex robots are marketed as advanced sex toys and sex dolls with artificial intelligence. Only a few considerations about the therapeutic use of sex robots in sexual therapy are debated in expert discussions.

Objective: The aim of this study was to conduct a first exploratory survey on the attitudes of sex therapists and physicians toward the therapeutic benefits of sex robots.

Methods: This study comprised a quantitative online survey and a qualitative interview study. A self-constructed questionnaire was used to survey the general attitudes of sex therapists and physicians regarding the benefits of sex robots in therapy. The qualitative study was designed to gain in-depth insight into the participants' beliefs and attitudes. Therefore, semistructured interviews were conducted. The quantitative data were evaluated by statistical analysis, and the interviews were transcribed and analyzed by using a grounded theory approach.

Results: A total of 72 sex therapists and physicians completed our self-constructed questionnaire (response rate 15%, 72/480). Only a few respondents (11%, 8/72) said that the use of sex robots was not conceivable for them, and almost half of all therapists and physicians could imagine recommending sex robots in therapy (45%, 33/72). The attitude toward sex robots as a therapeutic tool was very heterogeneous, with gender (P=.006), age (P=.03), and occupational differences (P=.05); female therapists, older therapists, and psychologists (in contrast to physicians) were more critical toward the therapeutic use of sex robots. The analysis of the 5 interviews identified 3 high-level core themes that were representative of the participants' responses: (1) the importance of the personal definition of sex robots for the assessment of their therapeutic benefits, (2) therapeutic benefits and dangers of sex robots, and (3) considerations on the quality of human-robot sexuality. Initial insights into the possible therapeutic use of sex robots in different disorders (eg, sexual dysfunction or pedophilia) and situations were gained from the perspective of sex therapists.

Conclusions: The results of this study provide a first overview of the potential therapeutic use of sex robots. Moral, ethical, and treatment-related issues in this context are still unresolved and need to be further researched. We suggest integrating the topic into the training of sex therapists to form opinions beyond media images and to show therapy possibilities. Scientists engaged in sexual research should be involved in the development of sex robots to design robots with positive effects on sexual education, sexual therapy, sexual counseling, and sexual well-being for interested groups.

(J Med Internet Res 2019;21(8):e13853) doi:10.2196/13853

KEYWORDS

robotics; sexual health; therapy

XSL•FO
RenderX

Introduction

Background

Robotics is an interdisciplinary field of research and practice, which is also relevant to sexuality because of the possibilities offered by human-machine interactions [1]. In the case of human-robot interaction, the emerging role of sex robots has piqued public interest. In the therapeutic debate about sex robots, it is important how psychologists and therapists define robot sex. The term technosexuality describes sexual activities that are combined with technology [2]. There are technosexual behaviors, such as internet pornography, that occur more frequently than others [3]. In the context of psychology, sexual activities with robots have mostly been described as objectophilia or robotic fetishism so far, whereby this definition is a rather pathologizing limitation. This is defined as a fetish attraction to humanoid or nonhumanoid robots, to people behaving like robots, or to people dressed in robot costumes [4]. In the most general and descriptive definition of robot sex, it can be defined as sexual use of the robot. This definition includes the use of special sex robots and the sexual use of other types of robots that are not specifically developed and marketed for sexual purposes [5].

Sex robots have triggered discussions in professional circles about robot design, social norms, and the status of human-robot sex in connection with human relationships as well as the possible benefits of sex robots [6]. For example, 1 benefit is in terms of using sex robots as a therapeutic tool in the treatment of sexual disorders. Different types of assistance and therapy robots have already been used for health care applications [7], for example, by people who are suffering from a stroke [8], dementia [9], autism [10], or physical disability [11]. A robotic assistant for health care applications can support users with training and rehabilitation programs that enable independent living. Although sex toys are used in sex therapy for the treatment of orgasm problems [12], there is no information about the opinion of sex therapists regarding sex robots as a tool in sex therapy. A robot designed for sex may have a different impact than other sex aids. Kerner [13] anecdotally reported that some sex therapists already have suggested a range of options that robots could help them with, including a variety of problems such as erectile dysfunction, ejaculatio praecox and social anxiety about having their first sexual encounter. In his book *Love and Sex with Robots*, David Levy [14] expressed a similar view when it comes to the potential psychosocial value of sex robots: “Many who would otherwise have become social misfits, social outcasts, or even worse will instead be better-balanced human beings.” Levy suggests that robotic sexual assistance contributes to health and well-being if it mitigates the exclusion of solo and partner sexuality associated with the impairment. Döring [15] adds that it would also be possible to use educational and therapeutic sex robots that allow certain exercise programs discreetly and without feelings of shame or guilt (eg, practice of safer sex techniques, treatment of orgasm disorders, and prevention of sexual assault). The statements quoted above are based on first considerations, but no therapists have been asked about their attitudes toward the therapeutic benefits of sex robots yet. First, we describe what types of sex robots are already available. Second, we discuss the existing controversy about sex robots, and finally, we summarize the current state of research.

Types of Sex Robots

The success of sexual gratification dolls that can be defined as material representations of the human body for sexual use paved the way for the design of robots in the field of sexuality [16]. Although sex robots are still at a very early stage of development, the sex industry is already offering a variety of products that include some kind of artificial intelligence software. Sex robots already exist in female, male, and transgender versions with corresponding primary and secondary genitalia. Current sex robots, as well as sex dolls, are made of silicone rubber and advertised by manufacturers as *warm to the touch*. The appearance, such as eye color, hair, skin, and makeup, can be determined by the customer himself or herself. Prices for the sex robots range from about $5000 to $15,000. The sex robots marketed so far look like sex dolls but are capable of having conversations and display certain preprogrammed emotions or personalities. Some robots are equipped with allover body sensors so that they can react to touch. The response is sometimes dependent upon the chosen personality trait of the sex robot. For example, Roxxxy Gold (TrueCompanion) has preprogramed personalities, such as *Frigid Farrah* that gives the impression of reserved shyness. However, *Wild Wendy* is programmed as an outgoing and adventurous personality. Some sex robots offer a number of mobility features. Suzie Software and Harry Harddrive by Sex Bot Company must be manually maneuvered into a sexual position and are then able to simulate sexual movements. Roxxxy Gold is advertised by the manufacturer as being capable of displaying orgasms, although it is not clear whether this is through motion, sound, or both. Harmony (Realbotix) is also advertised as having the ability to simulate orgasm. The developers advertise the sex robot with the following capabilities: it has neck articulation, facial expression, moving eyes, and the ability to lip sync with spoken audio [17].

Robot sex also involves software robots without a materially embodied counterpart, in which the counterpart is represented virtually in an immersive virtual reality (VR) application. Immersive VR applications can also be combined with teledildonics. These are sex toys that offer haptic stimulation of male or female genitalia, synchronized with VR application [18].

Controversy

In the international literature, the controversy about robot sex began 10 years ago, largely triggered by David Levy's monograph *Love and Sex with Robots* [14]. Since then, the topic sex robotics remains controversial; advocates interpret the development of sex robots as the next step in human-robot interaction and argue that robot sex will be a way to promote more openness in the context of sexuality. Sex robotics can be a health-promoting supplement and an expansion of partner sexuality. Another argument that emphasizes the benefits of sex robots is that people who cannot or do not want to commit (to a serious relationship) could benefit from it. The emotional attachment to a sex robot might be a helpful substitute that

XSL•FO
RenderX

reduces loneliness and promotes well-being. The opponents' side sees risks or, for example, conducts an ethical debate on whether human sexuality can be alienated in this way. A potential risk that critics often notice is that sex with robots could or will lead to social isolation [19]. The reasons vary as follows: spending time in a relationship with a robot could create an inability to form human friendships; real sexual relationships could become overwhelming as relations with robots are easier; robots do not meet the species-specific needs of humans; and sex robots could desensitize people to intimacy and empathy, which can only be developed by experiencing human interaction and consensual relationships.

Döring [5] explains that the polarization of the debate into pro and contrapositions (hype vs moral panic, utopian vs dystopian visions) is quite typical for dealing with technical innovations as there is no experience. Another problem with the public perception of sex robots is that the public is currently not well informed about robots in general. Sex robots are new, and information about them stem mainly from science fiction movies or books. Many of these science fiction stories focus on intimacy with robots, where robots (mostly female) are portrayed as sexual objects [20]. In a recent study, the media representations of intimate human-robot relationships were analyzed with the result that they reveal stereotypical gender roles, heteronormativity, and a focus on sexual versus emotional intimacy [21].

Theoretical Framework

This study is based on the uses and gratifications approach. Studies following this approach investigate how recipients actively deal with media [22]. The aim of this concept is to find out the motives and utilization scenarios of the media usage preferences of the consumers. The users decide on the basis of their interests (content, formats, and aesthetics) and needs (eg, escape from reality, information, and entertainment) whether and what kind of media they prefer to use. The use of media thus depends on their expectations and the satisfaction of the needs the media offer [23]. The uses and gratifications approach is less a theory than a research strategy. Further developments such as the expectancy-value theory help to give this approach a stronger theoretical foundation. The expectancy-value theory deals with the relationship between attitude and behavior. It provides an explanation of how the subjective attitude toward an object affects the probability of performing certain actions. According to Fishbein's model [24], it is believed that an attitude (A) toward an object (O) can be expressed in a function of beliefs (B) toward this object and the evaluations (E) of these expectations. According to this theory, there is a relationship between attitudes and the resulting behavior. Applied to our research project, this means that the expectation that sex robots possess a certain attribute, such as therapeutic potential, influences their use or recommendation. As sex robots is a new phenomenon, which is still in the development phase and experience reports have been missing so far, we wanted to survey the motives, possible use scenarios, and expectations of the therapeutic potential of therapists and physicians. This is of particular importance in the early phase of the development of sex robots so as to be able to contribute to this development from the perspective of the therapist.

Current State of Research

A few studies about the acceptance of and attitudes toward sex robots already exist, whereby the study participants were not therapists or physicians, but the sample comprised people from the general population.

In 2016, Scheutz and Arnold [25] surveyed a sample of 100 US internet users (43% women, average age of 33 years) by using online questionnaires in the United States. They reported the first findings about the appropriateness and evaluation of sex robots with the result that many types of sex robot use were approved by the survey participants (eg, sex robots as an alternative to prostitution, sex robots for people with disabilities, sex robots to prevent violence), whereas only a few options or possibilities were rejected (eg, child sex robots). In their survey, subjects viewed sex with a sex robot as a kind of masturbation or using a vibrator than having sex with a human. In contrast to the agreement of the study participants on the definition of sex robots, the authors found relevant gender differences; women consistently rated each respective use and possible robotic form as less appropriate than men and were much less likely to consider a sex robot use in the future. A further survey study on attitudes toward sex robots among 203 German internet users (70% women, average age of 31 years) showed that 82.3% of respondents supported the use of sex robots, especially in cases of physical disabilities, instead of prostitution and as a way to live out certain sexual fantasies. Over 80% of the respondents could imagine the use of sex robots to treat a sexual problem (eg, ejaculatio praecox) and over 55% could imagine using sex robots in a therapeutic context [26].

A study by Szczuka and Krämer [27] compared men's explicit and implicit evaluation of the (sexual) attractiveness of sex robots and women. The sample comprised 229 heterosexual men. It could be shown that social contacts, loneliness, fear of rejection, relationship status, and satisfaction with sexual life were not related to the evaluation of attractiveness. Instead, the authors found that the negative attitude toward robots was the main user characteristic that predicted the attractiveness ratings of sex robots. The study by Richards et al [28] with 133 participants also concluded that participants who generally viewed robots as negative, assessed the probability of having sex with a robot in the future as low.

The acceptance of sex robots among the general population is high. The same applies to the assessment of the potential benefits of sex robots for the treatment of sexual problems. As women and people with a generally negative attitude toward robots rate sex robots more negatively, it can be assumed that these groups of individuals may also have a lower acceptance of sex robots in the therapeutic context. The results lead us to the question whether sex therapists evaluate the use of sex robots similarly to the general population. The Foundation for Responsible Robotics [17] hypothesized that it is possible that the use of sex robots in some therapies could potentially help some people with sexual healing (eg, problems with sexual functioning or social anxiety). The lack of research and the simultaneous emerging interest in sex robots led us to the following questions, which we investigated in this study: What attitudes do sex therapists have toward the use of sex robots in the therapy of

XSL•FO
RenderX

various sexual disorders? Do sex therapists consider sex robots as a therapeutic tool? Which patients could benefit from sex robots from the point of view of therapists? Do the therapists differ in their opinion on sex, age, education, personality, and affinity to technology?

Methods

Study Design

This was an exploratory pilot study to examine the attitudes toward and acceptance of sex robots in sex therapy by sex therapists and physicians in Germany, Austria, and Switzerland. A questionnaire was used for the quantitative survey, followed by a qualitative interview study with semistructured interviews to deepen the results.

Ethics Approval

The study protocol was approved by the Ethics Committee of the Sigmund Freud University Vienna (approval number—electronic ID: LAWW6CYK@VV5BX86374).

Survey

The quantitative data were collected with 3 questionnaires. The therapeutic acceptance of sex robots and the conditions under which the use of sex robots in sexual therapy appears acceptable were determined by using a self-developed questionnaire. This questionnaire was combined with 2 standardized questionnaires: The Questionnaire on technical affinity-attitude towards and handling of electronic devices (TA-EG) [29] to assess the affinity to technology and the NEO Five Factor Inventory [30] to gather personality traits of the sample and to test whether these factors had an influence on attitudes. The development of our questionnaire was based on the findings from the few existing studies on the acceptance of sex robots in the general population as described above. The self-developed questionnaire comprised 25 items and was divided into 3 parts. The first part included questions referring to participants' characteristics such as age, employment, and education. In the second part, we provided the subjects with a short introductory text involving the definition of sex robots. As sex robots are new developments and the research field is so young, it can be assumed that the therapists themselves have little experience with robots. The participants were first asked about their knowledge of sex robots and subsequently about their personal attitudes. For example, we asked what use of sex robots would generally be conceivable and how therapists would categorize sexual activities with robots. With the questions in the third part, we wanted to assess the therapists' attitudes toward the therapeutic potential of sex robots. A total of 2 standardized questionnaires were used to identify further possible factors that could influence the acceptance of sex robots. The self-constructed questionnaire was pretested with 5 participants. The few remarks were analyzed, and the instrument was revised regarding its practicability, comprehensibility, and completeness of item formulation.

With the TA-EG we wanted to learn more about the experiences and attitudes of the participants regarding technical devices. The subscales of TA-EG are enthusiasm for technology, subjective competence with technology, positive consequences of technology, and negative consequences of technology.

In the third and last part of the questionnaire, the NEO-FFI was used to collect personality traits to find out if they influence the attitudes of the participants and to learn more about the target group that is open to the therapeutic use of sex robots. The participants received 60 short statements describing themselves and were asked to evaluate the statements according to whether they applied to them or not. NEO-FFI factors include neuroticism, extraversion, openness to experience, tolerability, and conscientiousness.

Interviews

Semistructured interviews were chosen for this purpose as they are considered as a valid and consistent method of data collection in qualitative research [31]. On the basis of the quantitative results, the interview guideline for the qualitative study was developed with the aim to further investigate open or controversially discussed aspects and determine possible therapeutic fields of implementation. As the quantitative survey showed that not all therapists had already heard of sex robots, the introductory question in the interview comprised what the interviewee knew about sex robots and where he or she obtained this knowledge from. A controversial result of the quantitative survey was the consideration of whether sex robots would be useful for the treatment of pedophile patients. On the basis of this result, we asked in the interview what the therapists' attitudes toward this consideration was. Another result of the first survey that we wanted to address in the interviews was the tension between the conceivable use of sex robots in therapy on the one hand and ethical concerns on the other. In the interviews, we asked how these contradictions could be resolved. The complete interview guide can be found in Multimedia Appendix 1. The interviews followed a general-to-specific approach. Interviews were piloted before the use to make some adaptions if necessary.

Sampling and Recruitment

For the quantitative survey, sex therapists and physicians were recruited through 4 professional associations: Institute for Sexual Therapy; German Society for Sexual Medicine, Sexual Therapy, and Sexual Science; Swiss Society of Sexology; and Austrian Society for Sexual Sciences. Cover letters were used to inform sex therapists and physicians about the survey, the implementation, the purpose of our investigation, and the exploitation of the results.

The theoretical sampling for the qualitative study was determined by the first results obtained from the quantitative data, which showed that sex therapists and physicians differ in their attitudes toward sex robots in gender, age, and education. Participants were sampled through Google searches. To obtain a broad spectrum of opinions in the interviews, female and male therapists with different ages and professional backgrounds were searched for. The information regarding age and education was found on the therapists' homepage.

XSL•FO
RenderX

Data Collection

The data for our online survey was collected using the Unipark software, which complies with all data protection regulations. The data collection took 4 weeks. A total of 480 therapists and physicians were contacted by email, which embedded a link to the online questionnaire. Of these, we received a total of 72 complete survey responses (response rate 15%). For the interviews, a total of 50 female and male therapists with medical, psychological, or social educational backgrounds of different ages were contacted by email. In total, 5 interviews were conducted by telephone and were digitally recorded. These ranged from 17 to 49 min, depending on the schedule of the participant and the number of issues they wanted to discuss.

Coding and Data Analysis

Statistical analyses were conducted using SPSS 18. Descriptive statistics were used to evaluate most of the items. Statistical correlations were calculated using appropriate statistical test procedures (cross table, Spearman correlation, Wilcoxon-Mann-Whitney test, and binary logistic regression analysis). The open questions of the questionnaire were analyzed by content analysis. All interviews were transcribed verbatim. The transcripts were evaluated according to grounded theory analysis using Microsoft Office Word. The grounded theory is, among other methods, one that is suitable for research that seeks to discover something new [32]. Our choice of using grounded theory lies in its capacity to discover participants' main concerns to consolidate the results of the quantitative survey. The analysis begins with the process of open coding. Open codes are, for example, certain words that occur recurrently in the data. The open codes are used to search for differences, similarities, behavioral patterns, etc, with the aim of forming categories [33]. The data were read and reviewed, and then a coding framework was developed to summarize the grouping of emerging issues from the data into core issues. The subsequent axial coding was about identifying the core phenomenon, causal conditions, resulting strategies (ie, related actions and interactions), context, and consequences. The collected data from each interview were further compared with each other and with the data from the other interviews to find similarities, repetitions, or differences in the emerging issues. Additional grouping codes were added as new topics emerged during the comparison process. Finally, all findings were extensively rereviewed by all authors to validate the data and to gain a high-level understanding of the collected information that would help to identify potential participant attitudes. The data collected in core themes are most relevant for the purposes of this study and are presented in the results.

Sample

Online Survey

The participants of the quantitative survey were all members of sexual associations in Germany, Austria, and Switzerland. A total of 72 sex therapists and physicians completed the self-constructed questionnaire. Table 1 summarizes the description of the sample according to the variables: gender, age, relationship status, education, and information on therapeutic work.

Interview Study

All interviewed therapists were very well informed about sex robots, had technical knowledge, attended advanced training courses on the subject, and were familiar with the various application areas of sex robots. In total, 3 women and 2 men were interviewed for the qualitative study. Table 2 below gives information about the participants.

XSL•FO
RenderX

Table 1. Sample description of quantitative study (N=72).

Variables	Statistics
Gender, n (%)	
Male	27 (38)
Female	45 (62)
Age (years)	
Range	32-80
Mean (SD)	51 (9.6)
Relationship status, n (%)	
Married	38 (52)
Long-term relationship	19 (26)
Divorced	7 (9)
Single	5 (6)
Other relationship	3 (4)
Main training (multiple answers were possible), n (%)	
Sexual therapeutic education	64 (89)
Sexual therapists	53 (83)
Sexual physicians	11 (17)
Psychotherapists	25 (35)
Psychologists	15 (21)
Physicians	15 (21)
Psychotherapeutic method (multiple answers were possible), n (%)	
Psychodynamic therapy	17 (68)
Systemic therapy	10 (40)
Cognitive behavioral therapy	8 (38)
Gestalt therapy	7 (28)
Therapeutic settings (multiple answers were possible), n (%)	
Individual therapy	55 (87)
Couple therapy	50 (79)
Outpatient setting	42 (66)
Group therapy	10 (15)
Mainly treated disorders (multiple answers were possible), n (%)	
Orgasmic disorders	56 (78)
Erectile dysfunction	51 (71)
Mental disorders	28 (40)

Table 2. Sample description of interview study.

Person	Gender	Age (years)	Profession
1	Male	71	Sex therapist, psychologist
2	Female	54	Sex therapist, physician
3	Female	50	Sex therapist
4	Male	53	Sex therapist
5	Female	77	Psychotherapist, psychologists, active in the academic field of sexual sciences

XSL•FO
RenderX

Results

Online Survey

Previous Knowledge of Sex Robots

To find out how well known the topic of sex robots is among therapists and physicians, we asked them whether they had already heard or read about sex robots or seen something about them. A total of 56 participants of the sample (77%, 56/72) had already heard of sex robots. The majority received their information via the internet (51%, 29/56). Approximately one-third (32%, 18/56) said they had heard about them at a further training course and another 30% (17/56) reported to have read about sex robots in a scientific journal.

General Attitudes Toward Sex Robots

The question "How positive do you generally rate the existence of sex robots?" was analyzed on the basis of frequencies. Participants were allowed to indicate their percentage score based on an interval from 0% to 100%. The mean was 32% (SD 29.27). Accordingly, the existence of sex robots was not rated as very positive. The majority of respondents believed that sex with a robot could not replace sex with a human being (90%, 65/72). More than half of the participants would define sex with a robot as masturbation (58%, 42/72).

The participants of this study were asked which use of sex robots would be imaginable for them. To be able to assess this, we provided various situations, motives, and robot use patterns, whereby the interviewees were able to indicate which use of sex robots would be conceivable. The responses (Table 3) showed that the sample (N=72) had different attitudes toward the general use of sex robots. Only 8 out of 72 respondents (11%, 8/72) stated that the use of sex robots was not conceivable for them. The majority of participants stated that they could imagine the use of sex robots for physically disabled people (65%, 47/72) and for living out certain sexual fantasies (61%, 44/72). However, the idea that sex robots can help to experience a trusting sexual relationship got the least approval by survey participants (5%, 4/72). In the open response category, 2 persons indicated that any use of sex robots would be imaginable, and 1 person suggested the use of sex robots in connection with sexuality for older people.

Potential of Sex Robots as a Therapeutic Tool

One part of the questionnaire related to the potential of sex robots as a tool in sex therapy as well as to the idea of recommending sex robots in the role of a practitioner (eg, "If you think about your practice and your experience: which use of sex robots would be conceivable for you as a practitioner for your patients?"). The participants were asked about their imaginable use of sex robots, judging by their work and experiences as a therapist.

At the time of the survey, none of the respondents had already recommended the use of sex robots to a patient. Some sex therapists recommended the use of sex toys to their patients, for example, as a couple exercise at home. Other therapists completely disagreed and declined such recommendations. Overall, almost half of all respondents (45%, 33/72) could imagine recommending sex robots in therapy. The entire sample was asked about therapeutic situations in which they would consider the use of sex robots conceivable. The answers to the question about imaginable situations for the general use of sex robots showed a similar frequency distribution as the question about imaginable situations from the point of view of the practitioner. Most of the respondents could imagine the use for people with physical disabilities (61%, 44/72), to live out sexual fantasies (48%, 35/72), and for people living in isolated environments, for example, prisons (44%, 32/72). In comparison with the question regarding which use of sex robots would be generally conceivable for the participants, it was noticeable that the respondents were less in favor of recommending sex robots in a therapeutic setting.

Use in Various Sexual Disorders

In the questionnaire, we asked sex therapists and physicians which diagnoses (eg, based on 10th revision of the International Statistical Classification of Diseases and Related Health Problems, ICD) they regarded as suitable for the use of sex robots and for which purposes they would recommend a sex robot to support therapy. To answer this question, diagnoses from the ICD-10 and possible problems were mentioned. One-third (33%, 24/72) rejected any use for their patients. Only 19% (14/72) could imagine using sex robots for patients who wanted to improve their sexual relationship. The most frequent use was conceivable in *patients with social anxiety that prevents a sexual life* (50%, 36/72), for patients who do not have a partner and still want to have a sex life without having to resort to prostitution or fleeting acquaintances (50%, 36/72), and in patients with *ejaculatio praecox* (47%, 34/72). Table 4 gives a detailed overview of the results.

In addition, the participants were asked via open questions for which other patient groups they considered the use of sex robots as useful. This option had been used 25 times. Overall, 2 therapists added that they could see benefits for older people and patients suffering from dementia. Furthermore, 4 people mentioned *pedophile patients* in the open response category. Another 4 emphasized the use in the context of social anxieties and 7 persons stated that the use of sex robots had to be decided on an individual basis. The remaining 8 people named other diagnoses, such as autism or ejaculatio praecox.

XSL•FO
RenderX

Table 3. Attitudes toward different uses of sex robots (N=72).

Imaginable use of sex robots	Statistics, n (%)
For physically handicapped persons	47 (65)
To be able to live out certain sexual fantasies	44 (61)
In isolated environments, for example, prisons, space stations, etc	36 (50)
To temporarily replace a human sexual partner	34 (47)
To be able to experience a sexual relationship when, for certain reasons, a sexual relationship with a person cannot arise	34 (47)
Instead of prostitution	31 (43)
To improve general psychological well-being	30 (41)
As a sex toy in the relationship	30 (41)
Out of sexual interest	30 (41)
To discover sexual pleasure for yourself (again)	28 (38)
To learn sexual practices	27 (37)
As a remedy for loneliness	24 (33)
Instead of cheating on the partner	23 (31)
For pornographic movies	23 (31)
To make forms of sexual harassment/sexual violence more tangible for training and prevention purposes	23 (31)
Out of technological interest	22 (30)
To expand your own sexual practices	20 (27)
To have sex regularly	19 (26)
As a remedy for boredom	18 (25)
For mixed group sex with humans and sex robots	14 (19)
To practice intimacy with someone else	12 (16)
To minimize the risk of sexually transmitted diseases	12 (16)
To build a sexual relationship	11 (15)
To facilitate the practice of religious/spiritual abstinence	8 (11)
To permanently replace a human sexual partner	5 (6)
To experience a trustful sexual relationship	4 (5)
No use at all would be conceivable	8 (11)

XSL•FO
RenderX

Table 4. Evaluation of therapeutic use in different diagnoses and situations (N=72).

Diagnoses and situations of sex robot use in therapy	Statistics, n (%)
For patients with social anxiety	36 (50)
For people who do not have a partner and still want to lead a sex life without having to resort to prostitution or fleeting acquaintances	36 (50)
Ejaculatio praecox	34 (47)
Erectile dysfunction	29 (40)
Psychoeducation	28 (38)
Orgasm disorders	27 (37)
Vaginismus	23 (31)
Paraphilias	22 (30)
Sexual aversion	17 (23)
Frigidity	16 (22)
Dyspareunia	15 (20)
Patients who want to improve their sexual relationship with their partner	14 (19)
Sexual maturity crisis	14 (19)
Sex addiction	11 (15)
Gender identity disorders	9 (12)
Not at all	24 (33)

Future Use of Sex Robots in Therapy

In addition, the respondents of the questionnaire were asked future-oriented questions such as “How likely do you think you’ll be using sex robots in therapy within the next year/the next 5 years/the next 25 years?” For this question, the participants could choose an answer on 4 scales, ranging from very probable to very unlikely. It was found that 90% (64/72) of the therapists thought that the use of sex robots in therapy within the next year was very unlikely or unlikely. Only 68% (49/72) thought that they were very unlikely or unlikely to recommend a sex robot in the next 5 years, whereas 32% (18/2) thought they would consider it. When asked what would happen in the next 25 years, therapists were more likely to consider a recommendation. Only 38% (27/72) thought that a recommendation would be (very) unlikely, whereas 62% (45/72) thought it would be highly likely.

The content analysis of the open question whether sex robots will change sexuality showed that the answers could be divided into 3 categories. Some therapists and physicians emphasized positive changes, such as the expansion of sexuality and therapy options. Others noted negative effects, such as the loss and replacement of real human relationships, and some statements can be described as neutral. In total, 46 people gave open answers with 49 units of sense being indicated. Table 5 gives an overview of the results.

Table 5. Future changes in sexuality caused by sex robots as predicted by surveyed therapists.

Category	Example answers	Number of units of senses
Positive aspects	“Expansion of therapy options”; “Relieve tension”; “Enhancing sexual possibilities”; “Experience sexual pleasure and emotional attention”	12
Neutral aspects	“Variant of sexuality”; “A new sex toy”; “Comparable to the internet”; “Another option”	17
Negative aspects	“Sensuality gets lost”; “Build pressure”; “Dehumanization of sexuality”; “Prohibiting practicing with partner”	20

Ethical Problems

In the questionnaire, we also asked therapists to indicate whether using sex robots could lead to ethical issues, with 62% (45/72) of the sample answering this question in the affirmative. To obtain detailed information on possible ethical problems, an open question was asked. In total, 30 persons gave open answers with 34 units of sense being indicated. The content analysis of the open question revealed 5 categories. The first category *dehumanization* described how people are dehumanized by the comparability with robots. However, spending time with sex robots could make people no longer distinguish between humans and robots. The second category *violence* collected responses expressing concerns that the use of sex robots could promote sexual violence. The third category *neglect of interpersonal relationships* referred to statements expressing concerns that interpersonal relationships were disregarded through the use of sex robots. The fourth category was designated as *narcissistic disorders/selfishness*. Some statements related to the concern that selfishness might increase and were summarized in this

XSL•FO
RenderX

category. In the last category, we summarized answers that do not address an ethical problem, such as *promotion of sex addiction*, which imply mental health topics. Isolated statements that did not fit into any of the categories were considerations about the (ethic) rights of robots, especially when artificial intelligence is so advanced that it resembles human consciousness. It was also stated that sex robots could lead to the misconception that sexual therapy was no longer needed. Table 6 gives an overview of the categories of ethical problems that have been identified.

Table 6. Ethical problems as predicted by surveyed therapists.

Category	Example answers	Number of units of senses
Dehumanization	"No more distinction between robots and human"; "Dehumanization"; "Confusion between human and machine"	10
Violence	"Performing trial offences"; "Risk of crossing borders in sexual contact with people"; "Sexual violence"; "Glorification of sexuality with children"	7
Neglect of interpersonal relationships	"Partner replacement, that is, no social relationship is established"; "Lack of understanding and flexibility in interpersonal relationships"; "Alienation from oneself, other people, and the world"	6
Narcissistic disorders/selfishness	"The tendency to use other people to satisfy one's own needs will increase"; "More narcissistic disorders"	4
Nonethical problems mentioned	"Sex addiction"; "Addiction development"; "Strengthening of sexual dysfunction"	5

Relationship Between Attitude and Sociodemographic Data

We also tested whether the therapists' and physicians' attitudes differ with regard to gender, age, and profession. The examined attitudes are the conceivable use of sex robots in certain situations, as well as the ability to recommend a sex robot as a practitioner in certain situations and the willingness to recommend a sex robot for certain diagnoses or sexual disorders. Only significant results are presented below.

Relationship Between Attitude and Gender

In the quantitative survey, male and female therapists differed in their attitudes toward sex robots. There was a significant difference between women and men regarding the variable *sex with a robot has therapeutic potential* (χ^2_1=7,5 N=72; *Phi*=−0.324; *P*=.006), to the effect that men affirmed this variable more often than women. In addition to the question of ethical problems, there was a gender difference, namely, female therapists more often assumed ethical problems than male therapists (χ^2_1=6,0, N=72; *Phi*=0.289; *P*=.01).

Relationship Between Attitude and Age

There were differences in attitudes toward sex robots in general and also in therapeutic practice between younger and older therapists. There was a significant difference in age with regard to the imaginability of the general use of sex robots (χ^2_2=6.4, N=72; *P*=.03). The post hoc tests in pairs showed that younger therapists (n=30, aged 32-50 years) and middle-aged therapists (n=31, aged 51-60 years) of this sample differed from older therapists (n=11, aged 61-80 years) with regard to the conceivability of a general use of sex robotics. Younger therapists (*U*=84, *Z*=−2.39, *P*=.02) and middle-aged therapists (*U*=90, *Z*=−2.31, *P*=.02) could imagine the use more frequently than older therapists. A significant age difference could also be observed with regard to the recommendation of sex robots for certain diagnoses (χ^2_2=7.2, N=72; *P*=.03). Younger therapists (*U*=96, *Z*=−2.04, *P*=.04) and middle-aged therapists (*U*=99, *Z*=−2.04, *P*=.04) were more open regarding the recommendation of sex robots.

Relationship Between Attitude and Education

The therapists also differed in their attitudes toward sex robots with regard to their profession. There was a clear correlation between the profession and the idea of recommending sex robots to patients. A medium positive correlation was found between the recommendation of sex robots in certain situations and the profession *physician* at rho=0.296 and *P*=.01, whereas the profession of sex therapist correlated negatively with this variable at rho=−0.233 and *P*=.049. There was another significant difference regarding *sex with a robot has therapeutic potential*. Physicians (χ^2_1=4.3, N=72; *Phi*=0.245; *P*=.03) confirmed this variable much more frequently than others. A significant difference could also be observed with regard to ethical problems. The variable *could the use of sex robots lead to ethical problems* was denied by physicians more frequently than by other occupational groups (χ^2_1=4.74, N=72; *Phi*=0.249; *P*=.03).

Relationship of Attitude With Affinity to Technology and Personality Traits

A large part of the sample (88%, 63/72) also completed the TA-EG and the NEO-FFI. Binary logistic regressions were calculated to check whether the general attitude toward sex robots as well as the assessment of the therapeutic potential could be predicted by the factors of NEO-FFI and the scales of TA-EG. However, no relationship could be established between these aspects and the attitudes toward the therapeutic use of sex robots.

Definition of Sex Robots: Consumer Products Versus Therapeutic Tools

One result of the qualitative study is that the subjective definition of sex robots influences the evaluation of them as a therapeutic tool. Two different positions became apparent. One definition understands sex robots as a consumer product that is

XSL•FO
RenderX

only used for physical satisfaction. "As consumer goods, sex robots are not a good development for sexuality and interpersonal relationships." According to the same therapist, sex robots differ from other technical aids (eg, vibrators). Furthermore, 2 people compared sex robots with the *artificial world* of pornography: "I think of sex robots just as critically as sex for sale or pornography—like anything that suggests an artificial world." Another therapist explains "Sex robots interfere with the development of pornography, sex toys, and internet-related sexuality." With regard to this definition, possible problems such as internet sex addiction are mentioned above all. Therapists who defined sex robots primarily as a consumer product assessed their existence and use as negative. In contrast, therapists who defined sex robots as technical or therapeutic devices assessed their existence more positively. One therapist explains "It is important to distinguish between the social use of sex robots, i.e. consumer goods, and sex robots as therapeutic tools in order to identify their therapeutic benefits." Therapists who defined sex robots as therapeutic tools described concrete ideas of how they should look like and work to actually be suitable for therapy. The skin of the robot was most frequently addressed in this context. A therapist described why skin sensation is important: "We know that the bonding hormone oxytocin is produced through skin contact between humans. The question would be if this also works for robots?" Another important point is that the robot body should resemble the human body. For therapists, this means that the robot body portrays an *imperfect* design to convey a healthy body image. The question "What kind of image of a woman is created by such a robot?" is also related to considerations about the optics of the robot. Another important issue was that sex robots should not be *conceived as slaves* but should have their own desires and needs. In addition, they should be able to express those needs, feelings, or desires.

Therapeutic Benefits and Dangers of Sex Robots

Many thoughts about the therapeutic benefits and dangers of sex robots from the point of view of sex therapists could be collected. Therapists who saw therapeutic benefits in sex robots also expressed ambivalent feelings toward them: "Even though I want to be open for this development, I have ambivalent feelings, for example, when I think of the loss of social skills as a possible consequence."

All therapists described the concern that the use of sex robots could lead to loneliness, *further autonomization of instincts*, and *loss of social skills and loss of interpersonal relationships*. These concerns were based on the therapists' experiences with the negative effects of excessive pornography consumption and on the assumption that *sex robots are part of this development*. The results of the quantitative survey, which showed the strongest agreement among therapists for the use of sex robots in physically handicapped people, in isolated environments, and instead of prostitution, could also be confirmed in the qualitative study. Even therapists who could not imagine any therapeutic use saw a general benefit of sex robots in these areas: "The only thing I could imagine is a benefit for physically handicapped people or even instead of prostitution so that fewer women have to suffer." The therapeutic benefit of sex robots was discussed in the context of different disorders.

Patients With Deviant Sexual Behavior

The use of sex robots for patients with deviant sexual behavior was discussed by all therapists. Sex robots could have the potential to reduce the sex drive of certain sexually active persons within the framework of therapy. "Whenever sexuality becomes dangerous, the use of sex robots is worth considering if it can protect a real human life." Therapists mention the use of sex robots in the context of sexual violence or rape and in the context of pedophile patients, with the strongest contrast of opinions being seen here. What seems important here is that pedophile patients must be treated differently. For some, an impulse control disorder is predominant, whereas others may be traumatized. Therapists point out that the benefits of sex robots must be decided individually for each specific case: "Pedophile patients are not all the same and it has to be decided here quite individually which patient could benefit from it." For some patients, it could be an opportunity to live out their sexuality with a sex robot. Then, they could discuss in therapy which fantasies were behind it (eg, not being able to cope with an adult). For some patients, the use of sex robots could be a kind of substitute. For others, the stimuli for the abuse of children might intensify. A therapist pointed out the following: "It should be considered that the neuronal connection could be intensified by living out the fantasies with child sex robots in the patients' brain." Another therapist assumed that the abuse would be intensified by the use of child sex robots and underlined "that the production of child sex robots is generally immoral." In contrast to this, another therapist argued that the patient's thoughts, for example during masturbation, could also lower the barrier to committing a crime and that prohibitions—important as they may be—do not necessarily reduce the number of criminal offences, but rather provide an additional attraction for many patients. The therapist argued as follows: "If a child can be protected, then it makes sense to torture a doll instead." Another therapist addressed one's own fear of *triggering* something in the patient by recommending sex robots to pedophile patients. The responsibility of the therapist was also addressed. Does a therapist want to take responsibility for recommending sex robots, even if the therapy with a sex robot turns out to be dangerous and the patient becomes violent? Finally, several therapists addressed the need for further research in this field: "It would need more applied research in this particular area to actually generate therapeutic benefits for pedophile patients."

Patients With Contact Disorders

The use of sex robots for people with contact disorders—the emotional interaction with another person is limited because of social anxiety—was also controversially evaluated. Those who completely reject the use of sex robots in therapy explained that sex robots are only a solution for loneliness for a short time. The use of sex robots could be seen as a kind of *substitute* or even *escape*, as those patients often experience so much fear to deal with a *real counterpart*. However, the use of sex robots could at least lead to a *1-dimensional satisfaction* or, depending on how sex robots are designed, could even help to practice social behavior and communication about sexuality, desires, needs, and borders.

XSL•FO
RenderX

Disorders of Female Sexuality

Some therapists discussed the use of sex robots in the context of the patient's gender, by referring to supposed differences between female and male sexuality, whereby male sexuality was described as more *animal instinctive*. Although all therapists could imagine the use of sex robots in therapy rather for male patients, we can also describe some application areas for female patients. In the context of female sexuality, the therapeutic benefits of sex robots regarding desire and orgasm disorders, vaginismus, and traumatic experiences were discussed: "I could imagine that traumatized women who can ride on a sex robot, for example, and who can do so without fear of being overwhelmed by their sex partner, can benefit from this experience and successively reduce their fears, or that penetration will perhaps only become possible again in the first place." Through a penetration-capable sex robot, women with traumatic experiences, such as sexual violence/rape, could reduce their fears, approach their own sexuality again, and regain access to their own bodies.

Quality of Human-Robot Sexuality

This category collects considerations relating to the quality of human-robot sexuality. The therapists explained what they understood by healthy sexuality and how sexuality had already changed because of pornography, the internet, and technical equipment. In this context, reflections were made on the question of how sexuality will be further changed by new developments such as sex robots. For therapists, sexuality has something to do with desire, eroticism, and communication. A therapist described the core of sexuality as something mental, but the physical part of it remains unmentioned here: "The core of the sexual event is the mental encounter." Sexuality must be negotiated, and it must go beyond the *technical satisfaction* that a sex robot can offer, "it is also about warmth, appreciation, and respect." Therapists postulated that the quality of a human relationship can never be achieved with a robot that, despite artificial intelligence, comprises inanimate material: "I'm sure a robot can never replace a real human relationship." Some therapists are critical about the future use of sex robots and fear that the quality of human relationships and sexuality could suffer from this development. However, others are certain that an *emotional peak* that humans strive for can never be achieved by robot sex. There is agreement that sex robots will play a growing role in sexuality and sexual therapy in the future: "I am sure that sex robots will be an important topic in the future—this development will come." All therapists argued that sex robots should not be seen as a substitute for human relationships and sexuality. Nevertheless, some therapists also see the potential of sex robots for sexuality. Sex robots could increase sexual satisfaction and provide an opportunity for more experimentation and sexual imagination.

Interview Study

Core Categories. The aim of the qualitative study was to gain more detailed insights into possible therapeutic uses of sex robots for different types of patients. Overall, 3 categories were generated as follows: (1) definition of sex robots: consumer goods versus therapeutic aids, (2) benefits and dangers of sex robots, and (3) quality of robot sexuality.

Discussion

Principal Findings

Summary Discussion

Following the uses and gratifications approach, this study is one of the first ones to systematically explore motives, possible application scenarios, and attitudes of sex therapists toward sex robots as a tool in sexual therapy. The presented results about the possible therapeutic use of sex robots (eg, for patients with contact disorders, patients with deviant sexual behavior, and patients with different sexual disorders) complement the few considerations about sex robots that have been existing in the scientific discourse so far. The participants in the quantitative study were to a large extent already familiar with sex robots (77%), but there were also therapists who stated that they had not heard or read anything about them yet. This is mainly because of the fact that sex robots are a very young field of research, and experiences in the therapeutic context have not been available yet.

Skepticism Versus Openness Toward Sex Robots

On the one hand, this study shows that therapists and physicians are generally skeptical toward sex robots. The existence of sex robots was not rated as very positive. On the other hand, only few respondents stated that the use of sex robots was not conceivable for them and almost half of all respondents could imagine recommending sex robots for therapy. The fact that therapists initially adopted a critical and cautious attitude toward the *introduction* of a new technology for therapeutic purposes was also known in the context of other electronic mental (e-mental) health implementations (eg, telephone or online therapy) [34]. In addition, sex robots are socially controversially discussed and have new fully automated options with artificial intelligence. On the basis of this background, the study shows that the surveyed therapists and physicians are relatively open toward this development.

In the meaning of expectancy-value theory, an attitude (A) toward an object (object O) can be expressed in a function of beliefs (B) toward this object and the evaluations (E) of these expectations. We would like to use this theoretical thinking to better understand the ambivalent attitude of therapists between skepticism and openness. The results of this study provide an insight into the beliefs and evaluations of therapists. It was shown that there are various ethical concerns. These beliefs influence the attitude toward sex robots. For example, it became clear that female therapists and physicians expressed more ethical concerns about sex robots in this survey and rated them more negatively than male therapists. We were able to determine that the ethical evaluation of sex robots strongly depends on whether they are classified as sexual aids, along other sex toys (and their use is normalized), or whether they are understood as a different category, namely, not as objects used *for* sex, but as *social actors* with whom one has sex. The results also show that not only the subjective definition of sex robots and ethical concerns influence the view of therapeutic benefit, but also expectations of the quality of human-robot sexuality. Furthermore, the attitudes toward sex robots as a therapeutic

XSL•FO
RenderX

tool were very heterogeneous (see *Relationship Between Attitude and Sociodemographic Data*).

Comparison With Prior Work

Gender Differences

A comparison with existing research results reveals interesting similarities and differences between the attitudes of therapists and the general population toward sex robots. Both the study of the general population and of the therapists indicated significant gender differences. Women rated sex robots more critically than men. Scheutz and Arnolds [25] assume that different judgments about the appropriateness between male and female participants in their survey could come from market and media forces that specifically address heterosexual men as customers and users. Another explanation could lie in the fact that heteronormative ideas of male hegemony are mirrored in the design of current sex robots. Especially in this context, representatives of the radical feminist extensively argue that promoting the development of sex robots reveals a compulsive attitude toward women's bodies [35]. In addition, Sullins [36] argues that sex robots "contribute to a negative body image." In the qualitative study, it became clear that sex therapists attach great importance to the physical design of sex robots when it comes to using them for therapeutic purposes. However, they clearly distinguish therapeutic robots from *pornographic sex robots*. Moreover, they advocate that sex robots should be available in different body shapes to promote a realistic and healthy body image. Kubes [37] assumes that the development of sex robots offers a great potential for reducing stereotypes and promoting diversity but current trends in sex robotics, however, do not explore these possibilities. In this study, it became clear that the therapists interviewed took a gender role-conform perspective as shown by the fact that a distinction was made between *male* and *female* sexuality. Moreover, therapists assumed that male rather than female patients could benefit from sex robots. Despite critical attitudes of women themselves and the gender-conform position previously described, possible benefits for female patients were also discussed.

Caution Among Therapists

With regard to the possible benefit of sex robots as a therapeutic tool, the general population [26] was more open-minded than the therapists surveyed in this study. Therapists were even more reluctant when it came to recommending sex robots in therapy. This shows a hesitant attitude toward taking responsibility for the use of sex robots in a therapeutic context. On the one hand, this reluctance could be related to the experiences of the therapists. Some therapists evaluate the development of sex robots in line with internet and pornography and mainly observe the negative effects of internet sex addiction in their daily practice. On the other hand, technical and ethical questions regarding treatment are not clarified, which can lead to uncertainty and restraint. A differentiated evaluation of the results in this context shows that not all therapists and physicians have this restraint. Psychologists and sex therapists were more critical concerning the recommendation of sex robots as a treatment tool than physicians. This may be associated with treatment-related considerations; especially, the majority of psychodynamic psychotherapists in this survey will probably also reflect the recommendation to patients from the perspective of transference and countertransference phenomena. Therapists could also ask themselves what bonding needs or experiences of early childhood could be projected onto sex robots by their patients. In the therapeutic context, the question of which early object relationships are represented by the robot may be also interesting. For psychoanalytically oriented therapists, the rule of abstinence may play a role as well when it comes to recommending sex robots in therapy. In contrast to systemically trained therapists, psychodynamically trained therapists in these interviews generally spoke out against a direct recommendation of sex toys in sex therapy. These considerations are supported by studies in which psychodynamic therapists were more critical of new media such as internet interventions than, for example, cognitive behavioral therapists [38].

Patients With Deviant Sexual Behavior

With regard to the treatment of pedophile patients, the results showed the opposite picture compared with attitudes in the general population. Although the general population is strongly against the use of sex robots in this context [25], it is controversially discussed by the therapists surveyed in this study. In this context, the consideration was expressed that the use of child sex robots could lead to the prevention of actual children's abuse. Similar thoughts have already been discussed in pornography research. However, studies have concluded that violent pornography is more likely to increase aggressiveness and therefore has no cathartic effects [39]. The considerations to live out sexual violence and sexual abuse with robots also lead to the question whether there are limits to how a robot should be handled. However, this is a question for the research of robot ethics [40].

Differences in Age

In the context of age effects, it would be obvious to assume that the younger generation's knowledge and openness to technology should also manifest itself in different attitudes toward sex robots. In line with this expectation, the survey found significant differences between younger and older therapists. Younger therapists were actually more open to the topic. Other studies in the field of e-mental health applications also found similar age differences between therapists, for example, in their attitudes toward the internet and online therapy [41].

Affinity to Technology

Other studies also showed that the negative evaluation of robots is generally a factor that influences attitudes toward sex robots [28]. In this study, no association with attitudes toward sex robots in relation to the evaluation of negative consequences of technology (TA-EG subscale) was found.

Personality Traits

Consistent with our findings that personality traits have no influence on respondents' attitudes toward sex robots, Szczuka et al [27] also found in their study that personality traits have no effect on the purchase of sex robots.

XSL•FO
RenderX

Limitations

There are some methodological limitations to this study. Owing to the small sample size, the results of this study need to be interpreted with caution. For this reason, the results are not representative. Comparable online survey studies with psychotherapists, however, show even lower response rates than ours [42]. One could assume that especially therapists interested in technology, who have a positive attitude toward robots, took part in this study. On the other hand, it also needs to be considered that mainly participants with a negative attitude might have participated. In this survey, we could not find any bias, as we got both positive and negative opinion patterns.

Another limitation of this study is that not all respondents shared the same level of knowledge about sex robots. The provision of a stimulus in the questionnaire such as images or film clips to define sex robots makes it possible to survey the same level of knowledge. In this survey, we refrained from using this kind of stimulus as the state of research is still so young that there is hardly any illustrative material available. In addition, media representations of intimate human-robot relationships show stereotypical gender roles and heteronormativity [21].

In spite of the small sample, the investigations provide first exploratory results.

Implications for Research and Practice

Further research on a larger sample of therapists is necessary to gain a more differentiated picture of the therapeutic potential. Similar to other authors before us, we conclude that ethical responsibility in the digital age cannot be perceived as a *critical distance* from technological development but is effective when sexual scientists play an active role in shaping a technology that can be used to promote sexual health, nonviolence, and sexual diversity [5]. The topic should also be integrated into the training of sex therapists to form opinions beyond media images and to point out therapeutic options. Instead of criticizing only dystopian visions of harmful sex robots, it is recommended to develop robots with positive effects on sexual education, sexual therapy, sexual counseling, and sexual well-being for interested groups. In future research, the different applications of robotic sex (eg, hardware robots and software robots) should be investigated in a differentiated way. The therapists' experiences with expert knowledge in robot technology and/or robot therapy should be included. The use of robots as a future tool in sex therapy still leaves many moral, ethical, and treatment-related questions unresolved, which need further research and evaluation.

Acknowledgments

The authors thank Hogrefe for providing the NEO-FFI and the authors of the TA-EG. The authors thank Jessica Huss, MSc, and Cornelia Küsel, MA, for their scientific support. The authors also gratefully acknowledge all participants for giving their time.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview Guide.
[PDF File (Adobe PDF File), 13KB - jmir_v21i8e13853_app1.pdf]

References

1. Cheok A, Levy D, Karunanayaka K, Morisawa Y. Love and sex with robots. In: Nakatsu R, Rauterberg M, Ciancarini P, editors. Handbook of Digital Games and Entertainment Technologies. Heidelberg: Springer; 2017:833-858.
2. Davis M. After the clinic? Researching sexual health technology in context. Cult Health Sex 2015;17(4):398-411. [doi: 10.1080/13691058.2014.928371] [Medline: 24955722]
3. Allen A, Kannis-Dymand L, Katsikitis M. Problematic internet pornography use: the role of craving, desire thinking, and metacognition. Addict Behav 2017 Dec;70:65-71. [doi: 10.1016/j.addbeh.2017.02.001] [Medline: 28214738]
4. Marsh A. Electronic Journal of Human Sexuality,Volume 13. 2010. Love among the Objetcum Sexuals URL: http://www.ejhs.org/volume13/ObjSexuals.htm
5. Döring N. [From internet sex to robotic sex: state of research and challenges for sexology]. Z Sex-Forsch 2017 Mar 22;30(01):35-57. [doi: 10.1055/s-0043-101471]
6. Grout V. Omics Online. Robot Sex: Ethics and Morality URL: https://www.omicsonline.org/open-access/robot-sex-ethics-and-morality-2090-9888-1000e104.php?aid=67397
7. Kuo I, Broadbent E, McDonald B. Designing a robotic assistant for healthcare applications. In: Proceedings of the Health Informatics New Zealand Conference. 2008 Oct 15 Presented at: Health Informatics New Zealand (HINZ); October 15, 2008; Rotorua, New Zealand.
8. Forrester L, Roy A, Hafer-Macko C, Krebs HI, Macko R. Task-specific ankle robotics gait training after stroke: a randomized pilot study. J Neuroeng Rehabil 2016 Jun 02;13(1):51 [FREE Full text] [doi: 10.1186/s12984-016-0158-1] [Medline: 27255156]

XSL•FO
RenderX

9. Cavallo F, Esposito R, Limosani R, Manzi A, Bevilacqua R, Felici E, et al. Robotic services acceptance in smart environments with older adults: user satisfaction and acceptability study. J Med Internet Res 2018 Dec 21;20(9):e264 [FREE Full text] [doi: 10.2196/jmir.9460] [Medline: 30249588]
10. So W, Wong MK, Lam W, Cheng C, Ku S, Lam K, et al. Who is a better teacher for children with autism? Comparison of learning outcomes between robot-based and human-based interventions in gestural production and recognition. Res Dev Disabil 2019 Mar;86:62-75. [doi: 10.1016/j.ridd.2019.01.002] [Medline: 30677695]
11. Bilyea A, Seth N, Nesathurai S, Abdullah HA. Robotic assistants in personal care: a scoping review. Med Eng Phys 2017 Dec;49:1-6. [doi: 10.1016/j.medengphy.2017.06.038] [Medline: 28734874]
12. Nappi R, Ferdeghini F, Abbiati I, Vercesi C, Farina C, Polatti F. Electrical stimulation (ES) in the management of sexual pain disorders. J Sex Marital Ther 2003;29(Suppl 1):103-110. [doi: 10.1080/713847129] [Medline: 12735094]
13. CNN. What the sex robots will teach us URL: http://edition.cnn.com/2016/12/01/health/robot-sex-future-technosexuality/
14. Levy D. Love and Sex with Robots : The Evolution of Human-Robot Relationships. New York, NY: Harper; 2007.
15. Döring N, Pöschl S. Sex toys, sex dolls, sex robots: our under-researched bed-fellows. Sexologies 2018 Jul;27(3):e51-e55 [FREE Full text] [doi: 10.1016/j.sexol.2018.05.009]
16. Ferguson A. The Sex Doll: A History. London: McFarland; Oct 01, 2010:2010.
17. Sharkey N, Wynsberghe AV, Robbins S, Hancock E. Foundation for Responsible Robotics. Our sexual future with robots URL: https://responsible-robotics-myxf6pn3xr.netdna-ssl.com/wp-content/uploads/2017/11/FRR-Consultation-Report-Our-Sexual-Future-with-robots-.pdf
18. Liberati N. Teledildonics and new ways of "Being in Touch": a phenomenological analysis of the use of haptic devices for intimate relations. Sci Eng Ethics 2017 Dec;23(3):801-823. [doi: 10.1007/s11948-016-9827-5] [Medline: 27757826]
19. Richardson K. Sex robot matters: slavery, the prostituted, and the rights of machines. IEEE Technol Soc Mag 2016 Jun;35(2):46-53. [doi: 10.1109/MTS.2016.2554421]
20. Ess C. What's love got to do with it? Robots, sexuality, and the arts of being human. In: Norskov M, editor. Social Robots: Boundaries, Potential, Challenges. Farnham: Ashgate Publishing Limited; 2016:57-79.
21. Döring N, Poeschl S. Love and sex with robots: a content analysis of media representations. Int J of Soc Robotics 2019 Feb 6:1-13. [doi: 10.1007/s12369-019-00517-y]
22. Katz E, Blumer J, Gurevitch M. Uses and gratifications research. Public Opin Q 1973;37:509-523. [doi: 10.1086/268109]
23. Blumer J. Academic Issues, Scholarly Inquiry. 2019. Uses and Gratifications Research URL: https://doi.org/10.1002/9781118841570.iejs0032
24. Fishbein M, Ajzen I. Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research. Boston, MA: Addison Wesley; 1975.
25. Scheutz M, Arnold T. Are We Ready for Sex Robots? In: Proceedings of the 11th ACM/IEEE Conference on Human-Robot Interaction. 2016 Oct 07 Presented at: 11th ACM/IEEE Conference on Human-Robot Interaction; March 7, 2016; New Zealand URL: https://hrilab.tufts.edu/publications/scheutzarnold16hri.pdf
26. Eichenberg C, Ostermaier E, Khamis M, Küsel C, Hübner L. [Sexrobotik: Two survey studies on acceptance and usage options in sex therapy]. 2018 Presented at: DGPPN-Kongress; November 29, 2018; Berlin.
27. Szczuka J, Krämer N. Not only the lonely—how men explicitly and implicitly evaluate the attractiveness of sex robots in comparison to the attractiveness of women, and personal characteristics influencing this evaluation. Multimodal Technologies Interact 2017 Feb 10;1(1):3 [FREE Full text] [doi: 10.3390/mti1010003]
28. Richards R, Coss C, Quinn J. Exploration of relational factorsthe likelihood of a sexual robotic experience. In: Cheok AD, Devlin K, Levy D, editors. Love and Sex with Robots: Second International Conference, LSR 2016, London, UK, December 19-20, 2016, Revised Selected Papers (Lecture Notes in Computer Science Book 10237). Cham: Springer International Publishing; 2016:97-103.
29. Karrer K, Glaser C, Clemens C, Bruder C. Capture technical affinity? the questionnaire TA-EG. In: Lichtenstein A, Stößel C, Clemens C, editors. [Man at the center of technical systems: 8th Berlin Workshop Human-Machine Systems, 7th to 9th October 2009]. Düsseldorf: VDI Verlag GmbH; 2009:196-201.
30. Borkenau P. In: Ostendorf F, editor. [NEO five-factor inventory (NEO-FFI) to Costa and McCrae]. Göttingen: Hogrefe-Verlag; 1993.
31. Liamputtong P, Ezzy D. Qualitative research methods. South Melbourne: Oxford University Press; 2005.
32. Strauss A, Corbin J. Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory. Thousand Oaks: Sage Publications; 1998.
33. Glaser B. Basics of Grounded Theory Analysis: Emergence Vs. Forcing. Mill Valley, CA: Sociology Pr; 1992:-.
34. Eichenberg C, Hübner L. The Vienna Psychoanalyst. 2017. Psychoanalysis online? A short overview of the current practice and the state of research URL: https://www.theviennapsychoanalyst.at/index.php?log=2&wbkatlog=8&wbidlog=766&lakat=
35. Danaher J. Robotic rape and robotic child sexual abuse: should they be criminalised? Crim Law Philos 2017 Mar 13;11(1):71-95. [doi: 10.1007/s11]
36. Sullins J. Robots, love, and sex: the ethics of building a love machine. IEEE Trans Affective Comput 2012 Oct;3(4):398-409. [doi: 10.1109/T-AFFC.2012.31]

XSL•FO
RenderX

37. Kubes T. New Materialist Feminist Perspectives on Sex Robots. 2019 Oct 05 Presented at: Manchester; 2019; UK p. 2019 URL: https://www.researchgate.net/publication/332465802_New_Materialist_Feminist_Perspectives_on_Sex_Robots
38. 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. Cognit Ther Res 2017;41(5):745-756. [doi: 10.1007/s10608-017?9850-0]
39. Oddone-Paolucci E, Genius M, Violato C. A meta-analysis of the published research on the effects of pornography. In: Violato C, Genius M, Oddone-Paolucci E, editors. The Changing Family and Child Development. Aldershot, UK: Ashgate; 2000:48-59.
40. Darling K. Extending Legal Rights to Social Robots. In: SSRN Journal. 2012 Apr 23 Presented at: We Robot Conference; 2016; Miami. [doi: 10.2139/ssrn.2044797]
41. Eichenberg C, Kienzle K. [Psychotherapists and the Internet: attitude to and use of online therapeutic offers in everyday treatment]. Psychotherapeut 2011 Jun 11;58(5):485-493. [doi: 10.1007/s00278-011-0833-4]
42. Schuster R, Pokorny R, Berger T, Topooco N, Laireiter A. The advantages and disadvantages of online and blended therapy: survey study amongst licensed psychotherapists in Austria. J Med Internet Res 2018 Dec 18;20(12):e11007 [FREE Full text] [doi: 10.2196/11007] [Medline: 30563817]

Abbreviations

e-mental: electronic mental
ICD: International Statistical Classification of Diseases and Related Health Problems
TA-EG: The Questionnaire on technical affinity-attitude towards and handling of electronic devices
VR: virtual reality

Edited by G Eysenbach; submitted 28.02.19; peer-reviewed by M Beutel, N Döring; comments to author 03.04.19; revised version received 23.05.19; accepted 11.06.19; published 20.08.19.

Please cite as:
Eichenberg C, Khamis M, Hübner L
The Attitudes of Therapists and Physicians on the Use of Sex Robots in Sexual Therapy: Online Survey and Interview Study
J Med Internet Res 2019;21(8):e13853
URL: https://www.jmir.org/2019/8/e13853/
doi:10.2196/13853
PMID:31432784

©Christiane Eichenberg, Marwa Khamis, Lisa Hübner. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 20.08.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

Effectiveness of Digital Education on Communication Skills Among Medical Students: Systematic Review and Meta-Analysis by the Digital Health Education Collaboration

Bhone Myint Kyaw[1], MBBS, MSc, PhD; Pawel Posadzki[1], PhD; Sophie Paddock[2], MCE, MBBS; Josip Car[1], MD, PhD, FRCP, FFPH; James Campbell[3], MSc, MPH; Lorainne Tudor Car[4,5], MSc, MD, PhD

[1]Centre for Population Health Sciences, Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore, Singapore

[2]Norfolk & Norwich University Hospital, Colney Lane, Norwich, United Kingdom

[3]Health Workforce Department, World Health Organization, Geneva, Switzerland

[4]Family Medicine and Primary Care, Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore, Singapore

[5]Department of Primary Care and Public Health, School of Public Health, Imperial College London, London, United Kingdom

Corresponding Author:
Lorainne Tudor Car, MSc, MD, PhD
Family Medicine and Primary Care
Lee Kong Chian School of Medicine
Nanyang Technological University
11 Mandalay Road
308232
Singapore,
Singapore
Phone: 65 69041258
Email: lorainne.tudor.car@ntu.edu.sg

Abstract

Background: Effective communication skills are essential in diagnosis and treatment processes and in building the doctor-patient relationship.

Objective: Our aim was to evaluate the effectiveness of digital education in medical students for communication skills development. Broadly, we assessed whether digital education could improve the quality of future doctors' communication skills.

Methods: We performed a systematic review and searched seven electronic databases and two trial registries for randomized controlled trials (RCTs) and cluster RCTs (cRCTs) published between January 1990 and September 2018. Two reviewers independently screened the citations, extracted data from the included studies, and assessed the risk of bias. We also assessed the quality of evidence using the Grading of Recommendations, Assessment, Development, and Evaluations assessment (GRADE).

Results: We included 12 studies with 2101 medical students, of which 10 were RCTs and two were cRCTs. The digital education included online modules, virtual patient simulations, and video-assisted oral feedback. The control groups included didactic lectures, oral feedback, standard curriculum, role play, and no intervention as well as less interactive forms of digital education. The overall risk of bias was high, and the quality of evidence ranged from moderate to very low. For skills outcome, meta-analysis of three studies comparing digital education to traditional learning showed no statistically significant difference in postintervention skills scores between the groups (standardized mean difference [SMD]=–0.19; 95% CI –0.9 to 0.52; I^2=86%, N=3 studies [304 students]; small effect size; low-quality evidence). Similarly, a meta-analysis of four studies comparing the effectiveness of blended digital education (ie, online or offline digital education plus traditional learning) and traditional learning showed no statistically significant difference in postintervention skills between the groups (SMD=0.15; 95% CI –0.26 to 0.56; I^2=86%; N=4 studies [762 students]; small effect size; low-quality evidence). The additional meta-analysis of four studies comparing more interactive and less interactive forms of digital education also showed little or no difference in postintervention skills scores between the two groups (SMD=0.12; 95% CI: –0.09 to 0.33; I^2=40%; N=4 studies [893 students]; small effect size; moderate-quality evidence). For knowledge outcome, two studies comparing the effectiveness of blended online digital education and traditional learning reported no difference in postintervention knowledge scores between the groups (SMD=0.18; 95% CI: –0.2 to 0.55; I^2=61%; N=2 studies [292 students]; small effect size; low-quality evidence). The findings on attitudes, satisfaction, and

XSL•FO
RenderX

patient-related outcomes were limited or mixed. None of the included studies reported adverse outcomes or economic evaluation of the interventions.

Conclusions: We found low-quality evidence showing that digital education is as effective as traditional learning in medical students' communication skills training. Blended digital education seems to be at least as effective as and potentially more effective than traditional learning for communication skills and knowledge. We also found no difference in postintervention skills between more and less interactive forms of digital education. There is a need for further research to evaluate the effectiveness of other forms of digital education such as virtual reality, serious gaming, and mobile learning on medical students' attitude, satisfaction, and patient-related outcomes as well as the adverse effects and cost-effectiveness of digital education.

(J Med Internet Res 2019;21(8):e12967) doi:10.2196/12967

KEYWORDS

randomized controlled trials; effectiveness; systematic review; communication skills; medical education

Introduction

Both qualitative and quantitative researchers have intensely studied the importance of communication between patients and doctors since the 1970s. Within health care, where an individual explores the unknown environment of one's own health and disease, effective communication skills can positively affect a number of health outcomes including better emotional and physical health, higher symptom resolution, improved pain control, greater treatment compliance, and enhanced patient satisfaction [1]. Furthermore, studies have reported reductions in emotional distress, levels of discomfort, concerns, fear, hopelessness, grief, depression, or health services utilization as a result of effective communication [2,3]. Communication involves respecting the persons' dignity, integrity, and autonomy [4,5] as well as an ability to explore and discuss their expectations or wishes in a warm, nonjudgmental, and friendly manner. Effective communication (verbal and nonverbal) includes traits such as empathy, understanding, active listening, and the ability to meet patients' needs and emotionally charged information [6]. In clinical practice, effective communication also requires features needed for effective symptom control such as honesty, open disclosure, an ability to gain trust [7], and influence over patient behavior [8]. These communication skills are essential in building the doctor-patient relationship or "therapeutic alliance." Finally, physicians have legal, ethical, and moral obligations to demonstrate a variety of communication skills including the ability to gather information, formulate an accurate diagnosis, provide therapeutic instructions and medical advice, communicate risk, and deliver health-related news to the patients [9,10].

Communication skills training is recognized as an important component of the curricula in undergraduate and postgraduate medical education and is endorsed, for example, by the UK General Medical Council, which states that students should be able to "communicate clearly, sensitively and effectively with patients, their relatives and colleagues" [11]. The optimal method of teaching and learning communication skills is considered a direct observation of the student's performance, followed by feedback from an experienced tutor [12,13]. This form of small-group teaching requires intensive planning and resources including simulated patients and experienced tutors. The lack of standardization within these patients and tutors can result in unequal learning outcomes.

Digital education encompasses a broad spectrum of didactic interventions characterized by their technological content, learning objectives/outcomes, measurement tools, learning approaches, and delivery settings. Digital education includes online digital education, offline digital education, massive open online courses, learning management systems, mobile digital education (mobile learning or m-learning), serious games and gamification, augmented reality, virtual reality, or virtual patient (VP) [14-17].

For medical students learning communication skills, digital education offers self-directed, flexible, and interactive learning (didactic); novel instructional methods; and the ability to simulate and rehearse different clinical scenarios (experiential learning) [18]. For instance, online digital education could be a potential method of delivering the theoretical concepts that underpin communication skills. Virtual patient simulations may also be useful in clinical scenarios that are difficult to replicate with standardized patients, such as communication with patients who have rare conditions, speech disorders, and neurological diseases. Digital education can be utilized flexibly and for an unlimited number of times alongside traditional methods such as role play with standardized patients, allowing students to practice their skills "interchangeably." For educators, digital education offers the potential to free up time, save manpower and space resources, automate evaluation and documentation of students' progress, and receive feedback from the students [19].

Given the shortage of trained and experienced health care educators to deliver communication skills training, digital education may be a novel, cost-effective modality. To the best of our knowledge, there is no similar systematic review assessing the effectiveness of digital education for medical students' communication skills training. The aim of this research was to evaluate the effectiveness of digital education compared with various controls in improving knowledge, skills, attitudes, and satisfaction of medical students learning communication skills. In doing so, we aim to fill an important gap in the literature.

Methods

For this systematic review, we adhered to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Guidelines and the Cochrane Handbook for Systematic Reviews

XSL•FO
RenderX

of Interventions [20]. For a detailed description of the methodology, please refer to the study by Car et al [21].

Eligibility Criteria

We considered studies eligible for inclusion if they were randomized controlled trials (RCTs) of any design and of any type of digital education including blended education (combination of digital education and traditional learning) for medical students (ie, preregistration); measuring any of the primary outcomes, ie, knowledge, skills, attitudes, satisfaction; or measuring secondary outcomes, ie, patient-related outcomes, adverse effects, or costs (economic evaluations). We included studies if the studies compared: digital education versus traditional learning, digital education versus other forms of digital education, digital education versus no intervention, blended digital education versus traditional learning, and blended digital education versus no intervention.

We did not exclude participants based on age, gender, or any other sociodemographic factor. If data within a study included both preregistration (undergraduate level) and postregistration (postgraduate level) students, the study was included if these data were presented separately. We did not impose any language restrictions. Nonrandomized studies or trials of postgraduates including continuous professional development; continuous medical education; and students of traditional, alternative, and complementary medicine were excluded.

Search Strategy and Data Sources

We searched the following databases from January 1, 1990, to September 20, 2018, for all relevant digital education trials: Cochrane Central Register of Controlled Trials (Wiley), Educational Resource Information Centre (Ovid), Embase (Elsevier), Cumulative Index to Nursing and Allied Health Literature (Ebsco), MEDLINE (Ovid), PsychINFO (Ovid), and Web of Science Core Collection. We also searched the two trials registers—International Clinical Trials Registry Platform and metaRegister of Controlled Trials—to identify unpublished trials. We selected 1990 as the starting year for our search because the use of computers was limited to very basic tasks prior to this year. There were no language restrictions. We searched reference lists of all the studies that we deemed eligible for inclusion in our review and relevant systematic reviews. For a detailed search strategy for MEDLINE, please see Multimedia Appendix 1.

Data Selection, Extraction, and Management

We merged the search results from the databases using EndNote software [computer software] (Version X.7.8. Philadelphia, PA: Clarivate Analytics) and removed duplicates of the same record. Three reviewers (PP, SP, and BK) independently screened titles and abstracts to identify potentially eligible articles. They then read the full-text versions of these studies and assessed them independently against the inclusion and exclusion criteria. Any disagreements about whether a study meets the eligibility criteria were resolved through discussion among the two review authors. A third review author's opinion was sought to resolve any disagreements between two review authors. If a study had more than one intervention group, for comparison, we chose the relevant digital education group (ie, more interactive intervention) against the least interactive controls. "Interactivity" referred to "the degree of control or adaptiveness a user might have with a system, without necessarily having to give a response" [22], and we applied this definition of "interactivity" throughout the review. For each of the included studies, two reviewers independently extracted data related to the characteristics of population, intervention, comparators, outcome measures, and study design, and any discrepant opinions were resolved by discussion.

Assessment of Risk of Bias

Three review authors (PP, SP, and BK) independently assessed the risk of bias of the included studies using the Cochrane Risk of Bias Tool [20]. Disagreements between the reviewers were resolved by discussion. We appraised the following domains: random sequence generation, allocation concealment, blinding (participants, personnel and outcome assessors), completeness of outcome data, selective outcome reporting, and other biases. Each item was judged as having high, low, or unclear risk of bias based on the definitions provided by Higgins and Green [20]. For cluster RCTs, the risk of bias assessment also focused on recruitment bias, baselines imbalance, loss of clusters, incorrect analysis, and comparability with individually randomized controlled trials [23]. We incorporated the results of the risk of bias assessment into the review using a graph and a narrative summary.

Data Synthesis and Analysis

For continuous outcomes, we reported postintervention mean scores and SD along with the number of participants in each intervention group. We reported postintervention mean outcome data to ensure consistency across the included studies, as most of the included studies (92%) reported postintervention data. We presented outcomes using postintervention standardized mean difference (SMD) and interpreted the effect size based on the Cohen rule of thumb (ie, with 0.2 representing a small effect, 0.5 representing a moderate effect, and 0.8 representing a large effect) [20,24]. If studies had multiple arms, we compared the most interactive intervention arm to the least interactive control arm and assessed the difference in postintervention outcomes.

For dichotomous outcomes, we summarized relative risks and associated 95% CIs across studies. Subgroup analyses were not feasible due to the limited number of studies within respective comparisons and outcomes. We used a random-effects model for meta-analysis. We used the I^2 statistic to evaluate heterogeneity, with I^2 <25%, 25%–75%, and >75% representing low, moderate, and high degree of inconsistency, respectively. The meta-analysis was performed using Review Manager 5.3 [25]. We reported the findings in line with the PRISMA reporting standards [26].

The three authors (SP, PP, and BK) independently assessed the overall quality of the evidence in accordance with the Grading of Recommendations, Assessment, Development and Evaluations criteria [27]. The following criteria were considered: limitations of studies (risk of bias), inconsistency, indirectness, imprecision and publication bias, and downgrading the quality where appropriate. We did this for all primary and secondary outcomes reported in the review. We rated the quality of

XSL•FO
RenderX

evidence for each outcome as "high," "moderate," and "low." We prepared "Summary of findings" tables for each comparison to present the findings and rated the quality of the evidence for each outcome (Multimedia Appendices 2-4). We were unable to pool the data statistically using meta-analysis for some outcomes (eg, attitude and satisfaction) due to high heterogeneity in the types of participants, interventions, comparisons, outcomes, outcome measures, and outcomes measurement instruments. We presented those findings in the form of a narrative synthesis. We used the standard method recommended by Higgins et al [20] to synthesize and represent the results.

Results

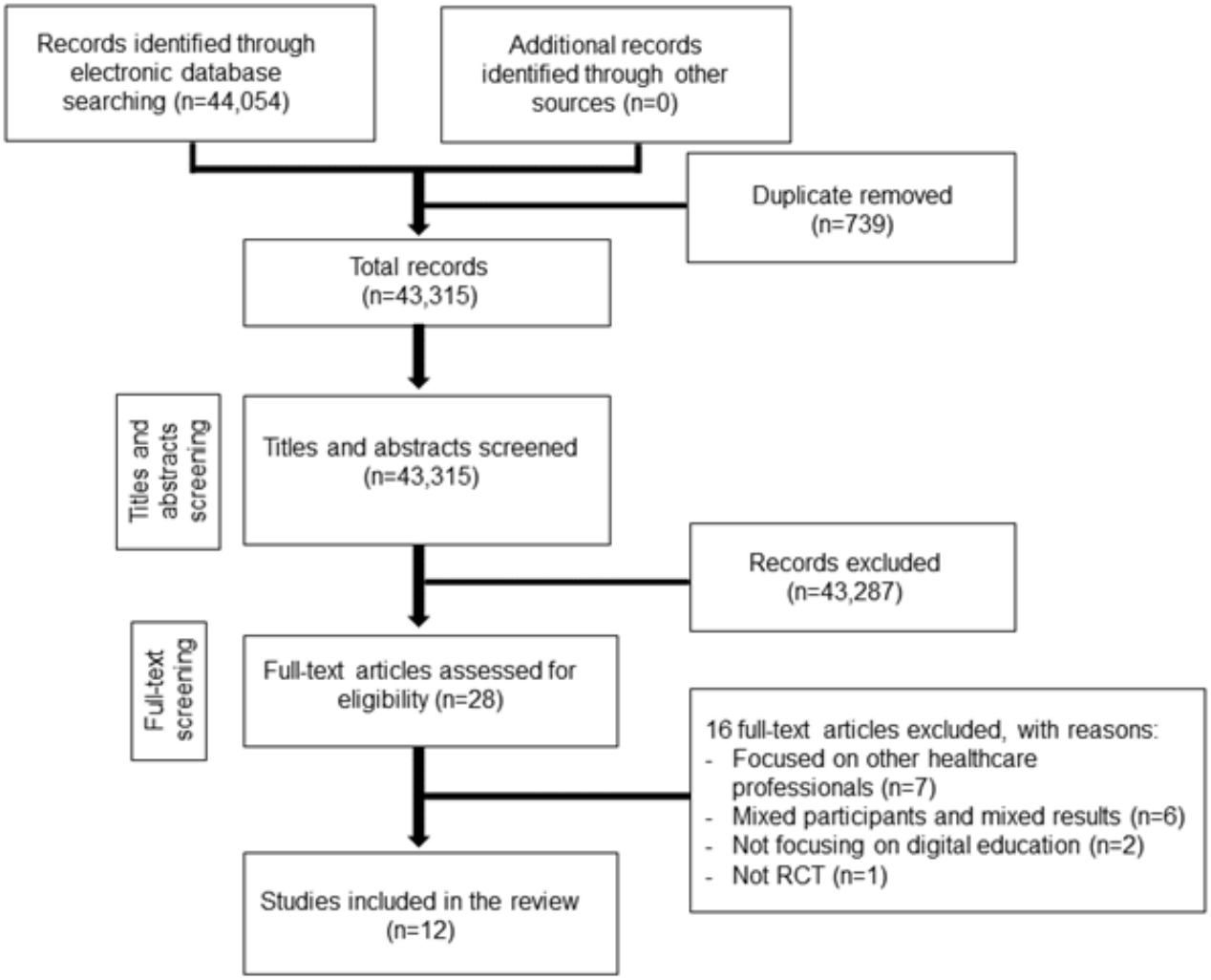
Overview

We identified 44,054 records overall from electronic database searches. We excluded 43,287 references after screening titles and abstracts and retrieved 28 studies for full-text evaluation, of which 12 studies met the inclusion criteria [28-39] and were included in the review (Figure 1). The total number of students was 2101.

We present details of the included trials in Table 1. The included studies were published between 2000 and 2018; of these, nine were RCTs, two were cluster RCTs [31,38], and one was a factorial-design RCT [30]. The studies originated from Australia [28], China [39], Germany [30,37], and the United States [29,31-36,38]. The sample sizes in the included studies ranged from 67 to 421 medical students, and they were in their first, second, third, and fourth year of studies. The included studies focused on different modalities of digital education. For instance, five studies (41.7%) [28,33,35,36,38] used VP, whereas the remaining seven studies (58.3%) used online modules; in addition, five studies (41.7%) used traditional learning in addition to digital education, that is, blended digital education [30,31,34,37,39]. Two studies (22.2%) had more than one intervention arm [29,30]. The content of those interventions also differed from history-taking and basic communication skills [28,30,33,36,37], cross-cultural communication [32], ethical reasoning [34], suicide risk management [35], interprofessional communication [38], ophthalmology-related communication skills training [39], and substance abuse–related communication [31] to end-of-life support [29]. Comparison groups ranged from other digital education such as virtual patient [28], online learning [38], traditional learning (written curriculum, didactic lecture, oral feedback, and standardized patient) [29,31-34,36,37,39], video group [35], or no intervention [30]. Outcomes were measured using a range of tools including scales, surveys, checklists, Likert scales, and Objective Structured Clinical Examination (OSCE), questionnaires; seven studies (58.3%) reported some type of validity evidence (ie, validity, reliability, and responsiveness for those measurement tools) [28,30,31,33-35,39].

XSL•FO
RenderX

Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses flow diagram. RCT: randomized controlled trial.



XSL•FO
RenderX

Table 1. Main characteristics of the included studies (N=12).

Study author names, year, country, and design	Number of participants (year)	Intervention and comparator	Outcome/measurement instrument (validity, reliability)	Results (postintervention): SMD[a] (95% CI)	Effect estimate
Digital education versus traditional learning					
Chittenden et al, 2013 [29] US, RCT[b]	121 (third)	IG1[c]: 45-min online module (multimedia) CG[d]: Traditional learning (written curriculum)	Skills: checklist, Likert scale	Skills: SMD=0.00 (–0.46 to 0.46)	Skills: Digital education=traditional learning
Deladisma et al, 2007 [33] US, RCT	84 (second)	IG: VP[e] simulation CG: Traditional learning (SP[f] consultation)	Skills: survey, Likert scale (Cronbach alpha=.87)	Skills: SMD=–0.89 (–1.35, –0.43)	Skills: Digital education<traditional learning
Gartmeier et al, 2015 [30] Germany, factorial RCT	168 (second and third)	IG1: 300-min online tutorial CG1: No intervention (wait list)	Skills: Likert scale (intraclass correlation=0.54)	Narrative presentation	Skills: Digital education>traditional learning
Kaltman et al, 2018 [36] US, RCT	99 (first)	IG: VP simulation (video-based) CG: Traditional learning (usual curriculum)	Skills: OSCE[g]	Skills: SMD=0.33 (–0.16 to 0.81)	Skills: Digital education=traditional learning
Blended digital education versus traditional learning					
Fleetwood et al, 2000 [34] US, RCT	173 (second)	IG: 45-min online module plus small group discussions (blended digital education) CG: 45-min traditional learning (group discussions)	Knowledge: MCQ[h] (exam) Skills: checklist (validated) Patient-related outcome (ie, patients' satisfaction): checklist (validated)	Knowledge: SMD=0.00 (–0.30 to 0.30) Skills: SMD=–0.20 (–0.50 to 0.10) Patients' satisfaction: SMD=–0.43 (–0.73 to –0.13)	Knowledge: Blended digital education=traditional learning Skills: Blended digital education=traditional learning Patient-related outcome (ie, patients' satisfaction): Traditional learning>blended digital education
Gartmeier et al, 2015 [30] Germany, factorial RCT	168 (second and third)	IG2: 300-min online tutorial and role play (blended digital education) CG2: Role play	Skills: Likert scale	Narrative presentation	Skills: Blended digital education>traditional learning
Lanken et al, 2015 [31] US, cluster RCT	370 (second and third)	IG: 1-h online module and small group discussion (blended digital education) CG: Traditional learning (usual curriculum)	Skills: survey Attitude: survey (Cronbach alpha=.89)	Skills: SMD=–0.08 (–0.28 to 0.13) Attitude (toward the outcome): SMD=0.05 (–0.15 to 0.26)	Skills: Blended digital education=traditional learning Attitude (toward the outcome): Blended digital education=traditional learning
Lee et al, 2015 [32] US, mixed method RCT	119 (third)	IG: 1-h online module (cultural competency and PACT[i] training plus standard curriculum) CG: Traditional learning (standard curriculum)	Knowledge: PACT (questionnaire) Skills: OSCE	Knowledge: SMD=0.38 (0.02-0.74) Skills: SMD=0.05 (–0.31 to 0.41)	Knowledge: Blended digital education>traditional learning Skills: Blended digital education=traditional learning
Ruesseler et al, 2017 [37] Germany, RCT	100 (fourth)	IG: 1.5-h video-assisted oral feedback (blended digital education: video recorded role play with video-assisted oral feedback) CG: Traditional learning (received direct oral feedback after role play)	Skills: OSCE	Skills: SMD=0.92 (0.51-1.33)	Skills: Blended digital education>traditional learning

XSL•FO
RenderX

Study author names, year, country, and design	Number of participants (year)	Intervention and comparator	Outcome/measurement instrument (validity, reliability)	Results (postintervention): SMD[a] (95% CI)	Effect estimate
Tang et al, 2017 [39] China, RCT	95 (fourth)	IG: Online video plus team discussion (blended online digital education) CG: Traditional learning (didactic lecture)	Attitude (toward the outcomes): Likert scale (three-point, validated) Students' satisfaction (with the intervention) Likert scale (three-point, validated)	Attitude (toward the outcomes): favored IG over CG (P=.04) Students' satisfaction (with the intervention): no difference (P=.61)	Attitude (toward the outcomes): Blended online digital education>traditional learning Students' satisfaction (with the intervention): Blended online digital education=traditional learning
Digital education (more interactive) versus digital education (less interactive)					
Bearman et al, 2001 [28] Australia, RCT	284 (not specified)	IG: 1-h problem-solving VP CG: 1-h narrative VP	Skills: scale (Cronbach alpha=.83)	Skills: SMD= –0.12 (–0.43 to 0.2)	Skills: Digital education (more interactive)=digital education (less interactive)
Chittenden et al, [29] US, RCT	121 (third)	IG1: 45-min online module (multimedia) IG2: 45-min online module (classic)	Skills: checklist, Likert scale	Skills: SMD=0.00 (–0.42 to 0.42)	Skills: Digital education (more interactive)=digital education (less interactive)
Foster et al, 2015 [35] US, RCT	67 (second)	IG: Online-based VP simulation CG: Video group (online module)	Skills: communication checklist and rapport subscale (Cronbach alpha=.97 and .84, respectively) Students' satisfaction: survey (Cronbach alpha=.89)	Skills: SMD=0.33 (–0.15 to 0.82) Students' satisfaction: P=.007)[j]	Skills: Digital education (more interactive)=digital education (less interactive) Students' satisfaction: Digital education (online-based video group)>digital education (online-based VP simulation)
Kron et al, 2017 [38] US, mixed method cluster RCT	421 (second)	IG: two VP simulations CG: online module (multimedia computer-based learning)	Skills: OSCE Attitude: survey	Skills: SMD=0.26 (0.06-0.45) Attitude (toward the intervention): SMD=0.71 (0.51-0.91)	Skills: Digital education (more interactive)>digital education (less interactive) Attitude (toward the intervention): Digital education (more interactive)>digital education (less interactive)

[a]SMD: standardized mean difference.

[b]RCT: randomized controlled trial.

[c]IG: intervention group.

[d]CG: control group.

[e]VP: virtual patient.

[f]SP: standardized patient.

[g]OSCE: Objective Structured Clinical Examination.

[h]MCQ: multiple-choice questionnaire.

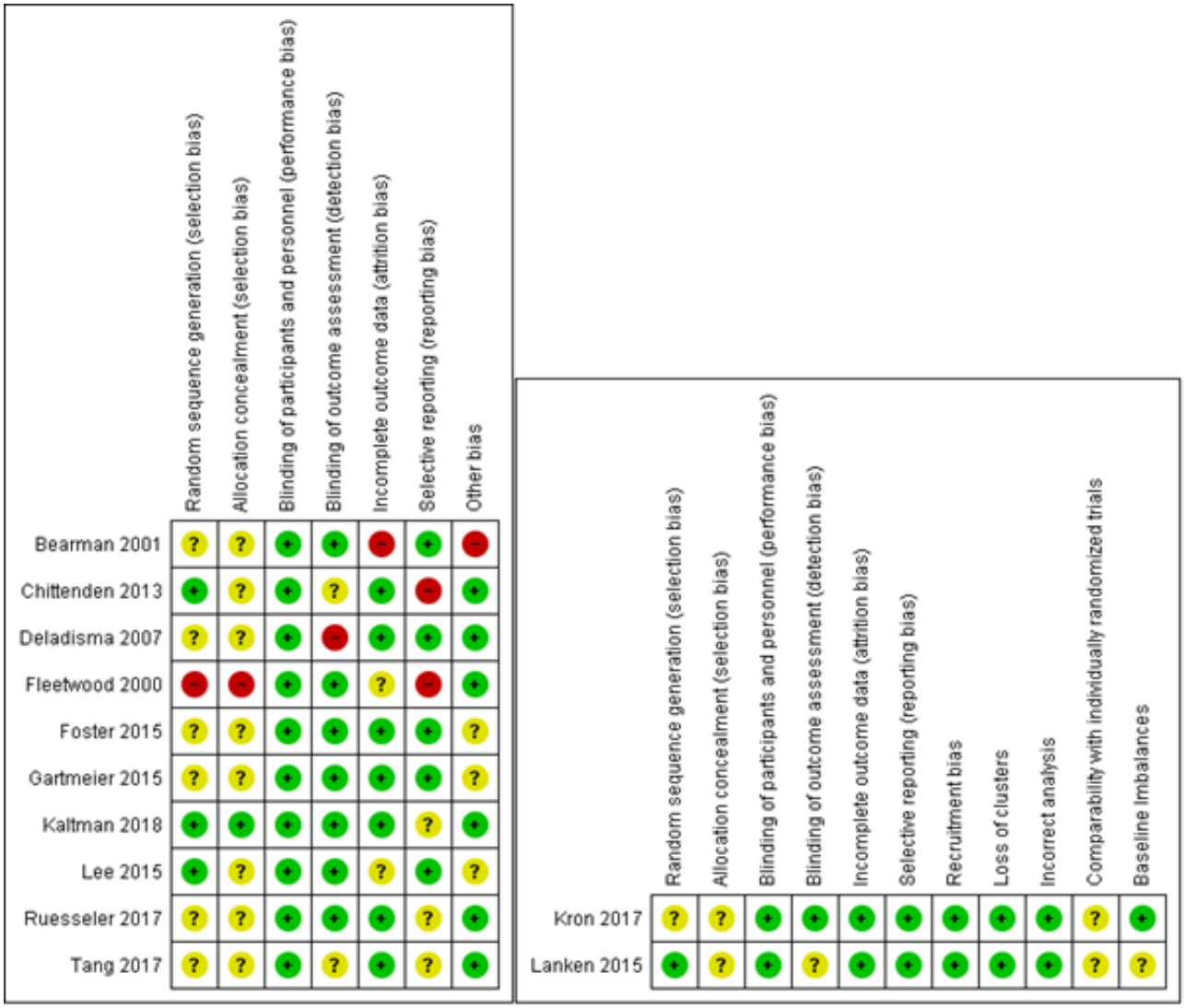
[i]PACT: Problem Affect Concern Treatment.

[j]Rating of the technology module overall.

In general, the risk of performance, detection, and attrition was predominantly low, and it was unclear or high for sequence generation bias, allocation concealment, and other bias. Reporting bias was judged as high in two (16.7%) of the studies. For two cRCTs, the overall risk of bias was low or unclear. Four of the 12 included studies (33.3%) were judged to have a high risk of bias in at least one domain (Figure 2). The quality of evidence ranged from moderate to very low due to study limitations, inconsistency, and imprecision across the studies.

XSL•FO
RenderX

Figure 2. Risk of bias summary: review authors' judgements about each risk of bias item for each included study.



Effect of the Interventions

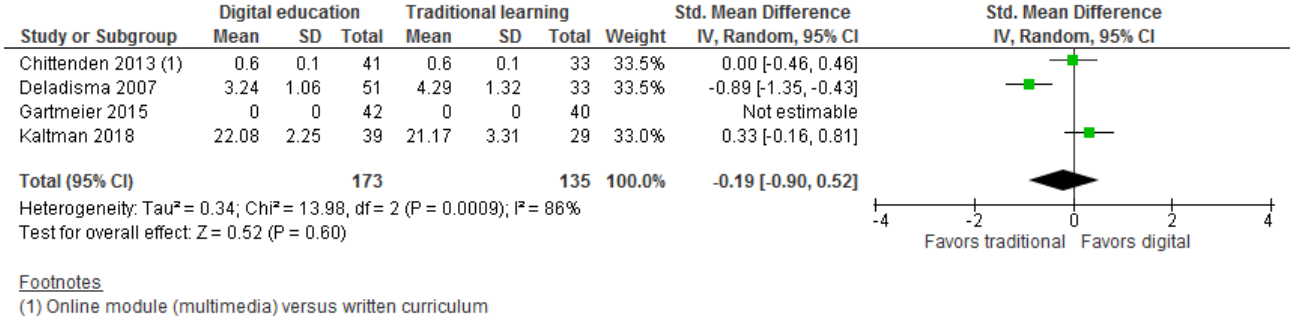
Digital Education Versus Traditional Learning

Four studies [29,30,33,36] compared the effectiveness of digital education and traditional learning and reported on postintervention skills, attitudes, and satisfaction outcomes. For skills, there was no statistically significant difference between the digital education group (ie, online modules, tutorials, and virtual patient simulation) and the traditional learning group at postintervention (SMD=–0.19; 95% CI –0.9 to 0.52; 3 studies (304 students); I^2=86%; low-quality evidence; Figure 3). However, this finding had high imprecision with wide CIs, which also included a large effect size in favor of traditional learning as well as a moderate effect size in favor of digital education. The high observed heterogeneity was largely driven by a study comparing the effectiveness of VP simulation to simulated patient training [33]. The remaining two studies compared the effectiveness of online modules or VP simulation with more passive forms of traditional learning such as written materials or usual curriculum [29,36]. Findings from one study [30] favoring online digital education over no intervention could not be pooled with the other studies due to the lack of comparable numerical data.

None of the studies reported on knowledge, attitudes, satisfaction, adverse effects, patient outcomes, or cost outcomes.

XSL•FO
RenderX

Figure 3. Forest plot of studies comparing digital education with traditional learning for postintervenion skill outcome, IV: inverse variance; random: random effect model.



Blended Digital Education Versus Traditional Learning

Six studies [30-32,34,37,39] compared the effectiveness of blended digital education (ie, blended online or offline [video-based] digital education) and traditional learning and assessed students' postintervention knowledge, skills, attitude, and patient-related outcomes (ie, patients' satisfaction). For skills, there was no statistically significant difference between the groups at postintervention (SMD=0.15; 95% CI –0.26 to 0.56; I^2=86%; 4 studies (762 students); small effect size; low-quality evidence; Figure 4). The reported findings were imprecise due to wide CIs including moderate effect sizes in favor of blended digital education. Three studies included in the meta-analysis compared a blend of online modules and a small group discussion or standard curriculum with standard curriculum or small group discussions only [31,32,34]. The high observed heterogeneity was largely driven by a study comparing role play and video-assisted oral feedback to role play with oral feedback only, favoring blended digital education [37]. Findings from one study favoring a blend of online tutorials and role play could not be included in the meta-analysis due to the lack of comparable outcome data [30].

For knowledge, two studies compared the effectiveness between blended online digital education and traditional learning and reported no statistically significant difference between the groups at postintervention (SMD=0.18; 95% CI –0.2 to 0.55; I^2=61%; 292 students; low-quality evidence; Figure 5). Wide CIs around the pooled estimate also included moderate effect size in favor of blended online digital education.

Two studies [31,39] assessed students' attitude toward the outcome (skills acquisition) at postintervention and reported no difference between the groups [31] or favored blended online education over traditional learning with didactic lectures (P=.04) [39].

One study also assessed students' satisfaction with the intervention at postintervention and reported no difference between the groups (P=.61) [39]. One study [34] reported patient-related outcomes (ie, patients' satisfaction) and compared a blend of online modules and small group discussions (ie, blended online digital education) with a control group of small group discussions only. The study reported slightly higher patients' satisfaction scores in the control group than in the blended online digital education (SMD=–0.43; 95% CI –0.73 to –0.13). None of the studies reported on the adverse effects or cost outcomes.

Figure 4. Forest plot of studies comparing blended digital education with traditional learning for postintervention skill outcome. IV: inverse variance; random: random effect model; DE: digital education.

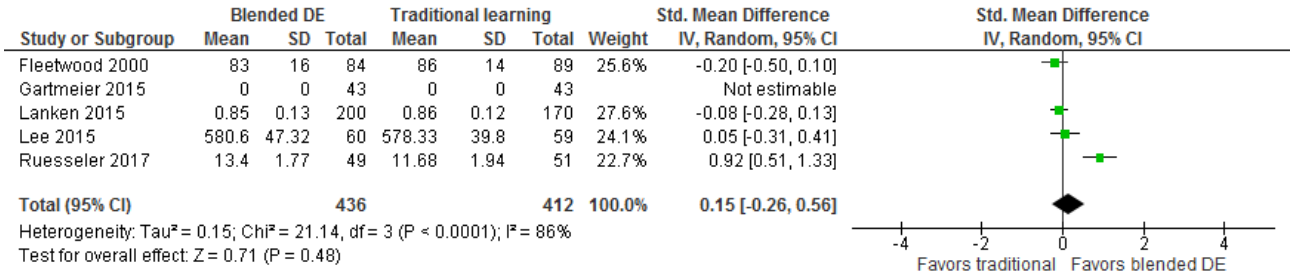
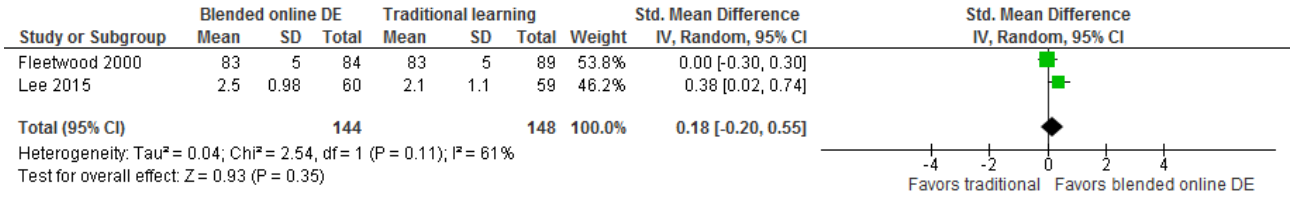


Figure 5. Forest plot of studies comparing blended online digital education with traditional learning for postintervenion knowledge outcome. IV: inverse variance; random: random effect model; DE: digital education.



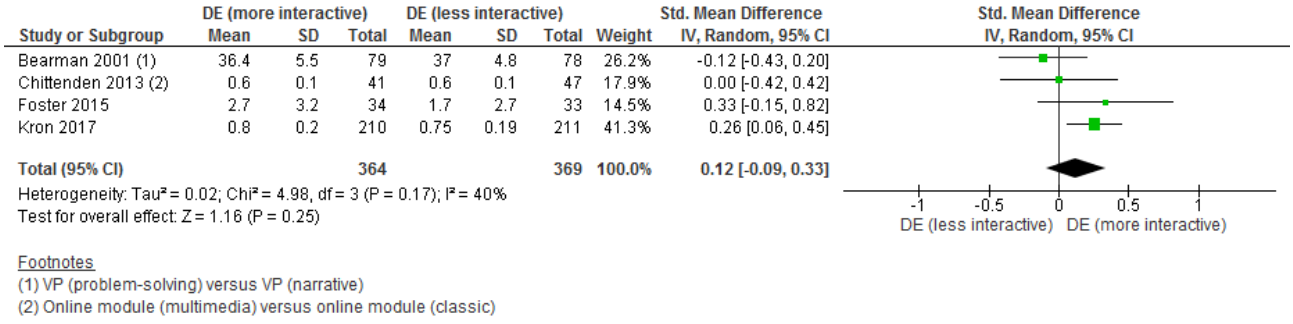
XSL•FO
RenderX

Digital Education (More Interactive) Versus Digital Education (Less Interactive)

Four studies [28,29,35,38] compared the effectiveness of more and less interactive digital education and assessed postintervention skills, attitudes, and satisfaction. More interactive forms of digital education (ie, problem solving, VP simulation, and online multimedia modules) reported similar effectiveness or no difference in postintervention skills compared to less interactive forms of digital education (ie, narrative virtual patient simulation, online video-based learning, and classic online modules) (SMD=0.12, 95% CI –0.09 to 0.33, I^2=40%, 4 studies [893 students], moderate-quality evidence; Figure 6).

One study [38] assessed students' attitude towards the intervention and reported moderate beneficial effect on postintervention attitude scores in the VP group compared to the online module group (SMD=0.71; 95% CI: 0.51-0.91). One study [35] assessed students' satisfaction and reported that students were more satisfied with VP simulation than the online-based video module (*P*=.007). None of the studies reported on knowledge, adverse effects, patient outcomes, or cost outcomes.

Figure 6. Forest plot of studies comparing digital education (more interactive) with digital education (less interactive) for postintervention skills outcome. IV: inverse variance; random: random effect model; DE: digital education.



Discussion

Principal Findings

This systematic review assessed the effectiveness of digital education on medical students' communication skills compared to traditional learning or other forms of digital education. We summarized and critically evaluated evidence for the effectiveness of digital education for medical students' communication skills training. Twelve studies with 2101 medical students met the eligibility criteria. We found low-quality evidence with wide CIs and high heterogeneity, showing no statistically significant difference between digital education and traditional learning in terms of communication skills. Blended digital education seems to be at least as effective as and potentially more effective than traditional learning for communication skills and knowledge. We also found no difference in postintervention skills between more and less interactive forms of digital education. Data on attitudes and satisfaction were limited and mixed. No study reported adverse or unintended effects of digital education nor conducted an economic evaluation. The majority of the studies (N=9, 75%) had a high risk of bias. The quality of evidence ranged from moderate to very low due to the study limitations, inconsistency, and indirectness (Multimedia Appendices 2-4).

The included studies differed considerably in terms of intervention, comparators, and outcome measures used, showing a wide scope of potential for the use of digital education in communication skills training for medical students. However, limited primary studies consisting of data with high heterogeneity prevent us from drawing strong conclusions on the topic. Furthermore, seven (58.3%) of the included studies failed to provide details of sample size or power calculations [28,29,32,35-37,39]. The included studies may have therefore been underpowered and unable to detect change in learning outcomes. Finally, the effect sizes were typically small. Other limitations pertained to the risk of bias. Overall, four of the 12 included studies (33.3%) were judged to have a high risk of bias in at least one domain.

The included evidence has some limitations. First, most of the studies were conducted in high-income countries (except one study that was conducted in China), which might further limit the transferability or applicability of the evaluated evidence in low- and middle-income countries. Second, the included studies focused only on specific forms of digital education such as online or offline digital education and VP simulation, and there is a need to explore the effectiveness of other forms of digital education such as virtual reality, serious gaming, mobile learning, and massive open online courses on the topic. Third, all included studies assessed short-term effectiveness of the interventions (ie, assessed effectiveness immediately after the intervention), and there is a need to assess long-term effectiveness of interventions through aspects such as knowledge retention and skills retention at 3-month or 6-month follow-ups. Lastly, the included studies mostly evaluated skills outcome, and there is limited evidence for other outcomes such as knowledge, attitude, satisfaction, adverse or untoward effects of the intervention, and patient and cost-related outcomes.

Implications for Future Research

We identified the need for further, more methodologically sound research that may lead to more conclusive findings. Studies identified in this review have many significant methodological weaknesses, from inadequate power to unclear theoretical underpinnings; insufficient description of educational interventions (complexity, duration and intensity); uncertainty of what constitutes a change (compared with baseline); little,

XSL•FO
RenderX

if any, description of technical features; skills retention (follow-up); and comparability of the content delivered digitally or traditionally. The use of validated and reliable measurement tools is paramount to advancing the field [40], as its transparent description on the level of trialists' involvement in instructions, outcome(s) in the background, usability testing, and data protection policies could affect the results of the outcomes. Other important factors that need further research include the availability of infrastructure, financial incentives for learners, previous experience in digital education, barriers or facilitators, cost evaluation, fidelity, adverse effects, and access to power supply. Finally, incorporation of evidence from low- and middle-income countries should increase generalizability and applicability in those settings.

Strengths and Limitations of the Review

Strengths of this study include comprehensive searches with no language limitations and robust screening, data extraction, risk of bias assessments, and a critical appraisal of the evidence. Nevertheless, some limitations must be acknowledged while interpreting the results of this systematic review. There was a considerable degree of methodological and clinical heterogeneity in pooled analyses, and the applicability of evaluated evidence might be limited due to high heterogeneity. Additionally, most of the included studies (92%) reported postintervention data, and we could not calculate pre-post intervention change scores. We also assumed that baseline characteristics including measure scores were adjusted before randomization. Finally, we were unable to obtain additional information from the study authors in six studies that reported mixed participants and mixed results.

Conclusions

The findings from this review suggest that digital education (standalone or blended with traditional learning) could be as effective as traditional learning (ie, didactic lectures, groups discussions, role play, or oral feedback) in improving postintervention communication skills for medical students. Similarly, more interactive forms of digital education have similar effectiveness or skills outcome compared to less interactive forms of digital education in terms of participant' skills. The overall risk of bias was high, and the quality of evidence ranged from moderate to very low for the reported outcomes. There is a need for further research assessing long-term effectiveness including knowledge or skills retention, other outcomes such as patient-related outcomes, and cost-effectiveness as well as other forms of digital education for medical students' communication skills training.

Acknowledgments

This review is conducted in collaboration with the Health Workforce Department, World Health Organization. We would also like to thank Mr Carl Gornitzki, Ms GunBrit Knutssön, and Mr Klas Moberg from the University Library, Karolinska Institutet, Sweden, for developing the search strategy and the peer reviewers for their comments. We gratefully acknowledge funding from Nanyang Technological University Singapore, Singapore, for e-learning for health professionals education grant. We would like to thank Associate Professor Josip Car and Dr James Campbell for providing valuable inputs on the manuscript.

Authors' Contributions

LC, JoC, and PP conceived the idea for the review. SP, PP and BMK wrote the review. LC provided methodological guidance and critically revised the review. JoC and JaC provided insightful comments on the review.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Medline (Ovid) Search Strategy.
[PDF File (Adobe PDF File), 28KB - jmir_v21i8e12967_app1.pdf]

Multimedia Appendix 2
Summary of the findings table for the effects of digital health education on communication skills.
[PDF File (Adobe PDF File), 15KB - jmir_v21i8e12967_app2.pdf]

Multimedia Appendix 3
Summary of findings table for the effects of blended digital health education on communication skills.
[PDF File (Adobe PDF File), 94KB - jmir_v21i8e12967_app3.pdf]

Multimedia Appendix 4
Summary of findings table for the effects of digital education (more interactive) compared to digital education (less interactive) on communication skills.
[PDF File (Adobe PDF File), 91KB - jmir_v21i8e12967_app4.pdf]

XSL•FO
RenderX

References

1. Wong SYS, Lee A. Communication Skills and Doctor Patient Relationship. Hong Kong Med Diary 2006 Mar;11(3):7-9.
2. Robinson JD, Hoover DR, Venetis MK, Kearney TJ, Street RL. Consultations between patients with breast cancer and surgeons: a pathway from patient-centered communication to reduced hopelessness. J Clin Oncol 2013 Jan 20;31(3):351-358 [FREE Full text] [doi: 10.1200/JCO.2012.44.2699] [Medline: 23233706]
3. Hasan S. A tool to teach communication skills to pharmacy students. Am J Pharm Educ 2008 Jun 15;72(3):67 [FREE Full text] [doi: 10.5688/aj720367] [Medline: 18698394]
4. Martínez Otero JM. Patients' autonomy and right to the information: from the recognition of rights, to the loss of confidence. Reflections on the new regulation about patients' rights at the end of life [in Spanish]. Cuad Bioet 2012;23(77):151-167 [FREE Full text] [Medline: 22548664]
5. Strasser F, Cerny T. Dealing with ethical dilemmas in oncological communication involving issues of culture and autonomy. Oncology (Williston Park) 2012 Jan;26(1):56, 62, 64 [FREE Full text] [Medline: 22393798]
6. King A, Hoppe RB. "Best practice" for patient-centered communication: a narrative review. J Grad Med Educ 2013 Sep;5(3):385-393 [FREE Full text] [doi: 10.4300/JGME-D-13-00072.1] [Medline: 24404300]
7. Bradley CT, Brasel KJ. Core competencies in palliative care for surgeons: interpersonal and communication skills. Am J Hosp Palliat Care 2007;24(6):499-507. [doi: 10.1177/1049909107310063] [Medline: 18182636]
8. Maeland JG. Changing patients' health behavior--consultation and physician-patient relationship [in Norwegian]. Tidsskr Nor Laegeforen 1993 Jan 10;113(1):47-50. [Medline: 8424251]
9. Berlin L. Informing patients about risks and benefits of radiology examinations utilizing ionizing radiation: a legal and moral dilemma. J Am Coll Radiol 2011 Nov;8(11):742-743. [doi: 10.1016/j.jacr.2011.06.024] [Medline: 22051453]
10. Gramma R, Parvu A, Enache A, Roman G, Ioan B. Truth or lie--some ethical dilemmas in the communication of a severe diagnosis. Rev Med Chir Soc Med Nat Iasi 2013;117(1):172-182. [Medline: 24505911]
11. General Medical Council. 2009. Domain 3 - Communication partnership and teamwork URL: https://www.gmc-uk.org/ethical-guidance/ethical-guidance-for-doctors/good-medical-practice/domain-3---communication-partnership-and-teamwork [accessed 2019-08-08]
12. Smith S, Hanson JL, Tewksbury LR, Christy C, Talib NJ, Harris MA, et al. Teaching patient communication skills to medical students: a review of randomized controlled trials. Eval Health Prof 2007 Mar;30(1):3-21. [doi: 10.1177/0163278706297333] [Medline: 17293605]
13. Choudhary A, Gupta V. Teaching communications skills to medical students: Introducing the fine art of medical practice. Int J Appl Basic Med Res 2015 Aug;5(Suppl 1):S41-S44 [FREE Full text] [doi: 10.4103/2229-516X.162273] [Medline: 26380210]
14. George PP, Papachristou N, Belisario JM, Wang W, Wark PA, Cotic Z, et al. Online eLearning for undergraduates in health professions: A systematic review of the impact on knowledge, skills, attitudes and satisfaction. J Glob Health 2014 Jun;4(1):010406 [FREE Full text] [doi: 10.7189/jogh.04.010406] [Medline: 24976965]
15. Kyaw BM, Saxena N, Posadzki P, Vseteckova J, Nikolaou CK, George PP, et al. Virtual Reality for Health Professions Education: Systematic Review and Meta-Analysis by the Digital Health Education Collaboration. J Med Internet Res 2019 Jan 22;21(1):e12959 [FREE Full text] [doi: 10.2196/12959] [Medline: 30668519]
16. George PP, Zhabenko O, Kyaw BM, Antoniou P, Posadzki P, Saxena N, et al. Online Digital Education for Postregistration Training of Medical Doctors: Systematic Review by the Digital Health Education Collaboration. J Med Internet Res 2019 Feb 25;21(2):e13269 [FREE Full text] [doi: 10.2196/13269] [Medline: 30801252]
17. Posadzki P, Bala MM, Kyaw BM, Semwal M, Divakar U, Koperny M, et al. Offline Digital Education for Postregistration Health Professions: Systematic Review and Meta-Analysis by the Digital Health Education Collaboration. J Med Internet Res 2019 Apr 24;21(4):e12968 [FREE Full text] [doi: 10.2196/12968] [Medline: 31017584]
18. Ruiz JG, Smith M, van Zuilen MH, Williams C, Mintzer MJ. The educational impact of a computer-based training tutorial on dementia in long term care for licensed practice nursing students. Gerontol Geriatr Educ 2006;26(3):67-79. [doi: 10.1300/J021v26n03_05] [Medline: 16446272]
19. Choules AP. The use of elearning in medical education: a review of the current situation. Postgrad Med J 2007 Apr;83(978):212-216 [FREE Full text] [doi: 10.1136/pgmj.2006.054189] [Medline: 17403945]
20. Higgins JPT, Green S. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0. 2011. URL: http://handbook-5-1.cochrane.org/ [accessed 2019-08-08]
21. Car J, Carlstedt-Duke J, Tudor Car L, Posadzki P, Whiting P, Zary N, Digital Health Education Collaboration. Digital Education in Health Professions: The Need for Overarching Evidence Synthesis. J Med Internet Res 2019 Feb 14;21(2):e12913 [FREE Full text] [doi: 10.2196/12913] [Medline: 30762583]
22. Schwan S, Riempp R. The cognitive benefits of interactive videos: learning to tie nautical knots. Learning and Instruction 2004 Jun;14(3):293-305. [doi: 10.1016/j.learninstruc.2004.06.005]
23. Puffer S, Torgerson D, Watson J. Evidence for risk of bias in cluster randomised trials: review of recent trials published in three general medical journals. BMJ 2003 Oct 04;327(7418):785-789 [FREE Full text] [doi: 10.1136/bmj.327.7418.785] [Medline: 14525877]

XSL•FO
RenderX

24. Cook DA, Hatala R, Brydges R, Zendejas B, Szostek JH, Wang AT, et al. Technology-enhanced simulation for health professions education: a systematic review and meta-analysis. JAMA 2011 Sep 07;306(9):978-988. [doi: 10.1001/jama.2011.1234] [Medline: 21900138]
25. RevMan. Review Manager 5.3. Copenhagen: The Nordic Cochrane Centre: The Cochrane Collaboration; 2014. URL: https://community.cochrane.org/help/tools-and-software/revman-5 [accessed 2019-08-08]
26. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JPA, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. J Clin Epidemiol 2009 Oct;62(10):e1-34 [FREE Full text] [doi: 10.1016/j.jclinepi.2009.06.006] [Medline: 19631507]
27. Schünemann HJ, Oxman AD, Higgins JP, Vist GE, Glasziou P, Guyatt GH. Cochrane Handbook for Systematic Reviews of Interventions. 2011. Presenting Results and 'Summary of Findings' Tables URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/9780470712184.ch11 [accessed 2019-08-15]
28. Bearman M, Cesnik B, Liddell M. Random comparison of 'virtual patient' models in the context of teaching clinical communication skills. Med Educ 2001 Sep;35(9):824-832. [Medline: 11555219]
29. Chittenden E, Anderson W, Lai C, O'Sullivan P. An evaluation of interactive web-based curricula for teaching code status discussions. J Palliat Med 2013 Sep;16(9):1070-1073. [doi: 10.1089/jpm.2012.0611] [Medline: 23937063]
30. Gartmeier M, Bauer J, Fischer MR, Hoppe-Seyler T, Karsten G, Kiessling C, et al. Fostering professional communication skills of future physicians and teachers: effects of e-learning with video cases and role-play. Instr Sci 2015 Apr 7;43(4):443-462. [doi: 10.1007/s11251-014-9341-6]
31. Lanken PN, Novack DH, Daetwyler C, Gallop R, Landis JR, Lapin J, et al. Efficacy of an internet-based learning module and small-group debriefing on trainees' attitudes and communication skills toward patients with substance use disorders: results of a cluster randomized controlled trial. Acad Med 2015 Mar;90(3):345-354. [doi: 10.1097/ACM.0000000000000506] [Medline: 25295964]
32. Lee A, Mader E, Morley C. Teaching cross-cultural communication skills online: a multi-method evaluation. Fam Med 2015 Apr;47(4):302-308 [FREE Full text] [Medline: 25853602]
33. Deladisma AM, Cohen M, Stevens A, Wagner P, Lok B, Bernard T, Association for Surgical Education. Do medical students respond empathetically to a virtual patient? Am J Surg 2007 Jun;193(6):756-760. [doi: 10.1016/j.amjsurg.2007.01.021] [Medline: 17512291]
34. Fleetwood J, Vaught W, Feldman D, Gracely E, Kassutto Z, Novack D. MedEthEx Online: a computer-based learning program in medical ethics and communication skills. Teach Learn Med 2000;12(2):96-104. [doi: 10.1207/S15328015TLM1202_7] [Medline: 11228685]
35. Foster A, Chaudhary N, Murphy J, Lok B, Waller J, Buckley P. The Use of Simulation to Teach Suicide Risk Assessment to Health Profession Trainees-Rationale, Methodology, and a Proof of Concept Demonstration with a Virtual Patient. Acad Psychiatry 2015 Dec;39(6):620-629. [doi: 10.1007/s40596-014-0185-9] [Medline: 25026950]
36. Kaltman S, Talisman N, Pennestri S, Syverson E, Arthur P, Vovides Y. Using Technology to Enhance Teaching of Patient-Centered Interviewing for Early Medical Students. Simul Healthc 2018 Jun;13(3):188-194. [doi: 10.1097/SIH.0000000000000304] [Medline: 29771814]
37. Ruesseler M, Sterz J, Bender B, Hoefer S, Walcher F. The effect of video-assisted oral feedback versus oral feedback on surgical communicative competences in undergraduate training. Eur J Trauma Emerg Surg 2017 Aug;43(4):461-466. [doi: 10.1007/s00068-016-0734-x] [Medline: 27730252]
38. Kron FW, Fetters MD, Scerbo MW, White CB, Lypson ML, Padilla MA, et al. Using a computer simulation for teaching communication skills: A blinded multisite mixed methods randomized controlled trial. Patient Educ Couns 2017 Apr;100(4):748-759 [FREE Full text] [doi: 10.1016/j.pec.2016.10.024] [Medline: 27939846]
39. Tang F, Chen C, Zhu Y, Zuo C, Zhong Y, Wang N, et al. Comparison between flipped classroom and lecture-based classroom in ophthalmology clerkship. Med Educ Online 2017;22(1):1395679 [FREE Full text] [doi: 10.1080/10872981.2017.1395679] [Medline: 29096591]
40. Law GC, Apfelbacher C, Posadzki PP, Kemp S, Tudor Car L. Choice of outcomes and measurement instruments in randomised trials on eLearning in medical education: a systematic mapping review protocol. Syst Rev 2018 May 17;7(1):75 [FREE Full text] [doi: 10.1186/s13643-018-0739-0] [Medline: 29776434]

Abbreviations

CG: control group
cRCT: cluster randomized controlled trial
GRADE: Grading of Recommendations, Assessment, Development and Evaluations
IG: intervention group
MCQ: multiple-choice questionnaire
OSCE: Objective Structured Clinical Examination
PACT: Problem Affect Concern Treatment
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

XSL•FO
RenderX

RCT: randomized controlled trial
SMD: standardized mean difference
SP: standardized patient
VP: virtual patient

Edited by A Marusic; submitted 21.12.18; peer-reviewed by M Gartmeier, V Dogas, M Vidak, C Bedard, A Giordano; comments to author 17.01.19; revised version received 30.01.19; accepted 10.07.19; published 27.08.19.

Please cite as:
Kyaw BM, Posadzki P, Paddock S, Car J, Campbell J, Tudor Car L
Effectiveness of Digital Education on Communication Skills Among Medical Students: Systematic Review and Meta-Analysis by the Digital Health Education Collaboration
J Med Internet Res 2019;21(8):e12967
URL: http://www.jmir.org/2019/8/e12967/
doi:10.2196/12967
PMID:31456579

©Bhone Myint Kyaw, Pawel Posadzki, Sophie Paddock, Josip Car, James Campbell, Lorainne Tudor Car. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 27.08.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Corrigenda and Addenda

Correction: The Search for Consumers of Web-Based Raw DNA Interpretation Services: Using Social Media to Target Hard-to-Reach Populations

Tiernan J Cahill[1], MA; Blake Wertz[1], MA; Qiankun Zhong[1], MA; Andrew Parlato[1], MA; John Donegan[1], MA, MBA; Rebecca Forman[1], MA; Supriya Manot[1], MA; Tianyi Wu[1], MA; Yazhu Xu[1], MA; James J Cummings[1], PhD; Tricia Norkunas Cunningham[2], MPH, MA; Catharine Wang[2], PhD, MSc

[1]Division of Emerging Media, Boston University College of Communication, Boston, MA, United States

[2]Department of Community Health Sciences, Boston University School of Public Health, Boston, MA, United States

Corresponding Author:
Catharine Wang, PhD, MSc
Department of Community Health Sciences
Boston University School of Public Health
801 Massachusetts Avenue
Boston, MA, 02118
United States
Phone: 1 617 358 1475
Email: clwang@bu.edu

Related Article:

Correction of: https://www.jmir.org/2019/7/e12980/

(J Med Internet Res 2019;21(8):e15735) doi:10.2196/15735

"The Search for Consumers of Web-Based Raw DNA Interpretation Services: Using Social Media to Target Hard-to-Reach Populations" (J Med Internet Res 2019;21(7):e12980) contained a typographical error in the Results section within the subsection "Demographic Comparisons Across Platforms and Tracking Methods" under the heading "Differences Between Tracking Methods". The proportion of female respondents among those recruited using click-tracking was reported as "(43/49, 73%)" but should have been "(43/59, 73%)" to be consistent with Figure 6. This error has been corrected and the sentence now reads as follows:

> *Female respondents made up the majority in both cases, but were more prevalent among those recruited using conversion-tracking (124/145, 85.5%), compared with those recruited using click-tracking (43/59, 73%).*

Additionally, incorrect sample sizes were noted in the second paragraph of the Discussion. The value for Facebook was listed as n=168 but should have been n=159; the value for Twitter was listed as n=170 but should have been n=167; and the value for Reddit was listed as n=114 but should have been n=112 to be consistent with values already reported in the Results section. The errors have been corrected and the sentence now reads as follows:

> *Nearly identical sample sizes were obtained via paid Facebook (n=159) and Twitter (n=167) advertising, as well as the sample obtained via unpaid posting on Reddit (n=112).*

The corrections will appear in the online version of the paper on the JMIR website on August 13, 2019, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article also has been resubmitted to those repositories.

XSL•FO
RenderX

Submitted 01.08.19; this is a non–peer-reviewed article;accepted 05.08.19; published 13.08.19.

Please cite as:
Cahill TJ, Wertz B, Zhong Q, Parlato A, Donegan J, Forman R, Manot S, Wu T, Xu Y, Cummings JJ, Norkunas Cunningham T, Wang C
Correction: The Search for Consumers of Web-Based Raw DNA Interpretation Services: Using Social Media to Target Hard-to-Reach Populations
J Med Internet Res 2019;21(8):e15735
URL: http://www.jmir.org/2019/8/e15735/
doi:10.2196/15735
PMID:31411145

©Tiernan J Cahill, Blake Wertz, Qiankun Zhong, Andrew Parlato, John Donegan, Rebecca Forman, Supriya Manot, Tianyi Wu, Yazhu Xu, James J Cummings, Tricia Norkunas Cunningham, Catharine Wang. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 13.08.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Corrigenda and Addenda

Figure Correction: Medium-Term Effects of a Tailored Web-Based Parenting Intervention to Reduce Adolescent Risk of Depression and Anxiety: 12-Month Findings From a Randomized Controlled Trial

Marie Bee Hui Yap[1,2], MPsych, PhD; Mairead C Cardamone-Breen[1], DPsych; Ronald M Rapee[3], PhD; Katherine A Lawrence[1], DPsych; Andrew J Mackinnon[4], PhD; Shireen Mahtani[1], DPsych; Anthony F Jorm[2], PhD, DSc

[1]School of Psychological Sciences and Turner Institute for Brain and Mental Health, Monash University, Melbourne, Australia

[2]Melbourne School of Population and Global Health, University of Melbourne, Melbourne, Australia

[3]Centre for Emotional Health, Macquarie University, Sydney, Australia

[4]Black Dog Institute, University of New South Wales, Sydney, Australia

Corresponding Author:
Marie Bee Hui Yap, MPsych, PhD
School of Psychological Sciences and Turner Institute for Brain and Mental Health
Monash University
18 Innovation Walk
Wellington Road, Clayton
Melbourne, 3168
Australia
Phone: 61 0399051250
Email: marie.yap@monash.edu

Related Article:

Correction of: https://www.jmir.org/2019/8/e13628

(J Med Internet Res 2019;21(8):e15915) doi:10.2196/15915

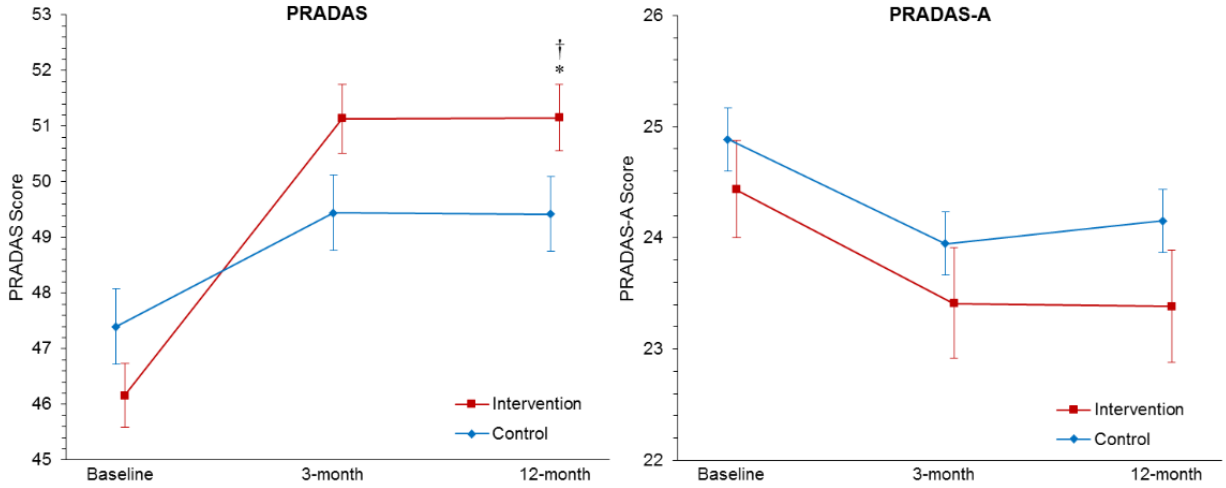
In "Medium-Term Effects of a Tailored Web-Based Parenting Intervention to Reduce Adolescent Risk of Depression and Anxiety: 12-Month Findings From a Randomized Controlled Trial" (J Med Internet Res 2019;21(8):e13628), Figure 2 was erroneously published with one graph missing from the image. The graph of Parenting to Reduce Adolescent Depression and Anxiety Scale (PRADAS) was not shown. The image has now been replaced with a version that includes two graphs:

1. Estimated marginal means for Parenting to Reduce Adolescent Depression and Anxiety Scale (PRADAS)
2. Estimated marginal means for Parenting to Reduce Adolescent Depression and Anxiety Scale—Adolescent report (PRADAS-A)

The correction will appear in the online version of the paper on the JMIR website on August 29, 2019, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article also has been resubmitted to those repositories.

XSL•FO
RenderX

Figure 2. Estimated marginal means for Parenting to Reduce Adolescent Depression and Anxiety Scale (PRADAS) and PRADAS—Adolescent (PRADAS-A) Report scores at baseline, postintervention (3-months postbaseline), and 12-month follow-up, estimated under the group × measurement-occasion mixed model. Error bars represent SEs. Higher scores on the PRADAS and PRADAS—Adolescent Report indicate greater concordance with the parenting guidelines (ie, more protective parenting factors and fewer parenting risk factors). Planned contrast of interaction (baseline to 12 months) was significant, $P<.001$. Pairwise comparison of group difference at 12-month follow-up was significant, $P=.04$.



Submitted 18.08.19; this is a non–peer-reviewed article;accepted 19.08.19; published 29.08.19.

Please cite as:
Yap MBH, Cardamone-Breen MC, Rapee RM, Lawrence KA, Mackinnon AJ, Mahtani S, Jorm AF
Figure Correction: Medium-Term Effects of a Tailored Web-Based Parenting Intervention to Reduce Adolescent Risk of Depression and Anxiety: 12-Month Findings From a Randomized Controlled Trial
J Med Internet Res 2019;21(8):e15915
URL: http://www.jmir.org/2019/8/e15915/
doi:10.2196/15915
PMID:31469076

©Marie Bee Hui Yap, Mairead C Cardamone-Breen, Ronald M Rapee, Katherine A Lawrence, Andrew J Mackinnon, Shireen Mahtani, Anthony F Jorm. Originally published in the Journal of Medical Internet Research (http://www.jmir.org), 29.08.2019. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research, is properly cited. The complete bibliographic information, a link to the original publication on http://www.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org

XSL•FO
RenderX