JMIR Formative Research

Impact Factor (2023): 2.0
Volume 3 (2019), Issue 4   ISSN 2561-326X   Editor in Chief: Amaryllis Mavragani, PhDc

Contents

Original Papers

An Electronic Health Tool to Prepare for the First Orthopedic Consultation: Use and Usability Study (e13577)
Aniek Claassen, Thea Vliet Vlieland, Vincent Busch, Henk Schers, Frank van den Hoogen, Cornelia van den Ende. 4

An Internet-Based Therapeutic Tool for American Indian/Alaska Native Adults With Posttraumatic Stress Disorder: User Testing and Developmental Feasibility Study (e13682)
Vanessa Hiratsuka, Laurie Moore, Jaedon Avey, Lisa Dirks, Barbara Beach, Denise Dillard, Douglas Novins. 15

Use of Patient-Reported Data to Match Depression Screening Intervals With Depression Risk Profiles in Primary Care Patients With Diabetes: Development and Validation of Prediction Models for Major Depression (e13610)
Haomiao Jin, Shinyi Wu. 33

Development of a Digital Decision Support Tool to Aid Participation of Children With Disabilities in Pediatric Rehabilitation Services: Explorative Qualitative Study (e14493)
Elin Vinblad, Ingrid Larsson, Maria Lönn, Emma Olsson, Jens Nygren, Petra Svedberg. 45

Adapting Mobile and Wearable Technology to Provide Support and Monitoring in Rehabilitation for Dementia: Feasibility Case Series (e12346)
Julia Thorpe, Birgitte Forchhammer, Anja Maier. 67

Passive Monitoring of Short-Acting Beta-Agonist Use via Digital Platform in Patients With Chronic Obstructive Pulmonary Disease: Quality Improvement Retrospective Analysis (e13286)
Jessica Chen, Leanne Kaye, Michael Tuffli, Meredith Barrett, Shelanda Jones-Ford, Tina Shenouda, Rahul Gondalia, Kelly Henderson, Veronica Combs, David Van Sickle, David Stempel. 85

Internet-Based Cognitive Behavioral Therapy via Videoconference for Patients With Bulimia Nervosa and Binge-Eating Disorder: Pilot Prospective Single-Arm Feasibility Trial (e15738)
Sayo Hamatani, Noriko Numata, Kazuki Matsumoto, Chihiro Sutoh, Hanae Ibuki, Keiko Oshiro, Mari Tanaka, Rikukage Setsu, Yohei Kawasaki, Yoshiyuki Hirano, Eiji Shimizu. 94

A Promising Food-Coaching Intervention Program to Achieve Optimal Gestational Weight Gain in Overweight and Obese Pregnant Women: Pilot Randomized Controlled Trial of a Smartphone App (e13013)
Ling-Jun Li, Izzuddin Aris, Wee Han, Kok Tan. 108

Accuracy of a Chatbot (Ada) in the Diagnosis of Mental Disorders: Comparative Case Study With Lay and Expert Users (e13863)
Stefanie Jungmann, Timo Klan, Sebastian Kuhn, Florian Jungmann. 118

Virtual Reality as a Therapy Adjunct for Fear of Movement in Veterans With Chronic Pain: Single-Arm Feasibility Study (e11266)

Christopher Fowler, Lisa Ballistrea, Kerry Mazzone, Aaron Martin, Howard Kaplan, Kevin Kip, Katherine Ralston, Jennifer Murphy, Sandra Winkler. 128

Web-Based Health Information Seeking Among Students at Kuwait University: Cross-Sectional Survey Study (e14327)

Hasan Ashkanani, Rabab Asery, Fajer Bokubar, Noor AlAli, Shahad Mubarak, Ali Buabbas, Abdullah Almajran. 143

A Mobile Phone App for the Provision of Personalized Food-Based Information in an Eating-Out Situation: Development and Initial Evaluation (e12966)

Katherine Appleton, Jeff Bray, Sarah Price, Gernot Liebchen, Nan Jiang, Ioannis Mavridis, Laure Saulais, Agnès Giboreau, Federico Perez-Cueto, Rebecca Coolen, Manfred Ronge, Heather Hartwell. 157

Formative Evaluation to Build an Online Parenting Skills and Youth Drug Prevention Program: Mixed Methods Study (e14906)

Lawrence Scheier, Karol Kumpfer, Jaynie Brown, QingQing Hu. 172

Feasibility and Acceptability of an Interactive Mental Well-Being Intervention for People With Intellectual Disabilities: Pilot Mixed Methods Study (e15190)

Leen Vereenooghe, Kristian Westermann. 186

The Role of Age, Education, and Digital Health Literacy in the Usability of Internet-Based Cognitive Behavioral Therapy for Chronic Pain: Mixed Methods Study (e12883)

Rosalie van der Vaart, Dorine van Driel, Kristel Pronk, Suzan Paulussen, Selma te Boekhorst, Judith Rosmalen, Andrea Evers. 198

A Patient Decision Aid App for Patients With Chronic Kidney Disease: Questionnaire Study (e13786)

Signe Therkildsen, Linda Hansen, Laura Jensen, Jeanette Finderup. 209

A Behavioral Activation Mobile Health App for Smokers With Depression: Development and Pilot Evaluation in a Single-Arm Trial (e13728)

Jaimee Heffner, Noreen Watson, Edit Serfozo, Kristin Mull, Laura MacPherson, Melissa Gasser, Jonathan Bricker. 218

Acceptability and Use of Interactive Voice Response Mobile Phone Surveys for Noncommunicable Disease Behavioral Risk Factor Surveillance in Rural Uganda: Qualitative Study (e15000)

Charles Ssemugabo, Elizeus Rutebemberwa, Dan Kajungu, George Pariyo, Adnan Hyder, Dustin Gibson. 233

Beliefs Related to Participation in a Large Web-Based Prospective Survey on Diet and Health Among Individuals With a Low Socioeconomic Status: Qualitative Study (e13854)

Mélina Côté, Annie Lapointe, Catherine Laramée, Simone Lemieux, Sophie Desroches, Ariane Belanger-Gravel, Benoît Lamarche. 242

Web-Based Module for the Collection of Electronic Patient-Reported Outcomes in People Living With HIV in Nouvelle Aquitaine, France: Usability Evaluation (e15013)

Diana Barger, Olivier Leleux, Valérie Conte, Vincent Sapparrart, Marie Gapillout, Isabelle Crespel, Marie Erramouspe, Sandrine Delveaux, Linda Wittkop, François Dabis, Fabrice Bonnet. 251

Understanding End-User Perspectives of Mobile Pulmonary Rehabilitation (mPR): Cross-Sectional Survey and Interviews (e15466)

Rosie Dobson, Pauline Herbst, Sarah Candy, Tamzin Brott, Jeffrey Garrett, Gayl Humphrey, Julie Reeve, Merryn Tawhai, Denise Taylor, Jim Warren, Robyn Whittaker. 262

Viewpoint

Understanding Health Behavior Technology Engagement: Pathway to Measuring Digital Behavior Change Interventions (e14052)
Heather Cole-Lewis, Nnamdi Ezeanochie, Jennifer Turgiss. 57

Review

A Comprehensive Framework to Evaluate Websites: Literature Review and Development of GoodWeb (e14372)
Rosalie Allison, Catherine Hayes, Cliodna McNulty, Vicki Young. 273

Original Paper

An Electronic Health Tool to Prepare for the First Orthopedic Consultation: Use and Usability Study

Aniek A O M Claassen[1], PhD, PT; Thea P M Vliet Vlieland[2], MD, PhD; Vincent J J F Busch[3], MD, PhD; Henk J Schers[4], MD, PhD; Frank H J van den Hoogen[1,5], MD, PhD; Cornelia H M van den Ende[1,5], PhD, PT

[1]Department of Rheumatology, Sint Maartenskliniek, Nijmegen, Netherlands
[2]Department of Orthopaedics, Rehabilitation and Physical Therapy, Leiden University Medical Center, Leiden, Netherlands
[3]Department of Orthopaedic Surgery, Sint Maartenskliniek, Nijmegen, Netherlands
[4]Radboud Institute for Health Sciences, Department of Primary and Community Care, Radboud University Medical Center, Nijmegen, Netherlands
[5]Department of Rheumatology, Radboud University Medical Center, Nijmegen, Netherlands

Corresponding Author:
Aniek A O M Claassen, PhD, PT
Department of Rheumatology
Sint Maartenskliniek
PO Box 9011
Nijmegen, 9700 GM
Netherlands
Phone: 31 243659145
Fax: 31 243659204
Email: a.claassen@maartenskliniek.nl

Abstract

Background: The use of electronic health (eHealth) technology to prepare patients with hip or knee osteoarthritis (OA) for their first orthopedic consultation seems promising. Exploration of the use and usability of an educational eHealth tool may highlight potential modifications that could increase patient engagement and effectiveness.

Objective: This study aimed to (1) identify the use and usability of a stand-alone educational eHealth tool for patients with suspected hip or knee OA, (2) explore whether the recorded questions in the eHealth tool were in line with an existing widely used question prompt list, and (3) investigate whether user characteristics are related to use and usability.

Methods: We used data from 144 participants in the intervention group of a randomized controlled trial, who were asked to use the educational eHealth tool to prepare for their upcoming first orthopedic consultation. We defined users and nonusers based on whether they had opened the tool at least once. Users were characterized as active or superficial depending on the extent of their use of the tool. The recorded questions for the consultation preparation were categorized into themes fitting 3 predefined questions or in a remaining category. Usability was measured using the System Usability Scale (SUS, 0-100). Data were collected including the patient demographic and clinical characteristics, knowledge of OA, and internet and smartphone usage in daily life. The characteristics associated with users and nonusers were analyzed using a multivariable logistic regression analysis.

Results: A total of 116/144 (80.6%) participants used the educational eHealth tool, of whom 87/116 (75.0%) were active users. Of the three components of the tool (information, my consultation, and medication), medication was the least used (34%). On the basis of recorded questions of the users, the fourth predefined question could be proposed. The mean (SD) SUS score was 64.8 (16.0). No difference was found between the SUS scores of superficial and active users (mean difference 0.04, 95% CI −7.69 to 7.77). Participants with a higher baseline knowledge of OA (odds ratio [OR] 1.2, 95% CI 1.0 to 1.4) and who used the internet less frequently in their daily life (OR 0.6, 95% CI 0.5 to 0.9) were more likely to use the educational eHealth tool. We found no differences between the demographics and clinical characteristics of the superficial and active users.

Conclusions: Based on the results of this study, it can be concluded that the use of an educational eHealth tool to prepare patients with hip and knee OA for the first orthopedic consultation is feasible. Our results suggest some improvements that should be made to the content of the tool to improve its usability. No clear practical implications were found to support the implementation of the educational eHealth tool in specific subgroups.

Trial Registration: Netherlands Trial Register NTR6262; https://www.trialregister.nl/trial/6262

(JMIR Form Res 2019;3(4):e13577) doi:10.2196/13577

KEYWORDS

osteoarthritis; patient education; consultation; eHealth; smartphone; use

Introduction

Background

Osteoarthritis (OA) is an age-related, degenerative joint disease and one of the most common causes of disability around the world [1]. International guidelines recommend nonsurgical treatments, such as lifestyle education, exercise therapy, weight loss if overweight, and pain medication, as a primary approach to manage hip or knee OA in the early stages [2,3]. Once these conservative treatment options have been adequately tried and failed, or in the case of diagnostic uncertainty, a referral to an orthopedic surgeon should be considered for further diagnostic evaluation and consideration of surgical interventions, for example, a total joint replacement (TJR) [2]. Patients with hip and knee OA often expect action to be taken when referred to an orthopedic surgeon [4], in particular, the planning of a TJR; however, only one-third to a half of referred patients are eligible for a TJR [5,6]. It is therefore conceivable that patients' expectations about the consultation may not always be met, resulting in patients being dissatisfied [7]. A solid preparation for the consultation might help to streamline patient expectations [8].

In general, educational interventions can help patients to be more prepared for a consultation by providing information on treatment options [6] and by assisting patients in reflecting on their own situation (eg, monitoring symptoms or recording medical history) [9-11]. Moreover, the use of self-prepared or provided question prompt lists for patients to ask or questions to be expected from the health care provider can facilitate the exchange of information during consultations [8,12,13]. Previous research has shown that the use of conventional educational tools to prepare patients for consultations and to aid treatment decision making in OA is associated with lower health care costs because it may postpone unnecessary early surgery [11,14].

The growing and emerging opportunities in the use of electronic health (eHealth) can be harnessed to further develop educational interventions with the potential to improve efficiency and lower costs [15]. To contribute to the emerging field of eHealth for OA and to support patients, an educational eHealth tool was developed to help hip and knee OA patients prepare for their first orthopedic consultation. This stand-alone smartphone and Web-based intervention provides information on treatment options for hip and knee OA, the option to prepare for a consultation by preparing questions, and enables patients to monitor their symptoms and medication use. A randomized evaluation of this educational eHealth tool showed that it did not influence patient satisfaction with their consultation, but it did have small effects on patient knowledge of OA and their treatment expectations (data not published yet). These results were less promising than expected; therefore, it seemed important to further explore the actual use of the intervention. Data on the usage of an intervention or its components, and its usability can provide information on potential intervention modifications that encourage engagement and, likely, effectiveness [16].

Objectives

The aim of this study was to identify the use and usability of the aforementioned educational eHealth tool. We therefore describe the user rates of different components of the tool and explore how the preparation component of the application is used (eg, which questions do participants prepare) and whether these questions are in line with an existing widely used question prompt list [17]. Our second aim was to investigate whether certain user characteristics are related to the use and usability of the educational eHealth tool to provide points of support for its implementation.

Methods

Design and Setting

The data for this study were collected as part of a randomized controlled trial (RCT; Dutch Trial Register NTR6262) evaluating the effect of an educational eHealth tool compared with standard care practices, which was carried out between March 2017 and May 2018 at the outpatient department for Orthopedic Surgery at the Sint Maartenskliniek, Nijmegen, the Netherlands. Baseline and follow-up data for the intervention group and data retrieved from the backend of the educational eHealth tool were used in this study. All patients gave their informed consent for participation. The Medical Ethics Committee on Research Involving Human Subjects (CMO) Region Arnhem-Nijmegen (study number 2016-3096) waived ethical approval because it is not required for this type of study under Dutch law.

Participants and Procedure

Patients who had a scheduled visit for a new treatment episode at the outpatient clinic of Orthopedic Surgery at the Sint Maartenskliniek, Nijmegen, were checked for their eligibility. The inclusion criteria were (1) age 18 years or above, (2) the referral letter of general practitioner or the referring specialist mentioning the (suspected) diagnosis of OA in the knee or hip, and (3) no previous visit to the department of Orthopedic Surgery at the Sint Maartenskliniek for that index joint. The exclusion criteria were (1) inability to read or understand Dutch, (2) not possessing a smartphone, computer, or tablet, or (3) not having an email address. Eligible patients were invited to participate through a letter providing information on the study. Patients who were willing to participate received further information about the study by email and were asked to fill in a baseline Web-based questionnaire 2 to 5 weeks before their consultation. Participants who were randomly assigned to the intervention group received an email with personal login details and an information flyer about the installation and use of the educational eHealth tool. The tool could be used during the 2 weeks before the scheduled consultation. One day after their consultation, the participants received a link to a follow-up Web-based questionnaire. All clinical data were collected using the electronic data capture and management program Castor EDC.

Intervention

The educational eHealth tool was developed on the initiative of patients, and in collaboration with patients with OA and health care professionals, using an 8-step method of persuasive design [18,19] followed by an iterative process of development involving 4 cycles of development, user-testing, adaptation, retesting, and finalizing. The tool was available as a mobile app (Android and iOS) and in a Web-based version (Comaxx, digital bureau). The tool consists of 3 parts, *information*, *my consultations*, and *medication*, covering the following functionalities: (1) short facts and information on OA and treatment modalities, based on a stepped-care strategy for OA [20], (2) preparation for the upcoming consultation consisting of predefined questions to answer (eg, "How long do your symptoms exist?" and "Do you have morning stiffness of the joint?"), and space to record additional questions the patient would like to ask the orthopedic surgeon, (3) the option to monitor pain and fatigue during the week before the consultation (4) a list of medication used (eg, dosage), with the option to set reminders for intake, and (5) the option to create a visual timeline with the scheduled consultation, assessments, and preparation. Users could earn 3 achievement awards while using the tool: one when they had scrolled through all information parts, one when an upcoming consultation was detailed (record date, place, reason of consultation and open question list) in the educational eHealth tool, and one when medication use is registered in the tool. For further details on the development and functionalities of the eHealth tool, see Multimedia Appendix 1.

Assessments

Use

On the basis of objective user data extracted from the backend of the educational eHealth tool, participants were classified as a *user* or *nonuser*. *Users* of the educational eHealth tool were defined as *opening the tool at least once*, while *nonusers* were those participants who did not open the educational eHealth tool at all. *Users* were further defined as *active* or *superficial*. If a participant had opened the tool and earned at least one achievement award, they were defined as an *active user*. Participants were defined as *superficial users* if they had used the tool but had not earned any achievements.

Preparation

The questions that patients recorded in the educational eHealth tool in preparation for their consultation were extracted from the backend of the tool. The questions were categorized into 3 themes, based on the 3 good preparation questions outlined by Shepherd et al: "What are my options?," "What are the possible benefits and harms of those options?," and "How likely are each of the benefits and harms to happen to me?" [18]. The Dutch versions of these 3 questions are implemented in several hospitals in the Netherlands. In this study, we chose to use the Dutch-implemented version, in which the last question is slightly adjusted and could better be translated as "What does this mean in my situation?" If the question did not fit 1 of the 3 themes, it was put in a *remaining* category, which was subsequently further defined based on the nature of questions assigned to that category. This categorization was performed independently by a research assistant and a researcher. Any disagreement was resolved by discussion, and if a consensus was still not reached, the third researcher was consulted.

Usability

In the follow-up questionnaire, the usability of the educational eHealth tool was assessed using the 10-item System Usability Scale (SUS) [21]. The items, which covered complexity, ease of use, and willingness to use the tool, among other factors, were scored on a 5-point Likert scale (*strongly disagree* to *strongly agree*). The final scores for the SUS could range from 0 to 100, where higher scores indicate better usability. The SUS is thought to be a robust, valid, and versatile questionnaire [22]. The extent to which patients were satisfied with the tool was measured by asking the patients to rate their satisfaction on a Numeric Rating Scale (NRS) ranging from 0 to 10, with higher scores indicating a higher satisfaction.

Demographic and Clinical Patient Characteristics

Demographic data were collected on the gender, age, body mass index (BMI), marital status, level of education, and work status of the patients. Clinical characteristics were collected on the OA location (hip or knee) and duration of symptoms (years). Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) scores were used for pain and function [23]. These scores were derived from completed Dutch Knee/Hip injury and Osteoarthritis Outcome Score questionnaires [24,25], and were presented as standardized scores (0-100), with higher scores indicating more pain and worse function. Fatigue during the past week was assessed on the NRS from 0 (*no fatigue*) to 100 (*extreme fatigue*). To record the use of pain medication, participants were asked (yes/no) whether they used pain medication in the past 3 months for their hip or knee symptoms.

Patient knowledge of OA (treatments) was assessed using a self-administered questionnaire. Based on the frequently asked questions on OA reported in a previous study [26], 22 statements could be scored on a 4-point scale (*totally disagree* to *totally agree*, with the additional option *I don't know*). A total score (maximum of 22) was calculated by awarding one point for each correct response. Each incorrect or undecided (*I don't know*) answer was scored as 0.

Technology usage (frequency of internet and smartphone use in daily life) was assessed using 2 subscales of the Media and Technology Usage and Attitudes Scale (MTUAS) [27]. These scales use a 10-point frequency scale (1=never, 5=several times a week, 10=all the time) to score possible activities on a smartphone (eg, *using apps* or *listening to music*) and searching activities on the internet (eg, searching for information). Mean scores can be calculated for each subscale.

Data Analysis

Descriptive Analysis

Baseline characteristics, user data, and usability were described descriptively using mean (SD), median (IQR), and number (%) where appropriate.

Statistical Analysis

The mean usability scores (SUS total scores and item scores) were compared between *superficial* and *active* users, using independent *t* tests (*P*<.05 for significance, two-tailed). The demographics and clinical characteristics were compared between *nonusers* and *users* and between *superficial* and *active users* of the educational eHealth tool using multivariable logistic regression analyses. First, the individual binominal associations between characteristics and the outcome variable (user or nonuser) were calculated. Variables with *P*≤.16 were selected for the multivariable logistic regression analyses [28]. By use of the Variance Inflation Factor (cutoff >10) statistic, the remaining variables were tested for collinearity [28]. A backward selection (*P*<.10 for removal) was used to generate the final model.

For the logistic regression analysis, multiple imputation using Imputation by Chained Equation was used to estimate the missing values. A total of 20 imputed datasets were combined using Rubin's rules [29]. All data were analyzed using Stata 13.1.

Results

Participants

A total of 144 patients with hip or knee OA were included in this study. Their mean (SD) age was 61.7 (10.4) years, and more women (57%) participated than men. The mean (SD) BMI of the participants was 27.9 (4.4) kg/m^2. A total of 57 (40%) participants had a low educational level (<12 years), and 58 (43%) had a paid job at the time of inclusion. The majority of participants had a scheduled consultation for a knee joint (80%). The duration of symptoms was less than 5 years for the majority of participants (58%). The participants had moderate-to-severe impaired functioning as reflected by the WOMAC (mean (SD): 50.6 (20.1) for pain and 55.1 (21.1) for functioning). Patient characteristics are shown in Table 1.

Table 1. Patient and clinical characteristics of users and nonusers of the educational electronic health tool.

Characteristics	Nonusers (n=28)	Users (n=116)
Gender (female), n (%)	14 (50.0)	67 (57.8)
Age (years), mean (SD)	59.4 (12.1)	62.2 (9.9)
BMI >25 kg/m^2, n (%)	18 (64.3)	78 (67.2)
Married, n (%)	20 (71.4)	86 (74.1)
Level of education (>12 years), n (%)	17 (60.7)	67 (57.8)
Work status (paid), n (%)	11 (39.3)	47 (40.5)
Osteoarthritis of the knee, n (%)	21 (75.0)	94 (81.0)
Duration of symptoms (>5 years), n (%)	13 (48.1)	46 (39.7)
Pain, WOMAC[a] (0-100), mean (SD)	58.1 (23.1)	49.1 (19.2)
Functioning, WOMAC (0-100), mean (SD)	62.4 (21.7)	53.5 (20.8)
Fatigue, NRS[b] (0-100), mean (SD)	40.1 (20.2)	48.5 (25.1)
Pain medication use, n (%)	21 (75.0)	93 (80.2)
Knowledge of osteoarthritis[c] (0-22), mean (SD)	9.7 (3.5)	11.4 (3.7)
Smartphone usage in daily life, MTUAS[d] (1-10), mean (SD)	6.1 (1.7)	5.0 (1.9)
Internet usage in daily life, MTUAS (1-10), mean (SD)	6.3 (1.9)	4.8 (1.7)

[a] WOMAC: Western Ontario and McMaster Universities Osteoarthritis Index.

[b] NRS: numeric rating scale.

[c] Self-administered.

[d] MTUAS: Media and Technology Usage and Attitudes Scale.

Use of the Educational Electronic Health Tool

Among the participants, 116 (81%) were users and 28 (19%) were nonusers of the eHealth tool (Figure 1). The group of 116 users, comprised 29 (25%) superficial users and 87 (75%) active users.

Among the users, 74 (64%) participants used the Android or iOS app, and 57 (49%) used the Web version of the educational eHealth tool, with 15 participants using both platforms. The 3 main components of the tool (*information, my consultation, and medication*) were all opened by the majority of users (91%-95%; Table 2), whereas the short facts on OA were opened by slightly fewer users (80%). The median number of opened components was 2 (IQR 1-3). The actual use (ie, earning an achievement award) was highest for the preparation for the consultation (57%), followed by reading all the information on OA treatments (35%) and listing one's medication (34%). A detailed look into the earned achievements revealed that participants who only earned one achievement received the award for the *information* or *my consultation* components. For those who earned 2 achievements, the combination of *my consultation* and *medication* was most frequently earned. Our results also show that of the 87 active users, 19 solely used the more passive component (ie, reading information), whereas more than half chose to use the interactive components alone (ie, preparing for a consultation only or in combination with listing medication).

Figure 1. Distribution of nonusers and users (superficial and active) among the study population.

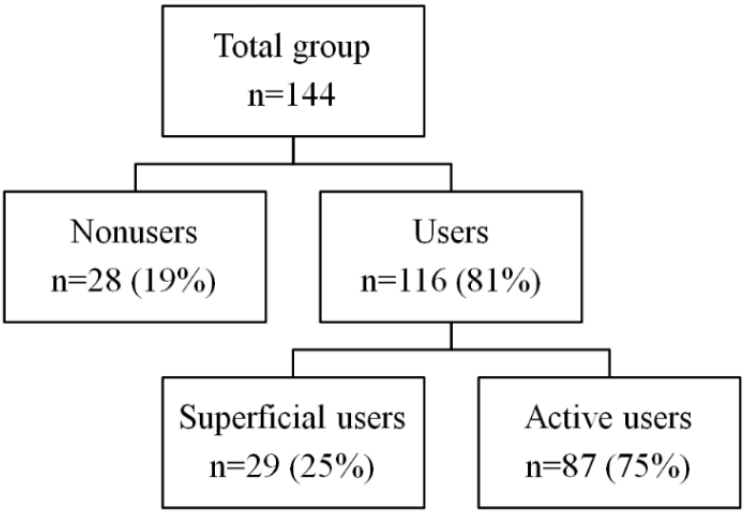


Table 2. Use of components of the educational electronic health tool among 116 users.

Component	Participants, n (%)	Frequency, median (IQR)	Earned achievement, n (%)
Information			41 (35.3)[a]
Opened *information*	110 (94.8)	4.5 (2-10)	—[b]
Read ≥1 short fact	93 (80.2)	7.5 (1.5-11)	—
My consultation			66 (56.9)[c]
Opened *my consultation*	109 (93.9)	6.5 (4-35)	—
Answered ≥1 preparation question	63 (54.3)	—	—
Recorded ≥ 1 questions for consultation	31 (26.7)	—	—
Scheduled pain and/or fatigue measurements	38 (32.8)	—	—
Medication			40 (34.5)[d]
Opened *medication*	105 (90.5)	6.5 (2.5-14.5)	—

[a] Visited all pages with information.

[b] Not applicable.

[c] Scheduled consultation date.

[d] Listed medication use.

Questions Prepared for the Consultation

About one-fourth of users recorded one or more questions in preparation for their consultation. A total of 75 questions were recorded in the tool. Disagreement about the categorization of 20 questions was resolved by discussion. In all, 2 questions were excluded from the categorization because they were formulated as notes rather than questions. A total of 46 questions were categorized into the 3 predefined themes "What are my options?," "What are the possible benefits and harms of those options?," and "What does this mean in my situation?" (Table 3). When discussing the remaining 27 questions, the fourth theme was identified, "What is my situation at this moment?," with 15 questions added to this category. The other 12 questions were grouped into the *remaining* category. These were mainly educational questions on how to deal with OA in daily life, some of which addressed the added value of experimental treatments.

Table 3. Categorization of questions prepared in the educational electronic health tool, with examples given for each theme.

Themes (number of questions)	Example questions
What is my situation at this moment? (15)	• How far has the osteoarthritis progressed? • What is the situation right now and what is the prognosis?
What are my options? (35)	• What is your advice in resolving the pain? • What are my treatment options?
What are the possible benefits and harms of those options? (5)	• What can I expect if I had surgery? • What is the recovery period of surgery?
What does this mean in my situation? (6)	• Is it still necessary to use orthopedic shoes? • Is it possible to get an injection in my knee one more time?
Remaining (12)	• At what level can I be physically active with regard to the wear and tear of my cartilage? • Is it possible to inject cartilage into the knee?

Usability

The mean (SD) usability score among users, as measured with the SUS, was 64.8 (16.0). Moreover, patient satisfaction with the educational eHealth tool was 6.9 (1.7) on a scale from 0 to 10. No differences were found in the SUS and satisfaction scores between *active* and *superficial* users (mean difference 0.04, 95% CI −7.69 to 7.77 and mean difference 0.3, 95% CI −0.50 to 1.11, respectively). The comparison of individual items of the SUS between *active* and *superficial* users did also not result in any differences (results not shown).

Subgroup Characteristics

Based on univariate binominal regression analyses, fatigue (*P*=.16), knowledge of OA (*P*=.04), and smartphone (*P*=.03) and internet (*P*=.009) use in daily life were included in a multivariable analysis. This analysis revealed that participants with a higher baseline knowledge of OA (OR 1.2, 95% CI 1.0 to 1.4) and who used the internet less frequently in their daily life (OR 0.6, 95% CI 0.5 to 0.9) were more likely to use the educational eHealth tool (Table 4). No statistically significant differences were found between the demographic and clinical characteristics of the superficial and active users.

Table 4. Results from the multivariable logistic regression analysis for differences between users and nonusers of the educational electronic health tool.

Variables	OR[a] (95% CI)	*P* value
Knowledge of OA[b] (0-22)	1.2 (1.0 to 1.4)	.02
Internet usage in daily life, MTUAS[c] (1-10)	0.6 (0.5 to 0.9)	.003

[a] OR: Odds ratio.

[b] Self-administered.

[c] MTUAS: Media and Technology Usage and Attitudes Scale.

Discussion

Principal Findings

This study explored the use and usability of a smartphone and Web-based educational eHealth tool. The educational eHealth tool was used by 81% of the patients with knee or hip OA who were offered it. Among users, 75% actively engaged with the tool and used at least one of the components, with *information* and *my consultation* being the most popular components. Questions that were recorded by participants in preparation for their consultation were mostly in line with a widely used question prompt list, although a considerable number remained, some of which could be categorized in a new additional theme ("What is my situation at this moment?"). Participants with a higher baseline knowledge of OA and who used the internet less frequently in their daily life were most likely to use the tool. No other statistically significant differences were found between users and nonusers of the educational eHealth tool.

Comparison With Previous Work

To our knowledge little is known about the use of stand-alone eHealth interventions in OA. Our finding that 81% of participants used the educational eHealth tool is in line with the results of De Vries et al, who evaluated adherence to a Web-based component of a blended care physical activity program for patients with hip or knee OA [30]. This study was part of a blended care intervention, involving interaction with a physiotherapist; therefore, it is not directly comparable with our study. One recently published RCT on the effectiveness of an educational smartphone and tablet app reported a 70% adherence rate [31]. Compared with other eHealth stand-alone interventions, these percentages are reasonably high [16,17]. The relatively high usage rate in our study could be explained by the short time frame in which the tool could be used (2 weeks before the upcoming consultation) and the specific objective of the tool.

In this study, we defined *users* and *nonusers* based on whether they opened the app. The users were defined as active when they opened the tool and earned at least one achievement award. In the development of the tool, all components included were indicated as important by stakeholders. However, we chose one achievement as cutoff for each individual user as there can be a difference in which components they want to use the app for. Using other cutoff points to define *use* might have resulted in different user rates and conclusions; however, considering the small differences we found and the lack of differences detected between the active and superficial users, it is not likely that changing the cutoff point would have resulted in additional findings of interest. However, this discussion does demonstrate the necessity of defining *use*. Currently, there is no consensus about how to define and appraise eHealth use; measures used to define use include the frequency of logging in or using a tool, the number of components used, or the time spent on the tool [32]. Often the threshold for *use* is drawn based on the concept that *more is better* or is not justified at all [33]. As every eHealth intervention has its own goal, it may not even be feasible or valuable to have one definition of *use*; however, we could strive for criteria that can be used to set a cutoff point for every intervention. This would help interpretation of eHealth *use* and facilitate comparisons between studies.

Regardless of the high user rates in our study, our results regarding usability and patient satisfaction about the tool show that there is still room for improvement. The mean usability score of our educational eHealth tool was 64.8 on a scale from 0 to 100, as measured with the SUS. Although this score corresponds to being fair to good [34], it does not reach the acceptable score (ie, 70) proposed by Bangor et al [35]. Previous studies on the usability of eHealth and mobile health (mHealth) tools had considerably higher SUS scores [36,37]. Scott et al reported a median (IQR) SUS score of 95 (86-98) immediately after providing instructions about a mobile app for daily postoperative self-reporting after colorectal surgery [36]; however, the majority of participants did not use the app after discharge or only used it once. This indicates that high usability alone is not sufficient to motivate people to use eHealth tools [36].

The Technology Acceptance Model states that the actual use of a technology system is determined by both the perceived usefulness (utility) and the perceived ease of use (usability) [38]. We found that the eHealth tool was used by a fair number of participants, despite the fact that its usability was somewhat disappointingly rated by users. This may indicate that patients see the benefits of using the educational eHealth tool in preparation for their consultation (utility). It is therefore conceivable that improvement of the content might increase the usability and ultimately optimize patient motivation to use the eHealth tool. Frie et al evaluated reviews of smartphone app for monitoring weight loss and found that users had a preference for apps with a limited number of features [39]. Our eHealth tool contained 3 components (*information*, *my consultation*, and *medication*), each of which contained multiple features (eg, monitoring pain and fatigue, recording questions, and answering predefined questions). For further improvement of the tool, the removal of the *medication* component should be considered, as this component was the least used by participants. Remarkably, the component *medication* was added on the request of the end users during the iterative development process; however, it is possible that we made the tool too complicated by taking too many requests into account.

An important part of our intervention was the preparation for the consultation, which involved listing questions to ask during the consultation. Currently, 3 standardized questions ("What are my options?," "What are the possible benefits and harms of those options?," and "How likely are the benefits and harms of each option to occur?") are used in several national campaigns in England, Australia, and the Netherlands [12]. However, it is not known to what extent these questions cover the essence of the questions patients want to ask. Here, we compared the listed questions with the 3 standardized questions. Our results showed that about a quarter of the questions listed by the participants did not fit these 3 themes. A considerable number of these remaining questions were focused on the current status/situation that patients were in, for example, "how far has my OA progressed?" and "what is the prognosis based on my current situation?" This shows that although prompting predefined questions may result in patients considering novel topics [40], it may also miss patient's individual information needs. This consideration is in line with a recent RCT performed by Bottachini et al, who compared the use of a question prompt list (predefined questions) with a question list in breast cancer patients and found that patients who used the prompt list were less satisfied with the information they received during their consultation [40]. Our results support the extension of the 3 standardized questions to 4, but also suggest the importance of finding ways to elicit the individual information needs of patients not covered in the predefined questions to optimize their preparation for consultations, for instance, by providing a space for a list of their own questions, as we did in our educational eHealth tool.

We found several differences between the characteristics of users and nonusers. Our results show that the baseline knowledge of OA was lower among nonusers than users, suggesting that some subgroups of patients may just not be as interested in learning more about their condition or are not able to [41]. In clinical practice, it is important to be aware of this subgroup of patients, which may need a different strategy to be educated. In addition, we found that users were less familiar with using the internet in their daily life than nonusers. Although only univariate, the same trend was seen for daily life smartphone usage (P=.03). A previous study on the determinants of adherence to a Web-based component of a physical activity program in OA qualitatively identified internet skills as important for optimal adherence [30]. On average, the users in our study indicated that they use the internet and smartphones several times a week, which suggests that they likely had significant skills in using these media. The literature also shows that younger people are less likely to adhere to eHealth tools [16,42]. Although not statistically significant, the nonusers in our study were slightly younger than the users. It is likely that younger adults, who in general use the internet and smartphones more often [43], have lost interest in new apps that are continuously being offered to them or already found similar apps or information about OA on the internet. Different strategies to target this subgroup should be explored.

Strengths and Limitations

Several limitations of this study need to be addressed. First, we do not know why some participants did not use the app or only used it in a superficial way. Qualitative research may provide additional insights into how we can further optimize the eHealth app. Second, we used the SUS questionnaire to obtain information on the usability of the app at one point in time. The SUS has been shown to be reliable and is the most widely used questionnaire for evaluating usability, making it easy to compare with other eHealth interventions [44]. Despite these benefits, the SUS was not originally developed with a focus on health care and therefore only provided us with a general idea of the tool's usability. Comparing the SUS scores with more objective outcomes like effectiveness and efficiency could be of added value in this regard. Also, measuring the SUS at more points in time, before and after the previously suggested adjustments to the tool, would provide a more specific picture of the tool's usability. Finally, it is important to note that our study sample consisted of patients willing to participate in an RCT evaluating an eHealth app. Although patients who were not willing to participate did not differ from those who did in terms of their age or gender, we do not know the extent of internet use by the patients not willing to participate; therefore, it is not possible to generalize the user characteristics we identified to the entire OA population. To consolidate our results in future research, we recommend the use of a study design in which every patient visiting for the first orthopedic consultation gets access to the educational eHealth tool.

Conclusions

On the basis of relatively high user rates and reasonable usability scores, it can be concluded that the use of an educational eHealth tool to prepare patients with hip and knee OA for their first orthopedic consultation is feasible; however, improvements to the content of the tool itself should be established to enhance its usability and user satisfaction. It is recommended that 44 predefined questions as mentioned in this study are included and that space is provided for patients to list additional questions to support their preparation for their consultation. Moreover, simplifying the tool may also increase its usability. We found no clear practical indications that specific subgroups should be targeted for implementation. The literature on the use of eHealth and especially mHealth technologies in patients with OA is scarce. The results of this study therefore provide new insights revealing that interventions using eHealth have the potential to reach this population and show that usage data can reveal how to further optimize the delivery of these interventions.

Acknowledgments

The authors would like to thank research assistants Nadia Koster and Anje Post for their efforts in patient recruitment and data collection and Jolanda Rubrech-van As for data management.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Development and functionalities of the educational eHealth tool.
[PDF File (Adobe PDF File), 492 KB - formative_v3i4e13577_app1.pdf]

References

1. Zhang Y, Jordan JM. Epidemiology of osteoarthritis. Clin Geriatr Med 2010 Aug;26(3):355-369 [FREE Full text] [doi: 10.1016/j.cger.2010.03.001] [Medline: 20699159]
2. Bennell KL, Hunter DJ, Hinman RS. Management of osteoarthritis of the knee. Br Med J 2012 Jul 30;345:e4934. [doi: 10.1136/bmj.e4934] [Medline: 22846469]
3. Fernandes L, Hagen KB, Bijlsma JW, Andreassen O, Christensen P, Conaghan PG, European League Against Rheumatism (EULAR). EULAR recommendations for the non-pharmacological core management of hip and knee osteoarthritis. Ann Rheum Dis 2013 Jul;72(7):1125-1135 [FREE Full text] [doi: 10.1136/annrheumdis-2012-202745] [Medline: 23595142]
4. Samsson KS, Bernhardsson S, Larsson ME. 'Take me seriously and do something!' - a qualitative study exploring patients' perceptions and expectations of an upcoming orthopaedic consultation. BMC Musculoskelet Disord 2017 Aug 24;18(1):367 [FREE Full text] [doi: 10.1186/s12891-017-1719-6] [Medline: 28838326]
5. Barten DJ, Smink A, Swinkels IC, Veenhof C, Schers HJ, Vliet Vlieland T, et al. Factors associated with referral to secondary care in patients with osteoarthritis of the hip or knee after implementation of a stepped-care strategy. Arthritis Care Res (Hoboken) 2017 Feb;69(2):216-225 [FREE Full text] [doi: 10.1002/acr.22935] [Medline: 27159735]
6. du Long J, Hageman M, Vuijk D, Rakic A, Haverkamp D. Facing the decision about the treatment of hip or knee osteoarthritis: what are patients' needs? Knee Surg Sports Traumatol Arthrosc 2016 May;24(5):1710-1716. [doi: 10.1007/s00167-016-3993-5] [Medline: 26831860]

7. Waters S, Edmondston SJ, Yates PJ, Gucciardi DF. Identification of factors influencing patient satisfaction with orthopaedic outpatient clinic consultation: A qualitative study. Man Ther 2016 Sep;25:48-55. [doi: 10.1016/j.math.2016.05.334] [Medline: 27422597]
8. Wetzels R, Harmsen M, van Weel C, Grol R, Wensing M. Interventions for improving older patients' involvement in primary care episodes. Cochrane Database Syst Rev 2007 Jan 24(1):CD004273. [doi: 10.1002/14651858.CD004273.pub2] [Medline: 17253501]
9. Kinnersley P, Edwards A, Hood K, Ryan R, Prout H, Cadbury N, et al. Interventions before consultations to help patients address their information needs by encouraging question asking: systematic review. Br Med J 2008 Jul 16;337:a485 [FREE Full text] [doi: 10.1136/bmj.a485] [Medline: 18632672]
10. Stacey D, Légaré F, Col NF, Bennett CL, Barry MJ, Eden KB, et al. Decision aids for people facing health treatment or screening decisions. Cochrane Database Syst Rev 2014 Jan 28(1):CD001431. [doi: 10.1002/14651858.CD001431.pub4] [Medline: 24470076]
11. Stacey D, Taljaard M, Dervin G, Tugwell P, O'Connor AM, Pomey MP, et al. Impact of patient decision aids on appropriate and timely access to hip or knee arthroplasty for osteoarthritis: a randomized controlled trial. Osteoarthritis Cartilage 2016 Jan;24(1):99-107 [FREE Full text] [doi: 10.1016/j.joca.2015.07.024] [Medline: 26254238]
12. Shepherd HL, Barratt A, Jones A, Bateson D, Carey K, Trevena LJ, et al. Can consumers learn to ask three questions to improve shared decision making? A feasibility study of the ASK (AskShareKnow) Patient-Clinician Communication Model(®) intervention in a primary health-care setting. Health Expect 2016 Oct;19(5):1160-1168 [FREE Full text] [doi: 10.1111/hex.12409] [Medline: 26364752]
13. Sansoni JE, Grootemaat P, Duncan C. Question Prompt Lists in health consultations: A review. Patient Educ Couns 2015 Jun 3. [doi: 10.1016/j.pec.2015.05.015] [Medline: 26104993]
14. Trenaman L, Stacey D, Bryan S, Taljaard M, Hawker G, Dervin G, et al. Decision aids for patients considering total joint replacement: a cost-effectiveness analysis alongside a randomised controlled trial. Osteoarthritis Cartilage 2017 Oct;25(10):1615-1622 [FREE Full text] [doi: 10.1016/j.joca.2017.05.022] [Medline: 28624294]
15. Elbert NJ, van Os-Medendorp H, van Renselaar W, Ekeland AG, 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]
16. Short CE, DeSmet A, Woods C, Williams SL, Maher C, Middelweerd A, et al. Measuring engagement in eHealth and mHealth behavior change interventions: viewpoint of methodologies. J Med Internet Res 2018 Nov 16;20(11):e292 [FREE Full text] [doi: 10.2196/jmir.9397] [Medline: 30446482]
17. Shepherd HL, Barratt A, Trevena LJ, McGeechan K, Carey K, Epstein RM, et al. Three questions that patients can ask to improve the quality of information physicians give about treatment options: a cross-over trial. Patient Educ Couns 2011 Sep;84(3):379-385. [doi: 10.1016/j.pec.2011.07.022] [Medline: 21831558]
18. Fogg BJ. Creating Persuasive Technologies: An Eight-step Design Process. In: Proceedings of the 4th International Conference on Persuasive Technology. 2009 Presented at: Persuasive'09; April 26-29, 2009; Claremont, California, USA. [doi: 10.1145/1541948.1542005]
19. 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]
20. Smink AJ, van den Ende CH, Vlieland TP, Swierstra BA, Kortland JH, Bijlsma JW, et al. 'Beating osteoARThritis': development of a stepped care strategy to optimize utilization and timing of non-surgical treatment modalities for patients with hip or knee osteoarthritis. Clin Rheumatol 2011 Dec;30(12):1623-1629. [doi: 10.1007/s10067-011-1835-x] [Medline: 21887488]
21. Brooke J. SUS: a quick and dirty usability scale. In: Jordan PW, Thomas B, McClelland IL, Weerdmeester B, editors. Usability Evaluation In Industry. London: Taylor and Francis; 1996.
22. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the system usability scale. Int J Human-Comput Interact 2008;24(6):574-594. [doi: 10.1080/10447310802205776]
23. Roorda LD, Jones CA, Waltz M, Lankhorst GJ, Bouter LM, van der Eijken JW, et al. Satisfactory cross cultural equivalence of the Dutch WOMAC in patients with hip osteoarthritis waiting for arthroplasty. Ann Rheum Dis 2004 Jan;63(1):36-42 [FREE Full text] [doi: 10.1136/ard.2002.001784] [Medline: 14672889]
24. de Groot IB, Favejee MM, Reijman M, Verhaar JAN, Terwee CB. The dutch version of the knee injury and osteoarthritis outcome score: a validation study. Health Qual Life Outcomes 2008 Feb 26;6:16 [FREE Full text] [doi: 10.1186/1477-7525-6-16] [Medline: 18302729]
25. de Groot IB, Reijman M, Terwee CB, Bierma-Zeinstra S, Favejee MM, Roos E, et al. Validation of the dutch version of the hip disability and osteoarthritis outcome score. Osteoarthritis Cartilage 2009 Jan;17(1):132 [FREE Full text] [doi: 10.1016/j.joca.2008.05.014] [Medline: 18620874]
26. Claassen AA, Kremers-van de Hei K, van den Hoogen FH, van der Laan WH, Rijnen WH, Koëter S, et al. Most important frequently asked questions from patients with hip or knee osteoarthritis: a best-worst scaling exercise. Arthritis Care Res (Hoboken) 2019 Jul;71(7):885-892. [doi: 10.1002/acr.23719] [Medline: 30055092]

27. Rosen LD, Whaling K, Carrier LM, Cheever NA, Rokkum J. The media and technology usage and attitudes scale: an empirical investigation. Comput Human Behav 2013 Nov 1;29(6):2501-2511 [FREE Full text] [doi: 10.1016/j.chb.2013.06.006] [Medline: 25722534]
28. Steyerberg E. Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating. New York: Springer; 2009.
29. Marshall A, Altman DG, Holder RL, Royston P. Combining estimates of interest in prognostic modelling studies after multiple imputation: current practice and guidelines. BMC Med Res Methodol 2009 Jul 28;9:57 [FREE Full text] [doi: 10.1186/1471-2288-9-57] [Medline: 19638200]
30. de Vries HJ, Kloek CJ, de Bakker DH, Dekker J, Bossen D, Veenhof C. Determinants of adherence to the online component of a blended intervention for patients with hip and/or knee osteoarthritis: a mixed methods study embedded in the e-exercise trial. Telemed J E Health 2017 Dec;23(12):1002-1010. [doi: 10.1089/tmj.2016.0264] [Medline: 28525310]
31. Timmers T, Janssen L, Pronk Y, van der Zwaard BC, Koëter S, van Oostveen D, et al. Assessing the efficacy of an educational smartphone or tablet app with subdivided and interactive content to increase patients' medical knowledge: randomized controlled trial. JMIR Mhealth Uhealth 2018 Dec 21;6(12):e10742 [FREE Full text] [doi: 10.2196/10742] [Medline: 30578185]
32. Mohr DC, Cheung K, Schueller SM, Brown C, Duan N. Continuous evaluation of evolving behavioral intervention technologies. Am J Prev Med 2013 Oct;45(4):517-523 [FREE Full text] [doi: 10.1016/j.amepre.2013.06.006] [Medline: 24050429]
33. 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]
34. Kortum PT, Bangor A. Usability ratings for everyday products measured with the system usability scale. Int J Human-Comput Interact 2012;29(2):67-76. [doi: 10.1080/10447318.2012.681221]
35. Bangor A, Kortum P, Miller J. Determining what individual SUS scores mean: adding an adjective rating scale. J Usability Stud 2009;4(3):114-123 [FREE Full text]
36. Scott AR, Alore EA, Naik AD, Berger DH, Suliburk JW. Mixed-methods analysis of factors impacting use of a postoperative mHealth app. JMIR Mhealth Uhealth 2017 Feb 8;5(2):e11 [FREE Full text] [doi: 10.2196/mhealth.6728] [Medline: 28179215]
37. Dekker-van Weering M, Jansen-Kosterink S, Frazer S, Vollenbroek-Hutten M. User experience, actual use, and effectiveness of an information communication technology-supported home exercise program for pre-frail older adults. Front Med (Lausanne) 2017;4:208 [FREE Full text] [doi: 10.3389/fmed.2017.00208] [Medline: 29250523]
38. Davis F, Bagozzi R, Warshaw P. User acceptance of computer technology: a comparison of two theoretical models. Manag Sci 1989 Aug;35(8):982-1003 [FREE Full text] [doi: 10.1287/mnsc.35.8.982]
39. Frie K, Hartmann-Boyce J, Jebb S, Albury C, Nourse R, Aveyard P. Insights from google play store user reviews for the development of weight loss apps: mixed-method analysis. JMIR Mhealth Uhealth 2017 Dec 22;5(12):e203 [FREE Full text] [doi: 10.2196/mhealth.8791] [Medline: 29273575]
40. Bottacini A, Goss C, Mazzi MA, Ghilardi A, Buizza C, Molino A, et al. The involvement of early stage breast cancer patients during oncology consultations in Italy: a multi-centred, randomized controlled trial of a question prompt sheet versus question listing. BMJ Open 2017 Aug 11;7(8):e015079 [FREE Full text] [doi: 10.1136/bmjopen-2016-015079] [Medline: 28801395]
41. Sørensen K, van den Broucke S, Fullam J, Doyle G, Pelikan J, Slonska Z, (HLS-EU) Consortium Health Literacy Project European. Health literacy and public health: a systematic review and integration of definitions and models. BMC Public Health 2012 Jan 25;12:80 [FREE Full text] [doi: 10.1186/1471-2458-12-80] [Medline: 22276600]
42. 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]
43. Centraal Bureau voor de Statistiek. Six Percent Never Use the Internet. URL: https://www.cbs.nl/en-gb/news/2019/01/six-percent-never-use-the-internet [accessed 2019-01-31] [WebCite Cache ID 75ppj4xMj]
44. Sousa VE, Lopez K. Towards usable e-Health. A systematic review of usability questionnaires. Appl Clin Inform 2017 May 10;8(2):470-490 [FREE Full text] [doi: 10.4338/ACI-2016-10-R-0170] [Medline: 28487932]

Abbreviations

BMI: body mass index
eHealth: electronic health
mHealth: mobile health
MTUAS: Media and Technology Usage and Attitudes Scale
OA: osteoarthritis
OR: odds ratio

RCT: randomized controlled trial
SUS: System Usability Scale
TJR: total joint replacement
WOMAC: Western Ontario and McMaster Universities osteoarthritis index

---

Edited by G Eysenbach; submitted 01.02.19; peer-reviewed by F Sieverink, M Valero, A Behmanesh; comments to author 28.04.19; revised version received 13.06.19; accepted 02.09.19; published 28.11.19.

Please cite as:
Claassen AAOM, Vliet Vlieland TPM, Busch VJJF, Schers HJ, van den Hoogen FHJ, van den Ende CHM
An Electronic Health Tool to Prepare for the First Orthopedic Consultation: Use and Usability Study
JMIR Form Res 2019;3(4):e13577
URL: http://formative.jmir.org/2019/4/e13577/
doi:10.2196/13577
PMID:31778119

Original Paper

An Internet-Based Therapeutic Tool for American Indian/Alaska Native Adults With Posttraumatic Stress Disorder: User Testing and Developmental Feasibility Study

Vanessa Y Hiratsuka[1], PhD, MPH; Laurie Moore[2], MPH; Jaedon P Avey[1], PhD; Lisa G Dirks[1], MLIS, MADM; Barbara D Beach[3], PhD; Denise A Dillard[1], PhD; Douglas K Novins[2], PhD

[1]Research Department, Southcentral Foundation, Anchorage, AK, United States

[2]Centers for American Indian and Alaska Native Health, University of Colorado Anschutz Medical Campus, Aurora, CO, United States

[3]Cherokee Nation Behavioral Health, Tahlequah, OK, United States

Corresponding Author:
Vanessa Y Hiratsuka, PhD, MPH
Research Department
Southcentral Foundation
4085 Tudor Centre Drive
Anchorage, AK, 99508-5919
United States
Phone: 1 907 729 8627
Email: vhiratsuka@scf.cc

Abstract

Background: Posttraumatic stress disorder (PTSD) is a major public health concern among American Indian and Alaska Native populations. Primary care clinics are often the first point of contact for American Indian and Alaska Natives seeking health care and are feasible locations for trauma-focused interventions.

Objective: Web-based therapeutic interventions have the potential to reduce PTSD symptoms by offering psychoeducation and symptom self-management tools. We investigated the feasibility of a culturally adapted Web-based therapeutic intervention in two American Indian and Alaska Native–serving primary care sites. We developed and tested a self-guided Web-based therapeutic intervention aimed at improving knowledge and awareness of, and provision of guidance, support, and symptom-management for, PTSD symptoms.

Methods: A community-based participatory research process was used to refine adaptations to the veteran's administration's PTSD Coach Online, to develop new content, and to guide and interpret the results of the feasibility pilot. This process resulted in a 16-guide intervention *"Health is Our Tradition: Balance and Harmony after Trauma"* website. The feasibility pilot included 24 American Indian and Alaska Natives aged 18 years and older who scored positive on a primary care PTSD screener. Enrolled participants completed a demographic questionnaire, an experience with technology questionnaire, and baseline behavioral health measures. Once measures were complete, research staff described weekly text messages, minimum study expectations for website use, and demonstrated how to use the website. Feasibility measures included self-reported website use, ratings of satisfaction and perceived effectiveness, and website metrics. Feasibility of obtaining measures for an effectiveness trial was also assessed to include behavioral health symptoms and service utilization through self-report instruments and electronic health record queries. Self-reported measures were collected at enrollment and at 6 and 12 weeks post enrollment. Electronic health records were collected from 12 months before study enrollment to 3 months following study enrollment. Changes between enrollment and follow-up were examined with paired *t* tests, analysis of variance or logistic regression, or the Wilcoxon signed rank test for nonnormally distributed data.

Results: The culturally adapted website and associated text message reminders were perceived as satisfactory and effective by participants with no differences by age or gender. The majority of participants (86%, 19/24) reported use of the website at 6 weeks and nearly all (91%, 20/22) at 12 weeks. At 6 weeks, 55% (12/22) of participants reported using the website at the recommended intensity (at least three times weekly), dropping to 36% (8/22) at 12 weeks. Participant use of modules varied from 8% (2/24) to 100% (24/24), with guide completion rates being greater for guides that were only psychoeducational in nature compared with guides that were interactive. There were no significant changes in patterns of diagnoses, screening, medications, or service utilization during exposure to the website.

Conclusions: *"Health is Our Tradition: Balance and Harmony after Trauma"* shows promise for an effectiveness pilot.

(JMIR Form Res 2019;3(4):e13682) doi:10.2196/13682

KEYWORDS

formative research; posttraumatic stress disorder; Web-based intervention; Indians, North American

Introduction

Background

American Indian (AI) and Alaska Native (AN) people experience disproportionally higher rates of acute, chronic, and intergenerational trauma than their non–AI/AN counterparts [1-4], with adverse impacts on physical and behavioral health [1,5-9]. As the frequency of traumatic events is significantly elevated, there is an increased likelihood that an AI/AN individual will experience multiple traumas compared with national samples [2,10-12,13]. Multiple lifetime traumas have additive risks to developing behavioral health disorders such as posttraumatic stress disorder (PTSD). AI/AN people have low rates of access to behavioral health services, especially those that are culturally responsive. Primary care clinics are often the first point of contact for AI/AN people seeking health care and are feasible locations for trauma-focused interventions [14-16].

Given the multitude of medical (eg, chronic pain, irritable bowel syndrome, and autoimmune disorders) [17-19] and behavioral (eg, depression, substance use disorders, eating disorders, somatization disorders, personality disorders, and PTSD) [20-23] health problems that occur following traumatic events, clinicians face multiple challenges in appropriately addressing these conditions and their comorbidities. This is especially critical as some research has suggested that the quality of medical outcomes for individuals with a history of trauma is related to their behavioral health outcomes [24,25]. Although behavioral health outcomes may be particularly important in ensuring good medical outcomes for individuals with a history of trauma, psychotherapeutic treatments for these conditions often involve the use of complex, manualized treatment that many behavioral health clinicians have not been trained to deliver (eg, cognitive processing therapy for PTSD and dialectical behavioral therapy for borderline personality disorder) [26,27]. Not all patient education programs are effective in improving health knowledge and treatment outcomes [28]. Research suggests that, to be effective, patient education programs must be intensive, focused on developing specific skills, and fostering a stronger sense of self-efficacy [28].

Health information technologies (HIT) offer innovative solutions to address patient access to critical health information and to reinforce key components of patient treatment plans. HITs offer multiple communication (eg, websites, apps for mobile phones, and clinical decision support systems) to improve the quality of service delivery [29-31], with interactive HIT shown to increase patient knowledge and related health outcomes [32,33] and reduce hospitalizations and health care costs [34]. Web-based therapeutic interventions (WBTIs) are self- or clinician-guided programs that are developed with the aim of positively improving knowledge, awareness, support, and treatment for health problems at low cost. WBTIs are a promising intervention modality for behavioral health treatment support among AI/AN and other indigenous people who have barriers to effective treatment because of access to care as WBTIs can be culturally tailored to address specific clinical issues and the unique characteristics of patient populations.

Objectives

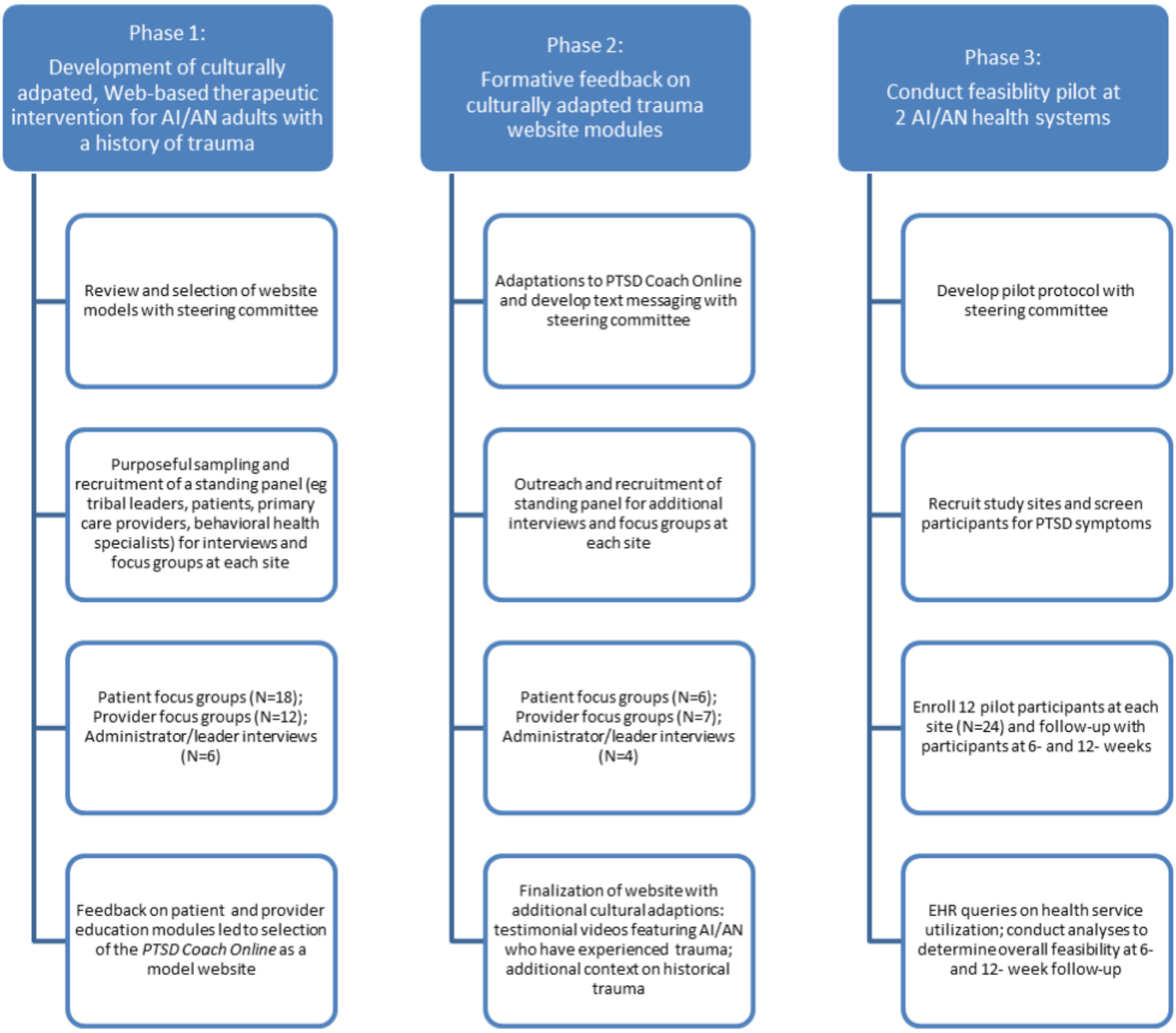
We undertook a feasibility pilot to develop and test a self-guided, WBTI aimed at improving knowledge and awareness of, and provision of guidance, support, and symptom-management for, PTSD symptoms. This manuscript describes our examination of the website-based intervention's feasibility in 2 AI/AN–serving primary care sites—the Cherokee Nation Health Services (CNHS) and Southcentral Foundation (SCF). Feasibility was assessed by website metrics, use reported by participants, and participant ratings of satisfaction and perceived effectiveness. We also assessed the feasibility of querying data from the electronic health record 12 months prior and 3 months after enrollment and obtaining self-reported PTSD and other behavioral health symptoms at enrollment, at 6 weeks (intervention midpoint), and at 12 weeks. Finally, patterns of change in behavioral health symptoms, diagnoses, and service utilization served as an additional feasibility indicator for an effectiveness pilot.

Methods

Steering Committee

An 11-person steering committee guided each step of the cultural adaptation of the website and associated text messages. A community-based participatory research process was used to refine adaptations and content, and to guide and interpret the results of the feasibility test of the resultant intervention (Figure 1). The steering committee was composed of individuals from the 2 tribal health study sites and the University of Colorado Denver (UCD). Members of the steering committee had multiple roles within their respective organizations. For instance, 5 individuals were clinical providers (3 at SCF, 2 at CNHS, and 1 at UCD), 6 individuals were community members at their study locations (4 at SCF and 2 at CNHS), 1 individual from SCF held an administrative leader role, and 6 individuals had extensive AI/AN health research backgrounds (4 at SCF and 2 at UCD). The committee included the cross-site Principal Investigator and study coordinator from UCD, and a site Principal Investigator at SCF and CNHS. Under advisement of the Steering Committee, we sought broad community feedback on the intervention content at both SCF and CNHS in 2 phases and verified the intervention materials in a second round of qualitative feedback using methods previously implemented by the research team [35]. Following development and AI/AN cultural adaptation of the website, the feasibility pilot was conducted at CNHS and SCF as described further.

Figure 1. Process for development and implementation of feasibility pilot. AI/AN: American Indian and Alaska Native; PTSD: posttraumatic stress disorder; EHR: electronic health record.



Setting and Participants

The project settings included 2 of the largest AI/AN–serving primary care facilities in the United States: CNHS and SCF. The Cherokee Nation operates a network of 8 health centers and 1 hospital in the tribe's jurisdictional boundaries, serving more than 130,000 AI patients. CNHS operates the W.W. Hastings Hospital in Tahlequah, Oklahoma, which provides primary care and behavioral health services. SCF is an AN nonprofit health care organization that provides a wide range of health services to over 65,000 AN/AIs from 231 federally recognized tribes in the Anchorage Service Unit. The Anchorage Service Unit is a geographical area extending from the Canadian border on the east to the entire Aleutian Chain and Pribilof Islands on the west, although most SCF patients live near Anchorage, Alaska. SCF operates the Anchorage Native Primary Care Center, which provides primary care and behavioral health services. Each site provides both primary care and behavioral health specialty services and has behavioral health staff embedded within primary care who screen for depression and substance use. The Centers for American Indian and Alaska Native Health at the University of Colorado—Anschutz Medical Campus provided overall coordination between sites. Before data collection, the Alaska Area, Cherokee Nation, and Colorado Multiple Institutional Review Boards approved study procedures. SCF provided tribal approval.

Cultural Adaptation

Website content was designed to align with national guidelines and local CNHS and SCF preferences for behavioral health interventions. To develop website and text message content, we first completed 2 cycles of 1-hour, semistructured interviews with key stakeholders (providers and clinical administrators) and 2-hour focus groups with providers/administrators and patients in each setting. Leaders and providers were concerned that exposure to website content would be triggering for participants and were concerned about additional health service utilization. Qualitative feedback indicated that a website intervention for PTSD would be welcome with several qualifications, including the emphasis should be on people in context of their community and families, patient modules should be accessible on a website rather than exclusively through mobile technology, and all website information should be consistent with existing clinical recommendations from primary care providers. The adapted content consisted of patient education on the website and the communication of weekly *tips* via text messaging.

On the basis of qualitative feedback, the research team revised the *PTSD Coach Online* [36,37], originally developed by the Veterans Administration to benefit those coping with stress following a traumatic event. The revision process included consultation with our Steering Committee and SCF's Family Wellness Warriors Initiative (FWWI) [38], an initiative that focuses on providing services for healing from trauma and abuse. The resultant website entitled *Health is Our Tradition: Balance and Harmony after Trauma* included patient education and interactive healing activities (Table 1) [39].

The *Health is Our Tradition: Balance and Harmony after Trauma* website included 16 sections referred to as guides [39]. Web page content was either educational only or educational with an interactive activity (eg, drag and drop), an observation/listening activity (eg, audio recorded relaxation exercises), or a written activity (eg, developing a plan for coping with trauma reminders). Video testimonials from AI/AN people who had experienced traumatic events were included in each guide. Web pages used text, programmed interactions, animations, and English language close-captioned videos to present the guide content. Select guides included an audio icon to click for listening to narration of the screen contents. In conjunction with the website, we developed strategic patient communications sent via weekly text-messaged *tips* (Table 2). Following the initial intervention adaptations, we conducted a second cycle of qualitative feedback with information used to further refine the adapted website. A professional website development organization was engaged to develop, test, and deploy the adapted website, whereas the research team planned for testing feasibility of its use in the participating primary care clinics (Figure 2).

Table 1. Intervention modules and associated adaptations based on qualitative feedback.

PTSD[a] Coach Online tools	Health is Our Tradition guides	Description	Adaptations
N/A[b]	Trauma and your Health	Patient education	Content developed from Phase 1 interviews, focus groups, and recommended existing SCF[c] and CNHS[d] patient education materials on the impact of trauma on overall health including testimonial videos from AI/AN[e] people with a history of trauma
Be in the moment	Be in the moment	Guided relaxation and grounding activities	Minor adaptation to add an estimation of activity duration and information before the guided relaxation activity acknowledging that some people may not have a quiet place to listen to the audio activity
Change feelings by changing thoughts	Change feelings by changing thoughts	Educational activities with written exercises to work toward changing negative beliefs about oneself to more helpful positive beliefs	Major adaptation to focus on false beliefs and shame messages common among AI/AN people who have experienced trauma including educational videos narrated by members of the AI/AN community; content developed based on Phase 1 interviews and focus groups, addition of testimonial videos from AI/AN people with a history of trauma
Change how you think about sleep	Change how you think about sleep	Educational activities with interactive content that suggests new ways of thinking about negative thoughts that impair sleep	Minor adaptation to replace the word nightmare with bad dream
Change negative thinking patterns	Change negative thinking patterns	Educational activities with written exercises to work identifying and changing negative thought patterns	No adaptation outside of those listed in Overall
Deal with trauma reminders	Cope with trauma reminders	Educational and guided relaxation activities with written exercises to learn to cope with reminders or triggers related to trauma	Minor adaptation to rename guide from Deal with trauma reminders to Cope with trauma reminders and to add additional examples of triggers
N/A	Develop healthy relationships	Educational and written activities to help identify how trauma may affect interactions with others and to change negative interactions to more positive ones	Content developed from Phase 1 interviews and focus groups and adapted from FWWI[f] curriculum materials including testimonial videos from AI/AN people with a history of trauma
N/A	Discover your story	Educational and written activities to encourage acceptance and sharing of one’s story that includes traumatic experiences and their effects on resilience and personal growth	Content developed from Phase 1 interviews and focus groups and adapted from FWWI curriculum materials including testimonial videos from AI/AN people with a history of trauma
Form good sleep habits	Form good sleep habits	Educational and interactive activity to learn about good sleep habits	Minor adaptation to replace a drawing of generic sleeping man with one more easily recognized as having AI/AN heritage; changed the drag and drop motion in the interactive activity to a mouse click motion
Identify your values and goals	Identify your values and goals	Educational and written activity to help identify values and goals to guide decisions in life	Minor adaptation in wording of text and omitting the calendar feature
Learn to be assertive	Express your feelings and wants	Educational, written, and checklist activities to foster effective communication skills	Major adaptation to replace all references to assertiveness with expressing oneself and to replace educational and interactive activities with content adapted from FWWI
Notice your thoughts and feelings	Notice your thoughts and feelings	Educational and mindfulness exercises to help user learn to live in the present moment	Minor adaptation to add an estimation of activity duration and information before the mindfulness activities acknowledging that some people may not have a quiet place to listen to the audio activity

PTSD[a] Coach Online tools	Health is Our Tradition guides	Description	Adaptations
Plan something enjoyable	Plan something fun	Educational and interactive activities to help users identify and plan fun activities	Minor adaptation to add or omit suggested activities that are common or uncommon in AI/AN communities; omitted the calendar function
Relax through breathing	Relax through breathing	Animated relaxation activity focused on deep breathing	No adaptation outside of those listed in Overall
Relax through visualization	Relax through visualization	Animated relaxation activity focused on picturing a calm or comfortable scene	Minor adaptation to add an estimation of activity duration and information before the guided relaxation activity acknowledging that some people may not have a quiet place to listen to the audio activity
Relax your body	Relax your body	Animated relaxation activity focused on releasing body tension	No adaptation outside of those listed in Overall

[a] PTSD: Posttraumatic stress disorder.

[b] N/A: not applicable.

[c] SCF: Southcentral Foundation.

[d] CNHS: Cherokee Nation Health Services.

[e] AI/AN: American Indian and Alaska Native.

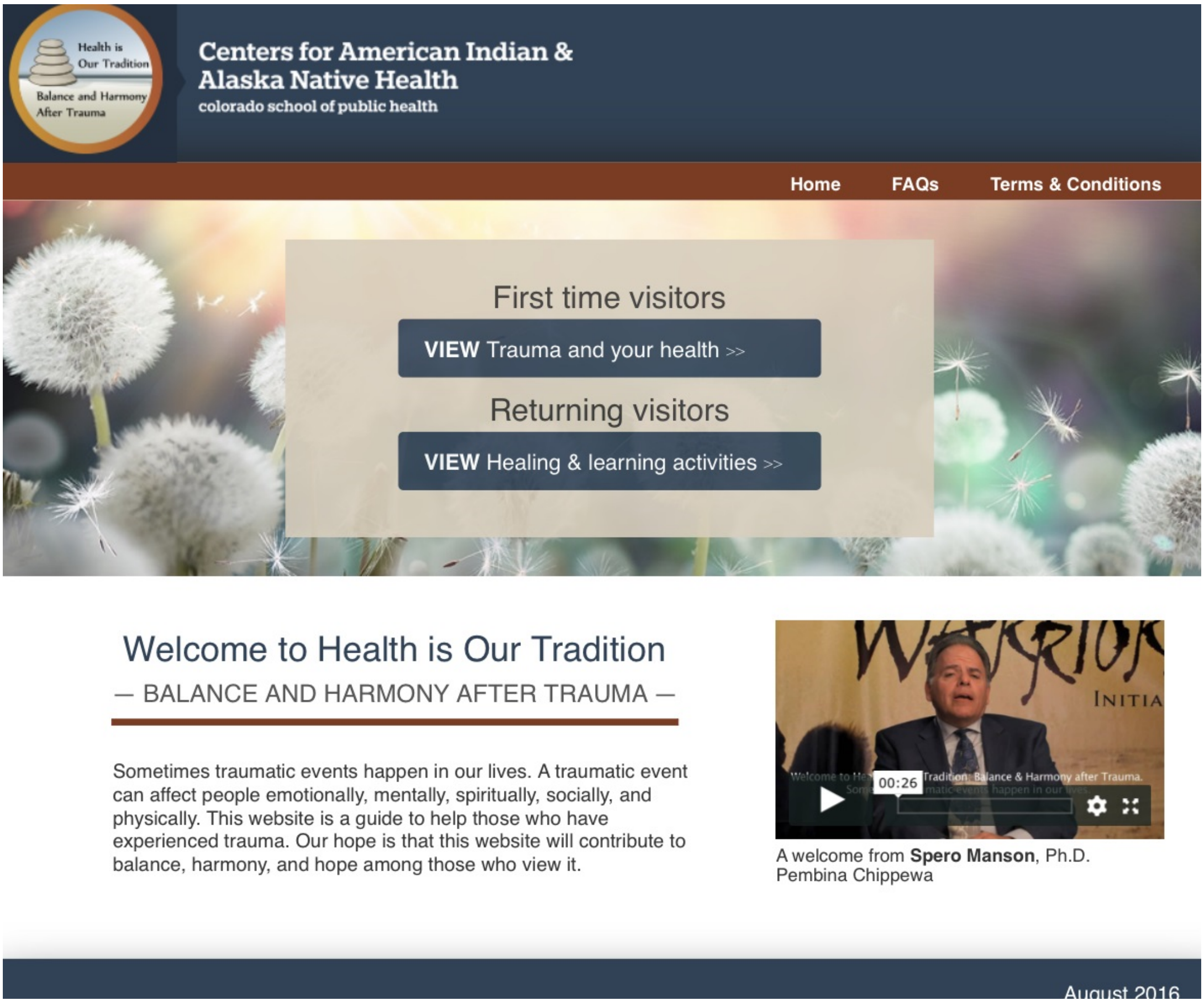
[f] FWWI: Family Wellness Warriors Initiative.

Table 2. Text reminders.

Timing	Text message	Related Health is Our Tradition guide(s)
Week 1	Did you log in to the website yet? If you need our help, call [site research assistant name] at [site phone number].	N/A[a]
Week 2	Have you ever wondered how a traumatic event can affect a person's health? This information and more can be found in the website guide called Trauma and your health	Trauma and your health
Week 3	Is there one best way to heal from trauma? The answer is no, there are many different things to try. Some may work better for you than others. Go to the guide called How to get back in harmony in the website to find out more.	How to get back in harmony
Week 4	What is a grounding activity and how can it help you feel more safe and stable while healing from trauma? This information and more can be found in the website guide called Be in the moment.	Be in the moment
Week 5	Having trouble falling asleep? The website guides Form good sleep habits and Change how you think about sleep will help you work on sleep problems.	Form good sleep habits; Change how you think about sleep
Week 6	We are looking forward to your midpoint check in visit. Until then, remember to be gentle with yourself. Learning new ways to help yourself heal from trauma can be difficult. Give yourself permission to heal at your own pace and on your own terms.	Identify your values and goals
Week 7	Feeling like you are alone on your healing journey? Throughout the website, there are real people telling their own stories of harm, healing, and their return to wellness. These stories may help you feel less alone with yours.	Discover your story
Week 8	Do you feel stuck with negative thoughts about your past and future? Learn more about ways to change your perspective in the Change feelings by changing thoughts guide.	Change feelings by changing thoughts
Week 9	Doing something you enjoy can help you feel better. Get some great ideas for planning and doing a fun activity in Plan something fun.	Plan something fun
Week 10	Feeling frazzled, stressed, or worried? Try a relaxation activity from any of the following guides: Relax through visualization, Relax your breathing, or Relax your body.	Relax through breathing; Relax through visualization; Relax your body
Week 11	Is there a situation where you feel your voice has been silenced? Working through the activities in the guide called Expressing your feelings and wants may give you a way to break the silence.	Express your feelings and wants
Week 12	We are looking forward to seeing you for your final research visit when we can celebrate the progress you have made toward balance and harmony.	N/A

[a] N/A: not applicable.

Figure 2. Health is Our Tradition homepage.



Welcome to Health is Our Tradition

— BALANCE AND HARMONY AFTER TRAUMA —

Sometimes traumatic events happen in our lives. A traumatic event can affect people emotionally, mentally, spiritually, socially, and physically. This website is a guide to help those who have experienced trauma. Our hope is that this website will contribute to balance, harmony, and hope among those who view it.

A welcome from **Spero Manson**, Ph.D.
Pembina Chippewa

Recruitment and Intervention Procedures

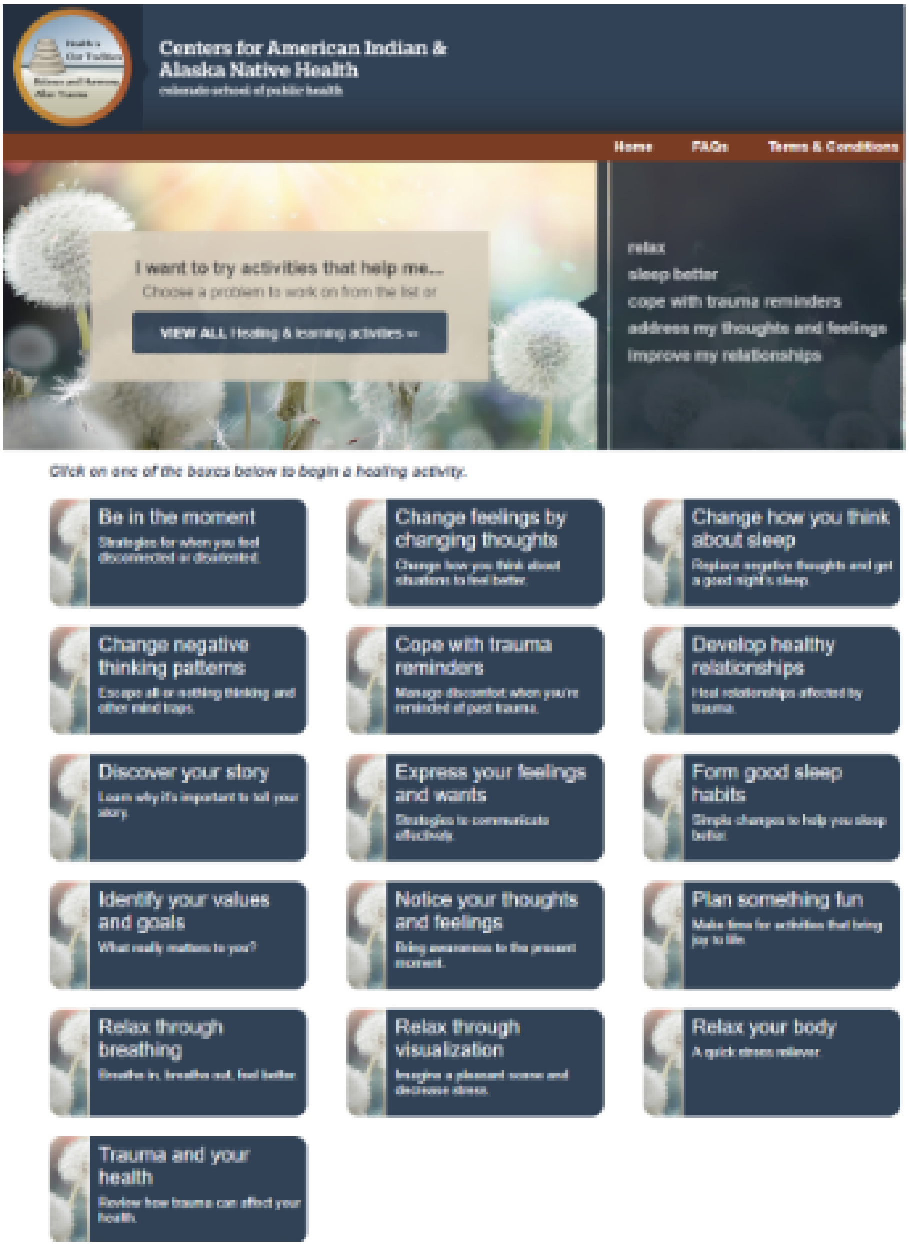
Research staff used lobby-based recruitment at both sites to recruit participants for the intervention. Staff recruited and enrolled 12 participants, at each site, as they waited for their scheduled appointments with a primary care provider. Recruitment occurred in winter 2016 through spring 2017.

Patients expressing interest at the lobby recruitment table were screened by trained staff for study eligibility and completed the informed consent process with eligible individuals. Inclusion criteria included ≥18 years of age, of AI/AN heritage, diagnosis of one or more chronic conditions (ie, heart disease, stroke, diabetes, kidney disease, arthritis, osteoporosis, cancer, asthma, depression, and chronic obstructive pulmonary disease), and 2 or more visits to the primary care clinic in the past 12 months. Exclusionary criteria included lack of access to a text message–capable mobile phone and a computer with an internet connection as well as cellular phone service for the 3 months following enrollment. Ineligible participants were thanked for their participation with a US $10 gift card and returned to the waiting room for their provider appointment.

After completing their primary care provider visit, participants met with a licensed, PhD level psychologist (CNHS) or a Master’s level Behavioral Health Consultant (SCF) to complete the Primary Care—PTSD Screen (PC-PTSD) [40] and an assessment assuring the participant was not currently in crisis. Any participants who were in crisis were counseled by the licensed psychologist/Behavioral Health Consultant and referred to additional behavioral health services, if needed. These participants and those who scored negative on the PC-PTSD were thanked for their participation with a US $10 gift card.

Participants who were not in crisis and scored positive on the PC-PTSD [39] completed a demographic questionnaire, an experience with technology questionnaire, and the following baseline behavioral health measures: Alcohol Use Disorders Identification Test (AUDIT) [41], Drug Abuse Screening Tool (DAST) [42], AI Symptom Inventory, PTSD Checklist—Civilian Version (PCL-C) [43], Patient Health Questionnaire Depression Scale (PHQ-9) [44], and Patient Health Questionnaire—Somatic, Anxiety, and Depressive Symptoms [45]. Once measures were complete, research staff described the weekly tips (text messages) and the website (Figure 3), including minimum study expectations for its use, and demonstrated how to use the website. Participants were asked to log in and use the website 5-10 min a minimum of 3 times a week over the 12-week intervention period. This expectation was a Steering Committee's recommendation based on clinical judgment as well as Phase 1 interview and focus group data. Finally, research staff scheduled the 6- and 12-week follow-up visits and thanked participants with a US $60 gift card.

Figure 3. Health is Our Tradition guides.



Centers for American Indian & Alaska Native Health
colorado school of public health

I want to try activities that help me...
Choose a problem to work on from the list or

relax
sleep better
cope with trauma reminders
address my thoughts and feelings
improve my relationships

Click on one of the boxes below to begin a healing activity.

Be in the moment
Strategies for when you feel disconnected or disoriented.

Change feelings by changing thoughts
Change how you think about situations to feel better.

Change how you think about sleep
Replace negative thoughts and get a good night's sleep.

Change negative thinking patterns
Escape all-or-nothing thinking and other mind traps.

Cope with trauma reminders
Manage discomfort when you're reminded of past trauma.

Develop healthy relationships
Heal relationships affected by trauma.

Discover your story
Learn why it's important to tell your story.

Express your feelings and wants
Strategies to communicate effectively.

Form good sleep habits
Simple changes to help you sleep better.

Identify your values and goals
What really matters to you?

Notice your thoughts and feelings
Bring awareness to the present moment.

Plan something fun
Make time for activities that bring joy to life.

Relax through breathing
Breathe in, breathe out, feel better.

Relax through visualization
Imagine a pleasant scene and decrease stress.

Relax your body
A quick stress reliever.

Trauma and your health
Review how trauma can affect your health.

A weekly tip via text message was sent by researchers to participants during the 12-week intervention (Table 2). Text message verbiage was devised to reinforce website use, complement website content, and remind participants about follow-up visits. Research staff also called each participant to remind them of follow-up appointments.

At each follow-up visit (6 weeks and 12 weeks), participants repeated the experience with technology questionnaire, all baseline measures, as well as a use, satisfaction, and perceived effectiveness questionnaire. Upon completion of each follow-up visit, research staff thanked participants and provided them with a $50 gift card. Finally, a qualified staff member, approved at each site to access electronic health records, did so after the last participant completed the 12-week follow-up. Measures from the electronic health record query were collected for the periods extending to 12 months before study enrollment and 3 months following study enrollment.

Measures

Demographic Questionnaire

At baseline, data were collected on participants' self-reported demographic characteristics. Racial/ethnic heritage was asked using a check all that apply list with the following response categories: AN, AI, white/European-American non-Hispanic, Filipino, other Asian, black/African American non-Hispanic, Hispanic any race, and other. Individuals were asked if they were male or female. Income was assessed by household yearly income before taxes with the following response categories: None, under US $9999 per year, US $10,000 to US $29,999 per year, US $30,000 to US $49,999 per year, US $50,000 to US $69,999 per year, US $70,000 to US $89,999 per year, and US $90,000 or more per year. Education level was asked using the question *How many years of school have you completed*? *(check all that apply)* with associated response categories: Some high school, but did not graduate, high school graduate or General Educational Development (GED), some college, 2-year college graduate (eg, community college), 4-year college graduate, some graduate school or postgraduate degree, and trade school. Number in household was determined using the question *How many people (including yourself) currently live in the home where you live?* with the following response categories: 1, 2-4, and 5 or more. Employment status was assessed using the question *Which of the following best describes your current status? (check all that apply)* with the following response categories: employed full-time, employed part-time, in the military, unemployed and looking for work, unemployed and not looking for work, retired, student, homemaker, disabled or too ill to work, living off the land/a subsistence lifestyle (hunting, fishing, and berry gathering), and other.

Use, Satisfaction, and Perceived Effectiveness Questionnaire

At 6 weeks and 12 weeks, participants completed 12 multiple choice and 2 open-ended questions with no time frame specified. Use was assessed by asking a frequency question *About how often did you access the Health is Our Tradition: Balance and Harmony After Trauma website?* with response categories: more than once daily, once daily, 2-3 times a week, once a week, less than once a week, and never. Website use duration was assessed by the survey question *On average, how much time did you spend using the website once you accessed it?* with response categories: never went on the website, 5 min or less, at least 10 min, at least 30 min, at least 1 hour, more than 1 hour, and can't recall. Barriers to use were assessed by asking *Did you experience any of the following difficulties using the website?* with response categories: I did not have access to a computer, my computer was not working, I forgot my password, I could not play the videos, I could not play the audio, I could not print, and other problems. Satisfaction was assessed by asking, *Overall, how satisfied are you with your experiences with Health is Our Tradition: Balance and Harmony after Trauma?* Responses to this item were: never went on the website, quite dissatisfied, indifferent or mildly satisfied, mostly satisfied, or very satisfied. Participants were asked if they would recommend the website to a friend with response categories: yes, definitely; yes, I think so; no, I don't think so; and no, definitely not. Participants were asked if they would use the website in the future in the same manner. The first perceived effectiveness item was *If you completed at least one website Guide, did it help you deal more effectively with your traumatic experience and its consequences?* Response categories to this item were: yes, it helped a great deal; yes, it helped; no, it really didn't help; and no, it seemed to make things worse. Perceived helpfulness of the text messages and testimonial videos were assessed in the same manner. Whether the testimonial videos were upsetting was also rated as: no, not upsetting; yes, slightly upsetting; yes, moderately upsetting; and yes, very upsetting.

Website Metrics

Employing industry-standard website analytic tools [46], we tracked visit patterns, frequency of use, and session duration. Types of interaction with each section (guide) in the website were measured (eg, advanced through an entire guide, played a video, entered personal data into an activity, completed a drag and drop activity, and completed a relaxation meditation) per individual.

Experience With Technology Questionnaire

This 16-item questionnaire, developed by the research team, assesses perceived skill and confidence using computers and peripherals as well as aspects of using the internet such as navigating between pages, installing software such as a flash player, managing popups, and downloading, saving, or printing files from the internet. Items assessing skills used the following stems *Can you…* or *Do you…* with binary (no/yes) response categories. Items assessing confidence of technology use had the following response categories: not confident at all; I usually need help; it takes me a while, but I can manage; and confident.

Posttraumatic Stress Disorder Screening

The PC-PTSD scale consists of 4 binary (no/yes) items assessing PTSD symptoms and included an introductory sentence to cue respondents to lifetime traumatic events. A positive screen is indicated and has been validated in primary care settings, when the sum of positive responses is 3 or higher [39].

Posttraumatic Stress Disorder Checklist—Civilian Version

The PCL-C assesses key symptoms of PTSD via 17 items on a 5-point rating scale: (1) not at all; (2) a little bit; (3) moderately; (4) quite a bit; or (5) extremely. The PCL-C is applied generally to any traumatic event with respondents considering how much they have been bothered by the PTSD symptom in the last month. Higher scores indicated more severe symptomatology.

Alcohol Use Disorders Identification Test

The AUDIT is a 10-item screener for alcohol use in the past year. Higher scores indicated higher levels of alcohol use and problems related to alcohol abuse. Possible responses included never (0), monthly (1), 2-4 times per month (2), 2-3 times a week (3), or 4 or more times a week (4).

Patient Health Questionnaire

The PHQ measures the severity of somatic, depression, and anxiety symptoms in the past 2 weeks. All items used rating scales. Somatic symptom severity was derived from the first 15 items (PHQ-15). Possible responses to these were not bothered (0), bothered a little (1), or bothered a lot (2). PHQ-15 scores of 5, 10, and 15 represented cut-points for low, medium, and high somatic symptom severity, respectively. The next 9 items (PHQ-9) measured depression symptoms with a 4-point rating scale: not at all (0), several days (1), more than half the days (2), nearly every day (3). These were followed by binary (no/yes) questions and 3-point ratings of frequency of anxiety symptoms: not at all (0), several days (1), and more than half the days (2). A final item assessed the severity of somatic, anxiety, and depression symptoms with possible responses of not difficult at all (0), somewhat difficult (1), very difficult (2), or extremely difficult (3). Higher scores indicated more severe anxiety/depression symptomatology.

Drug Abuse Screening Tool

The DAST is a 10-item screener for drug use in the past 12 months. Items were binary (no/yes); higher scores indicated higher levels of problems related to drug abuse.

American Indian Symptom Inventory

The American Indian Symptom Inventory was developed for this study [5], specifically for use with AI/AN people. Drawing on items from such common measures as the Symptom Checklist-90 [47], as well as items suggested by focus groups as highly relevant to psychological distress among AI/ANs, this 50-item measure assesses a wide range of symptomatology prefaced by the statement, “Here is a list of problems people may have. How much have any of these problems bothered or upset you during the past month, including today...Not at all (0), Some (1), or A lot (2).”

Electronic Health Records

Electronic health record data collected included date of birth to calculate age; any depression and substance use screening scores; diagnoses of behavioral and trauma-related physical health disorders (eg, International Classification of Diseases–10 codes); prescribed behavioral health medications; and number of visits to primary care, emergency room/urgent care, inpatient, other clinical/ambulatory, and behavioral health clinics.

Data Analysis

Using SAS version 9.4 (SAS Institute, Inc), analysts created dummy variables for *check all that apply* questions (race, education, and employment) and examined frequencies, means, and standard deviations for these demographic responses as well as sex, number in household, and income. Chi-square, Fisher exact, or *t* tests were calculated to compare participant demographic characteristics by site. Changes in use, feasibility, and perceived effectiveness between the 6-week and 12-week follow-ups were examined with paired *t* tests. To examine the potential relationship of the intervention with service utilization, analysts compared clinic visit frequency from the health record data in the 12 months before the intervention (divided by 4 for comparison) and the 3 months after it with paired *t* tests. Patterns of symptomatology over the 12-week intervention was explored through one-way analysis of variances for repeated measures or repeated measures logistic regression tests. The Wilcoxon signed rank test for nonparametric data was used to test significance for nonnormally distributed data.

Results

Participant Demographics

A total of 24 participants were enrolled in the feasibility pilot at baseline (Table 3). Their mean age was 49 (SD 14) years, and 71% (17/24) were women. Participant data from SCF (n=12) and CNHS (n=12) were compared. The 2 samples were similar in most regards, except CNHS participants were more likely to have attained a 4-year college degree or greater (P=.04; Fisher exact Test), and be employed full time or part time (X^2_1=4.2; P=.04). One individual at each site was lost to follow-up before the 6-week measures resulting in a retention rate of 92% (22/24).

Table 3. Pilot participant demographics (N=24).

Characteristics	Value
Women, n (%)	17 (71)
Age (years), mean (SD)	49 (14)
Education, n (%)	
Less than high school/completed high school and/or GED[a]	8 (33)
Trade school/some college/2-year college graduate	8 (33)
4-year college graduate/some graduate school/postgraduate degree	8 (33)
Employment, n (%)	
Employed (full time, part time)	11(46)
Military	0 (0)
Unemployed	6 (25)
Retired/student/homemaker/subsistence/disabled/too ill	7 (29)
Pretax income, n (%)	
<US $10,000	8 (33)
US $10,000-US $29,999	6 (25)
US $30,000- US $49,999	5 (21)
>US $50,000	4 (17)
Number of household members, n (%)	
1	7 (29)
2-4	12 (50)
>5	5 (21)

[a] GED: General Educational Development.

Intervention Use, Satisfaction, and Perceived Effectiveness

The majority of participants 86% (19/24) reported any use of the website at 6 weeks, and nearly all 91% (20/22) used it at 12 weeks. However, participants did not use the website with the intensity requested by the researchers. Twelve of 22 participants reported that they used the website at the recommended intensity (a minimum of 3 times per week) at 6 weeks. At 12 weeks, use at recommended intensity dropped to 36% (8/22) (Table 4). Twenty of 22 participants reported reading the weekly tip text messages at 6 weeks. Of these 20 participants, 75% (15/20) considered them as at least moderately helpful. Similarly, 96% (21/22) reported reading the messages at 12 weeks, and 91% (19/22) considered them at least moderately helpful. At both 6 and 12 weeks, the majority of participants were mostly or very satisfied with the website, would recommend it to a friend in need, and reported that the tools in the website helped them cope with their trauma and its consequences.

Testimonial videos were viewed by 77% (17/22) at 6 weeks and 82% (18/22) reported viewing at 12 weeks. In all, 77% (13/17) and 83% (15/18) reported feeling that the testimonials were at least moderately helpful at 6 and 12 weeks, respectively. Out of 17 participants, 4 (24%) noted that testimonial videos were moderately to very upsetting at 6 weeks, a rate that dropped to 0% at 12 weeks (t_{14}=2.36; *P*=.03). Out of 22 participants, 20 (91%) reported at each time point that they would use the website again in the future. Difficulties in using the website included loss of the website address and/or password and slow internet connection speeds. One participant reported computer access/performance issues. A research team member spoke with the participant in the first week of the intervention and was able to address these issues, allowing this participant to use the website.

Table 4. Feasibility at 6 weeks and 12 weeks following enrollment.

Feasibility question	6 weeks (N=22)[a], n (%)	12 weeks (N=22)[a], n (%)
Use		
How often did you access the website?		
Never	3 (14)	2 (9)
Less than once a week	3 (14)	4 (18)
Once a week	4 (18)	8 (36)
2-3 times weekly	9 (41)	6 (27)
Once daily	2 (9)	1 (5)
More than once daily	1 (5)	1 (5)
How much time was spent on the website each session?		
At least 10 min	9 (41)	8 (36)
At least 30 min	3 (14)	7 (32)
At least 1 hour	6 (27)	1 (5)
More than 1 hour	1 (5)	4 (18)
Can’t recall	0 (0)	1 (5)
Did you read the weekly tips sent to your email address or by text to your phone?		
Yes, read weekly tip texts/emails	20 (91)	21(96)
Did you use the audio option on any of the information screens or activities?		
Yes, used the audio option for any guide	13 (59)	13 (59)
Did you view any of the testimonial (real life story) videos?		
Yes, viewed 1+ testimonial videos	17 (77)	18 (82)
What difficulties did you experience using the website?		
No computer access	0 (0)	1 (5)
Computer not working	1 (5)	1 (5)
Forgot password	2 (9)	0 (0)
Could not play the videos	0 (0)	1 (5)
Could not play the audio	0 (0)	1 (5)
Could not print	2 (9)	1 (5)
Lost or confused about website address	3 (14)	1 (5)
Wanted to use it on my smartphone but it didn’t work	2 (9)	2 (9)
Satisfaction		
How satisfied are you with your experiences with website?		
Never went on site	3 (14)	1 (5)
Quite dissatisfied	1 (5)	1 (5)
Indifferent or mildly satisfied	3 (14)	0 (0)
Mostly satisfied	6 (27)	8(36)
Very satisfied	9 (41)	12 (55)
Would you use the website again in the future?		
Yes, definitely	12 (55)	10 (46)
Yes, I think so	8 (36)	10 (46)
No, I don’t think so	1 (5)	2 (9)
No, definitely not	0 (0)	0 (0)

Feasibility question	6 weeks (N=22)[a], n (%)	12 weeks (N=22)[a], n (%)
Would you recommend the website to a friend in need?		
Yes, definitely	13 (59)	13 (59)
Yes, I think so	7 (32)	9 (41)
No, I don’t think so	1 (5)	0 (0)
No, definitely not	0 (0)	0 (0)
Perceived effectiveness		
Did the website help you deal more effectively with your traumatic experience and its consequences?		
Yes, a great deal	7 (32)	6 (27)
Yes, it helped	11 (50)	16 (73)
No, it really didn’t help	1 (5)	0 (0)
No, it seemed to make things worse	0 (0)	0 (0)
How helpful did you find the testimonial videos?		
Did not watch	5 (24)	4 (18)
Not at all helpful	0 (0)	0 (0)
Slightly helpful	3 (14)	3 (14)
Moderately helpful	4 (18)	3 (14)
Very helpful	9 (41)	12 (55)
Did you find any of the testimonial videos upsetting?[b]		
Did not watch	5 (24)	4 (18)
No, not upsetting	9 (41)	13 (59)
Yes, slightly upsetting	3 (14)	5 (23)
Yes, moderately upsetting	2 (9)	0 (0)
Yes, very upsetting	2 (9)	0 (0)
Helpfulness of weekly tip texts/emails		
Not at all helpful	2 (9)	0 (0)
Slightly helpful	3 (14)	3 (14)
Moderately helpful	11 (50)	10 (46)
Very helpful	4 (18)	9 (41)

[a] A total of 24 participants were consented and enrolled, but 2 were lost to follow-up before the 6-week measures.

[b] Participants reported significantly lower levels of finding testimonial videos upsetting at 12 weeks compared with 6 weeks (t_{14}=2.36; P=.03).

Website Usage

A total of 158 unique website sessions were recorded in the 12-week study period. The average duration of a session per participant was 16 min. Although all guides in the website were completed at least once, some were completed by more participants (Table 5). The 4 guides labeled for *first time visitors* were completed at rates ranging from 42% (10/24 participants; What can you do for yourself while healing?) to 100% (24/24 participants; What is trauma?). All 4 *first time visitors* guides were only psychoeducational in nature and did not include interactive activities.

Among those for *returning visitors*, guides that did not include written exercises (*Be in the moment, Discover your story, Notice your thoughts and feelings, Relax through breathing, Relax through visualization*, and *Relax your body*) were completed by a minimum of 3 out of 24 participants (13%) and a maximum of 18 out of 24 participants (75%; Be in the moment). Guides that included written exercises (*Change feelings by changing thoughts, Change negative thinking patterns, Cope with trauma reminders, Develop healthy relationships, Express your feelings and wants*, and *Identify your values and goals*) were completed less frequently. The least frequently completed guide was Develop healthy relationships. Out of 24, 7 participants (29%) completed it. The most frequently completed guide was *Discover your story*, which was completed by 11 out of 24 participants (49%). Overall, completion rates among guides that were only psychoeducational in nature were greater (t_{18}=−2.90; P<.01) than those that were interactive (guided relaxation, checklists/click or drag activities, and written exercises). There were no significant differences in completion rates relative to the content length of the guide.

Table 5. Health is Our Tradition guides (N=24).

Guide name[a]	Guide completion, n (%)
Trauma and your health: what is trauma?	21 (88)
Trauma and your health: how to get back in harmony?	11 (46)
Be in the moment	11 (46)
Trauma and your health: how trauma affects health?	10 (42)
Trauma and your health: common reactions to trauma	9 (38)
Notice your thoughts and feelings	9 (38)
Trauma and your health: what to do to feel better	8 (33)
Change how you think about sleep	8 (33)
Change negative thinking patterns	7 (29)
Identify your values and goals	7 (29)
Cope with trauma reminders	6 (25)
Plan something fun	6 (25)
Change feelings by changing thoughts	5 (21)
Relax through visualization	5 (21)
Discover your story	5 (21)
Form good sleep habits	5 (21)
Express your feelings and wants	4 (17)
Relax through breathing	4 (17)
Relax your body	3 (13)
Develop healthy relationships	2 (8)

[a] Listed by greatest number of guide completions.

Feasibility of Measures and Indicators of Effectiveness

Active participants at both sites were able to complete all self-report measures at baseline, 6 weeks and 12 weeks. Analyses of patterns of symptoms found reductions in self-reported PTSD, depression/anxiety/panic, physical symptoms related to PTSD, and problematic alcohol use between baseline and the 12-week follow-up. Depression symptom severity decreased between baseline and the 12-week follow-up visit.

Both sites were also able to query all data elements from electronic health records. There were no significant changes in patterns of diagnoses, screening, medications, or service utilization during exposure to the website.

Discussion

Principal Findings

A culturally adapted website for AI/AN people reporting symptoms of PTSD was developed and assessed for feasibility in 2 large AI/AN primary care settings. The website was well received with no difference in use, satisfaction, perceived effectiveness, or technical skills needed for use by age or gender. However, only half of the participants reported using the website at the recommended intensity at the 6-week follow-up and just over one-third did so at the 12-week follow-up (Table 4). Although website usage was only moderate and decreased across the study period, participants reported that the website was helpful in coping with their PTSD symptoms and related problems, and some self-reported symptoms decreased between enrollment and 12 weeks. Leader and provider concerns voiced during Phase 1 interviews and focus groups that participants may be unduly triggered by the website content, clinically decompensate, or have other negative impacts from using the website on their own were not supported by the data.

Collecting measures via self-report and electronic health record queries was successful, and a larger effectiveness trial is also warranted and appears feasible. Our participant attrition of 8% (2/24) is notably lower than the sizable attrition rates reported in self-administered internet interventions for depression, anxiety, and PTSD [48,49]. The weekly text-messaged *tips* may have contributed to retention and intensity of use by reminding participants about the study, their follow-up appointments, and highlighting content that may have sparked their interest. Although internet connectivity and speed were concerns for this population, very few participants reported these items as barriers to website access and usage. Participant website usage occurred throughout the 12-week intervention period with psychoeducational and shorter interactive activities being completed by more participants compared with worksheet-based activities requiring data input by participants. These findings align with preliminary evaluation of the *PTSD Coach Online* app, where participants indicated psychoeducation and self-management components as moderately helpful with symptom self-management as the most useful functions of the intervention [37].

Limitations

As a small pilot study with a convenience sample intended to assess feasibility, we cannot determine whether the culturally adapted intervention is effective in reducing posttraumatic and related symptoms. Key limitations of this study include its small sample size, lack of a control group, and lack of clinician-administered outcome measures. Calculations of website use via Google Analytics was limited as we could not link individuals to specific usage data owing to confidentiality and feasibility constraints. Text message reminder delivery, reading, and potential impacts on website usage were not electronically tracked. Finally, in the review of the AUDIT and DAST measures, we note that these measures assessed past year use rather than current use. Thus, we recommend that these measures be edited or reconsidered in a future effectiveness study. Although not many participants found the video testimonials upsetting, additional exploration of the degree of upset and content provoking upset is indicated. Intensity of use required for maximum therapeutic benefit should be assessed as well as potential order effects. The last guide was specifically developed, given feedback from Phase 1 participants about how trauma impacts important relationships, yet viewed by only 2 participants. Future research may be enhanced by adding clinician support to compare the effectiveness of the self-management WBTI alone compared with clinician-supported use. Additional usage log data collection and analysis methods may also be employed to assess more detailed activity of participant website usage behavior. Assessment of the effectiveness of text message reminders to reinforce website usage and effectiveness is warranted.

Conclusions

Despite these challenges and limitations, the results indicate Web-based therapeutic support of AI/AN adults with posttraumatic symptoms is feasible and warrant a large-scale randomized control trial to assess its potential effectiveness. The website appears to be a promising intervention given feasibility testing in 2 large, real-world AI/AN health systems. The intervention could be easily implemented by behavioral health staff integrated into primary care settings and was satisfactory to and viewed as helpful by AI/AN adults in this study.

Acknowledgments

This study was funded by the National Center on Minority Health and Health Disparities of the National Institutes of Health (Grant number 2P60MD000507).

Conflicts of Interest

None declared.

References

1. Beals J, Manson SM, Whitesell NR, Spicer P, Novins DK, Mitchell CM. Prevalence of DSM-IV disorders and attendant help-seeking in 2 American Indian reservation populations. Arch Gen Psychiatry 2005 Jan;62(1):99-108. [doi: 10.1001/archpsyc.62.1.99] [Medline: 15630077]
2. Manson SM, Beals J, Klein SA, Croy CD, AI-SUPERPFP Team. Social epidemiology of trauma among 2 American Indian reservation populations. Am J Public Health 2005 May;95(5):851-859. [doi: 10.2105/AJPH.2004.054171] [Medline: 15855465]
3. Walls ML, Whitbeck LB. Distress among Indigenous North Americans: generalized and culturally relevant stressors. Soc Ment Health 2011 Jul 1;1(2):124-136 [FREE Full text] [doi: 10.1177/2156869311414919] [Medline: 22368789]
4. Heart MY, DeBruyn LM. The American Indian Holocaust: healing historical unresolved grief. Am Indian Alsk Native Ment Health Res 1998;8(2):56-78. [doi: 10.5820/aian.0802.1998.60] [Medline: 9842066]
5. Beals J, Novins DK, Whitesell NR, Spicer P, Mitchell CM, Manson SM. Prevalence of mental disorders and utilization of mental health services in two American Indian reservation populations: mental health disparities in a national context. Am J Psychiatry 2005 Sep;162(9):1723-1732. [doi: 10.1176/appi.ajp.162.9.1723] [Medline: 16135633]
6. Boyd-Ball AJ, Manson SM, Noonan C, Beals J. Traumatic events and alcohol use disorders among American Indian adolescents and young adults. J Trauma Stress 2006 Dec;19(6):937-947. [doi: 10.1002/jts.20176] [Medline: 17195969]
7. Buchwald D, Goldberg J, Noonan C, Beals J, Manson S, AI-SUPERPFP Team. Relationship between post-traumatic stress disorder and pain in two American Indian tribes. Pain Med 2005;6(1):72-79. [doi: 10.1111/j.1526-4637.2005.05005.x] [Medline: 15669952]
8. Ehlers CL, Gizer IR, Gilder DA, Ellingson JM, Yehuda R. Measuring historical trauma in an American Indian community sample: contributions of substance dependence, affective disorder, conduct disorder and PTSD. Drug Alcohol Depend 2013 Nov 1;133(1):180-187 [FREE Full text] [doi: 10.1016/j.drugalcdep.2013.05.011] [Medline: 23791028]
9. Ehlers CL, Gizer IR, Gilder DA, Yehuda R. Lifetime history of traumatic events in an American Indian community sample: heritability and relation to substance dependence, affective disorder, conduct disorder and PTSD. J Psychiatr Res 2013 Feb;47(2):155-161 [FREE Full text] [doi: 10.1016/j.jpsychires.2012.10.002] [Medline: 23102628]

10. Bassett D, Buchwald D, Manson S. Posttraumatic stress disorder and symptoms among American Indians and Alaska Natives: a review of the literature. Soc Psychiatry Psychiatr Epidemiol 2014 Mar;49(3):417-433 [FREE Full text] [doi: 10.1007/s00127-013-0759-y] [Medline: 24022752]
11. Beals J, Belcourt-Dittloff A, Garroutte EM, Croy C, Jervis LL, Whitesell NR, AI-SUPERPFP Team. Trauma and conditional risk of posttraumatic stress disorder in two American Indian reservation communities. Soc Psychiatry Psychiatr Epidemiol 2013 Jun;48(6):895-905 [FREE Full text] [doi: 10.1007/s00127-012-0615-5] [Medline: 23135256]
12. Beals J, Manson SM, Croy C, Klein SA, Whitesell NR, Mitchell CM, AI-SUPERPFP Team. Lifetime prevalence of posttraumatic stress disorder in two American Indian reservation populations. J Trauma Stress 2013 Aug;26(4):512-520 [FREE Full text] [doi: 10.1002/jts.21835] [Medline: 23900893]
13. Robin RW, Chester B, Rasmussen JK, Jaranson JM, Goldman D. Prevalence and characteristics of trauma and posttraumatic stress disorder in a southwestern American Indian community. Am J Psychiatry 1997 Nov;154(11):1582-1588. [doi: 10.1176/ajp.154.11.1582] [Medline: 9356568]
14. Duran B, Sanders M, Skipper B, Waitzkin H, Malcoe LH, Paine S, et al. Prevalence and correlates of mental disorders among Native American women in primary care. Am J Public Health 2004 Jan;94(1):71-77. [doi: 10.2105/ajph.94.1.71] [Medline: 14713701]
15. Parker T, May PA, Maviglia MA, Petrakis S, Sunde S, Gloyd SV. PRIME-MD: its utility in detecting mental disorders in American Indians. Int J Psychiatry Med 1997;27(2):107-128. [doi: 10.2190/C6FD-7QWB-KNGR-M844] [Medline: 9565718]
16. Regier DA, Narrow WE, Rae DS, Manderscheid RW, Locke BZ, Goodwin FK. The de facto US mental and addictive disorders service system. Epidemiologic catchment area prospective 1-year prevalence rates of disorders and services. Arch Gen Psychiatry 1993 Feb;50(2):85-94. [doi: 10.1001/archpsyc.1993.01820140007001] [Medline: 8427558]
17. Kendall-Tackett KA. Physiological correlates of childhood abuse: chronic hyperarousal in PTSD, depression, and irritable bowel syndrome. Child Abuse Negl 2000 Jun;24(6):799-810. [doi: 10.1016/s0145-2134(00)00136-8] [Medline: 10888019]
18. Seng JS, Graham-Bermann SA, Clark MK, McCarthy AM, Ronis DL. Posttraumatic stress disorder and physical comorbidity among female children and adolescents: results from service-use data. Pediatrics 2005 Dec;116(6):e767-e776. [doi: 10.1542/peds.2005-0608] [Medline: 16322133]
19. Roy-Byrne PP, Davidson KW, Kessler RC, Asmundson GJ, Goodwin RD, Kubzansky L, et al. Anxiety disorders and comorbid medical illness. Gen Hosp Psychiatry 2008;30(3):208-225. [doi: 10.1016/j.genhosppsych.2007.12.006] [Medline: 18433653]
20. Campbell DG, Felker BL, Liu C, Yano EM, Kirchner JE, Chan D, et al. Prevalence of depression-PTSD comorbidity: implications for clinical practice guidelines and primary care-based interventions. J Gen Intern Med 2007 Jun;22(6):711-718 [FREE Full text] [doi: 10.1007/s11606-006-0101-4] [Medline: 17503104]
21. Bromet E, Sonnega A, Kessler RC. Risk factors for DSM-III-R posttraumatic stress disorder: findings from the National Comorbidity Survey. Am J Epidemiol 1998 Feb 15;147(4):353-361. [doi: 10.1093/oxfordjournals.aje.a009457] [Medline: 9508102]
22. Shalev A, Bleich A, Ursano RJ. Posttraumatic stress disorder: somatic comorbidity and effort tolerance. Psychosomatics 1990;31(2):197-203. [doi: 10.1016/S0033-3182(90)72195-0] [Medline: 2330402]
23. Kessler RC, Sonnega A, Bromet E, Hughes M, Nelson CB. Posttraumatic stress disorder in the National Comorbidity survey. Arch Gen Psychiatry 1995 Dec;52(12):1048-1060. [doi: 10.1001/archpsyc.1995.03950240066012] [Medline: 7492257]
24. Torkelson RM, Jellinek HM, Malec JF, Harvey RF. Traumatic brain injury: psychological and medical factors related to rehabilitation outcome. Rehabil Psychol 1983;28(3):169-176. [doi: 10.1037/h0090992]
25. Michaels AJ, Michaels CE, Moon CH, Zimmerman MA, Peterson C, Rodriguez JL. Psychosocial factors limit outcomes after trauma. J Trauma 1998 Apr;44(4):644-648. [doi: 10.1097/00005373-199804000-00013] [Medline: 9555835]
26. Addis ME, Wade WA, Hatgis C. Barriers to dissemination of evidence-based practices: addressing practitioners' concerns about manual-based psychotherapies. Clin Psychol Sci Pract 1999 May 11;6(4):430-441. [doi: 10.1093/clipsy.6.4.430]
27. Beidas RS, Kendall PC. Training therapists in evidence-based practice: a critical review of studies from a systems-contextual perspective. Clin Psychol (New York) 2010 Mar;17(1):1-30 [FREE Full text] [doi: 10.1111/j.1468-2850.2009.01187.x] [Medline: 20877441]
28. Berkman ND, Sheridan SL, Donahue KE, Halpern DJ, Viera A, Crotty K, et al. Health literacy interventions and outcomes: an updated systematic review. Evid Rep Technol Assess (Full Rep) 2011 Mar(199):1-941. [Medline: 23126607]
29. Vismara LA, Young GS, Stahmer AC, Griffith EM, Rogers SJ. Dissemination of evidence-based practice: can we train therapists from a distance? J Autism Dev Disord 2009 Dec;39(12):1636-1651 [FREE Full text] [doi: 10.1007/s10803-009-0796-2] [Medline: 19582564]
30. Garg AX, Adhikari NK, McDonald H, Rosas-Arellano MP, Devereaux PJ, Beyene J, et al. Effects of computerized clinical decision support systems on practitioner performance and patient outcomes: a systematic review. J Am Med Assoc 2005 Mar 9;293(10):1223-1238. [doi: 10.1001/jama.293.10.1223] [Medline: 15755945]
31. Depompei R, Gillette Y, Goetz E, Xenopoulos-Oddsson A, Bryen D, Dowds M. Practical applications for use of PDAs and smartphones with children and adolescents who have traumatic brain injury. NeuroRehabilitation 2008;23(6):487-499. [doi: 10.1201/noe0849313622.ch18] [Medline: 19127002]

32. Kobb R, Hoffman N, Lodge R, Kline S. Enhancing elder chronic care through technology and care coordination: report from a pilot. Telemed J E Health 2003;9(2):189-195. [doi: 10.1089/153056203766437525] [Medline: 12855041]
33. Sanders DL, Aronsky D. Biomedical informatics applications for asthma care: a systematic review. J Am Med Inform Assoc 2006;13(4):418-427 [FREE Full text] [doi: 10.1197/jamia.M2039] [Medline: 16622164]
34. Darkins A, Ryan P, Kobb R, Foster L, Edmonson E, Wakefield B, et al. Care Coordination/Home Telehealth: the systematic implementation of health informatics, home telehealth, and disease management to support the care of veteran patients with chronic conditions. Telemed J E Health 2008 Dec;14(10):1118-1126. [doi: 10.1089/tmj.2008.0021] [Medline: 19119835]
35. Hiratsuka VY, Moore L, Dillard DA, Avey JP, Dirks LG, Beach B, et al. Development of a screening and brief intervention process for symptoms of psychological trauma among primary care patients of two American Indian and Alaska native health systems. J Behav Health Serv Res 2017 Apr;44(2):224-241 [FREE Full text] [doi: 10.1007/s11414-016-9519-6] [Medline: 27328846]
36. Kuhn E, Greene C, Hoffman J, Nguyen T, Wald L, Schmidt J, et al. Preliminary evaluation of PTSD Coach, a smartphone app for post-traumatic stress symptoms. Mil Med 2014 Jan;179(1):12-18. [doi: 10.7205/MILMED-D-13-00271] [Medline: 24402979]
37. Miner A, Kuhn E, Hoffman JE, Owen JE, Ruzek JI, Taylor CB. Feasibility, acceptability, and potential efficacy of the PTSD Coach app: a pilot randomized controlled trial with community trauma survivors. Psychol Trauma 2016 May;8(3):384-392. [doi: 10.1037/tra0000092] [Medline: 27046668]
38. Gottlieb K, Outten B. Family wellness warriors. Fam Community Health 2011;34(3):235-241. [doi: 10.1097/FCH.0b013e318219623b] [Medline: 21633216]
39. Health is Our Tradition. 2016. URL: https://healthtradition.ucdenver.edu/ [accessed 2019-06-25]
40. 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 Psych 2004 Jan 1;9(1):9-14. [doi: 10.1185/135525703125002360]
41. 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. [doi: 10.1111/j.1360-0443.1993.tb02093.x] [Medline: 8329970]
42. Skinner HA. The drug abuse screening test. Addict Behav 1982;7(4):363-371. [doi: 10.1016/0306-4603(82)90005-3] [Medline: 7183189]
43. Conybeare D, Behar E, Solomon A, Newman MG, Borkovec TD. The PTSD Checklist-Civilian Version: reliability, validity, and factor structure in a nonclinical sample. J Clin Psychol 2012 Jun;68(6):699-713. [doi: 10.1002/jclp.21845] [Medline: 22517497]
44. 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]
45. Kroenke K, Spitzer RL, Williams JB. The PHQ-15: validity of a new measure for evaluating the severity of somatic symptoms. Psychosom Med 2002;64(2):258-266. [doi: 10.1097/00006842-200203000-00008] [Medline: 11914441]
46. Unified Advertising and Analytics - Google Marketing Platform. Google Analytics URL: http://www.google.com/analytics/features/ [accessed 2016-12-01]
47. Derogatis LR, Melisaratos N. The Brief Symptom Inventory: an introductory report. Psychol Med 1983 Aug;13(3):595-605. [doi: 10.1017/S0033291700048017] [Medline: 6622612]
48. 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]
49. Kuester A, Niemeyer H, Knaevelsrud C. Internet-based interventions for posttraumatic stress: a meta-analysis of randomized controlled trials. Clin Psychol Rev 2016 Feb;43:1-16. [doi: 10.1016/j.cpr.2015.11.004] [Medline: 26655959]

Abbreviations

AI/AN: American Indian and Alaska Native
AUDIT: Alcohol Use Disorders Identification Test
CNHS: Cherokee Nation Health Services
DAST: Drug Abuse Screening Tool
FWWI: Family Wellness Warriors Initiative
GED: General Educational Development
HIT: health information technologies
PCL-C: PTSD Checklist—Civilian Version
PC-PTSD: Primary Care—PTSD Screen
PHQ-9: Patient Health Questionnaire depression scale
PTSD: Posttraumatic stress disorder
SCF: Southcentral Foundation
UCD: University of Colorado Denver

WBTI: Web-based therapeutic intervention

Edited by G Eysenbach; submitted 11.02.19; peer-reviewed by C Pearson, Z El-Khatib; comments to author 16.05.19; revised version received 05.07.19; accepted 31.08.19; published 13.11.19.

Please cite as:
Hiratsuka VY, Moore L, Avey JP, Dirks LG, Beach BD, Dillard DA, Novins DK
An Internet-Based Therapeutic Tool for American Indian/Alaska Native Adults With Posttraumatic Stress Disorder: User Testing and Developmental Feasibility Study
JMIR Form Res 2019;3(4):e13682
URL: http://formative.jmir.org/2019/4/e13682/
doi:10.2196/13682
PMID:31719027

Original Paper

Use of Patient-Reported Data to Match Depression Screening Intervals With Depression Risk Profiles in Primary Care Patients With Diabetes: Development and Validation of Prediction Models for Major Depression

Haomiao Jin[1,2], PhD; Shinyi Wu[1,2,3], PhD

[1]Suzanne Dworak-Peck School of Social Work, University of Southern California, Los Angeles, CA, United States

[2]Edward R Roybal Institute on Aging, University of Southern California, Los Angeles, CA, United States

[3]Daniel J Epstein Department of Industrial and Systems Engineering, University of Southern California, Los Angeles, CA, United States

Corresponding Author:
Haomiao Jin, PhD
Suzanne Dworak-Peck School of Social Work
University of Southern California
1150 S Olive St
Suite 1400
Los Angeles, CA, 90015
United States
Phone: 1 213 821 6441
Email: haomiaoj@usc.edu

Abstract

Background: Clinical guidelines recommend screening for depression in the general adult population but recognizes that the optimum interval for screening is unknown. Ideal screening intervals should match the patient risk profiles.

Objective: This study describes a predictive analytics approach for mining clinical and patient-reported data from a large clinical study for the identification of primary care patients at high risk for depression to match depression screening intervals with patient risk profiles.

Methods: This paper analyzed data from a large safety-net primary care study for diabetes and depression. A regression-based data mining technique was used to examine 53 demographics, clinical variables, and patient-reported variables to develop three prediction models for major depression at 6, 12, and 18 months from baseline. Predictors with the strongest predictive power that require low information collection efforts were selected to develop the prediction models. Predictive accuracy was measured by the area under the receiver operating curve (AUROC) and was evaluated by 10-fold cross-validation. The effectiveness of the prediction algorithms in supporting clinical decision making for six "typical" types of patients was demonstrated.

Results: The analysis included 923 patients who were nondepressed at the study baseline. Five patient-reported variables were selected in the prediction models to predict major depression at 6, 12, and 18 months: (1) Patient Health Questionnaire 2-item score; (2) the Sheehan Disability Scale; (3) previous problems with depression; (4) the diabetes symptoms scale; and (5) emotional burden of diabetes. All three depression prediction models had an AUROC>0.80, comparable with published depression prediction studies. Among the 6 "typical" types of patients, the algorithms suggest that patients who reported impaired daily functioning by health status are at an elevated risk for depression in all three periods.

Conclusions: This study demonstrated that leveraging patient-reported data and prediction models can help improve identification of high-risk patients and clinical decisions about the depression screening interval for diabetes patients. Implementation of this approach can be coupled with application of modern technologies such as telehealth and mobile health assessment for collecting patient-reported data to improve privacy, reducing stigma and costs, and promoting a personalized depression screening that matches screening intervals with patient risk profiles.

(JMIR Form Res 2019;3(4):e13610) doi:10.2196/13610

KEYWORDS

patient-reported data; patient-centered decision making; depression screening; depression; diabetes; health information technology; data analytics; predictive modeling; machine learning; data mining

Introduction

Depression is a common comorbid mental illness for many chronic conditions including diabetes [1-3]. About 10%-20% of adults with diabetes have major depressive disorders [1,2], but as high as 50% of these individuals are undiagnosed [3]. Comorbid depression and diabetes may significantly worsen the course of both disorders, leading to higher medical costs, reduced functioning and quality of life, increased risks of cardiovascular diseases, and increased mortality [4-9].

Depression screening is effective in identifying people with depression [10]. In primary care, validated tools such as the Patient Health Questionnaire (PHQ) are usually used for depression screening [11]. Results of the screening provide important information for decision making by clinicians and patients, often triggering referral of the patient to mental health professionals or an adjustment of care plans to better coordinate the physical and mental care.

Depression screening policy specifies the scope and frequency of screening and therefore plays a central role in implementing depression screening for large populations. Recent updates from the US Preventive Services Task Force (USPSTF) recommend depression screening for every adult, but they do not provide guidelines for screening frequency [10]. In practice, primary care providers may administer one-time depression screenings for all patients and must thereafter rely on practical strategies to determine the frequency [10,12].

One practical strategy recommends depression screening at a fixed frequency (eg, once a year). Providers may implement the fixed-frequency strategy by checking the time interval between the current clinical encounter and the last screening. They would administer a new screening if the time between screenings has been too long, according to a predefined requirement. This strategy has the benefit of guaranteeing at least a minimum screening frequency, but the major drawback is not differentiating between patients who are at high risk and those who are those at low risk. As a result, high-risk patients may receive insufficient screening and be undiagnosed for their depression, while low-risk patients may undergo unnecessary screening, which wastes clinical resources and patients' time.

Another practical strategy relies on the judgements of health professionals (eg, primary care physicians, nurse practitioners, case managers, or social workers) to determine the screening frequency [12]. For example, a depression screening may be administered if a patient talks about his/her persistent depressed mood or loss of interest in performing daily activities with health providers. The problems associated with this strategy are threefold. First, a primary care encounter may be brief, and there would need to be enough time to address multiple conditions and issues along with depression-related problems [13]. Second, there are stigmas surrounding depression, which can often discourage, if not inhibit, patients from talking about their affective problems [14]. Third, culture differences may influence the expression of depression-related problems. Racial minorities such as Latinos may be more likely to attribute signs of depression to physical symptoms and not have their depression diagnosed [15,16].

Increasing the availability of patient-generated and patient-reported data may provide novel opportunities to improve policy making for depression screening. Many patient-generated and patient-reported variables such as socioeconomic status, stress level, and functional disability are significantly associated with depression [17-21]. Leveraging these data to develop accurate predictive models may improve identification of patients at high risk for depression and enable providers to match screening frequency with patient risk profiles. The systematic collection of patient-generated and patient-reported data also establishes a mechanism to encourage patients to communicate their affective problems: The logistics of a questionnaire serve as a formal protocol to elicit information, guiding patients to report their depression-related symptoms and concerns more effectively.

A few studies have explored the development of depression risk predictive tools [22-25]. King et al [22] used a stepwise logistic regression method to develop a risk prediction model for major depression at 6 and 12 months by using a dataset from general practice attendees. Huang et al [23] developed a logistic regression model for major depression at 6 and 12 months using electronic health record data. In addition, using the logistic regression method, Wang et al [24] developed a risk prediction model of major depression in 2-3 years for the general adult population. Finally, Liu et al [25] developed a decision tree model for predicting poststroke depression in stroke survivors. The predictive accuracy, as measured by the area under the receiver operating curve (AUROC) of those studies, was approximately 0.8.

The objective of this paper is to develop tools for predicting depression risk by using clinical, patient-generated, and patient-reported data from primary care patients with diabetes and discuss the implications of applying the prediction tools for making policies and practices for depression screening. The study will examine an array of demographic, clinical, patient-generated, and patient-reported variables and select the most predictive ones to assemble accurate depression prediction models. Implementation of the prediction models for risk prediction and the implications for patient data collection and decision making for the depression screening interval will be discussed.

Methods

Data Source

We analyzed data from the Diabetes–Depression Care-Management Adoption Trial (DCAT) [13,16,19,26-31], which enrolled 1406 patients with type 2 diabetes from eight safety-net primary care clinics affiliated with the Los Angeles

County Department of Health Services (LAC-DHS), the second largest safety-net health care system in the United States. As described by Wu et al [13,16], the DCAT adopted a quasiexperimental comparative effectiveness design with three study groups to test an automated telephone depression screening and monitoring system, which was integrated with a diabetes disease management program in one group to facilitate the adoption of a collaborative depression care model. The other two groups were the LAC-DHS usual primary care clinics and the diabetes disease management program for adopting collaborative depression care, both of which did not include the automated telephone assessment system. The eight clinics were staffed by six teams of providers, two in each study group that were matched by geographic location and patient sociodemographics to form the three study groups. The patients were not randomly assigned; each patient was assigned to a study group based on the clinic from which he/she was recruited.

Patients were eligible for the DCAT if they were ≥18 years, had been diagnosed with type 2 diabetes, had a working phone number, spoke English or Spanish, and could read and understand the consent form. Patients with baseline possible suicidal ideation, cognitive impairment, alcohol abuse, or recent use of lithium or antipsychotic medication were ineligible for the trial. In all three trial groups, patients were assessed with the 9-item Patient Health Questionnaire (PHQ-9) [11], which has nine questions that are consistent with the nine criteria on which the 4th edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV) bases its diagnosis of depression. However, the PHQ-9 score was neither an inclusion criterion nor an exclusion criterion for trial enrollment. Hence, the DCAT sample comprised both patients with and those without depression. The trial was conducted from 2010 through 2013, with four waves of comprehensive assessments of patient-reported data through individual interviews at study baseline and 6-, 12-, and 18-month follow-ups.

Study Sample

Our study sample featured patients with a 2-item Patient Health Questionnaire (PHQ-2) score<3 measured at study baseline [32,33]. PHQ-2 is a brief depression assessment tool with two questions that inquire about the frequency of depressed mood and anhedonia over the past 2 weeks [32,33]. The scores range from 0 to 6, and a cutoff point of 3 was shown to have a sensitivity of 83% and specificity of 92% for detection of current major depression [32]. The PHQ-2 comprises the first two items of the PHQ-9. The PHQ-9 score ranges from 0 to 27, and a cutoff point of 10 has been suggested for the diagnosis of major depression [11]. Kroenke et al [11,32] suggested a “two-step” procedure to administer PHQ screening in primary care: (1) All patients should receive the easy-to-administer PHQ-2 screening, and patients with PHQ-2 scores<3 are ruled out for current depression. (2) Patients with PHQ-2 scores≥3 will receive the full PHQ-9 assessment for diagnosis and determination of the severity of depressive symptoms. We used PHQ-9 scores≥10 at 6-, 12-, and 18-month follow-up assessments of DCAT as our primary predicted outcomes.

The reason that the subsample of the DCAT is pertinent for this predictive analysis study is because no intervention was provided for patients who were ruled out for depression at baseline. The only exception is that the patients in the technology group were reached by the automated call system every 3 months for up to 12 months to undergo assessment for the PHQ-2, pain, and self-management activities, including regular physical and fun activities and patient requests to be contacted by a provider. The call lasted about 2 minutes, and the response rate was about 50%. Because the DCAT provided no intervention and limited contact, these patients remained in their natural course of health conditions.

Measurement of Patient-Generated and Patient-Reported Data

The predictive variables were selected from the extensive patient-reported data collected in the DCAT via in-person or telephone interviews, as described by Wu et al [16]. In addition to the PHQ-2 depressive symptomatology measurement described above, two standard questions from the structured clinical interview for DSM-IV were used to assess dysthymia [34]. The Sheehan Disability Scale was used to rate functional impairment on a 10-point Likert scale [35], which consisted of three questions on whether disease symptoms have disrupted the respondent’s work, social life, or family life. Chronic pain was defined as pain present most of the time for 6 months or longer during the past year and measured by the Short Form-12 [36] (one-item) pain impact questionnaire that asks respondents to rate the level of pain interference with normal work on a scale of 1-5 (1=*none* and 5=*extremely*). Anxiety was assessed by the Brief Symptom Inventory [37]. Health-related quality of life was assessed using the Medical Outcomes Study Short-Form Health Survey [36] and the Physical and Mental Component summaries. Patient satisfaction with diabetes care and emotional care was assessed with a single item on a 5-point Likert scale, ranging from very 1 (very dissatisfied) to 5 (very satisfied).

The assessment also included the Whitty 9-item questionnaire to assess diabetes symptoms [38], the brief two-item Diabetes Distress Scale to screen for distress [39], and the summary of Diabetes Self-Care Activities Questionnaire to assess self-reported adherence [40]. Moreover, we assessed social demographics, employment, self-reported weight and height (from which we calculated the body mass index), diabetes treatment and complications, comorbid medical illness, and socioeconomic stress.

Statistical Analysis

To select the most predictive variables and develop the prediction models, we used the Least Absolute Shrinkage and Selection Operator (LASSO), a regression-based data mining technique that performs both variable selection and regularization to concurrently enhance the simplicity and prediction accuracy of the statistical model it produces [41,42]. LASSO can achieve both of these goals by introducing a penalization parameter, lambda, to a standard regression to penalize the size of coefficient estimation. As the value of lambda increases, every coefficient estimation shrinks toward 0 but at varying speed. The shrinkage speed provides a way to rank the predictive power of each variable because those that approach 0 slowly would have a better chance of being selected for the final prediction model when manipulating lambda from small values to large values. In this way, LASSO has the advantage of seeking a model that predicts well and is parsimonious. [42]

Candidate variables examined in this study included 53 demographic, clinical, and patient-generated and patient-reported variables measured at the baseline of DCAT (Table 1). We entered these variables into three LASSO logistic regressions to predict the occurrence of major depression, measured by PHQ-9≥10, at 6, 12, and 18 months. Thereafter, the variables were order-ranked based on their coefficient estimation shrinkage speed. We further considered efforts of patients required in reporting data for the variables. Hence, the criteria for selecting variables in the final predictive models were that the selected predictors should have slow shrinkage speed and be among the top ranked variables; in addition, those that require lesser efforts for patients to report data (measured by number of items in a scale) are preferred over those that required greater efforts.

We performed 10-fold cross-validation to evaluate the predictive accuracy of the candidate predictive models based on selected variables. The 10-fold cross-validation procedure randomly partitioned the dataset into 10 subsets. The complete procedure involved 10 rounds of computation; in each round, 9 of the 10 subsets of data were used to fit the prediction model and the remaining subset was used to assess predictive accuracy, which was measured by the AUROC. The 10 rounds of evaluation generated 10 AUROC scores, and the overall predictive accuracy was computed by taking the average. The resulting AUROC has a range of 0-1, with a larger value indicating better predictive accuracy. We also evaluated the model calibration using the Brier score. Model calibration refers to whether the predicted probabilities or scores can be used to predict the actual class membership probabilities. The Brier score has a range from 0 to 1, with smaller values indicating better model calibration. We examined the predictive models using the selected variables to predict major depression at 6, 12, and 18 months. The derived risk prediction algorithms can be interpreted in the same way as ordinary logistic regression, that is, the linear models compute the log odds risk of major depression. The final risk scores were calculated by taking the exponential of the linear scores. The final scores have a range from 0 to positive infinite, with higher scores indicating a higher risk of major depression. We evaluated sensitivity and specificity corresponding to different cutoff points of the predicted risk scores.

To demonstrate prediction algorithms' effectiveness in supporting clinical decision making, we generated risk profiles for six "typical" types of patients: (1) the "median" patient, whose reported values in the selected predictors were set to the median values of our DCAT analysis sample; (2) the "average" patient, whose reported values in the selected predictors were set to the average values of our DCAT analysis sample; (3) patients with frequent diabetes symptoms, who reported experiencing a few days of diabetes-related symptoms (eg abnormal thirst, blurred vision, etc), determined by a Whitty-9 Diabetes Symptoms Scale score of 2.5, and had the rest of the predictors at the median values; (4) patients with some depressive symptoms, who had a PHQ-2 score of 2 and had the rest of the predictors at the median values; (5) patients who reported having previous depression problems but currently had no depressive symptoms; and (6) patients whose diseases affected their daily functioning, as measured by a Sheehan Disability Scale score of 4. We generated recommendations for clinical decision making based on the predicted depression risk scores for the 3 periods (ie, 6, 12, and 18 months) and chosen cutoff points. Depression screening was suggested for a period if the predicted depression risk score was equal to or higher than the chosen cutoff point. No screening was suggested for a period if the predicted depression risk scores up to that period were all lower than their cutoff points. If the predicted depression risk score for a period was lower than its cutoff point and at least one of the predicted risk scores for previous periods was higher than the other cutoff points, further clinical judgement was recommended to determine whether the patient needed depression screening. As shown below, we used a risk score of 8 as the cutoff point for the three periods. These cutoff points have sensitivities of 86%, 75%, and 90% and specificities of 64%, 71%, and 64% for 6, 12, and 18 months, respectively. Providers may choose different cutoff points based on their clinical needs. A lower cutoff point would increase sensitivity, reduce specificity, decrease the number of individuals with undiagnosed depression, and require more provider and patient time and resources to conduct more screenings.

Results

We identified 999 patients with baseline PHQ-2 scores<3 from the DCAT dataset. We excluded 76 patients from analysis due to incomplete data on the candidate baseline predictors as listed in Table 1, and thus, 923 patients were included in the training and validation of risk prediction models. Table 1 summarizes the sample included in the analysis. Among the analysis sample, 83/776 (10.7%), 72/741 (9.7%), and 77/625 (12.3%) had PHQ-9 scores≥10 at 6, 12, and 18 months, respectively. The retention rates are 84.1% (776/923), 80.3% (741/923), and 67.8% (741/923) at 6, 12, and 18 months, respectively.

The comparison of baseline candidate predictors between patients who were later depressed at 6, 12, or 18 months and patients who were not later depressed is shown in the Multimedia Appendix 1. Table 2 shows that the LASSO regression method produced three predictive models to forecast major depression at 6, 12, and 18 months, respectively. Among the 53 variables examined, 6 patient-generated and patient-reported variables consistently appeared as the top predictors with the slowest shrinkage speed in all the three models: (1) the PHQ-2 score, measuring the two core symptoms of depression (ie, depressed mood and anhedonia); (2) the Sheehan Disability Scale score, measuring interference of health issues to work, social, and family life; (3) patient-reported previous problems with depression; (4) patient-reported diabetes symptoms; (5) patient-reported emotional burden from diabetes; and (6) patient-reported stressors, measuring the total number of stressors using a 12-item survey. The total number of stressors was excluded from the final models since the survey is longer than the other five scales and its exclusion has little impact on predictive accuracy. All the three final five-predictor models have an AUROC larger than 0.80. Sensitivity and specificity of identifying the depressed cases vary by cutoff points. Increasing sensitivity would reduce specificity, decrease the number of undiagnosed depression, and require more provider and patient time and resources to conduct more screenings. Balanced sensitivity and specificity are often recommended in the literature, but providers may opt to choose higher sensitivity if time and resources are sufficient. The primary goal is to reduce undiagnosed depression. The calibration performance of all the three models is good, as indicated by the small values of Brier scores.

Table 3 shows the predicted depression risk profiles and screening suggestions for five "typical" types of patients. The suggestions are based on predicted depression risk scores and the chosen cutoff points. Depression screening is suggested for a period if its predicted depression risk score is equal to or higher than the chosen cutoff point. No screening is suggested for a period if the predicted depression risk scores up to that period are all lower than their cutoff points. If the predicted depression risk score for a period is lower than its cutoff point and at least one of the predicted risk scores for the previous periods is higher than its cutoff points, further clinical judgement is suggested to determine whether the patient needs depression screening. Based on a cutoff point of 8 for the three prediction periods, patients with median reported values in the five predictors in the DCAT samples had depression risk scores lower than the cutoff points in all three periods and were therefore not recommended to receive screenings. Patients with average reported values in the five predictors, frequent diabetes symptoms, or some depressive symptoms were predicted to be at risk for depression at 6 months. They were advised to receive a follow-up screening at 6 months, and providers were recommended to consider whether further screenings at 12 and 18 months were warranted. Depression risk for patients who reported having previous depression problems but currently had no depressive symptoms was predicted to increase with time and across the cutoff points at 12 and 18 months. Thus, those patients were recommended screening at 12 and 18 months. Finally, patients who reported that their daily functioning was impaired by health status were predicted to be at an elevated risk for depression in all three periods. Depression screenings were recommended every 6 months for those high-risk patients. Providers may choose cutoff points other than 8, based on their clinical needs. A lower cutoff point would increase sensitivity, reduce specificity, decrease the number of individuals with undiagnosed depression, and require more provider and patient time and resources to conduct more screenings.

Table 1. Baseline data from the Diabetes–Depression Care-Management Adoption Trial to train and validate the depression prediction models.

Variables	Value (n=923)
Demographic variables	
Age, mean (SD)	53.15 (9.52)
Female, n (%)	557 (60)
Latino, n (%)	818 (89)
Birth place in the United States, n (%)	120 (13)
Spanish the preferred language, n (%)	759 (82)
Less than high school education, n (%)	629 (68)
Married, n (%)	515 (56)
Patient-generated and patient-reported variables	
Two-item Patient Health Questionnaire, mean (SD)	0.62 (0.80)
Smoking, n (%)	62 (7)
Onset age of diabetes, mean (SD)	43.27 (10.49)
Family history of diabetes, n (%)	704 (76)
Diabetes self-care score, mean (SD)	4.48 (1.25)
Diabetes symptoms score, mean (SD)	1.50 (0.50)
Chronic pain, n (%)	163 (18)
Pain impact on normal work, n (%)	128 (14)
Pain impact on social life, n (%)	72 (8)
Bothered by thinking or dreaming of terrible things, n (%)	75 (8)
Six-item Brief Symptom Inventory, mean (SD)	0.50 (1.82)
Previous diagnosis of major depression, n (%)	41 (4)
Previous diagnosis of anxiety disorders, n (%)	7 (1)
Ever had a problem with depression, n (%)	128 (14)
Ever had a problem with anxiety, n (%)	31 (3)
Talking to someone about your depression, n (%)	44 (5)
Number of stressors, mean (SD)	2.04 (2.07)
Diabetes emotional burden, mean (SD)	2.56 (1.91)
Diabetes regimen distress, mean (SD)	2.45 (1.89)
Unemployed, n (%)	579 (63)
Doing work for extra income, n (%)	85 (9)
No health insurance, n (%)	68 (7)
Feeling that my financial situation is getting worse, n (%)	288 (31)
Having difficulty in paying bills, n (%)	615 (67)
Having money left over at the end of the month, n (%)	839 (91)
Financial worry score, mean (SD)	3.69 (2.07)
Sheehan disability scale, mean (SD)	1.21 (2.12)
Clinical variables	
Hemoglobin A_{1c}, mean (SD) (%)	9.21 (2.13)
Body mass index, mean (SD) (kg/m^2)	32.62 (7.19)
Number of diabetes complications, mean (SD)	1.11 (1.08)
Taking insulin, n (%)	493 (53)
On diabetes treatment - oral medication, n (%)	803 (87)

Variables	Value (n=923)
On diabetes treatment - nutritionist observation, n (%)	32 (4)
Had microalbumin test done in the past 6 months, n (%)	688 (75)
Taking pain medication, n (%)	131 (14)
Taking antidepressant, n (%)	53 (6)
Taking anxiety medication, n (%)	7 (1)
Number of ICD-9[a] diagnoses in the past 6 months, mean (SD)	6.93 (4.27)
Hospitalized overnight in past 6 months, n (%)	122 (13)
ICU[b] admission in the past 6 months, n (%)	23 (3)
ER[c] use in the past 6 months, n (%)	228 (25)
Number of primary care visits in the past 6 months, mean (SD)	9.09 (5.94)
Had missed medical appointment in the past 6 months, n (%)	92 (10)
Future health care cost, mean (SD)	6614.79 (3714.58)
Enrolled into disease management program, n (%)	643 (70)
Receiving automatic telephone screening and monitoring, n (%)	307 (33)

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

[b] ICU: intensive care unit.

[c] ER: emergency room.

Table 2. Predictive accuracy to forecast major depression at 6, 12, and 18 months among primary care patients with diabetes recruited in the Diabetes–Depression Care-Management Adoption Trial study.

Predicted outcome	Algorithm to predict log odds risk score of major depression	AUROC[a]	Brier score	Cutoff[b]	Sensitivity (%)	Specificity (%)
6-month depression	100×exp(–4.58 + 0.55×PHQ2[c] + 0.13×SDS[d] + 0.80×PBD[e] + 0.80×DSS[f] + 0.09×DEB[g])	0.80	0.07	22; 14; 8; 5	49; 70; 86; 90	88; 79; 64; 45
12-month depression	100×exp(–4.83 + 0.54×PHQ2 + 0.21×SDS + 1.26×PBD + 0.80×DSS + 0.05×DEB)	0.81	0.06	22; 14; 8; 5	56; 67; 75; 86	90; 84; 71; 56
18-month depression	100×exp(–4.53 + 0.46×PHQ2 + 0.16×SDS + 1.23×PBD + 0.62×DSS + 0.20×DEB)	0.83	0.07	22; 14; 8; 5	67; 76; 90; 91	87; 78; 64; 47

[a] AUROC: area under receiver operating curve.

[b] The cutoff points 22, 14, 8, and 5 correspond to 100×exp(–1.5), 100×exp(–2.0), 100×exp(–2.5), and 100×exp(–3.0), respectively.

[c] PHQ-2: Patient Health Questionnaire - 2-item.

[d] SDS: Sheehan Disability Scale.

[e] PBD: ever had a problem with depression.

[f] DSS: Diabetes Symptoms Scale.

[g] DEB: diabetes emotional burden.

Table 3. Examples of patient depression risk profiles and suggestions regarding follow-up depression screening.

Profile	PHQ-2[a]	SDS[b]	PBD[c]	DSS[d]	DEB[e]	Predicted depression risk score					
						6 mo		12 mo		18 mo	
						Score	Suggestion	Score	Suggestion	Score	Suggestion
The "median" patient	0	0	0	1.33	1	3	No Screening	2	No Screening	3	No Screening
The "average" patient	0.62	1.21	0.14	1.50	2.56	22	Depression Screening	6	Clinical judgement	8	Clinical judgement
Frequent diabetes symptoms	0	0	0	2.5	1	8	Depression Screening	6	Clinical judgement	6	Clinical judgement
Some depressive symptoms	2	0	0	1.33	1	10	Depression Screening	7	Clinical judgement	8	Clinical judgement
Had previous depression problems but currently no symptoms	0	0	1	1.33	1	7	No Screening	9	Depression Screening	10	Depression Screening
Diseases affect functionality	1	4	0	2.5	1	25	Depression Screening	25	Depression Screening	19	Depression Screening

[a] PHQ-2: Patient Health Questionnaire - 2-item.

[b] SDS: Sheehan Disability Scale.

[c] PBD: ever had a problem with depression.

[d] DSS: Diabetes Symptoms Scale.

[e] DEB: diabetes emotional burden.

Discussion

Principal Findings

This paper examined 53 demographic, clinical, and patient-generated and patient-reported variables from a large clinical trial dataset generated from an urban safety-net primary care setting. Using a regression-based data mining technique, five predictive variables, which are all patient-generated and patient-reported, were selected to develop three accurate prediction models to forecast major depression at 6, 12, and 18 months.

Compared with the fixed-frequency depression-screening policy, the prediction models enable providers to distinguish patients at high risk from those at low risk for depression and match screening frequency to patient risk profiles. The results recommend providers, who would follow USPSTF's guidelines to conduct depression screening for every adult patient, to collect additional patient-generated and patient-reported data from patients judged to be nondepressed; these data included information on functional disability, whether the patient ever had problems with depression, diabetes symptoms, and diabetes emotional burden. The risk scores computed from the initial depression-screening result and the additionally collected patient-generated and patient-reported data could provide valuable decision-support information for providers.

Compared with the clinical judgement–based depression-screening policy, the selected predictors specify a formal protocol to elicit targeted and useful information about depression risks. Rather than the occasion and less structured way of talking about depression during clinical encounters, a formal data collection protocol can guide patients to systematically review and disclose indicators of depressive vulnerabilities, such as current depressive symptoms and ever having had problems with depression, and stressors that may activate depression in the future like functional disability, various diabetes symptoms, and emotional burden from diabetes. Additionally, providers may carry out the predictor collection protocol in private and confidential ways, such as through a technology-facilitated assessment [31], to reduce patient stigma and encourage the disclosure of depression-related information.

It is worth mentioning that having clinical resources in place to ensure appropriate information collection and patient outreaching/follow-ups is indispensable for successful implementation of the prediction model–based screening policy. In the United States, such clinical resources are often implemented in the form of team-based collaborative depression care in a primary care setting [43,44], which can be effective but costly, especially for safety-net care systems [27]. Recent developments in telehealth and mobile health technologies for depression, such as automated telephone [31,45] and text-messaging [46] depression assessment, may facilitate more cost-effective implementation. Hay et al [27] reported that using an automatic telephonic assessment technology can reduce the costs of collecting information on depressive symptoms, medication, and functioning from US $35 per assessment by a health professional to US $2 per assessment via technology. Such automatic telephonic technology was also reportedly perceived as useful, private, and secure by safety-net primary care patients [29]. Patient-generated and patient-reported predictors for depression can also be collected via a text messaging assessment, which lowers costs even more than an automatic telephonic assessment [46]. In addition, providers may use technologies to proactively reach out to patients identified as high risk by the prediction models. Wu et al [13] reported that a proactive collaborative primary and depression care intervention based on automatic telephonic assessment is effective in reducing depressive symptoms, facilitating the diabetes care processes, and improving patient satisfaction.

Findings from the predictive modeling analysis in this paper may benefit providers from countries other than the United States that recommend a different approach for managing depression screening. The UK National Institute for Health and Care Excellence recommends targeted case identification rather than universal depression screening [47]. The Canadian Task Force on Preventive Health Care currently does not provide strong recommendations for depression screening but advises that clinicians be alert to the possibility of depression in patients with clinical clues, especially those at increased risk of depression, and implement treatment as appropriate when depression is diagnosed [48]. Findings of this paper identify top predictors for depression among patients with diabetes and suggest profiles of diabetic patients at an increased risk of depression. High-risk patients may include patients with frequent diabetes symptoms; patients with previous problems with depression; patients with some but not clinically significant depressive symptoms; and patients whose daily functioning in work, social, and family life is impaired by diseases.

Limitations

This study has several limitations. The analysis is based on a clinical trial dataset consisting of safety-net primary care patients with type 2 diabetes. Whether the results could be generalized to other patient populations with diabetes needs more validation. An advantage of the data mining technique used in this study is that it could be automated. This eases repetition of the current analysis on other datasets. This study is also limited by measuring the major depression by PHQ-9, which is a depression screening tool rather than a diagnostic tool. The use of scores of ≥10 to classify major depression in this paper may lead to false-positive depression cases. As described in Kroenke et al [49], PHQ-9<10 seldom occur in individuals with major depression, whereas scores of ≥15 usually signify the presence of major depression. In the gray zone of scores between 10 and 14 points, increasing the PHQ-9 scores is associated with increasing specificity and declining sensitivity. This paper tested only one predictive algorithm—LASSO. Although the algorithm satisfactorily achieved our goals of developing parsimonious prediction models with comparative predictive accuracy, other data mining algorithms may achieve similar or even better performance. Estimating standard errors or CIs and performing statistical tests are still challenging and unresolved issues for LASSO regression [50]. As a result, we are unable to show reliable estimates of the CIs for the LASSO estimates of regression coefficients. The dataset used to develop the prediction models covered an 18-month period, which may not be long enough to state recommendations about the screening periodicity. The DCAT intervention applied to the study sample may result in different relationships between the baseline predictors and future depression across different study groups, although our analysis has included the study group as one of the candidate predictors and did not determine this variable to be among the top predictive ones. Finally, there is a lack of external validation of the derived prediction models besides the cross-validation used in the paper. Future research should test the prediction models derived in this paper on an external dataset. It is also unknown whether primary care providers would accept use of the prediction models to guide policy making for depression screening. The success of actual implementation may be influenced by many factors such as predictive accuracy, leadership support, appropriate training, providers' experience and expectation of predictive analytics, and the costs and benefits associated with the implementation. It may be burdensome and challenging for providers to collect the patient-generated and patient-reported data required to run the prediction models. Providers may also be overwhelmed by responding to high-risk depression cases identified by the prediction models. As discussed above, integrating the prediction models with telehealth and mobile-health technologies could automate the data collection, risk prediction and patient outreach, and consequently relieve providers from time-consuming work. Testing the feasibility and effectiveness of such a prediction- and technology-facilitated depression-screening intervention would be worth further investigation.

Comparisons to Prior Works

This study adds to the growing body of research that utilizes predictive analytics to improve the forecasting of depression [19,22-25,28]. A strength of this study is that the derived prediction models use a relatively small number of predictors (ie, 5 compared with 10-20 in most prior studies) while still achieving comparable predictive accuracy as measured by the AUROC. This reduces the time and effort needed in clinical practices to collect necessary information for depression prediction and may therefore facilitate the implementation of the prediction model–based depression-screening policy. A second strength of this study is that it examined a broad array of candidate predictors that include both patient-generated and patient-reported variables and clinical variables. Finally, the prediction models are developed in a policy-making context and address an important policy issue (ie, unknown optimal depression screening intervals) as identified in the recent update of clinical guideline on depression screening from the USPSTF [10]. Surprisingly, few studies have addressed this important policy issue [10,12]. Our findings suggest that the collection of patient-generated and patient-reported data with the application of advanced data mining techniques may be a promising direction to match depression screening intervals with patients' depression risk profiles.

Conclusions

This study developed prediction models to improve identification of primary care patients with diabetes who are at high risk for depression and discusses the implications for policymaking on depression screening. The derived models rely on five patient-generated and patient-reported predictors to make accurate predictions. Implementation of the prediction models, especially when integrating with telehealth and mobile-health assessment technologies for data collection and patient outreach, may improve privacy, reduce stigma and costs, and promote a personalized depression-screening policy that matches screening frequency with patient risk profiles.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Comparison of baseline data from the Diabetes–Depression Care-Management Adoption Trial (DCAT) between patients who were later depressed versus those who were not depressed.

[PDF File (Adobe PDF File)111 KB - formative_v3i4e13610_app1.pdf]

References

1. Ducat L, Philipson LH, Anderson BJ. The mental health comorbidities of diabetes. JAMA 2014 Aug 20;312(7):691-692 [FREE Full text] [doi: 10.1001/jama.2014.8040] [Medline: 25010529]
2. Katon WJ. The comorbidity of diabetes mellitus and depression. Am J Med 2008 Nov;121(11 Suppl 2):S8-15 [FREE Full text] [doi: 10.1016/j.amjmed.2008.09.008] [Medline: 18954592]
3. Li C, Ford ES, Zhao G, Ahluwalia IB, Pearson WS, Mokdad AH. Prevalence and correlates of undiagnosed depression among U.S. adults with diabetes: the Behavioral Risk Factor Surveillance System, 2006. Diabetes Res Clin Pract 2009 Feb;83(2):268-279. [doi: 10.1016/j.diabres.2008.11.006] [Medline: 19111364]
4. Chen P, Chan Y, Chen H, Ko M, Li C. Population-based cohort analyses of the bidirectional relationship between type 2 diabetes and depression. Diabetes Care 2013 Feb;36(2):376-382 [FREE Full text] [doi: 10.2337/dc12-0473] [Medline: 23150281]
5. Egede LE. Diabetes, major depression, and functional disability among U.S. adults. Diabetes Care 2004 Feb;27(2):421-428. [Medline: 14747223]
6. Golden SH, Lazo M, Carnethon M, Bertoni AG, Schreiner PJ, Diez RAV, et al. Examining a bidirectional association between depressive symptoms and diabetes. JAMA 2008 Jun 18;299(23):2751-2759 [FREE Full text] [doi: 10.1001/jama.299.23.2751] [Medline: 18560002]
7. Murray CJ, Lopez AD. Alternative projections of mortality and disability by cause 1990-2020: Global Burden of Disease Study. Lancet 1997 May 24;349(9064):1498-1504. [doi: 10.1016/S0140-6736(96)07492-2] [Medline: 9167458]
8. Murray CJ, Lopez AD. Global mortality, disability, and the contribution of risk factors: Global Burden of Disease Study. Lancet 1997 May 17;349(9063):1436-1442. [doi: 10.1016/S0140-6736(96)07495-8] [Medline: 9164317]
9. Murray CJL, Lopez AD. Measuring the global burden of disease. N Engl J Med 2013 Aug 01;369(5):448-457. [doi: 10.1056/NEJMra1201534] [Medline: 23902484]
10. Siu AL, Bibbins-Domingo K, Grossman DC, Baumann LC, Davidson KW, Ebell M, et al. Screening for Depression in Adults: US Preventive Services Task Force Recommendation Statement. JAMA 2016 Jan 26;315(4):380-387. [doi: 10.1001/jama.2015.18392] [Medline: 26813211]
11. 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]
12. Nease DE, Maloin JM. Depression screening: a practical strategy. J Fam Pract 2003 Feb;52(2):118-124. [Medline: 12585989]
13. Wu S, Ell K, Jin H, Vidyanti I, Chou C, Lee P, et al. Comparative Effectiveness of a Technology-Facilitated Depression Care Management Model in Safety-Net Primary Care Patients With Type 2 Diabetes: 6-Month Outcomes of a Large Clinical Trial. J Med Internet Res 2018 Apr 23;20(4):e147 [FREE Full text] [doi: 10.2196/jmir.7692] [Medline: 29685872]
14. Brown C, Conner KO, Copeland VC, Grote N, Beach S, Battista D, et al. Depression stigma, race, and treatment seeking behavior and attitudes. J Community Psychol 2010 Apr;38(3):350-368 [FREE Full text] [doi: 10.1002/jcop.20368] [Medline: 21274407]
15. Lyness JM, Cox C, Curry J, Conwell Y, King DA, Caine ED. Older age and the underreporting of depressive symptoms. J Am Geriatr Soc 1995 Mar;43(3):216-221. [Medline: 7884106]
16. Wu S, Ell K, Gross-Schulman SG, Sklaroff LM, Katon WJ, Nezu AM, et al. Technology-facilitated depression care management among predominantly Latino diabetes patients within a public safety net care system: comparative effectiveness trial design. Contemp Clin Trials 2014 Mar;37(2):342-354. [doi: 10.1016/j.cct.2013.11.002] [Medline: 24215775]
17. Everson SA, Maty SC, Lynch JW, Kaplan GA. Epidemiologic evidence for the relation between socioeconomic status and depression, obesity, and diabetes. J Psychosom Res 2002 Oct;53(4):891-895. [Medline: 12377299]
18. Greenglass E, Fiksenbaum L, Eaton J. The relationship between coping, social support, functional disability and depression in the elderly. Anxiety, Stress & Coping 2006 Mar;19(1):15-31. [doi: 10.1080/14659890500436430]
19. Jin H, Wu S, Di Capua P. Development of a Clinical Forecasting Model to Predict Comorbid Depression Among Diabetes Patients and an Application in Depression Screening Policy Making. Prev Chronic Dis 2015 Sep 03;12:E142 [FREE Full text] [doi: 10.5888/pcd12.150047] [Medline: 26334714]
20. Pittenger C, Duman RS. Stress, depression, and neuroplasticity: a convergence of mechanisms. Neuropsychopharmacology 2008 Jan;33(1):88-109 [FREE Full text] [doi: 10.1038/sj.npp.1301574] [Medline: 17851537]

21. Torres L. Predicting levels of Latino depression: acculturation, acculturative stress, and coping. Cultur Divers Ethnic Minor Psychol 2010 Apr;16(2):256-263. [doi: 10.1037/a0017357] [Medline: 20438164]
22. King M, Walker C, Levy G, Bottomley C, Royston P, Weich S, et al. Development and Validation of an International Risk Prediction Algorithm for Episodes of Major Depression in General Practice Attendees. Arch Gen Psychiatry 2008 Dec 01;65(12):1368. [doi: 10.1001/archpsyc.65.12.1368]
23. Huang SH, LePendu P, Iyer SV, Tai-Seale M, Carrell D, Shah NH. Toward personalizing treatment for depression: predicting diagnosis and severity. J Am Med Inform Assoc 2014 Nov 01;21(6):1069-1075. [doi: 10.1136/amiajnl-2014-002733]
24. Wang J, Sareen J, Patten S, Bolton J, Schmitz N, Birney A. A prediction algorithm for first onset of major depression in the general population: development and validation. J Epidemiol Community Health 2014 Jan 03;68(5):418-424. [doi: 10.1136/jech-2013-202845]
25. Liu R, Yue Y, Jiang H, Lu J, Wu A, Geng D, et al. A risk prediction model for post-stroke depression in Chinese stroke survivors based on clinical and socio-psychological features. Oncotarget 2017 Sep 08;8(38):62891-62899 [FREE Full text] [doi: 10.18632/oncotarget.16907] [Medline: 28968957]
26. Di Capua P, Wu B, Sednew R, Ryan G, Wu S. Complexity in Redesigning Depression Care: Comparing Intention Versus Implementation of an Automated Depression Screening and Monitoring Program. Popul Health Manag 2016 Dec;19(5):349-356. [doi: 10.1089/pop.2015.0084] [Medline: 27028043]
27. Hay JW, Lee P, Jin H, Guterman JJ, Gross-Schulman S, Ell K, et al. Cost-Effectiveness of a Technology-Facilitated Depression Care Management Adoption Model in Safety-Net Primary Care Patients with Type 2 Diabetes. Value Health 2018 Dec;21(5):561-568. [doi: 10.1016/j.jval.2017.11.005] [Medline: 29753353]
28. Jin H, Wu S, Vidyanti I, Di Capua P, Wu B. Predicting Depression among Patients with Diabetes Using Longitudinal Data. A Multilevel Regression Model. Methods Inf Med 2015;54(6):553-559. [doi: 10.3414/ME14-02-0009] [Medline: 26577265]
29. Ramirez M, Wu S, Jin H, Ell K, Gross-Schulman S, Myerchin SL, et al. Automated Remote Monitoring of Depression: Acceptance Among Low-Income Patients in Diabetes Disease Management. JMIR Ment Health 2016 Jan 25;3(1):e6 [FREE Full text] [doi: 10.2196/mental.4823] [Medline: 26810139]
30. Vidyanti I, Wu B, Wu S. Low-income minority patient engagement with automated telephonic depression assessment and impact on health outcomes. Qual Life Res 2015 May;24(5):1119-1129 [FREE Full text] [doi: 10.1007/s11136-014-0900-8] [Medline: 25543270]
31. Wu S, Vidyanti I, Liu P, Hawkins C, Ramirez M, Guterman J, et al. Patient-centered technological assessment and monitoring of depression for low-income patients. J Ambul Care Manage 2014;37(2):138-147 [FREE Full text] [doi: 10.1097/JAC.0000000000000027] [Medline: 24525531]
32. Kroenke K, Spitzer RL, Williams JBW. The Patient Health Questionnaire-2: validity of a two-item depression screener. Med Care 2003 Nov;41(11):1284-1292. [doi: 10.1097/01.MLR.0000093487.78664.3C] [Medline: 14583691]
33. Löwe B, Kroenke K, Gräfe K. Detecting and monitoring depression with a two-item questionnaire (PHQ-2). J Psychosom Res 2005 Feb;58(2):163-171. [doi: 10.1016/j.jpsychores.2004.09.006] [Medline: 15820844]
34. First M, Spitzer R, Gibbon M, Williams J. Structured Clinical Interview for DSM-IV Axis I Disorders (SCID-I). Washington, DC: American Psychiatric Publishing, Incorporated; 1997.
35. Sheehan DV, Harnett-Sheehan K, Raj BA. The measurement of disability. Int Clin Psychopharmacol 1996 Jun;11 Suppl 3:89-95. [Medline: 8923116]
36. Ware J, Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. Med Care 1996 Mar;34(3):220-233. [Medline: 8628042]
37. Derogatis LR, Melisaratos N. The Brief Symptom Inventory: an introductory report. Psychol Med 1983 Aug;13(3):595-605. [Medline: 6622612]
38. Whitty P, Steen N, Eccles M, McColl E, Hewison J, Meadows K, et al. A new self-completion outcome measure for diabetes: is it responsive to change? Qual Life Res 1997 Jul;6(5):407-413. [Medline: 9290307]
39. 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]
40. Toobert DJ, Hampson SE, Glasgow RE. The summary of diabetes self-care activities measure: results from 7 studies and a revised scale. Diabetes Care 2000 Jul;23(7):943-950 [FREE Full text] [Medline: 10895844]
41. Tibshirani R. Regression shrinkage and selection via the LASSO. J R Stat Soc Ser B Methodol 1996;58(1):267-288. [doi: 10.1111/j.1467-9868.2011.00771.x]
42. Zhao P, YU B. On model selection consistency of Lasso. J Mach Learn Res 2003;3:1157-1182.
43. Wu B, Jin H, Vidyanti I, Lee P, Ell K, Wu S. Collaborative depression care among Latino patients in diabetes disease management, Los Angeles, 2011-2013. Prev Chronic Dis 2014 Aug 28;11:E148 [FREE Full text] [doi: 10.5888/pcd11.140081] [Medline: 25167093]
44. Katon W, Unützer J, Wells K, Jones L. Collaborative depression care: history, evolution and ways to enhance dissemination and sustainability. Gen Hosp Psychiatry 2010;32(5):456-464 [FREE Full text] [doi: 10.1016/j.genhosppsych.2010.04.001] [Medline: 20851265]

45. Kroenke K, Theobald D, Wu J, Norton K, Morrison G, Carpenter J, et al. Effect of telecare management on pain and depression in patients with cancer: a randomized trial. JAMA 2010 Jul 14;304(2):163-171 [FREE Full text] [doi: 10.1001/jama.2010.944] [Medline: 20628129]
46. Jin H, Wu S. Screening Depression and Related Conditions via Text Messaging Versus Interview Assessment: Protocol for a Randomized Study. JMIR Res Protoc 2019 Mar 29;8(3):e12392. [doi: 10.2196/12392] [Medline: 30924787]
47. National Institute for Health and Care Excellence. Depression in adults: recognition and management URL: https://www.nice.org.uk/guidance/cg90 [accessed 2019-06-19]
48. Joffres M, Jaramillo A, Dickinson J, Lewin G, Pottie K, Shaw E, Canadian Task Force on Preventive Health Care. Recommendations on screening for depression in adults. CMAJ 2013 Jun 11;185(9):775-782 [FREE Full text] [doi: 10.1503/cmaj.130403] [Medline: 23670157]
49. 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]
50. Kyung M, Gill J, Ghosh M, Casella G. Penalized regression, standard errors, and Bayesian lassos. Bayesian Anal 2010 Jun;5(2):369-411. [doi: 10.1214/10-BA607]

Abbreviations

AUROC: area under the receiver operating curve
DCAT: Diabetes–Depression Care-Management Adoption Trial
DEB: diabetes emotional burden
DSM-IV: Diagnostic and Statistical Manual of Mental Disorders
DSS: Diabetes Symptoms Scale score
ER: emergency room
ICD-9: International Classification of Diseases, Ninth Revision, Clinical Modification.
ICU: intensive care unit
LAC-DHS: Los Angeles County Department of Health Services
LASSO: Least Absolute Shrinkage and Selection Operator
PBD: ever had a problem with depression
PHQ: Patient Health Questionnaire
SDS: Sheehan Disability Scale score
USPSTF: US Preventive Services Task Force

Edited by G Eysenbach; submitted 03.02.19; peer-reviewed by J Wang, X Hu, R Morris, S Cho; comments to author 25.02.19; revised version received 10.06.19; accepted 31.08.19; published 01.10.19.

Please cite as:
Jin H, Wu S
Use of Patient-Reported Data to Match Depression Screening Intervals With Depression Risk Profiles in Primary Care Patients With Diabetes: Development and Validation of Prediction Models for Major Depression
JMIR Form Res 2019;3(4):e13610
URL: https://formative.jmir.org/2019/4/e13610
doi:10.2196/13610
PMID:31573900

Original Paper

Development of a Digital Decision Support Tool to Aid Participation of Children With Disabilities in Pediatric Rehabilitation Services: Explorative Qualitative Study

Elin Vinblad[1*], MSc; Ingrid Larsson[2*], PhD; Maria Lönn[3*], MSc; Emma Olsson[4*], BSc; Jens M Nygren[2*], PhD, Prof Dr; Petra Svedberg[2*], PhD, Prof Dr

[1]Child and Youth Rehabilitation Services, Region Skåne, Kristianstad, Sweden
[2]School of Health and Welfare, Halmstad University, Halmstad, Sweden
[3]Region Halland, Halmstad, Sweden
[4]Falkenbergs Municipality, Falkenberg, Sweden
[*]all authors contributed equally

Corresponding Author:
Petra Svedberg, PhD, Prof Dr
School of Health and Welfare
Halmstad University
Box 823
Halmstad, S-30118
Sweden
Phone: 46 35167693
Email: petra.svedberg@hh.se

Abstract

Background: Building a health care system in accordance with the rule of law requires child-centered care, where children and young people, regardless of ability, are allowed to participate in visits with their health care professionals. As part of an overall project focusing on developing and implementing a digital decision support tool to increase the participation of children with disabilities in pediatric rehabilitation, this study brings new knowledge as to how this specific patient group views participation.

Objective: The aim of this formative study was to explore the experiences of children and young people with disabilities concerning increasing their participation in the pediatric rehabilitation services.

Methods: The formative study had an explorative design, based on a latent qualitative content analysis with an inductive approach. Interviews were conducted with 20 children (6-17 years) and 8 young people (19-30 years) with disabilities about their experiences of participation in pediatric rehabilitation services.

Results: A total of 3 categories emerged reflecting the participants' possibilities of participation in the pediatric rehabilitation services: to feel involved, to feel independent, and to work in partnership. To feel involved meant being listened to and being connected, to feel independent meant being admitted and being enabled, and to work in partnership meant being supported and being able to entrust others with the decision making. With the overall theme *moving toward empowerment of children in pediatric rehabilitation*, a true feeling of participation can be experienced.

Conclusions: The views of children and young people with disabilities are that children should be given the prerequisites for empowerment by being allowed to feel involved and independent as well as to work in partnership to experience true participation in the pediatric rehabilitation services. This finding is essential in the design of a digital decision support tool based on the children's needs and perspectives.

(JMIR Form Res 2019;3(4):e14493) doi:10.2196/14493

KEYWORDS

child; child care; decision making; eHealth; disabled children; patient participation; rehabilitation; qualitative research

Introduction

Background

The Patient Act in Sweden [1] aims to protect the right of an individual to participate in health care decisions. Allowing patients to participate in decision making has been shown to improve patient-reported satisfaction, patient compliance, and reported quality of care [2]. The United Nation's Convention on the Rights of the Child (UNCRC) will be incorporated into Swedish law on January 1, 2020, thus supplementing the Patient Act [3]. The UNCRC states that all children have a right to participate in society, to have an opinion, and to be listened to in all matters affecting them [4]. The incorporation of the UNCRC into Swedish law, together with the Patient Act, further increases the pressure on pediatric rehabilitation services to involve the children in their care, as child participation becomes both a prerequisite for quality health care as well as an obligation by law.

The implementation of interventions that support children's participation in health care is still rare [5-8] and especially where children with disabilities are concerned [9]. The consequences of unsatisfactory participation are, however, particularly severe for children with disabilities as their needs for extensive care place greater demands on efficient interaction with professionals [10,11]. A core method for strengthening the patient's active role in health care is *shared decision making* (SDM), where the patient and staff work together to make decisions based on current evidence and the patient's experiences, while also considering different alternatives, goals, and preferences [12]. Although SDM aligns well with the principles of UNCRC, research has shown that child participation in health care and rehabilitation services is almost nonexistent [6,13,14], thus indicating a lack of strategies to promote child participation. This is true for most children, and especially where children with neuropsychiatric and intellectual disabilities are concerned [11,15].

The pediatric rehabilitation services in Sweden have a tradition of a family-centered approach, where the whole family is regarded as the client. Researchers have in recent years asserted the need for a conceptual move from a family-centered practice toward a child-centered practice within health care services. Taking the concept of child-centered practice one step further, researchers now stress the need to move from earlier guidelines of *child-focused* health care toward *the child's focus* within their own health care [16]. The importance of including the child´s perspective has been interpreted as the importance of inviting children to share their needs [17-20] and allowing the child to take a more active role in goal formulation and decision making [21,22]. The use of SDM is thus supported as well as promoting the fulfillment of the UNCRC. The child is involved and accepted as a partner in a child-centered way of working, thus strengthening the child and increasing the possibilities for gaining confidence and independence [23]. However, knowledge regarding how to move toward a more child-focused practice is at present scant, and research is thus needed. In a study aiming to identify barriers and facilitators for SDM in mental health services for children and young people, Gondek et al [13] found that SDM is not always consciously practiced, although staff considers that they fulfill the requirements upon learning the definition. The same study also found that, according to the caregivers, young people who are not included in decisions about their care, may be less willing to follow treatment protocols, thus indicating the need to involve children and young people in decisions regarding their health care [13]. It has been found in the literature that children's involvement in research that concerns them increases the relevance, content, and ethics of the research and results [24,25]. It is, thus, necessary for researchers to include the children that are affected when investigating how the transformation toward a child-centered practice can be achieved. However, it has been identified in the previous research that there are certain difficulties when it comes to including children with disabilities in research. The difficulties may be related to the research setting, such as ethics, having sufficient time and competent interviewers, or related to the child's specific difficulties, such as language, cognition, and motor skills [23,26-28]. This has resulted in the group being generally unrepresented within research that affects them [23,29], and it is, thus, even more essential to involve this target group in the research.

Although research projects incorporating children with disabilities as partners in research appear to be scarce, previous studies have included children from the other areas of health care. These studies, focusing on increasing child participation in health care, have shown that digital tools providing support for communication and participation have a positive outcome on the children's involvement and engagement in their own care (IM Carlsson et al, unpublished data, 2019; I Larsson et al, unpublished data, 2019) [30-33]. However, despite widespread optimism about the potential of such digital tools, research reveals that disparities remain in relation to health and well-being among those who are in vulnerable positions such as children with disabilities [34], and the evidence base for informing policy and practice in relation to this is insufficient [35,36].This study is part of an overall research project with the aim of strengthening children's participation in pediatric rehabilitation services by developing, validating, and evaluating an electronic health intervention that is based on a digital decision support tool for children with disabilities. The overall research project involves children with disabilities, their parents, young people with disabilities, and professionals in a user-centered design process, striving to mediate the children's voices and introducing a child-centered way of working in pediatric rehabilitation services.

Objectives

The aim of this formative study was to explore the experiences of children and young people with disabilities about increasing their participation in the pediatric rehabilitation services.

Methods

Design

This formative study had an explorative design, with the purpose of informing the development of an intervention that is based on a digital decision support tool for children with disabilities. Formative research helps the researcher to identify and understand interests, behaviors, and needs of target populations to inform the decisions and actions when designing and implementing an intervention [37]. The analysis was based on qualitative content analysis, with an inductive approach. Qualitative content analysis can be used to identify similarities and differences in text-bound material [38] and offers a flexible, pragmatic method for developing and expanding knowledge as well as understanding human experiences [38,39]. Qualitative content analysis has been widely used in health care research and was considered an appropriate approach for this study as the target group's views on participation were assumed to vary greatly.

Participants

The target group was 28 children and young people with disabilities. The inclusion criteria for the children were as follows: aged 6 to 18 years, with an established contact with the pediatric rehabilitation services in Southern Sweden. Furthermore, the children had to be able to participate in an adapted interview setting, as well as answering questions using their chosen mode of communication. A total of 20 children were included in the study. Most children had been using the rehabilitation services since early childhood, with an average of 6.3 years of contact. For the young people, inclusion criteria were as follows: aged 19 to 30 years, with previous established contact with pediatric rehabilitation services in Southern Sweden. A total of 8 young people were included in the study (Table 1).

Table 1. Sociodemographic data (n=28).

Characteristic	Number of participants
Age range (years)	
6-10	8
11-13	7
14-17	5
18-24	6
25-30	2
Sex	
Female	15
Male	13
Main disabilities	
Physical disability	14
Intellectual disability	5
Autism spectrum disorder	9

Data Collection

The recruitment of and individual interviews with the children with disabilities were carried out in 2017 and early 2018 by the first author (EV), who was an employee at a local pediatric rehabilitation center in Southern Sweden with an experience of working with children with disabilities. Each child was able to decide when and where the interview was to take place and whether they preferred another adult to attend. Only one child chose to have a parent present. The interviews varied from 20 min to 60 min in length and followed a semistructured interview guide with open-ended questions, resulting in an informal dialogue with the children in a conversational tone. The children were asked to expand certain statements and were guided through their experiences of participation. They were also encouraged to forward thoughts to the researcher that could arise after the interview, an offer which 2 participants used.

The recruitment of the young people with disabilities was carried out in 2017 by an established contact person from a pediatric rehabilitation service in Southern Sweden. The young people first participated in group interviews held by the third, fourth, and last authors (ML, EO, and PS); the group size varied between 2 and 5 people. Individual interviews were then carried out with 2 participants who were considered able to further contribute with essential data [40]. The data collection was carried out in semistructured interviews, starting with open questions such as "Can you describe what participation in your pediatric rehabilitation looks like to you?" To attain greater depth in the data, follow-up questions were used to further investigate the participants' views and experiences. Participants had full access to all augmentative tools needed throughout the interviews, and all the interviews were digitally recorded.

Data Analysis

Qualitative content analysis was used, and both the manifest and latent content was analyzed [38,41]. The interviews were transcribed verbatim by the interviewing authors (EV, ML, and EO) and were read several times to gain familiarity. The first (EV) and second (IL) authors conducted the initial analyses of the children's interviews, while the initial analysis of the young people's interviews was carried out by the third and fourth authors (ML and EO). In the initial analysis, meanings or phrases with information relevant to the object of the study were identified and extracted, together with the surrounding text, to preserve the content. The meaning units were then abstracted and coded. Then, the first (EV) and the second (IL) authors searched for differences and similarities in the total material from both children and young people before grouping the codes into subcategories and categories. A total of 6 subcategories and 3 categories, reflecting the core message in the interviews, constituted the manifest content. The content of the categories was abstracted into a theme reflecting the underlying meaning, constituting the latent content. The analysis was discussed in the research group to establish consensus. Representative quotations from the children and young people were used to illustrate the data in the categories. An example of the data analysis is shown in Table 2.

Table 2. Examples of meaning units, codes, subcategories, categories, and theme.

Meaning unit	Code	Subcategory	Category	Theme
I tell them [that I don't want to] and we do something else. [Child 9]	The professionals listen to the child.	Being listened to	Feeling involved	Moving toward empowerment of children in pediatric rehabilitation
But I know it best [what I need]. [Child 6]	The child is respected.	Being connected	Feeling involved	—[a]
It's like...I can understand what they are saying. [Child 4]	The child understands the information.	Being admitted	Feeling independent	—
To be allowed to take it at your own pace, to be allowed time depending on your challenges. [Young person 4]	Adjusting the child's pace.	Being enabled	Feeling independent	—
When I needed a new chair, dad printed all the papers because he knew we had to convince [the therapists of the need for a new chair]. [Young person 2]	The child has parental support.	Being supported	Working in partnership	—
I think it's better if they [the therapists] decide... [Child 11]	The child trusts adults in decision making.	Entrusting decision making	Working in partnership	—

[a] All the categories were abstracted into the overall theme moving toward empowerment of children in pediatric rehabilitation.

Ethical Considerations

This study conforms to the ethical principles for research on human beings as set out by the World Medical Association in the declaration of Helsinki [42] as well as the national guidelines on ethical principles [43]. The ethical approval was granted by the Regional Ethical Review Board at Lund University, Sweden (No: 2017/707). The children, young people, and parents received oral and written information about the study and the voluntary nature regarding participation and their right to withdraw at any time without explanation. Written informed consent was given from the parents of participating children as well as from the participating young people before their inclusion in the study. The children gave their consent either in writing or orally. The children and young people were, immediately after the interview, given the opportunity to discuss any emotions or thoughts that had emerged with staff who possessed the necessary knowledge to deal with their concerns.

Results

The results reflecting the children's and young people's positive experiences of participation in the pediatric rehabilitation services included the overall theme *Moving toward empowerment of children in pediatric rehabilitation* and 3 related categories: *Feeling involved*; *Feeling independent*, and *Working in partnership* (Textbox 1).

Textbox 1. Overview of theme, categories, and subcategories.

Theme: Moving toward empowerment of children in pediatric rehabilitation.

Categories and subcategories:

- Feeling involved
    - Being listened to
    - Being connected
- Feeling independent
    - Being admitted
    - Being enabled
- Working in partnership
    - Being supported
    - Entrusting decision making

Moving Toward Empowerment of Children in Pediatric Rehabilitation

The children's possibilities for moving toward empowerment in pediatric rehabilitation incorporated their experiences of participation in each of the categories. When children felt involved, independent, and were in partnership, they learned to trust that their voices would be heard, that they would be granted access, and that they would be able to make decisions or entrust others to make decisions for them. This, in turn, led to inner feelings of inclusion, autonomy, and self-efficacy that boosted empowerment.

Feeling Involved

The category *Feeling involved* incorporated the subcategories of *being listened to* and *being connected*. The category emerged from the importance of professionals addressing the child directly, asking questions and listening to the answers, encouraging the child to make suggestions and keeping the child at the center of their attention. When a child was used to feeling involvement, a higher level of self-esteem was developed, which was in turn related to the child's ability to demand participation for himself. When a child was not included, this led to exclusion, a low level of self-esteem and a lack of confidence in demanding participation.

Being Listened To

The subcategory of *being listened to* contained children's and young people's experiences of contributing with opinions that would directly affect the care they received as well as taking part in the decision making as compared with only being present at a meeting. Both children and young people could reflect that this is easier for children with a high level of self-esteem as these children are more likely to demand participation. Although a child simply expressed that "you have to be really good at saying what you want," a young person expressed a deeper understanding that children with varying abilities have varying possibilities of participation:

> *I have had the chance to discuss [my care], but that's because I'm an outgoing person as you can hear, and therefore, I have an advantage.* [Young person 1]

Both children and young people wanted to be listened to and be allowed to take part in the decision making concerning their care and thus feel involved. Regarding the right to make decisions, the young people leaned more toward allowing children to decide everything for themselves, whereas the children preferred to share the responsibilities:

> *Perhaps children shouldn't decide everything, but at least a little...So that everyone is happy.* [Child 17]

Both children and young people expressed an understanding that not everything is optional where necessary procedures regarding their rehabilitation were concerned. They did, however, wish for a child to be involved in briefings about the procedures to answer questions and capture any negative emotions that may arise. A child explains in the following quote why the inclusion of children in decision making can affect their motivation to participate in the care that the professionals deem necessary:

> *Because it's all about me. I'm the one who then has to do everything...Otherwise, it won't turn out as well, and it wouldn't be the way I want it to be, and then, it won't be as fun [doing the exercises].* [Child 8]

A topic that emerged in terms of being listened to, was the right to object to nonvital decisions. Although the young people emphasized the need to listen when a child says no, the children on the other hand simply assumed that their opinions, if expressed, would be respected. Even children who preferred not to be part of the decision making were confident that if they raised any objections to what the adults had decided, the decisions would be revised. A child was asked about what would happen if the child did not approve of the exercises that the therapist had decided:

> *I tell them that I don't want to. They'll listen to me.* [Child 10]

The young people wished that staff at the rehabilitation center would at an early stage encourage the child to speak its own mind while at the same time requesting the parents to not speak as much. Furthermore, the young people expressed a need for adults to focus completely on the child, and parents who did not cooperate would have to leave the room.

Being Connected

The subcategory of *being connected* encompassed a feeling of belonging in the different rehabilitation settings, as well as being respected as an individual. This could mean establishing a positive relationship with one's caregivers, feeling welcomed and being asked one's opinion, being familiar with one's condition, as well as being presented with choices. One child explained why it is important that the doctor not only puts questions to the parents but also to the child:

> *Because I have everything. Up here [points to forehead].* [Child 9]

Another child explained that it is important that adults involve children in discussions concerning their diagnosis and prognosis so that the child can not only explain the situation to classmates but also motivate themselves to do necessary exercises:

> *If I have to do something boring and I don't even know why I have to do it, then it will be even more boring...I have to understand it so that I'm not just showing up at the hospital and don't even know why.* [Child 8]

Both the children and the young people pointed out the need to understand their abilities to be more independent and self-assured. Being able to understand their needs and communicate them helped in feeling connected to the discussions around them:

> *I needed it on paper, written what I can and cannot handle...Having it on paper like this, I was taken more seriously and I like that.* [Young person 4]

The young people described that when they were children, they wanted to be involved and accepted as an individual and treated with the same respect as an adult. They also wanted their rehabilitation work to be based on their interests, so that they could be motivated to perform it.

Parents were sometimes described as barriers, in this category, as the professionals might turn directly to the parent instead of the child, thus resulting in the child being excluded and disconnected.

Feeling Independent

The category *Feeling independent* concerned the child's possibilities of accessing the services at the rehabilitation center as independently as possible, both physically and otherwise, which was manifested in the subcategories: *being admitted* and *being enabled.* Unnecessary barriers to participation could be removed by having sufficient access and through adjustments being made, their dependence on adults could be decreased, and the child could focus on the more permanent barriers that came with the disability. Both children and young people expressed ideas relating to developing their independence, thus striving toward adulthood where they would not have to rely on their parents for everything.

Being Admitted

When the participants experienced the possibility of being admitted, they described a feeling of being able to commute to the center, being able to get in contact when they need to, and understanding the information they received. Having unhindered access contributed to higher levels of both independence and participation.

The children expressed a strong desire to have access to the rehabilitation services in terms of being able to get in contact with them when they needed to, regardless of which day it was or what the time of day was. They also expressed a desire to be able to contact the services independently of their parents, making suggestions about having a direct chat link to their therapists, or a walkie-talkie mounted on their wheelchairs. The children expressed no concerns regarding weekends or holidays, and they instead appeared to believe that their staff lived at the rehabilitation center and would always be accessible if they only had a way to communicate with them. The young people also preferred a phone call instead of participating in meetings with several unknown adults:

> *No thank you, I'll call when I need the help.* [Young person 2]

Another concern among the young people was that of difficulties visiting the pediatric rehabilitation center and instead wanting to access activities closer to home:

> *I think, if you live in a smaller town, you should be able to get help there as well.* [Young person 2]

The young people talked of the distance between their home and the rehabilitation service influencing the help they received and how likely they were to request the help that could be supplied. The sense of there being long distances between the school, the pediatric rehabilitation services, the family, and the child was experienced as a barrier for the child's participation. In contrast with the young people, the children expressed beliefs that adults could simply make the logistics work.

Being Enabled

The subcategory of *being enabled* included experiences of having both physical and cognitive adjustments to be able to participate independently. Participants in wheelchairs expressed a great need for adjustments to freely experience their environment:

> *Sometimes I need help in controlling [the environment]...Like moving things, so I can get through with this one [points to wheelchair].* [Child 1]

Both children and young people were satisfied with their physical surroundings at the rehabilitation center, instead of focusing on cognitive adjustments to gain enablement. Adapted information was considered important by both children and young people. The children talked about how adults like to sit in meetings and talk, whereas they themselves preferred not to participate. Some children, but not all, preferred to have a child-adapted premeeting with only 1 member of the staff, where the child could communicate its needs more comfortably. In the quotation below, the child has explained that it is easier to talk to just 1 person and says how this usually works:

> *We plan together what they [the adults] should talk about next time. And then we have decided, she and I.* [Child 1]

Children and young people knew that complex matters need to be discussed sometimes and then emphasized the need for adjustments to enable the child's understanding of the matter. Being enabled to comprehend their rehabilitation increased the experience of independence. The young person in the quotation below explained how understanding the information helped increase participation while at the pediatric rehabilitation center:

> *They sat down and took their time to explain so that I could understand based on my prerequisites.* [Young person 3]

In the quote mentioned above, the young person mentions the importance of having enough time to comprehend. This thought was common among the young people, who expressed a need to have sufficient time with the professionals, otherwise, they felt that the meeting was stressful and would rather not participate in the rehabilitation work at all. This idea was quite the opposite among the children, who did not want their encounters with adults to be too long, or they would tire and lose their motivation.

Working in Partnership

The category *Working in partnership* captured the different enablers in the child's surrounding network who worked together with the child to increase his/her level of participation. Working in partnership included the subcategories of *being supported* and *entrusting decision making*. The facilitating partner was described as an adult who enabled the child's participation by bridging the gap between the child and the professionals, as well as an adult who could speak on behalf of the child when the child felt the need for it. The children wanted to work in partnership with adults, not in terms of receiving support based on the adults' perspectives, but rather based on what the children themselves define as important.

Being Supported

Both children and young people expressed similar opinions regarding the appreciation of good support from adults. If the children were put in a situation that they did not enjoy and were not capable of telling the staff, they leaned on parents to communicate for them. Not all children were comfortable expressing their opinions, explaining that they needed support to find courage:

> *My counselor tells me that I must be brave...I have some problems with talking about my emotions and needs at home, but I can talk to my counselor.* [Child 13]

The young people described how parents were necessary to support them when they did not have the energy to participate themselves, or when they did not understand the discussions about their care:

> *Sometimes the rehabilitation doctor asks questions that can be difficult for children to even understand...and then I look at them and ask "what are you talking about?"...Sometimes I want my parents to join me so that they can interpret and explain.* [Young person 3]

The young people could also reflect on the need for 1 adult to take charge of the cooperation between all the adults around the child so that the child did not have to take responsibility for keeping everyone informed. This reduced the pressure on the child, leaving him/her free to focus on other things. One young person explained that a child might sometimes also need direct support in handling all the contacts:

> *Like if you've got lower intelligence or difficulties coping with many things at the same time; then it'll all be too much...When you have to check this thing, and that thing, and call this one and then call that one...You probably need some extra help there.* [Young person 1]

Entrusting Decision Making

The category of working in partnership also captured a more indirect version of participation, presented in the subcategory of *entrusting decision making*. Children generally wanted to be given a choice of when to participate and when to delegate certain decisions to others whom they trusted. When asked to elaborate, the children replied that they are content with deciding some things, but not other things. The following child was asked about whether the therapists gave him enough space to decide about his individual gym practice:

> *I only want to decide about the football [the chosen reward after gym practice].* [Child 5]

Another child expressed a deep trust in the therapists and a desire not to intervene:

> *I think it's better if they [the therapists] decide...Because she does good things. So I think I should do what she says.* [Child 11]

The children seemed to think that the act of entrusting the decision making to a partner could both serve as a way to save their energy and leaving them free to focus on more pressing matters, as well as ensuring high-quality health care by allowing the professionals to determine the direction of the therapy. The young people, on the other hand, did not express any wish to pass on the act of decision making to others, but could see positive gains in having a trusted adult speak on their behalf:

> *Perhaps if there could have been a contact person there instead, who could pass it on [my needs and wishes], instead of me sitting there with a bunch of curious old ladies.* [Young person 2]

Having adult partners such as these, who could make decisions on behalf of the children or pass on important information, meant the children could participate on their own terms. By choosing when, where, and how to participate, they could adapt their level of inclusion to suit their current situation.

Discussion

Principal Findings

This study is part of an overall research project aimed at developing, implementing, and evaluating a digital decision support tool to increase participation for children in pediatric rehabilitation. The result from this formative study will further develop the forthcoming digital tool by contributing with ideas that can promote participation according to the target group themselves. The aim of this study was, thus, to explore the experiences of children and young people with disabilities about increasing their participation in the pediatric rehabilitation services. The participants define participation from their own point of view, resulting in the categories of feeling involved, feeling independent, and working in partnership. Combining the categories lead to the identified theme *moving toward empowerment in pediatric rehabilitation*. When a child is able to participate in a meeting where the child feels listened to, understands the conversation, and is able to choose whether to take part in the decision making or not, the whole meeting becomes centered around that child, where he/she is always able to interject if he/she wanted to. This, in turn, can increase the child's self-confidence and can be assumed to increase the chances that the child will want to claim the same rights at the next appointment at the pediatric rehabilitation center. This can be surmised as being able to lead to individuals being empowered and aware of their rights for inclusion and autonomy. The results thus indicate that the digital decision support tool can support children's self-efficacy and autonomy to create greater possibilities for participation in rehabilitation services. The digital communication tool needs to be constructed in such a way to meet these children's developmental, intellectual. and cognitive levels, so that they have the capacity to obtain and understand the information needed to participate in their own care. The desire to feel independent, involved, and being able to work in partnership with adults needs to be addressed and incorporated into the digital tool—a solution that will be covered in future studies. An aspect of particular interest, which would not have been considered without the input from the children, is the need to be able to consciously delegate decisions.

The results reveal that children and young people with disabilities experience participation through feelings of involvement. The participants in the present study express an awareness of the fact that involvement may be difficult to achieve, addressing matters relating both to their own person and to their surroundings. Participants who express a high level of belief in themselves also express greater independence and a willingness to take more responsibility in their rehabilitation, thus demanding to be included and involved. This relates to research stating that a person's internal characteristics such as capacity, ability, self-esteem, self-efficacy, and motivation are affected by previous experiences, which in turn influence future behavior, performance, and experiences of participation [11,19,44,45]. Research also concludes that relational factors such as effective communication between patients and staff is crucial for high-quality health care for people with disabilities [46,47] and that low communication skills negatively affect the possibilities for participation [48]. In this study, both children and young people with disabilities indicate the need for easier ways to communicate with professionals, whereas also reflecting that there are great demands on children to be forthright. This supports the concept of a decision support tool with simplified modes of communication that can be used to increase the child's possibilities of being involved and active in their rehabilitation, where the children feel that they are respected and that they are able to be part of the decisions concerning their care.

The children and young people reveal a high dependence on being able to access the rehabilitation facilities. Research indicates that the location of health care services impacts greatly on the patients' likelihood of attending [47] and should be considered an important factor regarding children with disabilities, who are very dependent on adults providing transportation. Although the concept of being independent can be linked to broader discussions of autonomy, the participants only expressed concerns regarding physical and cognitive access, indicating that these issues are more pressing. If these obstacles were more satisfactorily solved, the participant could perhaps be able to consider autonomy in a broader sense of the word. Having cognitive access, that is, being able to understand the information, is viewed as an important factor for the participants. Some appreciated being allowed enough time to comprehend, whereas others wished for shorter meetings so as not to lose interest, thus once more emphasizing the need for a child-centered health care that acknowledges the child's focus [16]. Considerations have to be made concerning the dual nature of the findings in the research literature in terms of the possibilities for understanding the interactions with professionals. There is a need for simplified communication to avoid health care barriers [47], and recognition of people with disabilities often feeling belittled by caregivers' inappropriate assumptions regarding their low levels of understanding [48]. It is important to carefully consider the individual needs for either a more simplified information or a more advanced exchange in the development of this digital decision support tool. The results suggested that a digital decision support tool needs to provide an individualized format that supports the children's possibilities for communicating their own perspectives of their situation and health. Furthermore, a digital tool that helps children to report their needs, problems, strengths, and experiences, could also be beneficial for staff in understanding the unique child's situation and provide a more child-centered rehabilitation. Greater consideration needs to be taken of individual experiences, and professionals need to consider that factors such as individual experiences cannot be objectively observed.

The results highlighted both the children's and the young people's appreciation of good support from and a partnership with adults. The term *gatekeeper* is used in the previous research to describe a person close to the child who can both facilitate and hinder the child's involvement in health care [23]. Participants agree that parents can serve as both facilitators and barriers depending on the setting and how the health care professionals treat the child in relation to the parent. The need to be seen as a partner in decision making is evident, where the child can choose to participate among other equals or to step aside knowingly trusting others. However, the children’s desires to be able to pass on certain decisions to adults, a thought that the young people do not seem to share, is a desire that needs to be carefully dealt with. It can on the one hand be a highly autonomous decision to be left out of the decision making, for example because of energy-saving concerns, but on the other hand it can represent the beginning of a pattern of low participation. It has been claimed in some previous research that participation competence can be actively trained and developed [49] but it seems reasonable to reverse the equation: If a child repeatedly wishes to entrust adults with their decision making, a passive attitude toward participation could be manifested. Research also states that each situation of delegation requires careful considerations, relating to the patient’s unique needs and preferences [50], thereby motivating the need for the future digital decision support tool that can provide possibilities for delegating the decision making while also urging the staff to be mindful of this request. It is, however, important to point out that participation is a right and not an obligation. It, thus, becomes important in the forthcoming digital tool for children to be able to choose which questions to take part in and which to entrust to others.

Methodological Considerations and Limitations

The interviews were conducted both individually and in pairs in this study to allow different settings for different needs, as recommended by Krol et al [20]. The individual interviews generated a profundity and became more personal, whereas group and pair interviews have the advantage of creating security in the group and giving a voice to those who cannot participate in an individual interview. The different interview forms had the purpose of facilitating the participation of all of the participants. As children can be hesitant to speak in very formal settings [51], allowing them to participate in a well-adapted situation contributes to securing their views and opinions in the research. Depending on each participant’s specific difficulties, the interviewer also had to consider potential challenges in terms of social skills, cognition, language skills, memory, as well as concentration and endurance. This, thus, entails a great demand on the interviewer to present appropriate questions with sufficient communication aids [26,27], using augmentative tools [28] as well as adapting the setting based on individual needs. The adaptations needed to complete the interviews also gave an insight into which adaptations can be needed in the digital decision support tool, such as having communication supported by pictures and manual signs.

This formative approach, with a constant seeking for the individual child’s perspective, can help contribute to a tool that can better capture the children’s experiences and needs. The children and young people in this study are included as partners in the research, where they evaluate certain aspects of their life situations themselves, without input from professionals. Previous studies have pointed out the need to include children in the research that affect them [30] as well as the need to include children with disabilities as partners in the research [23]. To our knowledge, there has been very little research that combines the 2: the inclusion of children with disabilities in research that is primarily aimed at improving their situation. A further novel aspect is the concept of including children with disabilities in research to improve their participation, and as cocreators of a digital decision support tool. We have not found any studies that have put this to practice, thus making the results presented in this study unique.

A weakness of this study, as with all retrospective studies, is that the young people look back at previous experiences that may not represent the current way of working in the pediatric rehabilitation services. The children, however, present their current views and provide statements that are highly important for the services to access. Nevertheless, all children have the right to participate and a need to be listened to which is confirmed in this study as well in others [13,18].

Conclusions

This formative study explored children’s and young people’s experiences of participation-enhancing factors in the pediatric rehabilitation services, with the aim of incorporating their experiences in the development of a digital decision support tool for increased participation. The findings concluded that children experienced participation through feeling involved, feeling independent, and working in partnership. When these requirements were met, the children were able to move toward empowerment in their pediatric rehabilitation. Experiencing participation through involvement meant feeling listened to and feeling connected, feeling independent meant being admitted and being enabled, and working in partnership meant feeling supported and being able to entrust the decision making to others. This definition of participation belongs to the participants themselves and can only to some extent be compared with other research. The participants’ opinions are core elements when developing an intervention with a digital decision tool to support the children’s participation, and their contribution will help in making the tool attractive, relevant, and in line with the UNCRC. However, to safeguard children’s right to participation, future research must explore potential obstacles to participation as identified by the target group themselves.

Acknowledgments

The authors would like to thank the children and young people who participated in this research.

Authors' Contributions

All authors (EV, IL, ML, EO, JMN, and PS) made significant contributions to the original paper, by together identifying the research question and design of the study. Data collection (the interviews) was carried out by (EV, ML, and EO). Data analysis was performed separately by (EV, IL, ML, and EO) and then compared and discussed with all authors (EV, IL, ML, EO, JMN, and PS). The manuscript was drafted by (EV and IL), whereas (PS and JN) provided critical revision of the paper in terms of important intellectual content. All authors (EV, IL, ML, EO, JMN, and PS) have read and approved the final submitted version.

Conflicts of Interest

None declared.

References

1. Riksdagen. Patientlag (2014:821) URL: https://www.riksdagen.se/sv/dokument-lagar/dokument/svensk-forfattningssamling/patientlag-2014821_sfs-2014-821 [accessed 2019-08-21]
2. Hughes TM, Merath K, Chen Q, Sun S, Palmer E, Idrees JJ, et al. Association of shared decision-making on patient-reported health outcomes and healthcare utilization. Am J Surg 2018 Jul;216(1):7-12. [doi: 10.1016/j.amjsurg.2018.01.011] [Medline: 29395026]
3. Riksdagens Oppna Data. 2017. Regeringens Proposition 2017/18:186: Inkorporering Av FN:s Konvention Om Barnets Rättigheter URL: https://data.riksdagen.se/fil/F99B084B-5202-4EDD-9E5D-D3B3D44DA1E9 [accessed 2019-08-21]
4. Office of the United Nations High Commissioner for Human Rights. 1989. Convention on the Rights of the Child URL: https://www.ohchr.org/documents/professionalinterest/crc.pdf [accessed 2019-08-21]
5. Coyne I, Harder M. Children's participation in decision-making: balancing protection with shared decision-making using a situational perspective. J Child Health Care 2011 Dec;15(4):312-319. [doi: 10.1177/1367493511406570] [Medline: 21828162]
6. Feenstra B, Boland L, Lawson ML, Harrison D, Kryworuchko J, Leblanc M, et al. Interventions to support children's engagement in health-related decisions: a systematic review. BMC Pediatr 2014 Apr 23;14:109 [FREE Full text] [doi: 10.1186/1471-2431-14-109] [Medline: 24758566]
7. Larsson I, Staland-Nyman C, Svedberg P, Nygren JM, Carlsson IM. Children and young people's participation in developing interventions in health and well-being: a scoping review. BMC Health Serv Res 2018 Jun 28;18(1):507 [FREE Full text] [doi: 10.1186/s12913-018-3219-2] [Medline: 29954392]
8. Virkki M, Tolonen TH, Koskimaa T, Paavilainen E. Children as decision-makers in health care-an integrative review. Clin Nurs Stud 2015;3(1):47-54. [doi: 10.5430/cns.v3n1p47]
9. McAnuff J, Brooks R, Duff C, Quinn M, Marshall J, Kolehmainen N. Improving participation outcomes and interventions in neurodisability: co-designing future research. Child Care Health Dev 2017 Mar;43(2):298-306. [doi: 10.1111/cch.12414] [Medline: 27766678]
10. Clavering EK, McLaughlin J. Children's participation in health research: from objects to agents? Child Care Health Dev 2010 Sep;36(5):603-611. [doi: 10.1111/j.1365-2214.2010.01094.x] [Medline: 20533922]
11. Imms C, Adair B, Keen D, Ullenhag A, Rosenbaum P, Granlund M. 'Participation': a systematic review of language, definitions, and constructs used in intervention research with children with disabilities. Dev Med Child Neurol 2016 Jan;58(1):29-38 [FREE Full text] [doi: 10.1111/dmcn.12932] [Medline: 26411643]
12. Charles C, Gafni A, Whelan T. Shared decision-making in the medical encounter: what does it mean? (or it takes at least two to tango). Soc Sci Med 1997 Mar;44(5):681-692. [doi: 10.1016/s0277-9536(96)00221-3] [Medline: 9032835]
13. Gondek D, Edbrooke-Childs J, Velikonja T, Chapman L, Saunders F, Hayes D, et al. Facilitators and barriers to person-centred care in child and young people mental health services: a systematic review. Clin Psychol Psychother 2017 Jul;24(4):870-886. [doi: 10.1002/cpp.2052] [Medline: 27910173]
14. Wyatt KD, List B, Brinkman WB, Lopez GP, Asi N, Erwin P, et al. Shared decision making in pediatrics: a systematic review and meta-analysis. Acad Pediatr 2015;15(6):573-583. [doi: 10.1016/j.acap.2015.03.011] [Medline: 25983006]
15. Cheng H, Hayes D, Edbrooke-Childs J, Martin K, Chapman L, Wolpert M. What approaches for promoting shared decision-making are used in child mental health? A scoping review. Clin Psychol Psychother 2017 Nov;24(6):O1495-O1511. [doi: 10.1002/cpp.2106] [Medline: 28752631]
16. Coyne I, Hallström I, Söderbäck M. Reframing the focus from a family-centred to a child-centred care approach for children's healthcare. J Child Health Care 2016 Dec;20(4):494-502. [doi: 10.1177/1367493516642744] [Medline: 27141084]
17. Arvidsson S, Gilljam BM, Nygren J, Ruland CM, Nordby-Bøe T, Svedberg P. Redesign and validation of Sisom, an interactive assessment and communication tool for children with cancer. JMIR Mhealth Uhealth 2016 Jun 24;4(2):e76 [FREE Full text] [doi: 10.2196/mhealth.5715] [Medline: 27343004]
18. Gilljam BM, Arvidsson S, Nygren JM, Svedberg P. Promoting participation in healthcare situations for children with JIA: a grounded theory study. Int J Qual Stud Health Well-being 2016;11:30518 [FREE Full text] [doi: 10.3402/qhw.v11.30518] [Medline: 27172512]
19. Söderbäck M, Coyne I, Harder M. The importance of including both a child perspective and the child's perspective within health care settings to provide truly child-centred care. J Child Health Care 2011 Jun;15(2):99-106. [doi: 10.1177/1367493510397624] [Medline: 21685225]
20. Krol M, Sixma H, Meerdink J, Wiersma H, Rademakers J. Exploring young patients' perspectives on rehabilitation care: methods and challenges of organizing focus groups for children and adolescents. Child Care Health Dev 2014 Jul;40(4):507-514. [doi: 10.1111/cch.12095] [Medline: 25028746]

21. Costa UM, Brauchle G, Kennedy-Behr A. Collaborative goal setting with and for children as part of therapeutic intervention. Disabil Rehabil 2017 Aug;39(16):1589-1600. [doi: 10.1080/09638288.2016.1202334] [Medline: 27385635]
22. Øien I, Fallang B, Østensjø S. Goal-setting in paediatric rehabilitation: perceptions of parents and professional. Child Care Health Dev 2010 Jul;36(4):558-565. [doi: 10.1111/j.1365-2214.2009.01038.x] [Medline: 20030659]
23. Bailey S, Boddy K, Briscoe S, Morris C. Involving disabled children and young people as partners in research: a systematic review. Child Care Health Dev 2015 Jul;41(4):505-514. [doi: 10.1111/cch.12197] [Medline: 25323964]
24. Davies HT, Phillips B, Preston J, Stones SR. Making research central to good paediatric practice. Arch Dis Child 2019 Apr;104(4):385-388. [doi: 10.1136/archdischild-2018-315117] [Medline: 30902887]
25. Phillips B, Davies HT, Preston J, Stones SR. Framework to help design and review research involving children. Arch Dis Child 2019 Jun;104(6):601-604. [doi: 10.1136/archdischild-2018-315119] [Medline: 31101643]
26. Boström P, Broberg M. Protection and restriction: a mixed-methods study of self-reported well-being among youth with intellectual disabilities. J Appl Res Intellect Disabil 2018 Jan;31(1):e164-e176. [doi: 10.1111/jar.12364] [Medline: 28544451]
27. Finlay WM, Lyons E. Methodological issues in interviewing and using self-report questionnaires with people with mental retardation. Psychol Assess 2001 Sep;13(3):319-335. [doi: 10.1037/1040-3590.13.3.319] [Medline: 11556269]
28. Preece D. Consultation with children with autistic spectrum disorders about their experience of short-term residential care. Br J Learn Disabil 2002 Sep;30(3):97-104. [doi: 10.1046/j.1468-3156.2002.00179.x]
29. Beresford B. Working on well-being: researchers' experiences of a participative approach to understanding the subjective well-being of disabled young people. Child Soc 2012 Apr 16;26(3):234-240. [doi: 10.1111/j.1099-0860.2012.00436.x]
30. Ruland CM, Slaughter L, Starren J, Vatne TM, Moe EY. Children's contributions to designing a communication tool for children with cancer. Stud Health Technol Inform 2007;129(Pt 2):977-982. [Medline: 17911861]
31. Stålberg A, Sandberg A, Söderbäck M. Child-centred care - health professionals' perceptions of what aspects are meaningful when using interactive technology as a facilitator in healthcare situations. J Pediatr Nurs 2018;43:e10-e17. [doi: 10.1016/j.pedn.2018.07.006] [Medline: 30056996]
32. Svedberg P, Arvidsson S, Larsson I, Carlsson IM, Nygren JM. Barriers and enablers for successful implementation of the ehealth service Sisom for improved child participation in pediatric care - a multi-center study. J Med Internet Res 2019:-(epub ahead of print)(forthcoming). [doi: 10.2196/14271]
33. Tsimicalis A, Rennick J, le May S, Stinson J, Sarkis B, Séguin K, et al. 'Tell it as it is': how Sisom prompts children and parents to discuss their cancer experience. Cancer Rep 2019 Apr 23;2(3):e1173. [doi: 10.1002/cnr2.1173]
34. Robards F, Kang M, Usherwood T, Sanci L. How marginalized young people access, engage with, and navigate health-care systems in the digital age: systematic review. J Adolesc Health 2018 Apr;62(4):365-381. [doi: 10.1016/j.jadohealth.2017.10.018] [Medline: 29429819]
35. Livingstone S, Lemish D, Lim SS, Bulger M, Cabello P, Claro M, et al. Global perspectives on children's digital opportunities: an emerging research and policy agenda. Pediatrics 2017 Nov;140(Suppl 2):S137-S141 [FREE Full text] [doi: 10.1542/peds.2016-1758S] [Medline: 29093049]
36. Veneri DA, Zdanis K. Are technology-informed methods better than traditional approaches in educating patients? A meta-analysis. Health Educ J 2018 Jan 9;77(3):261-276. [doi: 10.1177/0017896917741511]
37. Gittelsohn J, Steckler A, Johnson CC, Pratt C, Grieser M, Pickrel J, et al. Formative research in school and community-based health programs and studies: 'state of the art' and the TAAG approach. Health Educ Behav 2006 Feb;33(1):25-39 [FREE Full text] [doi: 10.1177/1090198105282412] [Medline: 16397157]
38. Graneheim UH, Lindgren BM, Lundman B. Methodological challenges in qualitative content analysis: a discussion paper. Nurse Educ Today 2017 Sep;56:29-34. [doi: 10.1016/j.nedt.2017.06.002] [Medline: 28651100]
39. Krippendorff K. Content Analysis: An Introduction to Its Methodology. Third Edition. Thousand Oaks, CA: Sage Publications; 2013.
40. DePoy E, Gitlin LN. Introduction to Research - E-Book: Understanding and Applying Multiple Strategies. Missouri: Elsevier Inc; 2013.
41. Graneheim UH, Lundman B. Qualitative content analysis in nursing research: concepts, procedures and measures to achieve trustworthiness. Nurse Educ Today 2004 Feb;24(2):105-112. [doi: 10.1016/j.nedt.2003.10.001] [Medline: 14769454]
42. WMA – The World Medical Association. 2013. WMA Declaration of Helsinki – Ethical Principles for Medical Research Involving Human Subjects URL: https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/ [accessed 2019-08-21]
43. Swedish Research Council. 2017. Good Research Practice URL: https://vr.se/download/18.5639980c162791bbfe697882/1523378510652/Good-Research-Practice_VR_2017.pdf: [accessed 2019-08-21]
44. Eakman AM. Person factors: meaning, sensemaking and spirituality. In: Christiansen C, Baum CM, Bass-Haugen J, Bass JD, editors. Occupational Therapy: Performance, Participation, and Well-Being. Thorofare, NJ, USA: Slack Incorporated; 2005:313-332.
45. Strauser DR. Applications of self-efficacy theory in rehabilitation counseling. J Rehabil 1995;61(1):7 [FREE Full text]
46. Ferguson M, Jarrett D, Terras M. Inclusion and healthcare choices: the experiences of adults with learning disabilities. Br J Learn Disabil 2011;39(1):73-83. [doi: 10.1111/j.1468-3156.2010.00620.x]

47. Lawthers AG, Pransky GS, Peterson LE, Himmelstein JH. Rethinking quality in the context of persons with disability. Int J Qual Health Care 2003 Aug;15(4):287-299. [doi: 10.1093/intqhc/mzg048] [Medline: 12930044]
48. Morris MA, Dudgeon BJ, Yorkston K. A qualitative study of adult AAC users' experiences communicating with medical providers. Disabil Rehabil Assist Technol 2013 Nov;8(6):472-481. [doi: 10.3109/17483107.2012.746398] [Medline: 23350909]
49. Adair B, Ullenhag A, Keen D, Granlund M, Imms C. The effect of interventions aimed at improving participation outcomes for children with disabilities: a systematic review. Dev Med Child Neurol 2015 Dec;57(12):1093-1104 [FREE Full text] [doi: 10.1111/dmcn.12809] [Medline: 26010935]
50. Sarkar T. Intellectual and developmental disability: transition to adulthood and decision making process. Int J Child Health Hum Dev 2015;8(4):517 [FREE Full text]
51. Vis SA, Strandbu A, Holtan A, Thomas N. Participation and health - a research review of child participation in planning and decision-making. Child Fam Soc Work 2011;16(3):325-335. [doi: 10.1111/j.1365-2206.2010.00743.x]

Abbreviations

SDM: shared decision making
UNCRC: United Nation’s Convention on the Rights of the Child

Edited by G Eysenbach; submitted 25.04.19; peer-reviewed by J Edbrooke-Childs, C Senesac; comments to author 29.05.19; revised version received 10.07.19; accepted 28.07.19; published 02.10.19.

Please cite as:
Vinblad E, Larsson I, Lönn M, Olsson E, Nygren JM, Svedberg P
Development of a Digital Decision Support Tool to Aid Participation of Children With Disabilities in Pediatric Rehabilitation Services: Explorative Qualitative Study
JMIR Form Res 2019;3(4):e14493
URL: https://formative.jmir.org/2019/4/e14493
doi:10.2196/14493
PMID:31579015

Viewpoint

Understanding Health Behavior Technology Engagement: Pathway to Measuring Digital Behavior Change Interventions

Heather Cole-Lewis[1*], MPH, PhD; Nnamdi Ezeanochie[1*], MD, DrPH; Jennifer Turgiss[1], MS, DrPH

Johnson and Johnson Health and Wellness Solutions, Inc, New Brunswick, NJ, United States

[*]these authors contributed equally

Corresponding Author:
Nnamdi Ezeanochie, MD, DrPH
Johnson and Johnson Health and Wellness Solutions, Inc
1 Johnson and Johnson Plaza
New Brunswick, NJ, 08933
United States
Phone: 1 7209330786
Email: nezeanoc@its.jnj.com

Abstract

Researchers and practitioners of digital behavior change interventions (DBCI) use varying and, often, incongruent definitions of the term “engagement,” thus leading to a lack of precision in DBCI measurement and evaluation. The objective of this paper is to propose discrete definitions for various types of user engagement and to explain why precision in the measurement of these engagement types is integral to ensuring the intervention is effective for health behavior modulation. Additionally, this paper presents a framework and practical steps for how engagement can be measured in practice and used to inform DBCI design and evaluation. The key purpose of a DBCI is to influence change in a target health behavior of a user, which may ultimately improve a health outcome. Using available literature and practice-based knowledge of DBCI, the framework conceptualizes two primary categories of engagement that must be measured in DBCI. The categories are health behavior engagement, referred to as “Big E,” and DBCI engagement, referred to as “Little e.” DBCI engagement is further bifurcated into two subclasses: (1) user interactions with features of the intervention designed to encourage frequency of use (ie, simple login, games, and social interactions) and make the user experience appealing, and (2) user interactions with behavior change intervention components (ie, behavior change techniques), which influence determinants of health behavior and subsequently influence health behavior. Achievement of Big E in an intervention delivered via digital means is contingent upon Little e. If users do not interact with DBCI features and enjoy the user experience, exposure to behavior change intervention components will be limited and less likely to influence the behavioral determinants that lead to health behavior engagement (Big E). Big E is also dependent upon the quality and relevance of the behavior change intervention components within the solution. Therefore, the combination of user interactions and behavior change intervention components creates Little e, which is, in turn, designed to improve Big E. The proposed framework includes a model to support measurement of DBCI that describes categories of engagement and details how features of Little e produce Big E. This framework can be applied to DBCI to support various health behaviors and outcomes and can be utilized to identify gaps in intervention efficacy and effectiveness.

(JMIR Form Res 2019;3(4):e14052) doi:10.2196/14052

KEYWORDS

engagement; user engagement; health behavior; health determinants; digital behavior change intervention; measurements

Introduction

Background

Globally, the creative integration of technology solutions to address health issues is growing [1]. This integration has been largely through the promotive uptake of healthy behaviors to achieve desired health outcomes via technology [2]. In this paper, these technologies will be referred to as a digital behavior change intervention (DBCI). Recent systematic reviews demonstrate that digital interventions supported by behavior science evidence hold the key to driving critical behavior change processes that lead to improved health behaviors and health outcomes [2]. Common examples of targeted behaviors include smoking cessation, increase in physical activity, improvements in dietary habits, medication adherence, and safe sexual practices [3].

However, the extent to which these innovations can deliver on the promise of demonstrable positive health outcomes depends on the successful utilization of interventions by users and the subsequent sustained performance of the intended health behaviors. This utilization has been generally referred to as "engagement" or "user-engagement." This term is largely recognizable, but at the same time, abstract and difficult to measure accurately in different settings and contexts. This measurement limitation can be mitigated by applying frameworks/models that guide the way engagement is considered, measured, and applied in analysis within an intervention. Precise definitions and measurements for engagement will allow for better estimation of DBCI effectiveness and provide accurate insights that inform relevant intervention recommendations [4].

Goal

The aims of this study are (1) to provide an analysis of current engagement definitions and models in the literature, (2) to propose a new model that builds on existing frameworks but addresses current limitations, and (3) to describe the implications for applying the proposed framework to measurements and analysis within DBCI.

Analysis of Current Definitions of Engagement in Literature

It is generally accepted that user engagement with DBCI is a precursor to improved health outcomes. Growing evidence supports that interaction with DBCI and adherence to intervention features improve the likelihood of reaching desired health outcomes [5]. There is currently no universally accepted and comprehensive definition for engagement, and this has two main implications. First, there is a limited and inconsistent understanding of the engagement's specific impact, particularly on DBCI effect size, attrition, and sustained health behaviors [4]. Second, there is also a lack of precision in DBCI measurement and evaluation, as the term "engagement" can be interpreted differently across industries (eg, marketing, psychology, and human-computer interaction) [6].

One definition of engagement evaluates esthetics and navigability, particularly how users interface with digital interventions [7]. This type of engagement typically observes usability, measuring interactions with features and functions of the digital solution. This definition is important because it provides insight into the level of use and interaction of a digital intervention. It also informs what components of the intervention users enjoy or use the most and opportunities to leverage these insights to inform future interventions.

Engagement on this level is an important element for the digital intervention scale. O'Brien and Tom [8] defined engagement as an assessment of a user experience (interaction and involvement) with an innovation or technology-based intervention [8]. They argue that this experience leads to the formation of sentiments that impact a user's likelihood to advocate the use of the innovation among others within their social networks. Furthermore, this definition of engagement helps researchers understand more about the type of users involved in their intervention (ie, data from this engagement addresses questions such as the emerging user-engagement hierarchy identified via subgroups based on user interaction with the features of the digital intervention) [9]. With an understanding of the baseline users of an intervention and their characteristics, researchers can build updates to digital interventions that positively support most types of users [8].

Nevertheless, there are limitations with this type of engagement definition. Definitions based exclusively on user interactions assume that engagement with the digital innovation ultimately leads to the intended intervention outcome (ie, higher engagement and adoption lead to desired health outcomes). This assumption is problematic because of the current limited interpretational challenges user interaction data present, such as the difficulty in distinguishing online versus offline interactions and the implications for intervention exposure, or the absence of frameworks that identify unique meaningful interaction sequences with digital features that translate to specific behavior change techniques (BCTs) embedded in a DBCI [10]. Therefore, this level of engagement definition is insufficient in meeting the need for better and more precise measurement of engagement to inform the effectiveness of DBCI.

Some researchers conceptualize engagement in terms of a user performing the intended health behaviors within an intervention [11]. This type of engagement seeks to understand the relationship between using the digital intervention and behavior change, drawing on evidence-based principles and behavioral theories to evaluate changes in health outcomes [11,12]. The advantage of this definition is that it is ultimately concerned with whether an individual performed the desired health behavior. This definition can closely tie the effectiveness of an intervention to a measure of engagement. Definitions of engagement based on behavior performance parameters only do not provide details of user interactions with components of a digital intervention. Thus, an understanding of what specific digital features are linked with behavior engagement (ie, what features are effective) is absent. This is a significant limitation because data needed to update features of the intervention are absent; thus, future optimization and cost-saving processes informed by evidence will be nearly impossible [13]. Therefore, one is left with a *potentially* effective DBCI and no understanding of what features contributed to their efficacy, thus lacking insights to inform intervention optimization and scale.

The absence of a standardized DBCI framework for defining and measuring engagement inhibits the ability to understand the mechanisms of action of an intervention. Therefore, current approaches are limited in providing accurate and detailed measurement data necessary to demonstrate effectiveness of DBCI.

A proposed model should include definitions of engagement related to DBCI feature interaction as well as performance of the desired health behavior [14]. This approach can inform which specific features of a DBCI influence performance of a health behavior. Essentially, an integrated model should be able to address the relationship between user interaction with the DBCI and the user health behaviors, thus answering the question, “*What level of user-engagement with the DBCI, and by which users, leads to a desired health?*”

Furthermore, this model should ideally delineate and identify unique, meaningful interaction sequences that represent digital features informed by specific behavior change theory–informed components of the intervention (eg, BCTs) [13]. This hierarchical representation of user feature interaction and intervention exposure to BCTs informs a robust dataset that contains explanatory variables that enable a deeper DBCI analysis in real-world settings and an explanation of the mechanism of action with regard to change processes [13]. For this model to be complete, we need to account for behavioral determinants that mediate or moderate the association between digital features engagement and their impact on associated intervention health behavior outcomes. These determinants should include determinants related to user engagement with DBCI [15]. Examples include technology self-efficacy; satisfaction with the DBCI; intervention usability [15]; and technology-associated determinants informed by the technology acceptance model (TAM), such as perceived usefulness, perceived ease of use, and perceived compatibility [16]. In addition, psychosocial determinants rooted in behavioral theories relevant to the behaviors of interest such as the social ecological model [17]; Capability, Opportunity, Motivation - Behavior (COM-B) model [18]; and social cognitive theory [19] must be included in the proposed engagement model. A model housing both engagement definitions (engagement with DBCI and desired health behavior, technology determinants, psychosocial determinants, and outcomes) is robust enough to inform datasets that support rigorous evaluation and measurements of DBCI.

The Johnson and Johnson Approach to Health Engagement: Definitions and Framework

Overview

Following the abovementioned recommendation outlining a proposed model for defining and measuring engagement, we present a definition and model for engagement. Several steps were taken to achieve this goal.

First, several definitions of engagement were reviewed from the literature and practice-based knowledge surrounding DBCI [7,11-15]. These definitions were analyzed to identify current limitations.

Second, new definitions for engagement were proposed. The intention here was to have new sets of definitions that addressed the observed limitations of current definitions and to standardize terminologies around DBCI engagement with the intention of creating better DBCI measurements.

Defining “Big E” and “Little e” Engagement

“Big E” is engagement with the targeted health behavior, hereafter referred to as health behavior engagement, and is the primary outcome of a DBCI. However, the goal is to attain health outcomes through the health behavior. Thus, it is important that engagement with the health behavior is achievable and measurable to determine if a DBCI is successful.

“Little e” is engagement with the digital behavior change intervention. This is sometimes referred to as user interaction with the digital solution. Hereafter, the term DBCI engagement will be used to represent engagement with the digital solution. DBCI engagement is comprised of two types of interactions:

- User interactions with the DBCI features and the context in which those interactions happen (little e_{UI})
- Interactions with behavior change intervention components/active ingredients specifically designed to influence the behavioral determinants which in turn influence the health behaviors (little e_{BCT})

The success of health behavior engagement is dependent upon DBCI engagement. If users do not interact with DBCI features, exposure to the behavior change intervention components will be less likely to influence behavioral determinants and lead to health behavior engagement across a broad population. Health behavior engagement is also dependent upon the quality and relevance of the behavior change intervention components within the solution. Therefore, user interactions and behavior change intervention components combine to create Little e; this in turn influences Big E. Figure 1 illustrates how these definitions of engagement inform the overall intervention measurement.

Figure 1. Diagram illustrating the Johnson and Johnson approach to health engagement. UI: user interaction; Ux: user experience; BCT: behavior change technique.

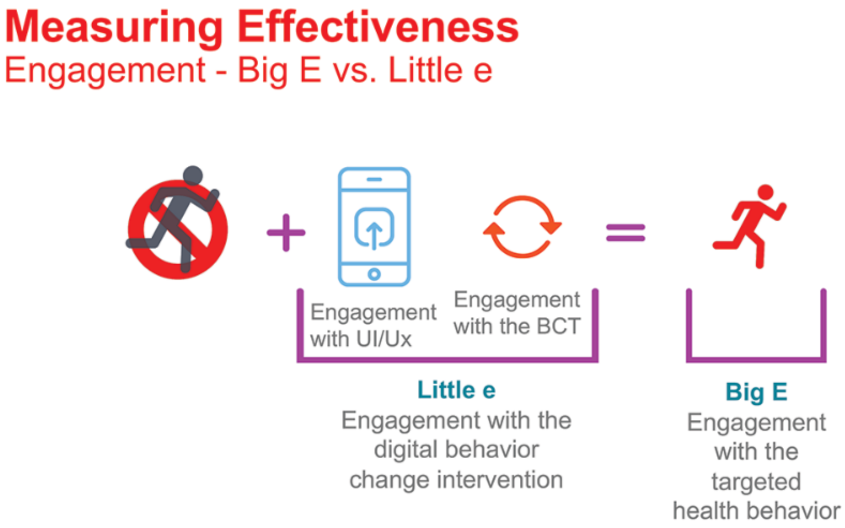
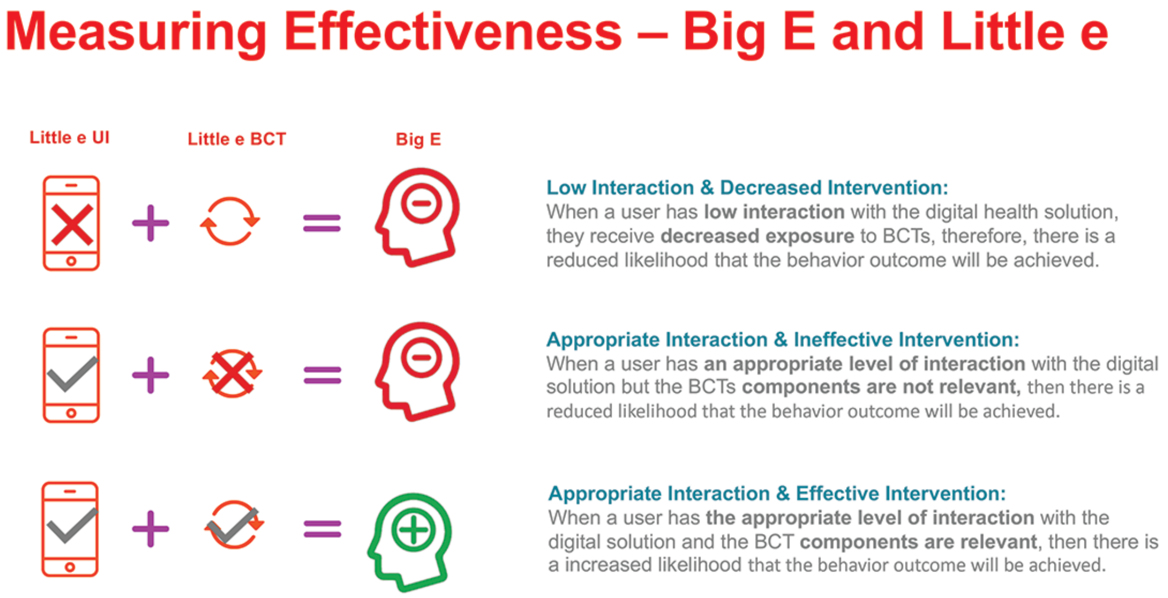


Figure 1 indicates that a sedentary individual will need to interact with the appropriate level of the DBCI features (little e_{UI}) to be exposed to the right and effective level of behavior change intervention components (little e_{BCT}). This should ideally lead to changes in determinants that support engagement with the health behavior. Inadequate levels of little e_{UI} interaction or ineffective exposure to appropriate BCTs will not lead to health behavior engagement (Figure 2). In some cases, appropriate exposure to a BCT is based on what is clinically relevant to influence a health behavior and by extension, a health outcome. For example, physical activity (steps) tracking is a BCT (self-monitoring of a behavior), but attaining *x* number of steps will be needed to begin seeing a decrease in *y* amount of weight (health outcome).

Figure 2. Diagram illustrating relationship engagement measurements. UI: user interaction; BCT: behavior change technique.



Low Interaction & Decreased Intervention: When a user has **low interaction** with the digital health solution, they receive **decreased exposure** to BCTs, therefore, there is a reduced likelihood that the behavior outcome will be achieved.

Appropriate Interaction & Ineffective Intervention: When a user has **an appropriate level of interaction** with the digital solution but the BCTs **components are not relevant,** then there is a reduced likelihood that the behavior outcome will be achieved.

Appropriate Interaction & Effective Intervention: When a user has **the appropriate level of interaction** with the digital solution and the BCT **components are relevant**, then there is a increased likelihood that the behavior outcome will be achieved.

Principles from program planning and intervention design literature [20] were leveraged to develop a framework model unique for building DBCI. This model is a hybrid that uses the layout of a theory of change conceptual model and a traditional logic model (Table 1) to provide the structure upon which we systematically embedded the new engagement definitions and then tested the robustness of these definitions in order to produce the right amount of data to inform complete and accurate measurements of DBCI effectiveness.

Table 1. The Johnson and Johnson approach to health proposed engagement definitions and framework. As with a traditional logic model, this table should be read from left to right. The underlying assumption is that exposure to an intervention should lead to changes in determinants that in turn influences health behaviors, health outcomes, and organizational outcomes. For each phase of the logic model, the table illustrates examples of the engagement category, measurement category, sample metrics, and the metrics/data source.

Logic model category	Exposure to intervention		Determinants	Health behavior	Health outcomes	Organizational outcomes
Engagement category	Little e engagement with the DBCI[a]		—[b]	Big E engagement in health behaviors	Outcome	Outcome
Measurement category	UI[c] interactions Contextual information	BCT[d] intervention components	Capability Opportunity motivation	Health behaviors	Health outcomes at individual level	Organizational outcomes
Sample metrics	Clicks Swipes Location data	Goal setting Restructuring thoughts	Change in: Skills Self-efficacy to perform physical activity Access to food store with health meal options	Sufficient level of physical activity Targeted nutrition behaviors	Improved A1C Maintain ideal weight	Descrease demand of health care system
Metric/data source	In-app	In-app	In-app/out of app	In-app/out of app	In-app/out of app	In-app/out of app

[a] DBCI: digital behavior change interaction.

[b] Not applicable.

[c] UI: user interface.

[d] BCT: behavior change technique.

This framework provides an engagement category that details how features of Little e inform Big E and builds upon components that align with an intervention logic model, such as exposure to the intervention, determinants, health behaviors, health outcomes, and organizational outcomes. There were five steps that informed the process of building the model and its components:

1. Utilization of market research, economic evaluation, and subject matter expertise in therapeutic areas to inform our choice of the organizational, individual health, and behavior outcomes (ie, we knew what specific behaviors needed to be performed to lead to the outcomes of interest and by what effect size).
2. Review of the behaviors and created subbehaviors and specific performance indicators associated with them.
3. Examination of the behavior science literature to deduce the factors (mediator or moderators) that influence the performance of the behaviors among individuals in the target population. These factors were grouped using various frameworks such as (a) enabler and barriers; (b) mediators or moderators; (c) factors personally and environmentally informed by the socioecological model [17]; and (d) capability, motivational, or opportunity factors [18].
4. Use of the understanding and learnings from several theories, such as the classical and operant processes, social cognitive theory, social learning theory, and self-determination theory [19,21-23], to determine which behaviors fall within the involuntary and voluntary process. Thereafter, we used this guidance to determine antecedent and consequence behavioral determinants (ie, BCTs that make up components of the intervention).
5. Use of the understanding from the intervention mapping literature [20] to determine specific change processes that illustrated our best-informed assumptions about which specific BCTs drive factors influence performance of the behaviors. This helped define and guide the digital product strategy where each behavior change process is represented by a set of user experience/user interaction (UI) and content features.

The measurement category provides additional detail of what factors should be measured in each engagement category. Finally, the example metric category provides examples of metrics used during evaluation.

Implications of the Engagement Framework for Interventions Measurements

Overview

The framework provides a structure to discuss implications for the measurement and analysis of DBCI. These implications will be discussed under the categories of the engagement framework. They include exposure to the intervention, determinants, behavior, and outcomes. These categories are organized in a manner that provides information on the potential causal pathway of change between intervention exposure and outcomes, that is, exposure to the intervention (Little e) influences determinants that influences behaviors and ultimately leads to health outcomes. Insights on approaches to DCBI analysis using engagement data will also be discussed in this paper.

Exposure to the Intervention (Little e)

This category is comprised of two engagement types: (1) user interaction with the DBCI features (little e_{UI}) and (2) interactions with the behavior change interventions components (little e_{BCT}).

User Interaction With the Digital Behavior Change Interventions

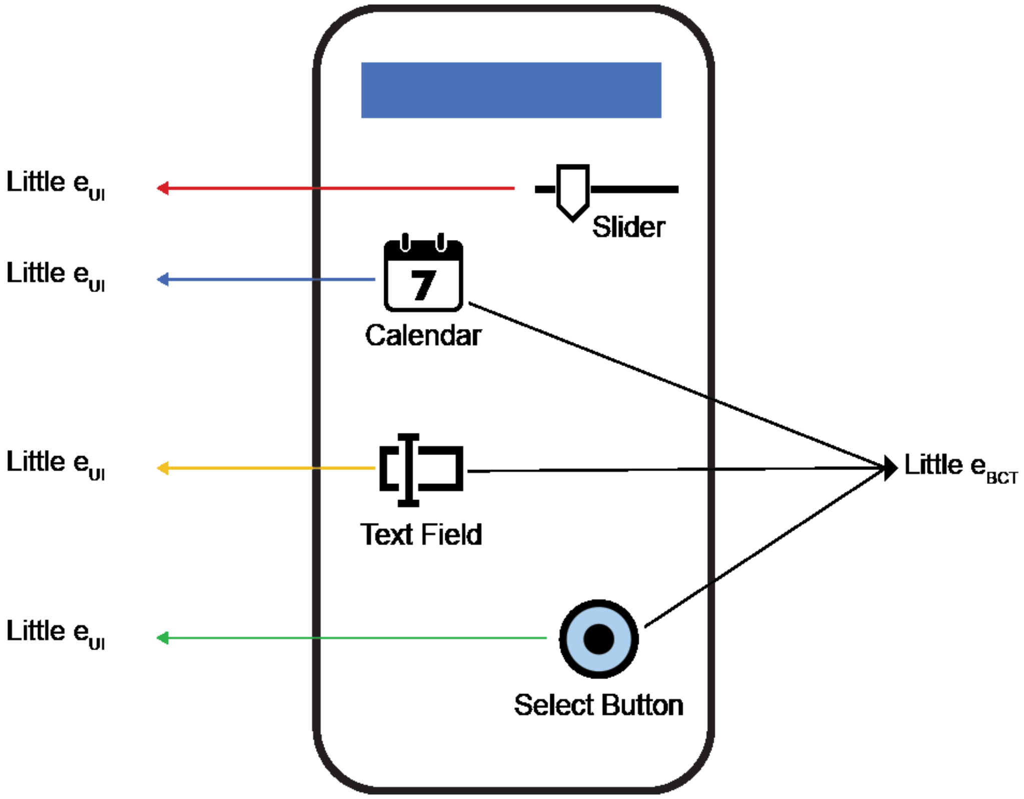
This refers to user interaction with the technology. Exemplary categories of digital solution interactions that could be measured include number of logins, clicks, swipes, time spent interacting with each feature, order of interactions (path taken), and real-time feedback/assessment of user experience. These interactions can be captured within the solution at the backend by using combinations of product analytics data, representing user-level individual interactions data over time (ie, longitudinally). It is critical to collect individual level, time-series data and specify this request upon the development of a DBCI solution. Often, however, out-of-box analytical programs do not have this capability as the default setting.

Interactions With Behavior Change Intervention Components

Interaction with, and subsequent exposure to, the behavior change intervention components (little e_{BCT}) is designed specifically to influence the determinants of behavior [15]. Using theory-informed models [17,19-23], we can detail which user little e_{BCT} interactions are associated with specific health behaviors. In traditional intervention design, these are the key components of behavior change theory that are used to influence changes in determinants, and ultimately, health behaviors.

Behavior change interventions are most effective when developed using behavior change theory [2]. Therefore, we have to understand exposure to behavior change techniques to determine specifically which behavior change intervention components are important to design and then measure in the DBCI. Behavior change techniques are the smallest component parts of widely used theories of health behavior, sourced from various fields and through consensus of experts [10]. BCTs are used because they enable researchers to focus on specific components of an intervention in order to determine how much each aspect of the intervention is contributing to the desired outcome. In other words, use of behavior change techniques, the smallest component part of a behavior change theory, allows us to parse out the effect of each behavior change technique, thus developing an understanding of which technique works for which population and outcome. When these insights are developed in a real-world setting across a range of users, they can be translated for a larger group of users to obtain better outcomes. These BCTs are often experienced as features in the user experiences (Figure 3) and, in some cases, inform how the technology solutions may function.

Figure 3. Diagram illustrating unique constellation of user interaction/user interface features that represent behavior change techniques. UI: user interaction; BCT: behavior change technique.



Further, BCTs arranged in a taxonomy, or ontology, allow for organization of the components in technology design by making behavior change theory concepts machine readable (eg, hierarchically categorized and mutually exclusive), which enables behavior science to converge with data science [24]. Examples of BCTs include goal setting, action planning, feedback, monitoring, and problem solving. A full list of behavior change techniques commonly used in various fields of behavior science can be found in The Behavior Change Technique Taxonomy (v1) of 93 Hierarchically Clustered Techniques developed by Michie et al [10]. This approach enables the identification of active ingredients of the DBCI (ie, the behavior change intervention components that are most critical to improve behavior) and ultimately determine a dose-response relationship between the DBCI and the desired health behavior and health outcome.

Measurement of behavior change intervention components (little e_{BCT}) can be represented by a combination of user interface interactions. Some of the same interactions captured for understanding user interaction with the system (little e_{UI}) can be combined to create measures for DBCI components/active ingredients (Figure 3).

In this illustration, the app user is exposed to an “Action Planning” BCT (little e_{BCT}) via a combination of calendar text features and a select button (little e_{UI}) to schedule an exercise routine.

Determinants

Determinants refers to the influence of factors that shape individual decisions and actual performance of the behaviors. Prior work on digital health technology engagement models advocate for examination of two main level of determinants: (1) determinants of user engagement with the intervention (eg, DBCI) and (2) psychosocial determinants [15]. Determinants of user engagement focus on the influences that support user interaction with a DBCI. For example, technology self-efficacy, satisfaction with the DBCI, intervention usability, and navigability and technology perceived compatibility and ease of use [15,16]. Psychosocial determinants refer to influence of factors (usually social) on individual/group decision and performance of behaviors (desirable or not) [25]. These factors are informed by various behavior science theories such as the COM-B model, socioecological model, the classical and operant processes, social cognitive theory/social learning theory, and self-determination theory [17-19,22,23].

Both types of determinants can be grouped as mediators or moderators and help tie both definitions of engagement (Big E and Little e) together. According to the COM-B model, behavior change occurs when motivation meets capability and opportunity. Specifically, for a behavior to happen, the person must have the capability, opportunity, and motivation to perform the behavior [18]. In our Johnson and Johnson approach to health engagement framework, behavior is defined as an observable action that a person does, capability is a physical and psychological ability to perform the behavior, opportunity is an immediate environmental factor(s) that promotes or enables a behavior, and motivation refers to the mental resolve to direct a behavior [18].

Within the DBCI, determinants can be captured in app through assessments, although the measurement of determinant is not critical. These assessments could be periodic or based on a predefined logic (for example, business rules informing a DBCI). In some instances, this predefined logic is built around specific ecological and environment contexts that are timed or random. These types of assessments are generally referred to as an ecological momentary assessment (EMA) [26]. In the context of a study or research, determinants could be measured using a survey. Determinants are important to measure the effectiveness of intervention components on behavior and outcomes. Therefore, it is important to ensure they can be measured within DBCI during product testing to understand the full mechanism of action. Once a DBCI is released, testing of determinants is unnecessary if no new audiences or population subgroups are introduced. When choosing a data capture method, it is critical to weigh the value of the data to be obtained with user burden expected. This is because a tedious user experience on data capture (eg, more burden on the user to actively track data) can, in turn, influence user engagement determinants negatively [27].

Health Behavior (Big E)

In this model, we categorize health behaviors as lifestyle behaviors, health care–related behaviors, and one-off behaviors. Lifestyle behaviors refer to activities performed very frequently, such as physical activity, sleeping, and eating. Examples of health care–related behaviors include adherence to medication and treatment plans, while getting a flu shot is an example of a one-off health behavior. Health behaviors can be measured though a variety of ways (in app, EMAs, surveys, etc). However, when possible, behaviors should be captured passively using in-app or out-of-app sensors to ensure accuracy of the data and reduce user burden on the people to track and monitor their behavior (except in cases when the intervention specifically calls for it).

Analysis of Digital Behavior Change Interventions Using Engagement Data

Overview

Data from the engagement framework ultimately aims to describe specific relations between Little e and Big E and explains the pathway of these relationships through the influence of determinants. In other words, DBCI analysis seeks to understand which BCT exposures lead to which health behavior and through the effects of which determinants. This framework informs the development of a specific dataset that addresses these questions. There are two broad approaches to consider when analyzing such data: an a priori theory approach and a grounded theory approach.

A Priori Theory Approach

In this approach, DBCI interactions and health behavior engagement can be identified or assigned as the solution has been designed. This is largely informed by theory and learnings from behavior-based intervention implementation literature, such as intervention mapping [20]. A priori definition helps determine if specific assumptions or theoretical concepts utilized in the design of the DBCI solution hold true. For example, if a DBCI feature was designed to help users address a specific motivational determinant of behavior, and the feature was never utilized, you would not expect the motivational determinant to be influenced through that feature/conceptual approach. Similarly, if the feature was utilized, but the person’s motivation did not change, you may question whether this component was an effective way to address this motivational factor. The critical limitation of this approach, however, is that researchers are restricted to the number of behavior change pathways (ie, from Little e to Big E) that were originally accounted for. Emerging new change processes could be potentially missed if they were not initially built in (ie, the analysis of the intervention is only as good as the change processes and pathways originally captured).

A Grounded Theory Approach

In this approach, DBCI interactions and health behaviors are interpreted in real time or after the solution has been used to develop new learnings in order to make the DCBI more effective, if needed. In this instance, methods such as cluster analysis or path analysis can help understand how users experience the intervention, identify useful features, and determine various user pattern groups (that might indicate which path is most successful for specific types of users) [28,29]. This approach helps identify patterns and interactions that the user takes. In some cases, the emerging user engagement patterns and change processes were not intuitive or originally intended. These findings can be used for further exploration and to inform other types of analyses. However, there could be a challenge in capturing and interpreting the data and thus running the risk of spurious associations and findings if they are not vetted by evidence or do not fit the schema of an existing theory [28].

In summary, there is potential value in each approach. However, the pros and cons of both approaches should be carefully considered when making analytic decisions.

Conclusion and the Promise of Engagement Data to Inform Measurement of Digital Behavior Change Interventions

Measurement of both little e_{UI} and intervention BCT components (little e_{BCT}) are necessary to understand the effectiveness of DBCI on Big E (health behavior). DBCI guided by the Johnson and Johnson approach to health engagement framework helps clarify exactly which actions a user takes (path) and the context in which such interaction and health behavior occur. Due to advancements in technology, it is possible to collect data over time regarding Little e and Big E and to consider additional determinants' data on dynamic contextual information (eg, location, time of day, and biometric characteristics like heart rate). Other determinant information (including demographics) are also important and should be leveraged to build user profiles that are robust and describe behavior change processes from Little e to Big E and influence our understanding of behavior change theories [15].

The full potential of digital health solutions can be achieved when insights from human behaviors can be observed and measured discretely to enhance and improve our existing models of behavior change theory, which were developed mostly using static measures of behaviors. Given the power of technology, we are now able to capture fluid behaviors of an individual and are better equipped to address the question of which intervention works best for whom and for what outcome.

It is also possible to measure mechanisms that influence adoption and maintenance of health behaviors to build product efficiencies and effectiveness (relevant to the greatest audience subgroup as feasible). Such mechanisms include Little e (interaction with UI + interaction with BCT components=health behavior via a determinant). Equipped with data about user interactions with technology, exposure to intervention BCT components, and execution of health behavior, we can answer questions of not only *if* the intervention, but also *how* the intervention worked and for *whom*.

We can identify and capitalize on efficiencies and observed insights to optimize a user's path to the health behavior and increase the likelihood of sustaining that health behavior. Evidence has shown that tailored and individualized interventions are successful for improving health behavior and health outcomes [30]. Utilizing all data available (little e_{UI}, little e_{BCT}, Big E), statistical and computational analyses can be utilized to understand patterns of human behavior, identify determinants of human behavior, and personalize interventions specifically for the person receiving the intervention [29,30]. By studying little e_{UI}, little e_{BCT}, and health behaviors in tandem and separately, we can address questions such as the following:

- What is the best user experience for users with particular characteristics?
- What is (are) the most appropriate set of behavior change technique(s)/intervention components, given a user has a certain set of characteristics?
- What is the next appropriate intervention component to provide a user, given they have had positive technology interactions with these components previously?

Essentially, through these data, we become equipped to tailor and personalize interventions without requiring a static battery of psychosocial and preference questions in the initial interaction.

Finally, to reach the full potential of technology applied to behavior change (DBCI), there is a need to draw from diverse populations and observe and learn how DBCI works for different people. Implementing an engagement framework–guided DBCI at scale enables the gathering of a diverse set of data while simultaneously diversifying the intervention appropriately at scale, rather than starting in small pockets and building health technologies that are inflexible.

Conflicts of Interest

HC-L, NE, and JT are full-time employees of Johnson & Johnson Health and Wellness Solutions, a center of excellence (COE) group within Johnson & Johnson.

References

1. Yamey G. Scaling up global health interventions: a proposed framework for success. PLoS Med 2011 Jun;8(6):e1001049 [FREE Full text] [doi: 10.1371/journal.pmed.1001049] [Medline: 21738450]

2. Bull S, Ezeanochie N. From Foucault to Freire Through Facebook. Health Educ Behav 2015 Sep 17;43(4):399-411. [doi: 10.1177/1090198115605310]
3. Riley WT, Rivera DE, Atienza AA, Nilsen W, Allison SM, Mermelstein R. Health behavior models in the age of mobile interventions: are our theories up to the task? Transl Behav Med 2011 Mar;1(1):53-71 [FREE Full text] [doi: 10.1007/s13142-011-0021-7] [Medline: 21796270]
4. 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]
5. Lustria MLA, Noar SM, Cortese J, Van Stee SK, Glueckauf RL, Lee J. A meta-analysis of web-delivered tailored health behavior change interventions. J Health Commun 2013;18(9):1039-1069. [doi: 10.1080/10810730.2013.768727] [Medline: 23750972]
6. Bouvier P, Lavoué E, Sehaba K. Defining Engagement and Characterizing Engaged-Behaviors in Digital Gaming. Simulation & Gaming 2014 Nov 06;45(4-5):491-507. [doi: 10.1177/1046878114553571]
7. O'Brien HL, Toms EG. What is user engagement? A conceptual framework for defining user engagement with technology. J Am Soc Inf Sci 2008 Apr;59(6):938-955. [doi: 10.1002/asi.20801]
8. O'Brien HL, Toms EG. The development and evaluation of a survey to measure user engagement. J Am Soc Inf Sci 2009 Oct 19;61(1):50-69. [doi: 10.1002/asi.21229]
9. Evans L. Search Engine Watch. 2009. Can social media be measured? URL: http://searchenginewatch.com/sew/news/2064649/can-social-media-be-measured [accessed 2019-04-15] [WebCite Cache ID 76M8Inzlm]
10. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. Ann Behav Med 2013 Aug;46(1):81-95. [doi: 10.1007/s12160-013-9486-6] [Medline: 23512568]
11. 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]
12. Gruman J, Rovner MH, French ME, Jeffress D, Sofaer S, Shaller D, et al. From patient education to patient engagement: implications for the field of patient education. Patient Educ Couns 2010 Mar;78(3):350-356. [doi: 10.1016/j.pec.2010.02.002] [Medline: 20202780]
13. West R, Michie S. A guide to development and evaluation of digital behavior interventions in healthcare. Surrey, United Kingdom: Silverback Publishing; 2016.
14. 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;16(6):e146 [FREE Full text] [doi: 10.2196/jmir.3077] [Medline: 24905070]
15. Short C, Rebar A, Plotnikoff R, Vandelanotte C. Designing engaging online behavior change interventions: A proposed model of user engagement. The European Health Psychologist 2015;17(1):32-38.
16. Holden RJ, Karsh B. The technology acceptance model: its past and its future in health care. J Biomed Inform 2010 Feb;43(1):159-172 [FREE Full text] [doi: 10.1016/j.jbi.2009.07.002] [Medline: 19615467]
17. Sallis J, Owen N, Fisher E. Ecological models of health behavior. In: Glanz K, Rimer BK, Viswanath K, editors. Health behavior: Theory, research, and practice. San Francisco, CA: Jossey-Bass; 2015:43-64.
18. Jackson C, Eliasson L, Barber N, Weinman J. Applying COM-B to medication adherence: a suggested framework for research and interventions. European Health Psychologist 2014;16(1):7-17.
19. Bandura A. Social foundations of thought and action: A social cognitive theory. Englewood Cliffs, NJ: Prentice-Hall; 1986.
20. Eldredge L, Markham C, Ruiter R, Kok G, Fernandez M, Parcel G. Planning health promotion programs: an intervention mapping approach. Hoboken, NJ: John Wiley & Sons; 2016.
21. Guerra LGGC, Silva MTA. Learning processes and the neural analysis of conditioning. Psychology & Neuroscience 2010 Jul;3(2):195-208. [doi: 10.3922/j.psns.2010.2.009]
22. Bandura A, Walters R. Social learning theory. Englewood Cliffs, NJ: Prentice-hall; 1977.
23. Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. American Psychologist 2000;55(1):68-78. [doi: 10.1037/0003-066X.55.1.68]
24. Larsen KR, Michie S, Hekler EB, Gibson B, Spruijt-Metz D, Ahern D, et al. Behavior change interventions: the potential of ontologies for advancing science and practice. J Behav Med 2016 Aug 01. [doi: 10.1007/s10865-016-9768-0] [Medline: 27481101]
25. Bolívar J, Daponte A, Rodríguez M, Sánchez JJ. The Influence of Individual, Social and Physical Environment Factors on Physical Activity in the Adult Population in Andalusia, Spain. IJERPH 2010 Jan 05;7(1):60-77. [doi: 10.3390/ijerph7010060]
26. Shiffman S, Stone AA, Hufford MR. Ecological Momentary Assessment. Annu Rev Clin Psychol 2008 Apr;4(1):1-32. [doi: 10.1146/annurev.clinpsy.3.022806.091415]
27. Suh H, Shahriaree N, Hekler E, Kientz J. Developing and validating the user burden scale: A tool for assessing user burden in computing systems. 2016 May 12 Presented at: Proceedings of the CHI conference on human factors in computing systems; 2016; San Jose, CA p. A-3999. [doi: 10.1145/2858036.2858448]

28. Nunnally J. Psychometric Theory. New York City, NY: McGraw-Hill; 1978.
29. Anderson JC, Gerbing DW. Predicting the performance of measures in a confirmatory factor analysis with a pretest assessment of their substantive validities. Journal of Applied Psychology 1991 Oct;76(5):732-740. [doi: 10.1037/0021-9010.76.5.732]
30. Teyhen DS, Aldag M, Centola D, Edinborough E, Ghannadian JD, Haught A, et al. Incentives to Create and Sustain Healthy Behaviors: Technology Solutions and Research Needs. Military Medicine 2014 Dec;179(12):1419-1431. [doi: 10.7205/MILMED-D-14-00111]

Abbreviations

BCT: behavior change technique
COM-B: Capability, Opportunity, Motivation - Behavior
DBCI: digital behavior change intervention
EMA: ecological momentary assessment
TAM: technology acceptance model
UI: user interaction

Edited by P Santana-Mancilla; submitted 19.03.19; peer-reviewed by V Guzman, J Contreras-Castillo; comments to author 18.04.19; revised version received 23.07.19; accepted 14.08.19; published 10.10.19.

Please cite as:
Cole-Lewis H, Ezeanochie N, Turgiss J
Understanding Health Behavior Technology Engagement: Pathway to Measuring Digital Behavior Change Interventions
JMIR Form Res 2019;3(4):e14052
URL: https://formative.jmir.org/2019/4/e14052
doi:10.2196/14052
PMID:31603427

Original Paper

Adapting Mobile and Wearable Technology to Provide Support and Monitoring in Rehabilitation for Dementia: Feasibility Case Series

Julia Thorpe[1], PhD; Birgitte Hysse Forchhammer[2], PhD; Anja M Maier[1], PhD

[1]Engineering Systems Design, DTU Management, Technical University of Denmark, Kongens Lyngby, Denmark

[2]Department of Neurology, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark

Corresponding Author:
Anja M Maier, PhD
Engineering Systems Design
DTU Management
Technical University of Denmark
Diplomvej, Building 317
Kongens Lyngby, 2800
Denmark
Phone: 45 45256045
Email: amai@dtu.dk

Abstract

Background: Mobile and wearable devices are increasingly being used to support our everyday lives and track our behavior. Since daily support and behavior tracking are two core components of cognitive rehabilitation, such personal devices could be employed in rehabilitation approaches aimed at improving independence and engagement among people with dementia.

Objective: The aim of this work was to investigate the feasibility of using smartphones and smartwatches to augment rehabilitation by providing adaptable, personalized support and objective, continuous measures of mobility and activity behavior.

Methods: A feasibility study comprising 6 in-depth case studies was carried out among people with early-stage dementia and their caregivers. Participants used a smartphone and smartwatch for 8 weeks for personalized support and followed goals for quality of life. Data were collected from device sensors and logs, mobile self-reports, assessments, weekly phone calls, and interviews. This data were analyzed to evaluate the utility of sensor data generated by devices used by people with dementia in an everyday life context; this was done to compare objective measures with subjective reports of mobility and activity and to examine technology acceptance focusing on usefulness and health efficacy.

Results: Adequate sensor data was generated to reveal behavioral patterns, even for minimal device use. Objective mobility and activity measures reflecting fluctuations in participants' self-reported behavior, especially when combined, may be advantageous in revealing gradual trends and could provide detailed insights regarding goal attainment ratings. Personalized support benefited all participants to varying degrees by addressing functional, memory, safety, and psychosocial needs. A total of 4 of 6 (67%) participants felt motivated to be active by tracking their step count. One participant described a highly positive impact on mobility, anxiety, mood, and caregiver burden, mainly as a result of navigation support and location-tracking tools.

Conclusions: Smartphones and wearables could provide beneficial and pervasive support and monitoring for rehabilitation among people with dementia. These results substantiate the need for further investigation on a larger scale, especially considering the inevitable presence of mobile and wearable technology in our everyday lives for years to come.

(JMIR Form Res 2019;3(4):e12346) doi:10.2196/12346

KEYWORDS

dementia; cognitive rehabilitation; mobility; activity; mHealth; uHealth; pervasive health care; P4 health care; health care design

Introduction

Background

New approaches are needed to respond to the growing dementia challenge as the population ages [1]. A global action plan recently issued by the World Health Organization calls for solutions to improve the lives of people with dementia and their caregivers and to reduce the impact the condition has on communities [2]. One promising approach is through rehabilitation, defined as a problem-solving process aimed at optimizing social engagement and well-being [3,4]. Rehabilitation among people with cognitive impairment, or cognitive rehabilitation, is characterized by a personalized collaborative approach, which involves setting and working toward individual goals within an everyday life context [5]. Following initial assessment to identify needs and set goals, the rehabilitation process involves two key elements: provision of support to help the person attain their goals and monitoring or evaluation to inform the care strategy for further iterations [4].

Technological advancement over the last decade has led to mobile and wearable technology that could provide both support and monitoring in rehabilitation. The application of personal devices, such as smartphones and wearables, for rehabilitation presents several potential advantages. Regarding support in everyday life, these include the following: wide availability, convenience and familiarity of an existing device, less stigma than specialized assistive technology, and a modular nature that allows for personalization (eg, custom device configurations). Regarding monitoring, sensing capabilities of the latest personal devices offers a scalable and connected approach toward ongoing evaluation. Therefore, this work investigates how smartphones and smartwatches might be applied toward rehabilitation among people with dementia by offering both personalized support in everyday life and objective continuous monitoring of mobility and activity as a means to evaluate function and engagement.

Mobile and Wearable Technology for Support and Monitoring Among People With Dementia

Various forms of information and communications technology (ICT) have been used for people with dementia to provide functional support in everyday life, improve safety, target psychosocial needs, and support caregivers [6,7]. In the future, a user's smartphone could provide a single platform through which to offer any combination of similar support features. A host of existing tools for communication, scheduling, reminders, navigation, social, and leisure purposes are already available from off-the-shelf mobile apps. An increasing proportion of elderly people will rely on such tools prior to dementia onset as current users age. Already today, many people with mild cognitive impairment and dementia are using ICT [8], and studies have described interest among members of this population in using wearables for support [9].

Mobile and wearable devices are also packed with sensors that can be used to gather information about users' lifestyles and behaviors, such as their mobility and activity levels or patterns. Monitoring behavior among people with dementia can provide valuable indicators for functional performance and well-being to inform care strategies [10] and thereby support the rehabilitation process.

Mobile technology has successfully been applied to fulfill various functions related to rehabilitation among people with dementia or other neurological diagnoses, such as providing memory support following traumatic brain injury [11]; monitoring activity, mobility [12,13], and goal setting [14]; and for rehabilitation after stroke [14]. We build upon this work by extending and combining support and monitoring capabilities of personal devices and by implementing this among the target population of people with dementia.

Aim and Objectives

The primary aim of this work is to evaluate the feasibility of using mobile and wearable technology to support rehabilitation for dementia. We have implemented a technological setup combining both personalized support and behavioral monitoring among people with dementia in a real-life context in a series of 6 in-depth case studies to address three main objectives:

1. Examine the technical feasibility of sensor-based behavioral measurement using smartphones and smartwatches among people with dementia.
2. Compare participants' subjective perceptions of their behavior with objective, sensor-derived measures.
3. Evaluate user acceptance focusing on usefulness and health efficacy.

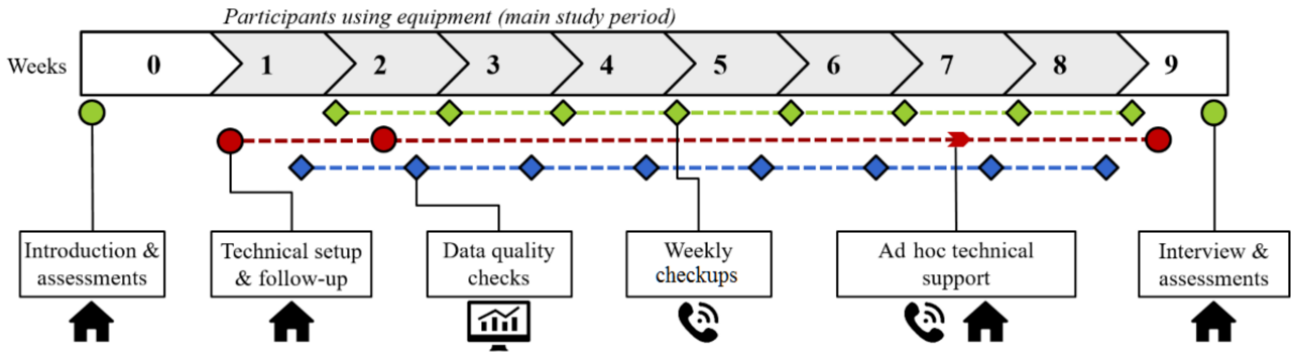
Through fulfilling these objectives, this work will contribute new evidence to the potential for smartphones and wearables to benefit rehabilitation in practice, identifying areas for further research on a larger scale. Addressing new opportunities presented by mobile devices and the data these generate makes an important step toward the wider vision of data-driven health care interventions that enable predictive, preventative, personalized, and participatory (P4) health care [15,16].

Methods

Study Design

A longitudinal study design comprising 6 case studies was employed. Each case included both a main participant (ie, person with dementia) and their caregiver over a period of at least 8 weeks. During the study, participants used smart devices for support in everyday life, answered daily reports on their mobility and activity levels, and followed an individual goal. Behavioral data was recorded from device sensors throughout participation. The study procedure is outlined in Figure 1, which shows the series of participant interactions through which both quantitative and qualitative data were collected throughout the process.

Figure 1. Procedure for each case study showing interaction points and data collection.



Recruitment and Participants

Recruitment was carried out through the dementia and memory clinic at a Danish hospital. The population of interest is community-dwelling adults in the early stages of dementia, which includes people with mild-to-moderate cognitive impairment. Participants were required to live with their primary caregiver, due to the supportive role the caregiver plays in the study (eg, providing information). Participants were not required to have any prior experience using smart technology. Exclusion criteria included any disability that would either affect the participant’s ability to use smart technology (eg, severe vision or hearing loss and apraxia) or cause extremely limited mobility and activity levels (eg, home- or bed-bound participants).

A total of 6 participants—2 female (33%) and 4 male (67%)—between 65 and 78 years of age completed the study. A further 3 participants who enrolled dropped out due to illness or feeling daunted at the prospect of using the devices. A summary of the participants’ demographic and clinical backgrounds is provided in Table 1. Cognitive assessment scores were collected from the clinic through which participants were recruited. While certain scores were within the normal range for cognitive function, behavioral and executive symptoms in the early stages of dementia were not always captured by the Mini-Mental State Examination (MMSE) and Addenbrooke's Cognitive Examination (ACE). All participants were diagnosed with mild-to-moderate cognitive impairment based on specialized evaluation at the clinic. A brief introduction to the participants in each case study is provided in the Results section.

Table 1. Summary of participant demographics and assessment scores.

Participant information, measure	Participant					
	1	2	3	4	5	6
Age in years	78	70	65	68	70	67
Years retired	16	2	3	2	10	2
Gender	Male	Male	Female	Male	Female	Male
Cognitive impairment						
MMSE[a]	27	27	24	26	27	23
ACE[b]	75	75	71	77	88	81
Quality of life						
QoL-AD[c]						
1[d]	33	44	41	51	45	50
2[e]	38	42	46	49	49	48
Functional performance						
FAQ-IADL[f]						
1	4	5	10	0	5	10
2	6	9	5	0	5	12
Caregiver burden						
ZBI[g]						
1	11	26	0	N/A[h]	N/A	20
2	1	28	0	18	3	13
Mobility						
LSA[i]						
1	76	84	50	66	100	66
2	74	84	84	90	84	74
Activity						
GPAQ[j] (hours/week)						
1	27.50	10.00	20.00	39.50	70.25	7.00
2	37.50	14.00	34.75	32.75	37.00	14.50

[a] MMSE: Mini-Mental State Examination.

[b] ACE: Addenbrooke's Cognitive Examination.

[c] QoL-AD: Quality of Life in Alzheimer's Disease scale.

[d] Prestudy measurements.

[e] Poststudy measurements.

[f] FAQ-IADL: Functional Assessment Questionnaire-Instrumental Activities of Daily Living.

[g] ZBI: Zarit Burden Interview.

[h] N/A: not applicable.

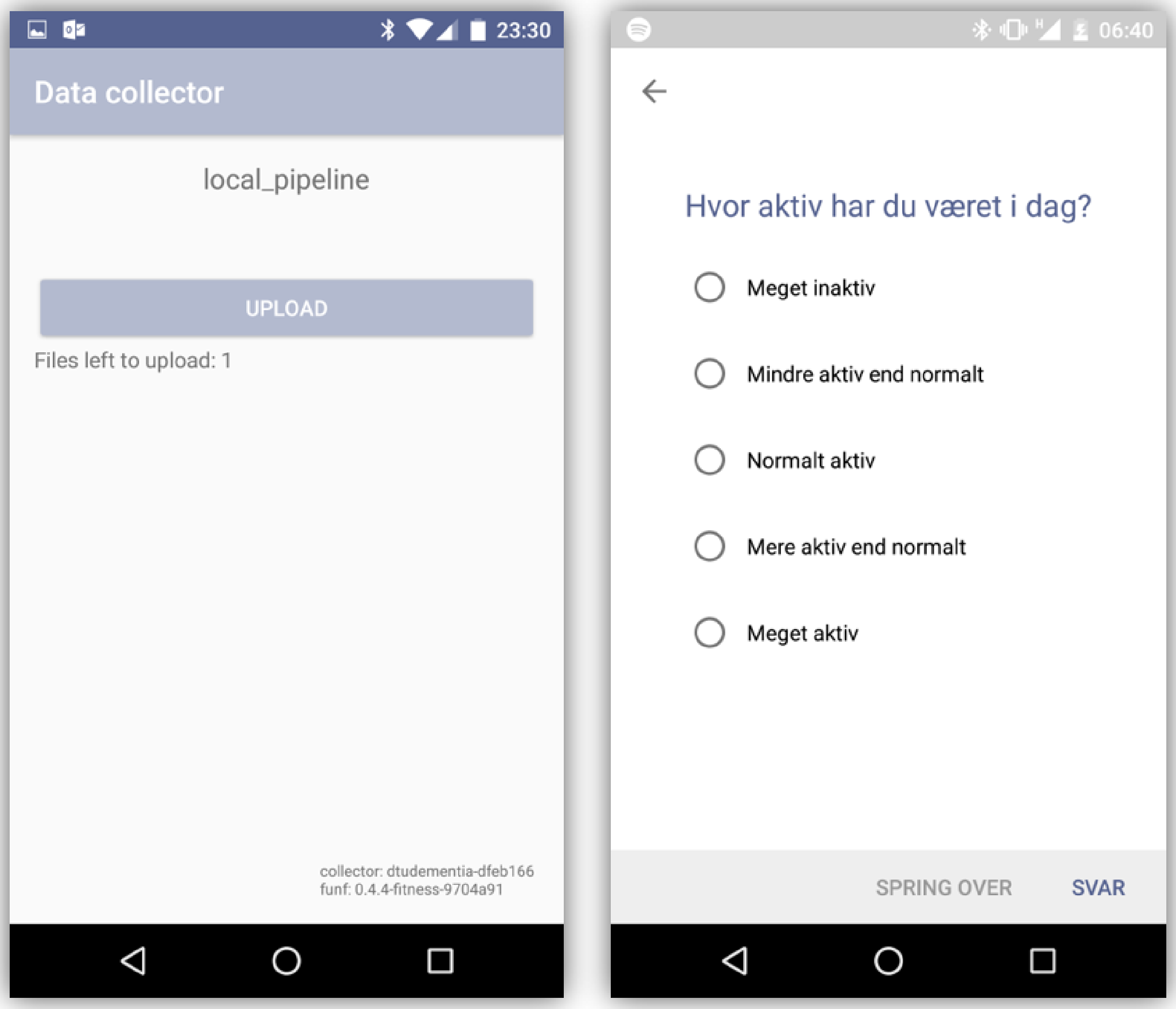
[i] LSA: Life-Space Assessment. The LSA, adapted, was used: the score was calculated assuming independent travel.

[j] GPAQ: Global Physical Activity Questionnaire. Results show total active time for work, including household, leisure, and transport (ie, walking or cycling).

Technical Setup for Support and Monitoring

Mobile devices were employed to support participants in their everyday lives and to collect behavioral data. The technological setup comprised a smartphone (ie, Nexus 5 running Android OS v6.0.1.) paired with a smartwatch (ie, Sony SmartWatch 3 running Android Wear), a mobile self-reporting module, and an app for secure collection of sensor data and logs from the devices (see Figure 2). A detailed description of the system is available in Thorpe et al [17].

Figure 2. The data collector custom app (left) runs in the background collecting device sensor and log data. The mobile self-reporting app (right) prompts users to evaluate their daily activity and mobility levels.



Data collector

local_pipeline

Files left to upload: 1

collector: dtudementia-dfeb166
funf: 0.4.4-fitness-9704a91

Hvor aktiv har du været i dag?

○ Meget inaktiv
○ Mindre aktiv end normalt
○ Normalt aktiv
○ Mere aktiv end normalt
○ Meget aktiv

A technical orientation meeting was held at the start of each case study to introduce the devices and to personalize device setup in collaboration with the participant and their caregiver. As a standard setup, the smartphone home screen displayed the time; upcoming appointments; and, based on interest from all participants, a daily step count. All participants were provided with Google Calendar and shown how this could be used to remember tasks or appointments, since this was shown to be useful and not too challenging for participants in an earlier study [18]. This standard setup was then further adapted by adjusting settings or adding apps to fit each participant according to individual needs, including what they already used. A follow-up visit 1 week after orientation was carried out to resolve early issues and repeat instructions. Participants were also provided with an illustrated manual and offered technical support throughout the study over the phone and during in-home visits, as required.

For collection of behavioral data, two custom apps were deployed: (1) a data collector app that runs in the background recording activity and location data from device sensors and (2) a self-reporting app that prompts input from users regarding their perceived mobility and activity according to a predefined schedule.

Goal Setting and Following

Participants set individual goals together with a member of the research team trained in psychology, which they followed throughout participation. Goal setting is used extensively in rehabilitation and was, therefore, emulated in the case studies specifically within the themes of mobility and activity and according to participants' own views of what was important to their lifestyle.

Goals included getting out of the house each day (participants 3 and 6), walking longer daily (participant 2), driving less and riding a bicycle more (participant 4), and maintaining an existing activity schedule (participants 1 and 5). Cognitive impairment can affect a person's ability to function in everyday life and remain socially engaged, leading to depressive symptoms and reclusiveness that can further aggravate cognitive decline. Such mobility- and activity-related goals are, therefore, intended to encourage participants to remain active and engaged, eliciting support as required to achieve their goals. Participants were asked to evaluate their own goal attainment in weekly phone calls, which provided an opportunity for them to further qualify their answers.

Data Collection

Quantitative and qualitative data were collected throughout the studies, as outlined in Table 2 and in Figure 1 showing the protocol steps.

Background and Assessments

Background information and most recent cognitive impairment test results were collected at the start of the study from the clinic. A set of questionnaire-based assessments was performed at the start and end of participation. These were collected as a reference for participants' profiles and in the case of marked change that may influence the participants' experiences; they were not used for pre-post analyses. Questionnaires included assessments for quality of life, function in daily living activities, caregiver burden, and mobility and activity (see Table 2). Mobility and activity questionnaires were adapted to fit the study purposes; independence level was excluded from the Life-Space Assessment (LSA) used to measure mobility [19], and activity intensity grading was excluded from the sections selected from the Global Physical Activity Questionnaire (GPAQ) [20].

Table 2. Summary of data collected from the case studies.

Category of data, frequency of data collection	Description
Background information	
At study start	Demographic: age, gender, education, and occupation Cognitive impairment severity: MMSE[a] and ACE[b]
Questionnaires	
At study start and end	Quality of life: QoL-AD[c] Functional performance: FAQ-IADL[d] Caregiver burden: ZBI[e], short form Mobility: LSA[f], adapted Activity: GPAQ[g], adapted
Device data	
Continuous	Location Activity Step count Battery status Screen on and off
Mobile self-reports	
Daily (evening)	Perceived mobility (daily) Perceived activity level (daily)
Meetings and interviews	
Weekly	Phone calls: perceived goal attainment score and supplementary notes
At study end	Semistructured interview on experiences and outcomes
Ad hoc	Support interactions

[a] MMSE: Mini-Mental State Examination.

[b] ACE: Addenbrooke's Cognitive Examination.

[c] QoL-AD: Quality of Life in Alzheimer's Disease scale.

[d] FAQ-IADL: Functional Assessment Questionnaire-Instrumental Activities of Daily Living.

[e] ZBI: Zarit Burden Interview.

[f] LSA: Life-Space Assessment.

[g] GPAQ: Global Physical Activity Questionnaire.

Sensor Data From Mobile Devices

Data from device sensors and logs were collected throughout study participation. Location and activity data, including steps and recognized activities, were recorded and used to calculate a set of mobility and activity metrics. The set of algorithms used to measure mobility and activity behavior was described in detail in an earlier study by Thorpe et al [17]; this set combined a range of metrics used for similar monitoring purposes in related literature [21-24].

Location data merged from the phone and watch was first used to extract a series of stays and moves throughout each day.

These, together with the raw global positioning system (GPS) coordinates, were then used to calculate a set of spatial, temporal, and frequency-based mobility measures, including the following:

1. Minimum convex polygon (MCP) or mobility envelope (ie, area of the polygon constructed around location data).
2. Action range or home range (ie, furthest straight-line distance travelled).
3. Total distance covered, out of home only.
4. Time spent out of home.
5. Time spent moving between locations.
6. Number of places visited.
7. Number of trips.

Recognized activities from the phone only, which were accessed using Google's activity recognition application programming interface, were used to extract bouts of activity within the categories: still, on foot, bicycle, and vehicle. Total daily steps were recorded independently from the phone and watch. This data were used to calculate the following activity measures:

1. Active time: sum of durations of all activity bouts on foot and bicycle.
2. Active bouts: total count of all activity bouts on foot and bicycle.
3. Still time: sum of duration of still bouts.
4. Total steps.

Self-Reported Activity, Mobility, and Goal Attainment

Mobile self-reports were issued daily asking participants about their perceived activity and mobility that day relative to their normal daily level on a 5-point scale. The five responses were as follows: *much less than normal*, *less than normal*, *normal*, *more than normal*, or *much more than normal*. These were recorded as values within the set {0, 0.25, 0.5, 0.75, 1}, where 0.5 equates to *normal* for interpretation as a median.

Participants evaluated their own goal attainment weekly in a regular phone call. Goal attainment was scored along a range from -2 to 2, where a score of 0 was assigned if the goal is met, with 2 levels in either direction for under- or overachievement, as has been described by Turner-Stokes [25].

Qualitative Data From Participant Interactions and Semistructured Interviews

A semistructured interview with each participant at the end of the study was used to gather qualitative data on their experiences, particularly in relation to technology acceptance factors. Interviews included questions on the following:

1. Pre-existing tools and coping strategies.
2. Experiences using the technology, including how it was used and for which purposes as well as the difficulties and benefits of its use.
3. Adequacy of the technical support and participants' own knowledge and skills for operating the devices.
4. Expectations and whether these were met, which needs were not met, and any desired but absent functionality.
5. Impact on participants' everyday lives and on their health.

Notes from the weekly phone calls as well as technical support logs were also collected.

Data Analysis

Overview

Three main analyses were performed to fulfill the study objectives. The first examines the availability and utility of the data generated by the smart devices; the second examines agreement between participants' subjective reports of their behaviors and objectives with sensor-based measures; and the third examines user acceptance factors.

Analysis 1: Data Availability and Utility

The aim of this analysis was to determine whether the smart devices, as used in a real-life context, generate adequate data for the intended purpose of monitoring behavior. This depends both on the technology functioning correctly and on the participants using them sufficiently (ie, keeping the devices charged and connected, carrying them around with them, and answering mobile self-reports). We examined the following:

1. Data availability: the proportion of the study period for which data were available from each of the devices and from the mobile self-reports.
2. Data utility: visual inspection of the device interaction and behavioral data to evaluate whether data quality and quantity were sufficient for behavioral patterns to be evident.

Analysis 2: Self-Reported Versus Sensor-Based Measurement of Behavior

Overview

The aim of this analysis was to determine whether the behavioral insights gained from sensor data agree with the participants' own perceptions of their behavior. Since current rehabilitation approaches rely heavily on input from participants and caregivers to assess progress and outcomes, it is also relevant to investigate how such subjective reports differ from the proposed sensor-based approach. Identifying where these differ can guide further investigation into whether this is due to inaccurate perceptions or technical failures. Sensor-derived measures were compared with participants' mobile self-reports of activity and mobility levels and their weekly goal attainment ratings.

Daily Activity and Mobility

Sensor-based measures of activity and mobility vary in units of measurement and were, therefore, ranked as percentiles using the empirical cumulative distribution function for comparison with self-reports (5-point scale). The self-reported and sensor-based measures were then plotted together for comparison within each case study.

Weekly Goal Attainment

The comparison between perceived goal attainment and sensor-based measures required individual analysis according to each participant's defined goals. In most cases, this involved selecting relevant activity and/or mobility measures and aggregating these by week. Other methods included detecting visits to a specified location (eg, a training center), which required a priori knowledge of the location's GPS coordinates. Goal attainment ratings ranged from -2 to 2, where 0 corresponded to having met the goal; there were 2 levels for under- or overachievement in either direction. Following the ranking of sensor-based measures as percentiles, as in the comparison with mobile self-reports, these were shifted by 0.5 such that the median lay on the zero line for comparison with goal attainment ratings.

Analysis 3: Qualitative Analysis of User Acceptance—Usefulness and Health Efficacy

The aim of the third analysis was to evaluate potential acceptance of the devices for support. Usability and usefulness are two important factors influencing technology acceptance, as proposed in the technology adoption model and numerous adaptations thereof [26]. We were interested in whether participants benefited from support selected from a broad range of available tools; therefore, usefulness was more appropriate than usability, which was evaluated for a similar smartphone and smartwatch setup among the population of interest in Thorpe et al [18]. Within health care technology, health efficacy is a further important consideration [27]. This analysis, therefore, focused on usefulness and health efficacy in terms of impact on aspects of quality of life, such as function in everyday life and social engagement. Interview recordings were transcribed and analyzed to identify the following:

1. Support offered to the participants by the mobile devices, the extent to which this was beneficial, and which support needs were met.
2. Perceived impact of using the technology on aspects of quality of life, such as function, independence, behavior, mood, or social engagement.

Results

Overview

Results are presented for three analyses on the following: data availability and utility, comparison between sensor-based and self-reported behavioral measurement, and user acceptance. As further background, a brief introduction to the 6 participants in the case studies is provided in Textbox 1.

Textbox 1. Brief introduction to the 6 case study participants.

Participant 1 (male, 78 years) lives with his wife and son. He has been retired for 16 years at the time of participation but follows an active weekly schedule involving diverse sports, hobbies, and interests that he often travels to by bicycle. This is the only participant who used his own smartphone for the study; other participants already using smartphones owned iPhones, which were incompatible with the custom apps for data collection.

Participant 2 (male, 70 years) lives with his wife, whom he depends on heavily for support in everyday life. She manages his schedule and they agree that she is "his most important aid." He struggles with fatigue and low energy, some days sleeping most of the day, but enjoys going out for walks when able. He owns an iPhone and, though he could not get used to the study phone, was able to operate the devices adequately for completing the study.

Participant 3 (65 years, female) lives with her husband. She has battled with psychosocial consequences of symptoms, such as anxiety and feeling unsafe leaving the house due to fear of not finding her way home. She keeps physically active doing housework, going out for walks, and doing physical training at home. She has used an iPhone but successfully learned to use the study phone to the same extent.

Participant 4 (male, 68 years) lives with his cohabiting partner. He describes the greatest impact his diagnosis has had as being the loss of his driver's license, which he had regained for 2 years just prior to participation. He keeps active by taking his dogs out, visiting a local center aimed at retired members of the community, and riding his bicycle. He also enjoys hobbies such as gardening and playing music. While he owns an iPhone that he uses extensively, he describes a disinterest in technology generally. He did not use the study phone for anything further than pairing it to the watch and answering mobile self-reports, as he did not feel comfortable learning to use a new device.

Participant 5 (female, 70 years) lives with her husband and has been retired for 10 years. She is highly active, with a physical training schedule including sessions 5 days per week. She is not experienced with smart technology and, though they own a household iPhone, she prefers to use a basic Nokia. She, therefore, did not use the study phone for other functions besides answering the mobile self-reports. She found the watch uncomfortable and did not use it for the study. A decrease in motivation during the study was reported and attributed to psychosomatic symptoms (ie, pain and fatigue).

Participant 6 (male, 67 years) lives with his wife. He retired earlier than planned due to health complications and tries to keep active and engaged by doing light housework, shopping, or going for walks. He has no prior experience with a smartphone. He has a basic Nokia that he can use for calls and for reading text messages but not for writing them. Despite his limited experience, he did not feel it was too difficult to use the devices for the basic purposes required (eg, turning them on and off and reading his step count) but needed help from his wife answering mobile self-reports.

Analysis 1: Data Availability and Utility

Here we examine whether adequate data were generated for the intended purpose of monitoring behavior among the population of interest. Table 3 shows the duration in days for participants in each case study along with how many days data were recorded at all, from the smartwatch, and from mobile self-reports. For mobile self-reports, a value of 0.5 was assigned for each of the two questions per day.

While data were recorded on nearly all days, only a small proportion was generated by the watch. Participant 5 did not wear the watch. Several participants had difficulty maintaining the connection between the watch and phone, including noticing when they were disconnected. A further known cause was a technical fault, whereby watch data were not transferred to the phone. This necessitated a manual upload and reset of both devices, which took days or weeks to resolve depending on the participants' availability. Missing self-reports were mostly attributed to technical failures and usability issues rather than adherence. Participant 1 only started receiving the questions 7 weeks into the study and, consequently, agreed to extend his participation. Participant 6 could not answer self-reports without assistance from his caregiver. Limited availability of watch data and self-reports did not necessarily inhibit the potential for behavioral monitoring in practice; all data types were recorded from the phone sensors, from which data were available throughout the study, and self-reports were for research purposes only. One noteworthy limitation was regarding step count, which was expected to be more reliable from the worn watch than from a phone that needed to be carried.

Visual inspection of the generated data indicates that this could indeed be adequate for behavioral monitoring. Figure 3 shows location data generated throughout the study for participants 4 and 5, indicating distance from home as a color scale. These demonstrate that the devices were carried with them and show the potential for resulting data to reveal patterns in behavior. For example, a daily rhythm is evident for participant 5 (ie, movement from 6 am to 8 am near home, further movement around noon, and less-consistent activity in the late afternoon), which contrasts with the less-structured movement for participant 4. Figure 4 shows activity data for participant 2, again showing that the phone was used enough to generate activity bouts throughout the study. This figure reveals the habit of walking during the evenings that the participant reported having started during the study out of motivation to increase his daily step count.

The three examples above are representative of all cases in terms of availability of data. This was sufficient despite variation in levels of interaction with the devices. Figure 5 shows screen-on time as an indication of device interaction for participant 3 who used the phone, participant 4 who did not use the phone, and participant 6 who appears to have attempted to use the phone at the start, with interaction diminishing over time.

Table 3. Data availability.

Participant	Participant duration (days), N	Data recorded at all (days), n (%)	Data recorded by smartwatch (days), n (%)	Data recorded by self-reports (days), n (%)
1	98	98 (100)	57 (58)	39.0 (40)
2	82	81 (99)	4 (5)	59.5 (73)
3	60	60 (100)	10 (17)	53.5 (89)
4	72	72 (100)	8 (11)	39.0 (54)
5	87	86 (99)	0 (0)	49.5 (57)
6	57	54 (95)	2 (4)	46.0 (81)

Figure 3. Location data for 2 participants throughout the study period. The time of day is shown along the horizontal axis and increasing date is shown along the vertical axis.

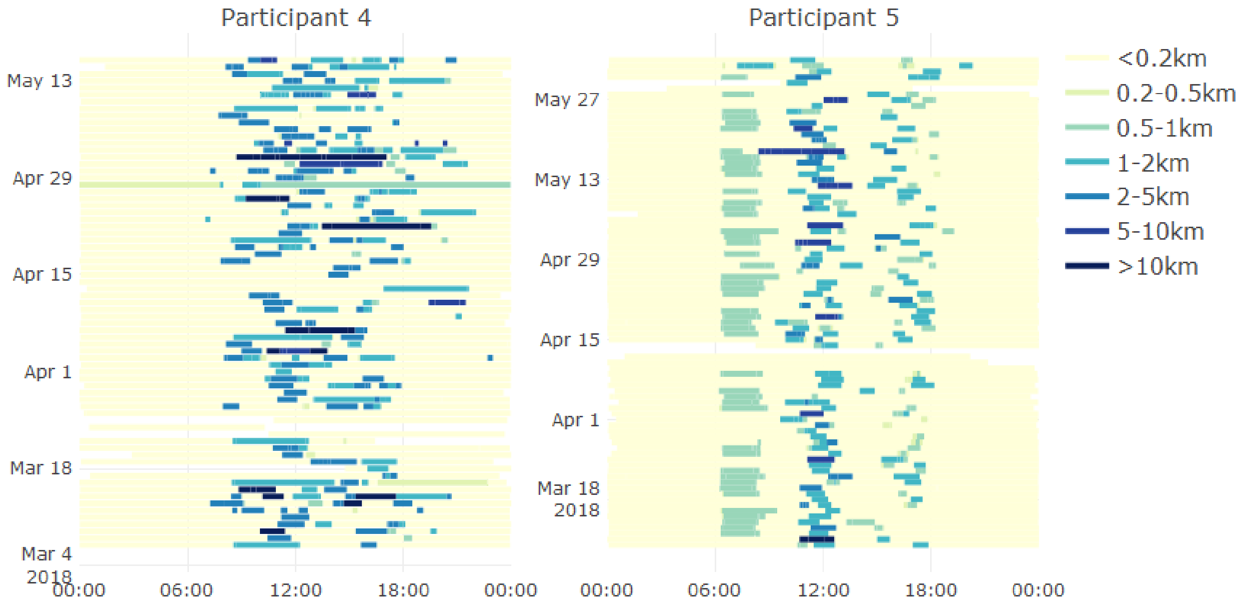


Figure 4. Activity data for participant 2 throughout the study period. The time of day is shown along the horizontal axis and increasing date is shown along the vertical axis.

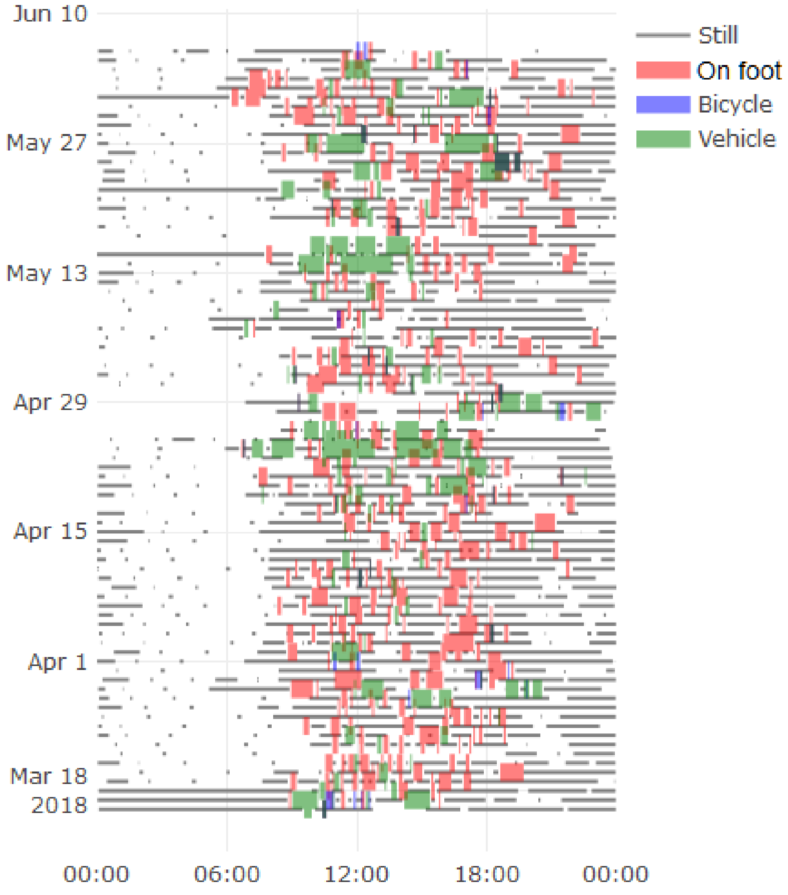
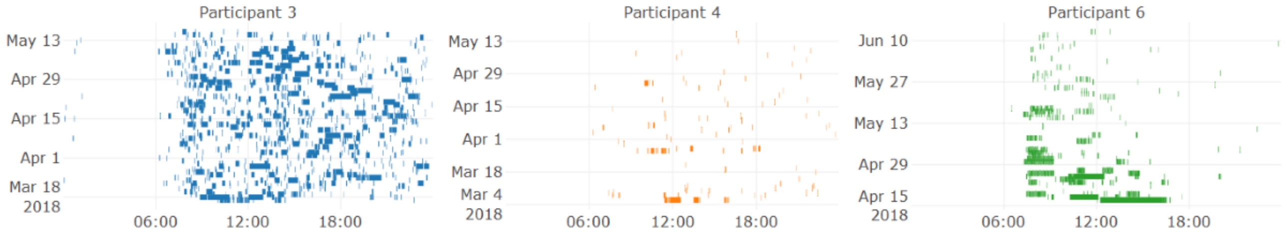


Figure 5. Screen-on time for 3 participants over their entire participation period. The time of day is shown along the horizontal axis and increasing date is shown along the vertical axis.



Analysis 2: Self-Reported Versus Sensor-Based Measurement of Behavior

Daily Activity and Mobility

Self-reported daily levels of mobility and activity were compared with sensor-derived measures. Measures were selected for analysis from the complete set based on several factors. A total of 4 participants out of 6 (67%) traveled out of town during participation, which affected measures such as *action range* or *time out of home* for all days away. For mobility, the measures *MCP*, *time spent moving*, and *number of places visited* were therefore included, which were only elevated on the actual days of travel. For activity, the measures *active time*, *active bouts*, and *steps* were included. Steps were taken only from the phone due to the limited watch data and potential for large differences in the count from different sources that could skew results if combined. A combined measure for mobility and activity was calculated as the average of the three measures for each.

By definition, most days were expected to be reported as *normal* and show sensor measurements near the median, inevitably resulting in agreement between the signals. Therefore, we focused on deviations from normal to determine whether the sensor-based measures followed subjectively reported fluctuations in behavior. Where they did not, we proposed sources of such disagreement and inferred potential advantages or disadvantages of each modality. Four illustrative results are included in Figures 6-9.

Figure 6 shows that the sensor-based measures closely aligned with frequent (ie, day-to-day) fluctuations in perceived activity levels for participant 3, whereas a slower trend (ie, increase over the first month) in the sensor-based measures for participant 5 was not reflected in her self-reported measures (see Figure

7). This could indicate that changes that were more gradual were not as easily perceived by the individual.

The mobility measures for participant 1 (see Figure 8) again show alignment in the direction of fluctuations, only this time with some distinct deviations, most notably in mid-April. This participant felt unsure about how to describe his mobility in relation to his schedule, noting that “normal for a Tuesday is different from normal for a Friday.” This could explain why higher mobility might have been reported as normal, requiring careful consideration of seasonality for detecting deviations in behavioral signals.

In the mobility comparison for participant 3 (see Figure 9), subjective reports lay predominantly in the upper range (ie, above normal) and were, thus, consistently higher than the sensor-based data. This participant described the substantial impact that the smartphone-based support had on her mobility and showed an increase in both mobility and activity assessment results from pre- to poststudy. Therefore, it is possible that these would be reasonable perceptions given a longer period. Alternatively, the positive impact on her mood and confidence could have exaggerated her perception of her own mobility. Other possible sources of disagreement, particularly for isolated deviations, were mistakes in using the mobile self-reports and forgetting to carry the phone.

Figure 6. Activity measures and self-reports for participant 3.

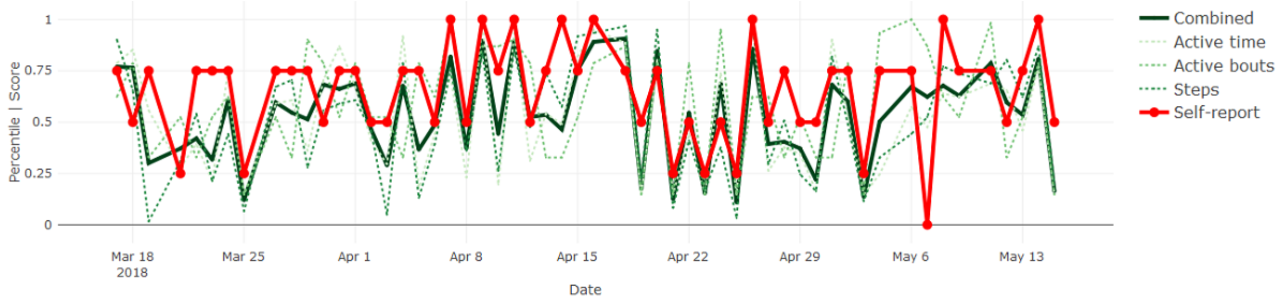


Figure 7. Activity measures and self-reports for participant 5.

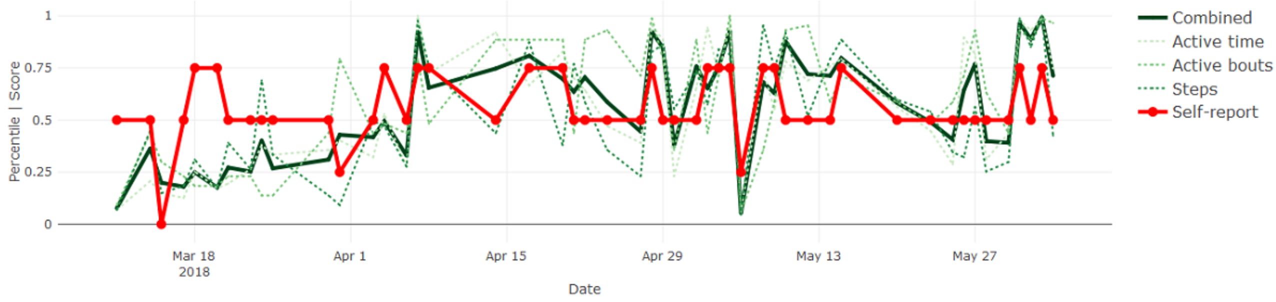


Figure 8. Mobility measures and self-reports for participant 1. MCP: minimum convex polygon.

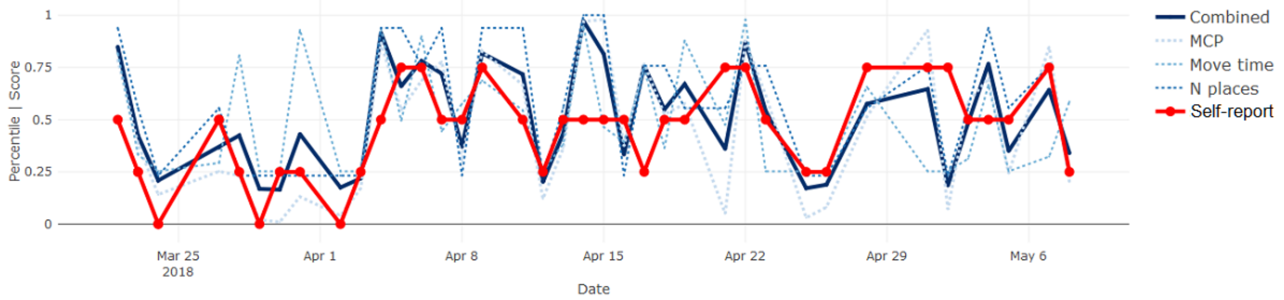
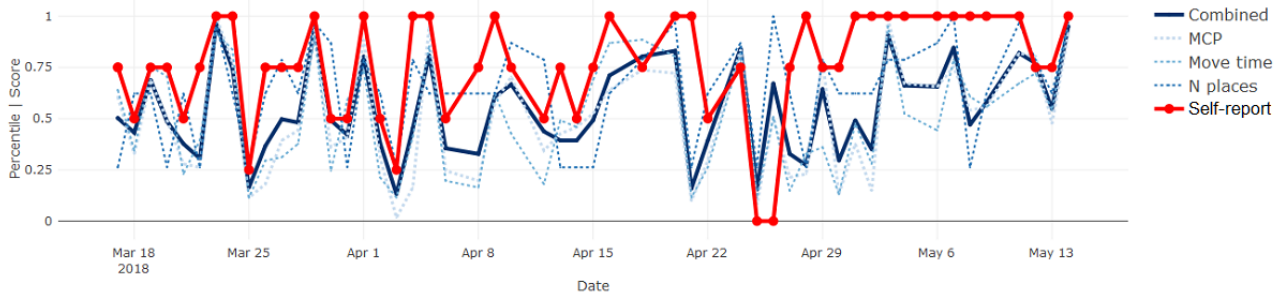


Figure 9. Mobility measures and self-reports for participant 3. MCP: minimum convex polygon.



A final observation from the results is that the combined measures appeared to agree better with participants' perceptions than each on their own. These were also potentially more robust against algorithm errors (eg, misclassification of activities or travel trajectories).

Weekly Goal Attainment

Overview

Participants' goals can be grouped into three categories:

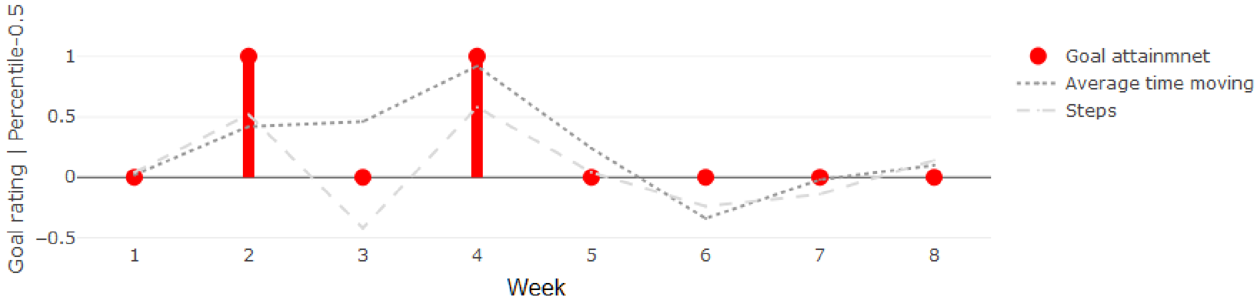
1. Target frequency of getting out of the home.
2. Increase or decrease in an activity type.
3. Adherence to an activity schedule.

Here we use examples from each category to examine whether sensor-based measures could be used to evaluate goal attainment, comparing these with participants' perceptions.

Target Frequency of Getting Out of the Home

Participants 3 and 6 followed the goal of getting out of their homes every day. This can be measured directly from the location data as the number of days per week that the participant went out at least once. Both participants went out 6-7 days per week during most of the study and each rated their goal as achieved or overachieved (score of 0 or 1) on those weeks. Participant 3 reported 1 week of underachieving the goal (score of -1), which was also the only week she went out for only 5 days, the minimum for her participation period. Participant 6 reported overachievement twice. While these weeks did not show any noticeable difference in number of days per week that he got out, they did show higher-than-usual step counts and time spent out moving between stays. This is demonstrated in Figure 10, which shows the average for these behavioral measures over each week along with goal attainment ratings.

Figure 10. Weekly goal attainment (red) and relevant average weekly measures (ie, shifted cumulative distribution function probabilities) for participant 6.



Increase and Decrease in an Activity Type

A total of 2 participants out of 6 (33%) defined goals of increasing an activity type: participant 2 aimed to walk more and participant 4 aimed to cycle more and drive less. This can be calculated from the extracted activity bouts as the amount of time spent in the relevant activity. For walking, step count was also included. Figure 11 shows the weekly goal attainment scores for participant 2 along with weekly averages for total daily steps and time spent on foot. The figure shows that the sensor-based measures fell under the median on those weeks where the participant perceived underachievement, over the median for overachievement, and close to the median for achievement. This indicates that there was agreement between the sensor-based measures and the participant's perceptions, should goal achievement correspond to normal behavior.

A similar approach was used to measure bicycle and vehicle activity for participant 4. However, when rating his goal attainment, he usually explained that he had not been riding his bike and discussed other activities instead. Around halfway through his participation, he first reported having started to ride his bicycle. This transition is evident in Table 4, which shows a marked increase in the total time spent cycling per week from the point at which he first reported cycling (week 5). The cycling time prior to week 5 was quite possibly the result of misclassifications in the activity recognition from Google, since confusion between vehicle and bicycle activities is a known issue [17].

Figure 11. Weekly goal attainment (red) and relevant average weekly measures (ie, shifted cumulative distribution function probabilities) for participant 2.

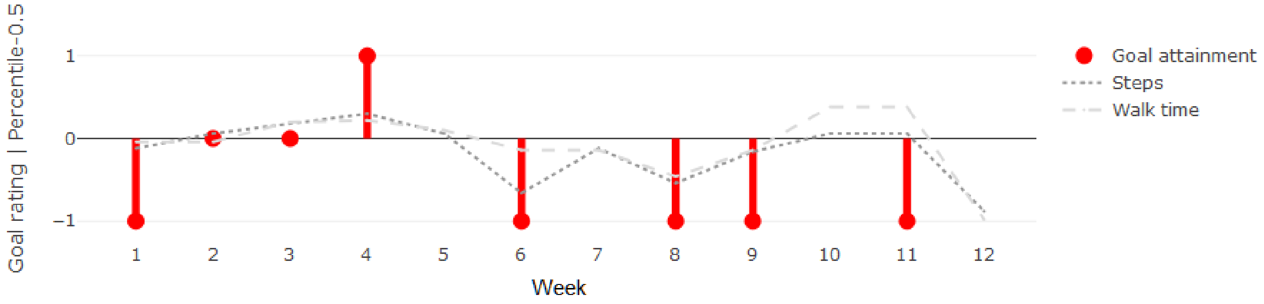


Table 4. Weekly goal attainment for participant 4.

Participant	Participant duration (days), N	Data recorded at all (days), n (%)	Data recorded by smartwatch (days), n (%)	Data recorded by self-reports (days), n (%)
1	98	98 (100)	57 (58)	39.0 (40)
2	82	81 (99)	4 (5)	59.5 (73)
3	60	60 (100)	10 (17)	53.5 (89)
4	72	72 (100)	8 (11)	39.0 (54)
5	87	86 (99)	0 (0)	49.5 (57)
6	57	54 (95)	2 (4)	46.0 (81)

[a]Missing information.

Adherence to an Activity Schedule

Activity schedule goals were detected using information provided by the relevant participants about their schedules and locations. Participant 5 trained 5 days per week at a training center, for which the location was provided. Visits to the center were detected based on distances between its location and the centroids of all stay events in the dataset. A summary of the number of days per week that she visited the training center and her reported goal attainment scores is provided in Table 5, which shows little agreement between her perceptions and sensor-based measures. The notes from weekly calls offer some explanation for the disconnect: her answers tended to be substantiated by how active she was, generally, rather than her training schedule, with reference to long walks or step counts. Furthermore, she twice rated her goal as achieved with an explanation that the training center was closed due to holidays, suggesting that since this was beyond her control and she remained active nevertheless, it was not perceived as underachievement on her part.

Generally, while in most cases a single sensor-based measurement of the goal did not correspond to participants' ratings, several related measures could sufficiently describe the behavior motivating these. This encourages the use of a multimodal approach to sensor-based monitoring. In the next analysis, we shift focus from monitoring to support from the devices for people with dementia.

Table 5. Weekly goal attainment for participant 5.

Goal information	Week							
	1	2	3	4	5	6	7	8
Reported goal attainment, score	-1	-1	0	X[a]	-1	0	0	0
Total time cycling, minutes	23	0	26	12	78	65	50	40

[a]Goal rating of 0 is *achieved*, -1 is *underachieved*, and 1 is *overachieved*.

Analysis 3: Qualitative Analysis of User Acceptance—Usefulness and Health Efficacy

Overview

Here we present results from the qualitative analysis of interview data to evaluate how useful support from the devices was and the impact this support had on quality of life.

Pre-existing support from smartphones was taken into consideration and mostly continued alongside the personalized support offered in the studies. Participants 1, 2, 3, and 4 used smartphones prior to the study and described using their smartphone daily or carrying it on them. Caregivers of participants 1 and 4 remarked on the dependence their spouse had on their phone, explaining that it is their "lifeblood" (participant 4) and how they might interrupt a meal to capture information on their phone before forgetting it (participant 1). Examples of purposes given for which the phones were used included calendar and reminders, taking and reviewing pictures, social, leisure, news and weather information apps, and to make payments.

Usefulness

Results from the analysis of poststudy interviews regarding usefulness of the personalized support offered to participants are summarized in Table 6. Interview recordings were analyzed to extract information about how the participants benefited from the support and to what extent. The same app might fulfill different purposes for different participants. For example, a location tracker was used by participant 5 to review her routes after a journey for fun (ie, leisure) and to remember where she has been (ie, memory); the same tracker was used by participant 3's caregiver to locate her (ie, safety). The picture-dialing feature, whereby participants added pictures to contact details in their phonebooks, offered memory support to participant 1, whereas participant 2 reported feeling safer when the phone rang if a picture was shown.

Table 6. Personalized support offered to participants based on individual needs.

Goal information	Week							
	1	2	6	7	8	9	10	11
Goal rating[a], score	0	1	0	0	-1	1	0	0
Training days	5	5	4	5	4	4	5	2

[a] Responses are shown with + or - to indicate positive or negative perceptions of benefit from the support, respectively; stronger positive or negative responses are indicated with ++ or --, respectively.

Responses to the smartwatch varied substantially. A total of 2 out of 6 participants (33%) did not like it at all; participant 5 explained that "it's big and it's heavy, and I must take care of it and have to take it everywhere with me," and participant 6 found the rubber strap to be uncomfortable, especially in warm weather. Other participants found the smartwatch enjoyable or convenient; participant 4 added that he did not notice it at all and preferred it to checking the time on his phone.

Overall, the support was perceived as beneficial. Two cases stand out as being particularly negative or positive:

1. Participant 1 was provided with Google Fit based on an interest in information about his activity; however, he misunderstood that he should manually enter all of his activities into Google Fit as a means of data collection in the study. This caused considerable effort on his part, led to anxiety over mistakes in recording the activities, and was challenging to clarify and resolve.
2. Participant 3's experience was overwhelmingly positive. She described her participation in the study as having "given her a new life," primarily attributed to personalized support selected to help her navigate home and for her caregiver (ie, husband) to track her location. She explains that "I have a new life... my husband is completely calm... Can you imagine, I can go anywhere! ... [with the find home feature] I have peace of mind and inner calm." Her caregiver describes no longer being scared of her getting lost and notes the impact on her: "We [family members] can indeed notice a huge difference, really. She is completely changed. She has become super positive, and has her good humor back."

Health Efficacy

Participants had varying views on the impact that the technology and broader intervention had on their everyday functioning, health, and well-being. Participants 1 and 5 perceived no benefits, both stating that they were active to begin with and were satisfied with their existing coping strategies. Participants 2, 4, and 6 all reported that they found tracking their step count to be motivating; participants 2 and 6 believed they had increased their activity as a result; for both, their pre- and postactivity assessments also showed an increase. Participant 3 perceived considerable impact on both functioning in everyday life and health status, reporting reduced anxiety, improved mood, increased activity levels, and having implemented more effective schedule management using the study devices.

To assess overall acceptance, participants were asked whether they would be interested in using the support beyond the study. Only participant 5, who had no previous experience with smartphones, expressed no interest at all; she explained that the benefits were not worth the effort and she felt satisfied with her pre-existing coping strategies. For participants who used smartphones prior to the study, interest was mostly dependent on being able to implement the setup with their own devices, with the exception of participant 3 who wished to continue with the study setup. All participants, including those without a personal interest in using the technology, felt that it could benefit other people with dementia.

Discussion

Principal Findings

This work describes 6 in-depth case studies to evaluate the feasibility of using mobile and wearable technology (ie, smartphones and smartwatches) to support rehabilitation among people with dementia. Results demonstrated the potential for data from personal devices used in a real-life context among the target population to reveal behavioral patterns, even with limited device interaction. Comparisons between objective, sensor-based measurement and participants' self-reports showed that sensor-based measures of mobility and activity reflected participants' perceived fluctuations in behavior, may reveal gradual trends not detected by the participants themselves, and provided insight into goal attainment for a range of related goals. Qualitative analysis of user acceptance indicated that personalized support offered by smart technology addressed functional, memory, safety, leisure, and psychosocial needs, where many of these depend on familiarity of the device and platform for user acceptance. Of 6 participants, 4 (67%) perceived this support to positively impact their health, mostly regarding motivation to be active, with 1 participant further describing considerable positive impact on perceived anxiety, independence, activity, and caregiver burden.

These findings demonstrate potential for smartphones and wearables to offer pervasive support that could benefit people with dementia, while also generating rich data to monitor behavior for evaluating function and engagement.

Implications for Clinical Practice

Here we discuss several implications of the findings of this research, particularly for implementation in practice.

Remote Behavioral Monitoring

This study shows high data availability from smartphones used in a real-world setting by the target population, including for low device interaction. This indicates that a sensor-based behavioral monitoring approach is not restricted to users inclined toward higher smartphone use, thus broadening the potential target group.

The data were used to derive a set of mobility and activity measures comparable to participants' own perceptions of their behavior. This result implies that the monitoring approach could replace traditional methods for gathering information about behavior (eg, questionnaires or interviews) with several advantages:

1. Scaled-up data output provides higher-resolution behavioral patterns and insight into day-to-day changes in behavior to enable more predictive and preventative care strategies.
2. Remote, passive measurement reduces burden on both the health care professional and care recipient and is, therefore, less resource-heavy than traditional approaches.
3. Data quality is not dependent on the user's ability to recall information.
4. Availability of objective information that can be visualized and shared between the care recipient and provider can facilitate collaborative care.

Goal Setting and Following

The use of goals in rehabilitation is well established [4,28]; however, this raised several challenges in this study. A total of 2 participants who were satisfied with their current lifestyle followed goals to maintain their schedule. While such a goal is relevant for people with dementia due a risk of decline in functional capacity, it was sometimes difficult for these participants to conceptualize or feel motivated by a goal to maintain rather than improve upon their status quo. Participants also generally found it difficult to recall their goal over the duration of the study, in some cases providing irrelevant information when asked about goal attainment.

This study indicates that sensor-based behavioral measures could aid in the following of goals by generating rich insights into behavior related to individual goals. This information could provide feedback regarding goal attainment to make it easier for care recipients to recall and follow goals and to motivate goal attainment.

Support in Everyday Life

Smartphones were found to be a feasible platform for offering support in everyday life. Indeed, most participants already used their own smartphones extensively as memory support, among other purposes, prior to participation. This study has further shown how leveraging support in everyday life from one's personal mobile device could be enhanced through a systematic approach (ie, identifying relevant support tools based on individual needs as in a typical rehabilitation process). This was exemplified by participant 3; though she already used her phone for support, the introduction of tools such as location tracking in the study considerably improved the impact this had on her quality of life.

Both familiarity and personalization were found to be important factors for acceptance in this study, echoing earlier work. These learnings offer some guidance for enhancing adoption and use in practice (eg, employing already-in-use devices, operating systems, and apps as far as possible, and adding to these only based on specific individual needs and preferences). Participant 5 showed the lowest acceptance, with several possible contributing factors, including familiarity (ie, no previous experience with smartphones) and an absence of perceived need, since she was satisfied with both her level of function and engagement (ie, following a highly active schedule) and her available support. One envisioned scenario might be that a decline in condition might motivate greater interest, though it is noted that learning to use a new tool would be increasingly challenging with worsening cognitive impairment, highlighting the importance of early adoption to enhance potential benefit.

Future Work: Technical Development and Clinical Research

Opportunities for future work are discussed in terms of further technical development and recommended next steps for clinical research.

Two technical limitations encountered in this study are noted as key issues to address in further development:

1. Incompatibility of the monitoring system with iPhones prevented 3 participants from being able to use their own familiar device, introducing the challenge of learning to use the Android platform.
2. Unreliable data acquisition from a connected wearable limited the availability of activity data from a worn device.

A further recommendation is the development of infrastructure for feedback and sharing of behavioral data.

The early evidence presented in this study indicates feasibility of using mobile and wearable technology for support and monitoring in rehabilitation in early-stage dementia as a means toward maintaining quality of life. Recommended next steps for further research are, therefore, to conduct larger, controlled studies to provide evidence needed to confirm and quantify the impact of the described system on quality of life outcomes, such as independence, social engagement, or psychosocial symptoms of dementia such as depression. Regarding the potential for sensor-based mobility and activity monitoring to reveal patterns in behavior, this further motivates larger-scale and longer-term studies to gather the data necessary to develop algorithms for detecting and predicting changes in condition status.

Conclusions

This work provides some of the first evidence describing the dual role of personal devices in rehabilitation for dementia by offering personalized support in everyday life and by monitoring activity and mobility behavior. Results show promise for smartphones and wearables to drive P4 approaches to dementia care. Core contributions include the following: results demonstrating that data gathered under real-life conditions are adequate for revealing behavioral patterns, initial evidence showing potential advantages of sensor-based measurement over self-reported behavior (eg, for detecting gradual trends or providing multifaceted insights into goal attainment), and qualitative reports from participants describing usefulness of the support and its impact on their health and well-being.

Acknowledgments

We would like to acknowledge the Dementia and Memory Clinic, Rigshospitalet-Glostrup, for their support with recruitment; Radu Gatej, Luna S Hansen, and Maria Özden for their assistance with implementing the case studies; the Copenhagen Centre for Health Technology (CACHET); and Videncenter for velfærdsteknologi (VihTek), all of whom helped make this research possible.

Conflicts of Interest

None declared.

References

1. The Lancet Neurology. Response to the growing dementia burden must be faster. Lancet Neurol 2018 Aug;17(8):651. [doi: 10.1016/S1474-4422(18)30256-4] [Medline: 30033050]
2. World Health Organization. Global Action Plan on the Public Health Response to Dementia: 2017-2025. Geneva, Switzerland: World Health Organization; 2017. URL: https://apps.who.int/iris/bitstream/handle/10665/259615/9789241513487-eng.pdf?sequence=1 [accessed 2019-10-08]
3. Wade DT, de Jong BA. Recent advances in rehabilitation. BMJ 2000 May 20;320(7246):1385-1388 [FREE Full text] [doi: 10.1136/bmj.320.7246.1385] [Medline: 10818031]
4. Wade D. Describing rehabilitation interventions. Clin Rehabil 2005 Dec;19(8):811-818. [doi: 10.1191/0269215505cr923ed] [Medline: 16323380]
5. Bahar-Fuchs A, Clare L, Woods B. Cognitive training and cognitive rehabilitation for mild to moderate Alzheimer's disease and vascular dementia. Cochrane Database Syst Rev 2013 Jun 05(6):CD003260. [doi: 10.1002/14651858.CD003260.pub2] [Medline: 23740535]
6. Martínez-Alcalá CI, Pliego-Pastrana P, Rosales-Lagarde A, Lopez-Noguerola JS, Molina-Trinidad EM. Information and communication technologies in the care of the elderly: Systematic review of applications aimed at patients with dementia and caregivers. JMIR Rehabil Assist Technol 2016 May 02;3(1):e6 [FREE Full text] [doi: 10.2196/rehab.5226] [Medline: 28582258]
7. Lauriks S, Reinersmann A, Van der Roest HG, Meiland FJ, Davies RJ, Moelaert F, et al. Review of ICT-based services for identified unmet needs in people with dementia. Ageing Res Rev 2007 Oct;6(3):223-246. [doi: 10.1016/j.arr.2007.07.002] [Medline: 17869590]
8. 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]
9. Maier A, Özkil A, Bang M, Hysse FB, Forchhammer B. Remember to remember: A feasibility study adapting wearable technology to the needs of people aged 65 and older with mild cognitive impairment (MCI) and Alzheimer's dementia. In: Proceedings of the 20th International Conference on Engineering Design (ICED15). 2015 Presented at: 20th International Conference on Engineering Design (ICED15); July 27-30, 2015; Milan, Italy p. 331-340 URL: https://orbit.dtu.dk/files/107110242/Remember_to_remember.pdf
10. Stucki RA, Urwyler P, Rampa L, Müri R, Mosimann UP, Nef T. A Web-based nonintrusive ambient system to measure and classify activities of daily living. J Med Internet Res 2014 Jul 21;16(7):e175 [FREE Full text] [doi: 10.2196/jmir.3465] [Medline: 25048461]
11. Evald L. Prospective memory rehabilitation using smartphones in patients with TBI. Disabil Rehabil 2018 Sep;40(19):2250-2259. [doi: 10.1080/09638288.2017.1333633] [Medline: 28589735]
12. Liddle J, Ireland D, McBride SJ, Brauer SG, Hall LM, Ding H, et al. Measuring the lifespace of people with Parkinson's disease using smartphones: Proof of principle. JMIR Mhealth Uhealth 2014 Mar 12;2(1):e13 [FREE Full text] [doi: 10.2196/mhealth.2799] [Medline: 25100206]
13. Zylstra B, Netscher G, Jacquemot J, Schaffer M, Shen G, Bowhay AD, et al. Extended, continuous measures of functional status in community dwelling persons with Alzheimer's and related dementia: Infrastructure, performance, tradeoffs, preliminary data, and promise. J Neurosci Methods 2018 Apr 15;300:59-67 [FREE Full text] [doi: 10.1016/j.jneumeth.2017.08.034] [Medline: 28865985]
14. Dicianno BE, Henderson G, Parmanto B. Design of mobile health tools to promote goal achievement in self-management tasks. JMIR Mhealth Uhealth 2017 Jul 24;5(7):e103 [FREE Full text] [doi: 10.2196/mhealth.7335] [Medline: 28739558]
15. Hood L, Balling R, Auffray C. Revolutionizing medicine in the 21st century through systems approaches. Biotechnol J 2012 Aug;7(8):992-1001 [FREE Full text] [doi: 10.1002/biot.201100306] [Medline: 22815171]
16. Sagner M, McNeil A, Puska P, Auffray C, Price ND, Hood L, et al. The P4 health spectrum: A predictive, preventive, personalized and participatory continuum for promoting healthspan. Prog Cardiovasc Dis 2017;59(5):506-521. [doi: 10.1016/j.pcad.2016.08.002] [Medline: 27546358]
17. Thorpe J, Forchhammer B, Maier A. Development of a sensor-based behavioral monitoring solution to support dementia care. JMIR Mhealth Uhealth 2019 May 30;7(6):e12013 [FREE Full text] [doi: 10.2196/12013] [Medline: 31199304]

18. Thorpe JR, Rønn-Andersen KV, Bień P, Özkil AG, Forchhammer BH, Maier AM. Pervasive assistive technology for people with dementia: A UCD case. Healthc Technol Lett 2016 Dec;3(4):297-302 [FREE Full text] [doi: 10.1049/htl.2016.0057] [Medline: 28008366]
19. Baker P, Bodner EV, Allman RM. Measuring life-space mobility in community-dwelling older adults. J Am Geriatr Soc 2003 Nov;51(11):1610-1614. [doi: 10.1046/j.1532-5415.2003.51512.x] [Medline: 14687391]
20. Bull FC, Maslin TS, Armstrong T. Global physical activity questionnaire (GPAQ): Nine country reliability and validity study. J Phys Act Health 2009 Nov;6(6):790-804. [Medline: 20101923]
21. Allet L, Knols RH, Shirato K, de Bruin ED. Wearable systems for monitoring mobility-related activities in chronic disease: A systematic review. Sensors (Basel) 2010;10(10):9026-9052 [FREE Full text] [doi: 10.3390/s101009026] [Medline: 22163393]
22. Canzian L, Musolesi M. Trajectories of depression: Unobtrusive monitoring of depressive states by means of smartphone mobility traces analysis. In: Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing. New York, NY: ACM; 2015 Presented at: ACM International Joint Conference on Pervasive and Ubiquitous Computing; September 7-11, 2015; Osaka, Japan p. 1293-1304. [doi: 10.1145/2750858.2805845]
23. Boissy P, Brière S, Hamel M, Jog M, Speechley M, Karelis A, EMAP Group. Wireless inertial measurement unit with GPS (WIMU-GPS): Wearable monitoring platform for ecological assessment of lifespace and mobility in aging and disease. Conf Proc IEEE Eng Med Biol Soc 2011;2011:5815-5819. [doi: 10.1109/IEMBS.2011.6091439] [Medline: 22255662]
24. Hirsch J, Winters M, Clarke P, McKay H. Generating GPS activity spaces that shed light upon the mobility habits of older adults: A descriptive analysis. Int J Health Geogr 2014 Dec 12;13:51 [FREE Full text] [doi: 10.1186/1476-072X-13-51] [Medline: 25495710]
25. Turner-Stokes L. Goal attainment scaling (GAS) in rehabilitation: A practical guide. Clin Rehabil 2009 Apr;23(4):362-370. [doi: 10.1177/0269215508101742] [Medline: 19179355]
26. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Q 1989 Sep;13(3):319. [doi: 10.2307/249008]
27. Bardram JE. CACHET Unified Methodology for Assessment of Clinical Feasibility. Copenhagen, Denmark: Copenhagen Center for Health Technology; 2018. URL: http://www.cachet.dk/-/media/Sites/Cachet/Layout-og-arkiv/Arkiv/cumacf/cumacf.ashx?la=da&hash=097004CD61CE0EBA12AFABAF46D79AA17290483D [accessed 2019-10-09]
28. Watermeyer TJ, Hindle JV, Roberts J, Lawrence CL, Martyr A, Lloyd-Williams H, et al. Goal setting for cognitive rehabilitation in mild to moderate Parkinson's disease dementia and dementia with Lewy bodies. Parkinsons Dis 2016:8285041-8285048 [FREE Full text] [doi: 10.1155/2016/8285041] [Medline: 27446628]

Abbreviations

ACE: Addenbrooke's Cognitive Examination
CACHET: Copenhagen Centre for Health Technology
FAQ-IADL: Functional Assessment Questionnaire-Instrumental Activities of Daily Living
GPAQ: Global Physical Activity Questionnaire
GPS: global positioning system
ICT: information and communications technology
LSA: Life-Space Assessment
MCP: minimum convex polygon
MMSE: Mini-Mental State Examination
P4: predictive, preventative, personalized, and participatory
QoL-AD: Quality of Life in Alzheimer's Disease scale
VihTek: Videncenter for velfærdsteknologi
ZBI: Zarit Burden Interview

Edited by G Eysenbach; submitted 27.09.18; peer-reviewed by B Chaudhry, M Karunanithi; comments to author 02.02.19; revised version received 07.06.19; accepted 21.07.19; published 17.10.19.

Please cite as:
Thorpe J, Forchhammer BH, Maier AM
Adapting Mobile and Wearable Technology to Provide Support and Monitoring in Rehabilitation for Dementia: Feasibility Case Series
JMIR Form Res 2019;3(4):e12346
URL: http://formative.jmir.org/2019/4/e12346/
doi:10.2196/12346
PMID:31625951

Original Paper

Passive Monitoring of Short-Acting Beta-Agonist Use via Digital Platform in Patients With Chronic Obstructive Pulmonary Disease: Quality Improvement Retrospective Analysis

Jessica Chen[1], MD; Leanne Kaye[2], PhD; Michael Tuffli[2], BA; Meredith A Barrett[2], PhD; Shelanda Jones-Ford[3], BSc; Tina Shenouda[3], MD; Rahul Gondalia[2], PhD; Kelly Henderson[2], MPH; Veronica Combs[4], BA; David Van Sickle[5], PhD; David A Stempel[2], MD

[1]ChenMed, Miami, FL, United States
[2]Propeller Health, San Francisco, CA, United States
[3]JenCare Senior Medical Center, Louisville, KY, United States
[4]Content Strategy Solutions, Louisville, KY, United States
[5]Propeller Health, Madison, WI, United States

Corresponding Author:
Leanne Kaye, PhD
Propeller Health
47 Maiden Lane
3rd Floor
San Francisco, CA
United States
Phone: 1 608 251 0470
Email: leanne.kaye@propellerhealth.com

Abstract

Background: Digital health programs assist patients with chronic obstructive pulmonary disease (COPD) to better manage their disease. Technological and adoption barriers have been perceived as a limitation.

Objective: The aim of the research was to evaluate a digital quality improvement pilot in Medicare-eligible patients with COPD.

Methods: COPD patients were enrolled in a digital platform to help manage their medications and symptoms as part of their routine clinical care. Patients were provided with electronic medication monitors (EMMs) to monitor short-acting beta-agonist (SABA) use passively and a smartphone app to track use trends and receive feedback. Providers also had access to data collected via a secure website and were sent email notifications if a patient had a significant change in their prescribed inhaler use. Providers then determined if follow-up was needed. Change in SABA use and feasibility outcomes were evaluated at 3, 6, and 12 months.

Results: A total of 190 patients enrolled in the pilot. At 3, 6, and 12 months, patients recorded significant reductions in daily and nighttime SABA use and increases in SABA-free days (all $P<.001$). Patient engagement, as measured by the ratio of daily active use to monthly active use, was >90% at both 6 and 12 months. Retention at 6 months was 81% (154/190). Providers were sent on average two email notifications per patient during the 12-month program.

Conclusions: A digital health program integrated as part of standard clinical practice was feasible and had low provider burden. The pilot demonstrated significant reduction in SABA use and increased SABA-free days among Medicare-eligible COPD patients. Further, patients readily adopted the digital platform and demonstrated strong engagement and retention rates at 6 and 12 months.

(JMIR Form Res 2019;3(4):e13286) doi:10.2196/13286

KEYWORDS

chronic obstructive pulmonary disease; telemedicine; quality improvement, feasibility; nebulizers and vaporizers; health services

Introduction

Chronic obstructive pulmonary disease (COPD) affects more than 65 million adults worldwide and is expected to become the third leading cause of death by 2030 [1]. COPD is associated with progressive loss of lung function and has significant impacts on both physical and mental well-being [2].

The use of inhaled long-acting muscarinic antagonist, long-acting beta-agonists alone or in combination with a long-acting muscarinic antagonists, or an inhaled corticosteroid–long-acting beta-agonist combination has been associated with improved symptom control and decreases in exacerbations associated with COPD [3]. Short-acting beta-agonists (SABA) are prescribed to manage acute symptoms. Knowledge of actual SABA use and its signal of increased risk may be used to adjust medication regimens and identify worsening symptoms and determine risk or occurrence of an exacerbation [4-6]. Historically, information on patient use of SABA has been self-reported, leading to challenges of recall bias through incorrect recollection or fabrication [7-9].

To address the burden and morbidity of COPD [10,11], novel approaches for disease management are warranted [12]. Web-based self-management, smartphone apps, electronic medication monitors, and text-based interventions have been used to track, automate, and provide real-time feedback on medication use and clinical status. These digital health interventions have demonstrated success for chronic respiratory diseases including asthma and COPD [13-17]; however, these programs may be limited by nonengagement and attrition [18,19].

Emerging research suggests that digital health programs integrating providers [20] or human coaches may improve clinical outcomes, rates of engagement, and retention [21,22]. However, there has been little effort to engage COPD patients in these programs due to perceived challenges with patient age and comfort with technology [18,23,24], as well as perceived increases in provider burden. This analysis aimed to determine the feasibility of integrating a digital quality improvement pilot in a clinical setting to improve outcomes in a COPD Medicare population.

Methods

Recruitment and Eligibility

Patients were recruited from three JenCare Senior Medical Center clinics in Louisville, Kentucky, from December 2015 to December 2016. JenCare clinics are managed care facilities that provide high-touch care to low- to moderate-income Medicare-eligible patients. Inclusion criteria were kept broad to include as many patients as possible in the quality improvement pilot. Patients with a physician diagnosis of COPD, using a sensor-compatible SABA medication, and fluent in English or Spanish were eligible for enrollment. An in-person enrollment process was led by clinical team members and two respiratory therapists. Eligible patients were required to accept Propeller Heath’s Terms of Use [25] prior to participation, which permits the use of aggregated, de-identified data in analyses. This analysis was determined to be exempt by the Copernicus Independent Review Board.

Study Design

Patients were provided with an electronic medication monitor (EMM; Propeller Health) that attached to their compatible SABA (Figure 1). The EMM was used to objectively monitor the date and time of SABA use. The EMM was paired via Bluetooth to a smartphone, which transmitted the information to Health Insurance Portability and Accountability Act–compliant servers. Patients without smartphones received a wireless hub to transmit EMM data passively. Lost or malfunctioning EMMs and hubs were replaced.

Figure 1. Electronic medication monitor that captures inhaler medication use.

The EMM provided is part of a US Food and Drug Administration–cleared digital health platform comprising a mobile app, Web dashboards, and short message service and email modalities [13,14,26]. Patients had access to a Web or mobile platform that promoted self-management with information about their SABA use trends, potential environmental triggers, and guidelines-based education [3]. Further, because this was a quality improvement pilot, there was no formal retention plan such that patients were not required to have a final visit or last sensor sync. No incentives were provided.

Health Care Providers and Electronic Notifications

Patients authorized their health care providers to view collected data and summary reports through a secure Web dashboard. This information could be used to inform clinical treatment such as medication adjustments or early intervention at the sign of increasing SABA use. In addition, at-risk notifications were sent to providers if a patient was considered at increased risk for an exacerbation (defined as use of ≥10 SABA puffs in a 24-hour period or SABA use above the patient's personal baseline level for two or more consecutive days). Based on these notifications, providers could determine appropriate follow-up with a patient via telephone or an in-person visit as needed. A respiratory therapist assisted patients with technical questions by phone and in person.

Feasibility Outcomes

We evaluated patient retention at 3, 6, and 12 months (defined as an EMM sync with at least 3, 6, and 12-month data points, respectively). Syncs demonstrate that the EMM is capable of collecting and transmitting data via a smartphone or hub. While syncs do not require actuation of the medication, they do require that the patient maintain the digital platform. For example, patients must ensure the hub remains charged and connected to an outlet and that the EMM is transferred to any refilled inhaler prescriptions. As such, to evaluate patient engagement, we examined the proportion of days that a sync was recorded at 3, 6, and 12 months. We also compared the percentage of daily active users with the percentage of monthly active users to determine level of daily engagement. Finally, we assessed provider notification burden by averaging the number of at-risk notifications sent to the provider per patient over the 12-month program.

Short-Acting Beta-Agonist Use

Change in mean SABA use and percentage of SABA-free days was assessed at 3, 6, and 12 months. The date, time, and number of SABA puffs were recorded as mean puffs per day. Nighttime SABA use was calculated as a subcategory of daily use, defined as use between 10:00 pm and 6:00 am. Percentage of SABA-free days were calculated as the percentage of days (unique 24-hour periods) in which a patient did not use their SABA inhaler.

Statistical Analyses

Descriptive statistics were completed for the feasibility outcomes at 3, 6, and 12 months. SABA use outcome analyses compared the change in the mean number of SABA uses per day, mean number of nighttime SABA uses, and percentage of SABA-free days from baseline to specified end points. The baseline period was defined as week 1 (days 1-7), with day 1 as the first EMM sync. We evaluated outcomes at baseline and compared against three end points: 3 months (week 13), 6 months (week 26), and 12 months (week 52) among those patients with sufficient follow-up. To account for potential biases due to attrition during the 12-month study period, we calculated stabilized inverse probability of attrition weights for each participant with 3 months of data, conditional on age, race, gender, season of enrollment, smartphone or hub use, baseline SABA use, and syncing using logistic regression [27]. We then estimated changes in SABA use parameters from baseline to 3, 6, and 9 months using attrition-weighted longitudinal, mixed-effects linear models adjusting for age, race, gender, and season of enrollment. All analyses were conducted in R version 3.4 (R Foundation for Statistical Computing).

Results

Population

A total of 190 patients enrolled in the digital quality improvement pilot. Patients were primarily white (99/190, 52.1%) or African American (85/190, 44.7%), female (124/190, 65.3%), and over age 60 years (153/190, 80.6%) with a mean age of 68.0 (SD 9.2) years (Table 1). Almost all patients transmitted data using the wireless hub device (184/190, 96.8%).

Table 1. Patient demographics (n=190).

Characteristics	n (%)
Age in years	
40-49	1 (0.5)
50-59	36 (18.9)
60-69	67 (35.3)
70 and older	86 (45.3)
Gender, female	124 (65.3)
Race	
White	99 (52.1)
African American	85 (44.7)
Other	4 (2.1)
Unknown	2 (1.1)
Device type	
Hub	184 (96.8)
Smartphone	6 (3.2)

Feasibility Outcomes

Patient retention was 90.5% (172/190) at 3 months, 81.0% (154/190) at 6 months, and 63.1% (120/190) at 12 months (Table 2). There were no significant differences in age, race, gender, baseline SABA use, or sync history between those patients who completed 12 months and those who did not (Table 3). The percentage of daily active users to monthly active users remained above 90% for all three end points. Similarly, the mean proportion of days patients synced at 3, 6, and 12 months was 91%, 90%, and 88%, respectively. Providers were sent 397 at-risk notifications over 12 months, a mean of 2.1 notifications per patient.

Table 2. Feasibility and process outcomes.

Characteristics	Value
Patient retention, n (%)	
3 months	172 (90.5)
6 months	154 (81.0)
12 months	120 (63.1)
Patient engagement (proportion of days with at least one sync), %	
3 months	91
6 months	90
12 months	88
Daily active user/monthly active user	90.5
Provider at-risk alerts, n (mean per patient)	397 (2.1)

Table 3. Comparison of baseline characteristics between patients with complete and noncomplete follow-up at 12 months.

Characteristic	Patients who completed 12 months (n=120)	Patients who did not complete 12 months (n=70)	*P* value
Age, n (%)	—	—	.38
40-49	1 (1)	0 (0)	—
50-59	18 (15)	17 (24)	—
60-69	44 (37)	24 (34)	—
70 and older	57 (48)	29 (41)	—
Race, n (%)	—	—	.35
White	61 (51)	38 (54)	—
African American	56 (47)	29 (41)	—
Other	3 (3)	1 (1)	—
Unknown	0 (0)	2 (3)	—
Female, n (%)	82 (68)	42 (60)	.31
SABA[a] at baseline, puffs/day, mean (SD)	3.1 (2.3)	4.0 (1.6)	.18
Days with at least one sync at baseline, %	94	88	.09

[a] SABA: short-acting beta-agonist.

Change in Short-Acting Beta-Agonist Use

In crude analyses, decreases in mean SABA use and nighttime SABA use and increases in SABA-free days were observed over the study period in patients with 3, 6, and 12 months of follow-up (Multimedia Appendix 1). After accounting for potential biases due to attrition and adjusting for potential confounders, mean SABA use, nighttime SABA use, and SABA-free days were 3.2 puffs per day, 1.1 puffs per day, and 33%, respectively. SABA use decreased on average by 1.4 (95% CI –1.6 to –1.2), 1.6 (95% CI –1.7 to–1.2), and 1.9 (95% CI –2.1 to –1.7) puffs per day, from baseline to 3, 6, and 12 months, respectively (Figure 2). Similarly, nighttime SABA use decreased by 0.5 (95% CI –0.6 to –0.4), 0.5 (95% CI –0.7 to –0.4), and 0.8 (95% CI –0.9 to –0.7) and the weekly proportion of SABA-free days increased by 25% (95% CI 22 to 28), 31% (95% CI 27 to 34), and 36% (95% CI 33 to 39) points from baseline to 3, 6, and 12 months, respectively (Figures 3 and 4).

Figure 2. Attrition-weighted and adjusted mean daily short-acting beta-agonist use at 3, 6, and 12 months.

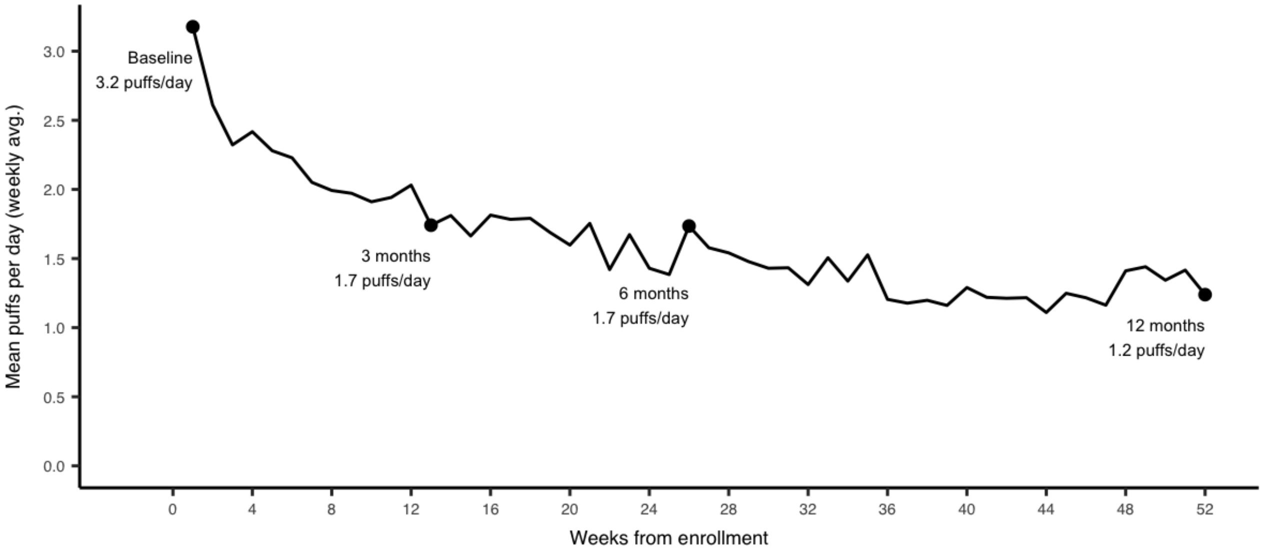


Figure 3. Attrition-weighted and adjusted nighttime short-acting beta-agonist use at 3, 6, and 12 months.

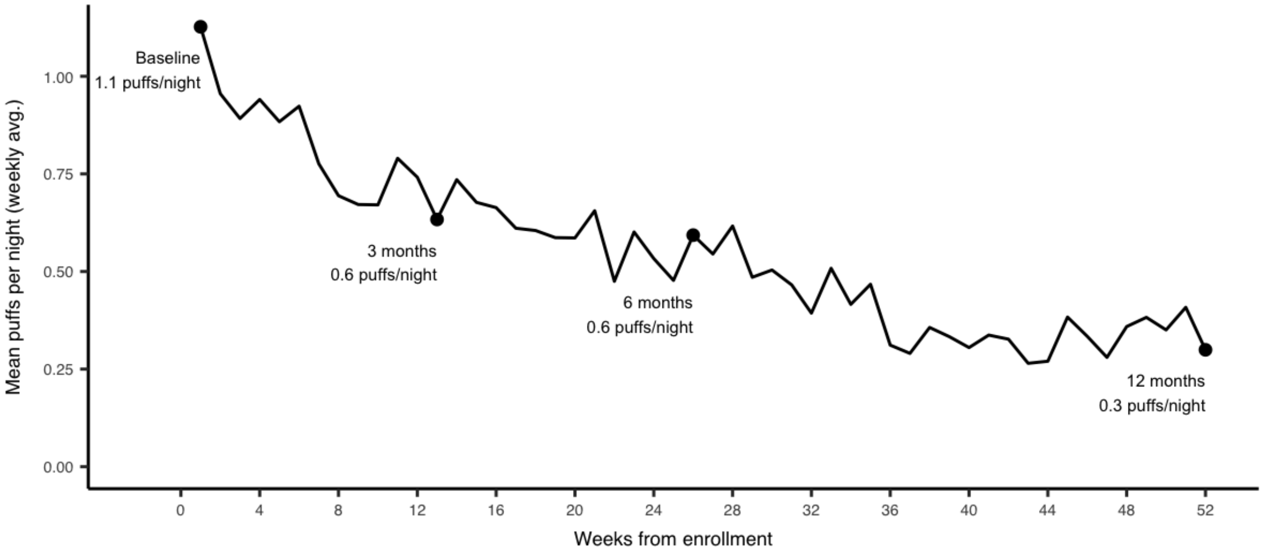
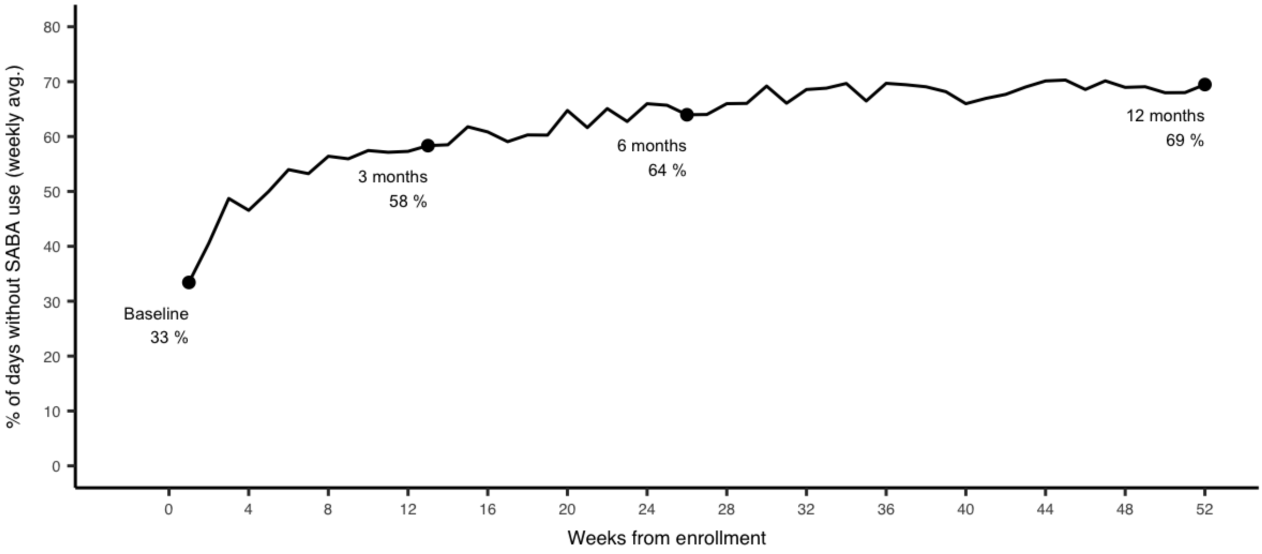


Figure 4. Attrition-weighted and adjusted percent of short-acting beta-agonist–free days at 3, 6, and 12 months.



Discussion

Principal Findings

This study evaluated the feasibility of a clinically integrated digital quality improvement pilot in a Medicare-eligible COPD population. Patients demonstrated significant improvements from baseline to 3, 6, and 12 months for reduced daily and nighttime SABA use and an increased percentage of SABA-free days. Further, this study demonstrated that a digital health program was feasible in this COPD population.

Providers had access to the patient's clinical status and real-time use of SABA. Providers received at-risk notifications at the patient level to alert them to increased SABA use and provide the opportunity for earlier intervention. While we did not formally record changes to medication regimens, other studies have shown that real-time identification of increased SABA use enables more timely provider intervention and may result in lower exacerbation rates and acute health care utilization including hospitalizations and emergency department visits [4,5]. Additional randomized controlled studies are needed to confirm this effect.

Comparison With Prior Work

Previous studies have demonstrated that engagement rates for digital health applications may peak around 3 months and decline rapidly thereafter [28,29], with approximately 10% of participants engaging at 6 months [30]. In this study, we saw 81% retention and 90% of days with a sync at 6 months. The three recruitment centers are characterized by high-touch, engaged providers. It is possible that the clinical improvement and engagement in this study may be explained in part by the frequency and quality of the provider engagement. For example, a previous meta-analysis exploring the impact of provider communication on patient engagement found a positive correlation between provider feedback and medication adherence rates [31]. Additional research shows that active provider engagement may improve retention rates [32,33]. Patient-provider data sharing and transparency may also improve outcomes and engagement through shared decision-making [15,34].

Digital health interventions may also increase provider burden [23]. However, this analysis suggests that passively tracking medication use and integrating this information into a comprehensive clinical care model may result in stronger patient engagement without excessive provider burden. For example, only 2.1 at-risk notifications were sent on average per patient throughout the 12-month study period. Future studies should address provider burden and satisfaction in a more robust manner.

EMMs may reduce patient burden. Measures of medication use in clinical studies frequently rely on self-report, placing the burden on patients and increasing the likelihood of recall bias [28]. By objectively and passively collecting medication use data and transmitting it to the provider, both patient burden and reporting accuracy may be improved.

Limitations

This study is limited, in part, by its design as a single-arm study and regression to the mean for SABA use should be considered. Further, controller medication use was not tracked due to a noncompatible formulary, limiting the opportunity for controller medication adherence promotion and possible impact on SABA use. It is also possible that patients used additional SABA inhalers without attaching an EMM. Measurement of engagement and clinical outcomes was limited and future programs should consider more definitive measures a priori. Finally, factors beyond the EMMs and platform may have contributed to the observed outcomes, and future randomized studies are needed to confirm the findings.

Conclusion

This quality improvement pilot demonstrates that a digital health program with integrated clinical care is feasible in a Medicare-eligible population of COPD patients. The results demonstrated strong retention rates at 12 months, with reduced daily and nighttime SABA use and more SABA-free days. With the increasing medical resources being used for COPD treatment [10], novel approaches for disease management are warranted [12].

Acknowledgments

Special thanks to the JenCare providers who participated: Drs Ali, Anderson, Gross, Kiriakos, and Mahin, as well as Shannon Goeldi, Melissa Williams, and Jennifer Morgan of Propeller Health for assistance with patient enrollment and outreach. We would also like to acknowledge the broad network of local partners that made the broader AIR Louisville program possible and thank the City of Louisville and all participants for supporting the program. Additional thanks to the Institute for Healthy Air, Water, and Soil and the Community Foundation of Louisville. We thank the Robert Wood Johnson Foundation for its generous funding (grant number #71592) and guidance. The views expressed here do not necessarily reflect the views of the Foundation.

Conflicts of Interest

JC receives salary and equity as an employee of ChenMed. SJ-F and TS are employees of JenCare. LK, MT, MB, RG, KH, DVS, and DS receive salary as employees of Propeller Health. DVS and MB have patents pending related to their work at Propeller Health, but not directly related to this manuscript. VC declares no conflict of interest.

Multimedia Appendix 1
Changes in short-acting beta-agonist use.
[PDF File (Adobe PDF File), 121 KB - formative_v3i4e13286_app1.pdf]

References

1. World Health Organization. Chronic respiratory diseases: burden of COPD URL: http://www.who.int/respiratory/copd/burden/en [accessed 2019-09-16]
2. Putcha N, Drummond MB, Wise RA, Hansel NN. Comorbidities and chronic obstructive pulmonary disease: prevalence, influence on outcomes, and management. Semin Respir Crit Care Med 2015 Aug;36(4):575-591 [FREE Full text] [doi: 10.1055/s-0035-1556063] [Medline: 26238643]
3. Vogelmeier CF, Criner GJ, Martinez FJ, Anzueto A, Barnes PJ, Bourbeau J, et al. Global strategy for the diagnosis, management, and prevention of chronic obstructive lung disease 2017 Report. GOLD Executive Summary. Am J Respir Crit Care Med 2017 Mar 01;195(5):557-582. [doi: 10.1164/rccm.201701-0218PP] [Medline: 28128970]
4. Fan VS, Gylys-Colwell I, Locke E, Sumino K, Nguyen HQ, Thomas RM, et al. Overuse of short-acting beta-agonist bronchodilators in COPD during periods of clinical stability. Respir Med 2016 Dec;116:100-106 [FREE Full text] [doi: 10.1016/j.rmed.2016.05.011] [Medline: 27296828]
5. Sumino K, Locke ER, Magzamen S, Gylys-Colwell I, Humblet O, Nguyen HQ, et al. Use of a remote inhaler monitoring device to measure change in inhaler use with chronic obstructive pulmonary disease exacerbations. J Aerosol Med Pulm Drug Deliv 2018 Jun;31(3):191-198. [doi: 10.1089/jamp.2017.1383] [Medline: 29035120]

6. Sharafkhaneh A, Altan AE, Colice GL, Hanania NA, Donohue JF, Kurlander JL, et al. A simple rule to identify patients with chronic obstructive pulmonary disease who may need treatment reevaluation. Respir Med 2014 Sep;108(9):1310-1320 [FREE Full text] [doi: 10.1016/j.rmed.2014.07.002] [Medline: 25130680]
7. Bourbeau J, Bartlett SJ. Patient adherence in COPD. Thorax 2008 Sep;63(9):831-838 [FREE Full text] [doi: 10.1136/thx.2007.086041] [Medline: 18728206]
8. Rand C. "I took the medicine like you told me, doctor": self-report of adherence with medical regimens. In: Stone A, editor. The Science of Self-Report: Implications for Research and Practice. Mahwah: Lawrence Erlbaum Associates Publishers; 2000:257-276.
9. DiMatteo MR. Variations in patients' adherence to medical recommendations: a quantitative review of 50 years of research. Med Care 2004 Mar;42(3):200-209. [Medline: 15076819]
10. Khakban A, Sin DD, FitzGerald JM, McManus BM, Ng R, Hollander Z, et al. The projected epidemic of chronic obstructive pulmonary disease hospitalizations over the next 15 years. a population-based perspective. Am J Respir Crit Care Med 2017 Dec 01;195(3):287-291. [doi: 10.1164/rccm.201606-1162PP] [Medline: 27626508]
11. Jinjuvadia C, Jinjuvadia R, Mandapakala C, Durairajan N, Liangpunsakul S, Soubani AO. Trends in outcomes, financial burden, and mortality for acute exacerbation of chronic obstructive pulmonary disease (COPD) in the United States from 2002 to 2010. COPD 2017 Dec;14(1):72-79 [FREE Full text] [doi: 10.1080/15412555.2016.1199669] [Medline: 27419254]
12. Ford ES, Murphy LB, Khavjou O, Giles WH, Holt JB, Croft JB. Total and state-specific medical and absenteeism costs of COPD among adults aged ≥ 18 years in the United States for 2010 and projections through 2020. Chest 2015 Jan;147(1):31-45. [doi: 10.1378/chest.14-0972] [Medline: 25058738]
13. Barrett MA, Humblet O, Marcus JE, Henderson K, Smith T, Eid N, et al. Effect of a mobile health, sensor-driven asthma management platform on asthma control. Ann Allergy Asthma Immunol 2017 Dec;119(5):415-421. [doi: 10.1016/j.anai.2017.08.002] [Medline: 29150069]
14. Merchant RK, Inamdar R, Quade RC. Effectiveness of population health management using the propeller health asthma platform: a randomized clinical trial. J Allergy Clin Immunol Pract 2016;4(3):455-463. [doi: 10.1016/j.jaip.2015.11.022] [Medline: 26778246]
15. Charles C, Whelan T, Gafni A. What do we mean by partnership in making decisions about treatment? BMJ 1999 Sep 18;319(7212):780-782 [FREE Full text] [Medline: 10488014]
16. Foster JM, Usherwood T, Smith L, Sawyer SM, Xuan W, Rand CS, et al. Inhaler reminders improve adherence with controller treatment in primary care patients with asthma. J Allergy Clin Immunol 2014 Dec;134(6):1260-1268. [doi: 10.1016/j.jaci.2014.05.041] [Medline: 25062783]
17. McKibben S, De Simoni A, Bush A, Thomas M, Griffiths C. The use of electronic alerts in primary care computer systems to identify the excessive prescription of short-acting beta-agonists for people with asthma: a systematic review. NPJ Prim Care Respir Med 2018 Apr 16;28(1):14 [FREE Full text] [doi: 10.1038/s41533-018-0080-z] [Medline: 29662064]
18. Tabak M, Brusse-Keizer M, van der Valk P, Hermens H, Vollenbroek-Hutten M. A telehealth program for self-management of COPD exacerbations and promotion of an active lifestyle: a pilot randomized controlled trial. Int J Chron Obstruct Pulmon Dis 2014;9:935-944 [FREE Full text] [doi: 10.2147/COPD.S60179] [Medline: 25246781]
19. Greenhalgh T, Wherton J, Papoutsi C, Lynch J, Hughes G, A'Court C, et al. Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. J Med Internet Res 2017 Nov 01;19(11):e367 [FREE Full text] [doi: 10.2196/jmir.8775] [Medline: 29092808]
20. Merchant R, Szefler SJ, Bender BG, Tuffli M, Barrett MA, Gondalia R, et al. Impact of a digital health intervention on asthma resource utilization. World Allergy Organ J 2018;11(1):28 [FREE Full text] [doi: 10.1186/s40413-018-0209-0] [Medline: 30524644]
21. Dennison L, Morrison L, Lloyd S, Phillips D, Stuart B, Williams S, et al. Does brief telephone support improve engagement with a web-based weight management intervention? Randomized controlled trial. J Med Internet Res 2014;16(3):e95 [FREE Full text] [doi: 10.2196/jmir.3199] [Medline: 24681761]
22. Yardley L, Ware LJ, Smith ER, Williams S, Bradbury KJ, Arden-Close EJ, et al. Randomised controlled feasibility trial of a web-based weight management intervention with nurse support for obese patients in primary care. Int J Behav Nutr Phys Act 2014;11:67 [FREE Full text] [doi: 10.1186/1479-5868-11-67] [Medline: 24886516]
23. Howard S, Lang A, Sharples S, Shaw D. What are the pros and cons of electronically monitoring inhaler use in asthma? A multistakeholder perspective. BMJ Open Respir Res 2016;3(1):e000159 [FREE Full text] [doi: 10.1136/bmjresp-2016-000159] [Medline: 27933181]
24. Anderson M, Perrin A. Tech adoption climbs among older adults. Washington: Pew Internet and American Life Project; 2017. URL: https://www.pewinternet.org/wp-content/uploads/sites/9/2017/05/PI_2017.05.17_Older-Americans-Tech_FINAL.pdf [accessed 2019-09-16]
25. Propeller Health: User Agreement. 2018. URL: https://my.propellerhealth.com/terms-of-service [accessed 2019-01-02]
26. Van Sickle D, Magzamen S, Truelove S, Morrison T. Remote monitoring of inhaled bronchodilator use and weekly feedback about asthma management: an open-group, short-term pilot study of the impact on asthma control. PLoS One 2013;8(2):e55335 [FREE Full text] [doi: 10.1371/journal.pone.0055335] [Medline: 23460785]

27. Howe CJ, Cole SR, Lau B, Napravnik S, Eron JJ. Selection bias due to loss to follow up in cohort studies. Epidemiology 2016;27(1):91-97. [doi: 10.1097/ede.0000000000000409]
28. Birnbaum F, Lewis D, Rosen RK, Ranney ML. Patient engagement and the design of digital health. Acad Emerg Med 2015 Jun;22(6):754-756 [FREE Full text] [doi: 10.1111/acem.12692] [Medline: 25997375]
29. Farago P. App engagement: the matrix reloaded. 2012 Oct 22. URL: http://flurrymobile.tumblr.com/post/113379517625/app-engagement-the-matrix-reloaded [accessed 2019-09-16]
30. Dorsey ER, Yvonne CY, McConnell MV, Shaw SY, Trister AD, Friend SH. The use of smartphones for health research. Acad Med 2017 Feb;92(2):157-160. [doi: 10.1097/ACM.0000000000001205] [Medline: 27119325]
31. Zolnierek KBH, Dimatteo MR. Physician communication and patient adherence to treatment: a meta-analysis. Med Care 2009 Aug;47(8):826-834 [FREE Full text] [doi: 10.1097/MLR.0b013e31819a5acc] [Medline: 19584762]
32. Bryant J, McDonald VM, Boyes A, Sanson-Fisher R, Paul C, Melville J. Improving medication adherence in chronic obstructive pulmonary disease: a systematic review. Respir Res 2013 Oct 20;14:109 [FREE Full text] [doi: 10.1186/1465-9921-14-109] [Medline: 24138097]
33. Farmer A, Williams V, Velardo C, Shah SA, Yu L, Rutter H, et al. Self-management support using a digital health system compared with usual care for chronic obstructive pulmonary disease: randomized controlled trial. J Med Internet Res 2017 May 03;19(5):e144 [FREE Full text] [doi: 10.2196/jmir.7116] [Medline: 28468749]
34. Tapp H, Shade L, Mahabaleshwarkar R, Taylor YJ, Ludden T, Dulin MF. Results from a pragmatic prospective cohort study: shared decision making improves outcomes for children with asthma. J Asthma 2017 May;54(4):392-402. [doi: 10.1080/02770903.2016.1227333] [Medline: 27813670]

Abbreviations

COPD: chronic obstructive pulmonary disease
EMM: electronic medication monitor
SABA: short-acting beta-agonist

Edited by G Eysenbach; submitted 02.01.19; peer-reviewed by M Eakin, T Clavier, B McKinstry; comments to author 25.03.19; revised version received 26.07.19; accepted 19.08.19; published 23.10.19.

Please cite as:
Chen J, Kaye L, Tuffli M, Barrett MA, Jones-Ford S, Shenouda T, Gondalia R, Henderson K, Combs V, Van Sickle D, Stempel DA
Passive Monitoring of Short-Acting Beta-Agonist Use via Digital Platform in Patients With Chronic Obstructive Pulmonary Disease: Quality Improvement Retrospective Analysis
JMIR Form Res 2019;3(4):e13286
URL: http://formative.jmir.org/2019/4/e13286/
doi:10.2196/13286
PMID:31647471

Original Paper

Internet-Based Cognitive Behavioral Therapy via Videoconference for Patients With Bulimia Nervosa and Binge-Eating Disorder: Pilot Prospective Single-Arm Feasibility Trial

Sayo Hamatani[1,2], PhD; Noriko Numata[1], PhD; Kazuki Matsumoto[1], PhD; Chihiro Sutoh[3], MD, PhD; Hanae Ibuki[1], MA; Keiko Oshiro[1], PhD; Mari Tanaka[3], MMSc; Rikukage Setsu[3,4], MD, PhD; Yohei Kawasaki[5], PhD; Yoshiyuki Hirano[1], PhD; Eiji Shimizu[1,3,6], MD, PhD

[1]Research Center for Child Mental Development, Chiba University, Chiba, Japan
[2]Japan Society for the Promotion of Science, Tokyo, Japan
[3]Department of Cognitive Behavioral Physiology, Graduate School of Medicine, Chiba University, Chiba, Japan
[4]Koutokukai Sato Hospital, Yamagata, Japan
[5]Biostatistics Section, Clinical Research Center, Chiba University Hospital, Chiba, Japan
[6]Cognitive Behavioral Therapy Center, Chiba University Hospital, Chiba, Japan

Corresponding Author:
Noriko Numata, PhD
Research Center for Child Mental Development
Chiba University
1-8-1, Chuo-ku
Chiba
Japan
Phone: 81 43 226 207
Email: n_numata@chiba-u.jp

Abstract

Background: A major problem in providing mental health services is the lack of access to treatment, especially in remote areas. Thus far, no clinical studies have demonstrated the feasibility of internet-based cognitive behavioral therapy (ICBT) with real-time therapist support via videoconference for bulimia nervosa and binge-eating disorder in Japan.

Objective: The goal of the research was to evaluate the feasibility of ICBT via videoconference for patients with bulimia nervosa or binge-eating disorder.

Methods: Seven Japanese subjects (mean age 31.9 [SD 7.9] years) with bulimia nervosa and binge-eating disorder received 16 weekly sessions of individualized ICBT via videoconference with real-time therapist support. Treatment included CBT tailored specifically to the presenting diagnosis. The primary outcome was a reduction in the Eating Disorder Examination Edition 16.0D (EDE 16D) for bulimia nervosa and binge-eating disorder: the combined objective binge and purging episodes, objective binge episodes, and purging episodes. The secondary outcomes were the Eating Disorders Examination Questionnaire, Bulimic Investigatory Test, Edinburgh, body mass index for eating symptoms, Motivational Ruler for motivation to change, EuroQol-5 Dimension for quality of life, 9-item Patient Health Questionnaire for depression, 7-item Generalized Anxiety Disorder scale for anxiety, and Working Alliance Inventory–Short Form (WAI-SF). All outcomes were assessed at week 1 (baseline) and weeks 8 (midintervention) and 16 (postintervention) during therapy. Patients were asked about adverse events at each session. For the primary analysis, treatment-related changes were assessed by comparing participant scores and 95% confidence intervals using the paired *t* test.

Results: Although the mean combined objective binge and purging episodes improved from 47.60 to 13.60 (71% reduction) and showed a medium effect size (Cohen *d*=–0.76), there was no significant reduction in the combined episodes (EDE 16D –41; 95% CI –2.089 to 0.576; *P*=.17). There were no significant treatment-related changes in secondary outcomes. The WAI-SF scores remained consistently high (64.8 to 66.0) during treatment.

Conclusions: ICBT via videoconference is feasible in Japanese patients with bulimia nervosa and binge-eating disorder.

Trial Registration: UMIN Clinical Trials Registry UMIN000029426; https://upload.umin.ac.jp/cgi-open-bin/ctr_e/ctr_view.cgi?recptno=R000033419

(JMIR Form Res 2019;3(4):e15738) doi:10.2196/15738

KEYWORDS

bulimia nervosa; binge-eating disorder; cognitive behavioral therapy; internet-based cognitive behavioral therapy; videoconference

Introduction

Eating disorders characterized by hyperphagia include bulimia nervosa and binge-eating disorder [1]. The prevalence of bulimia nervosa in Japan is estimated to be 1.9% to 2.9% [2]. Bulimia nervosa is a psychiatric disorder characterized by repeated overeating and inappropriate compensatory behavior [1]. Like bulimia nervosa, binge-eating disorder is characterized by repeated episodes of hyperphagia but without the repetitive inappropriate compensatory behavior seen in bulimia nervosa [1]. Several meta-analyses of randomized controlled trials (RCTs) that included wait list controls have reported that cognitive behavioral therapy (CBT) is effective for bulimia nervosa and binge-eating disorder [3,4-7]. Our research group has recently conducted a single-arm study that confirmed the effectiveness of guided self-help CBT for Japanese patients with bulimia nervosa [8]. Although previous studies have shown the effectiveness of CBT for bulimia nervosa and binge-eating disorder, the patients who reside in rural areas or far from specialized treatment institutions cannot receive evidence-based therapy such as CBT. Comorbidities such as panic disorder and agoraphobia make it difficult for patients to go out, and patients with severe depression are easily tired, making it difficult for patients to attend CBT sessions once a week. Therefore, it is necessary to devise methods to provide the approved treatment at home for all patients.

Recently, CBT for patients with eating disorders has been increasingly provided via the internet [9-13]. Most internet-based CBT (ICBT) clinical trials provided to patients are guided by the therapist via email, but not in real time. A systematic review of 8 studies in patients with bulimia nervosa, binge-eating disorder, or eating disorder not otherwise specified (EDNOS) suggested that ICBT without videoconference has the effect of reducing the frequency of overeating [14]. Further, a systematic review including 22 RCTs, 2 controlled studies, and 16 uncontrolled studies found evidence for the efficacy of guided ICBT for bulimia nervosa patients; however, only one RCT included ICBT via videoconference [15]. Hence, more research is needed to determine the efficacy and limitations of ICBT via videoconference. There have been 2 case series and one clinical trial of ICBT via videoconference reported specifically in patients with eating disorders [16-18]. A case series with ICBT via videoconference was performed on two patients with bulimia nervosa in the United States [16]. A case series of ICBT via videoconference was conducted in Scotland and included 12 patients: bulimia nervosa (n=5), anorexia nervosa (n=1), and EDNOS (n=6, including 3 with binge-eating disorder) [17]. A total of 67% of these patients rated ICBT via videoconference favorably because of the convenience of receiving treatment in their area, feeling more in control, and feeling less intimidated than in a face-to-face session. An RCT was performed in the United States and included 128 patients with bulimia nervosa and EDNOS who were assigned to face-to-face CBT or ICBT via videoconference [18]. The posttreatment symptom rating scale score and symptomatic improvement showed no significant difference between ICBT via videoconference and face-to-face CBT.

The evidence for the effectiveness of ICBT via videoconference for bulimia nervosa and binge-eating disorder appears promising but is limited and has only been demonstrated in Western populations. No clinical trials of ICBT via videoconference have been conducted in non-Western populations, including in Japan. The aim of this study was to evaluate the feasibility of ICBT via videoconference for Japanese patients with bulimia nervosa and binge-eating disorder by evaluating the posttreatment change in symptom evaluation scores.

Methods

Study Design

This prospective single-arm open trial was performed in the outpatient clinic at the Cognitive Behavioral Therapy Center of Chiba University Hospital between October 2017 and March 2019. Because of the deadline for grant support, we could not extend the recruitment period. Since this trial was the first to employ an individual ICBT via videoconference intervention design against bulimia nervosa and binge-eating disorder in Japan, a single-arm trial examining feasibility rather than effectiveness was considered to be an appropriate design.

Ethics Statement

The study was approved by the institutional review board of Chiba University Hospital (reference number G29028) and registered at University Hospital Medical Information Network Clinical Trials Registry [UMIN000029426]. Written informed consent was obtained from all patients who participated in the study after they had been fully informed about the study protocol.

Study Participants

Patients were enrolled in the study if they met the following inclusion criteria: assessment by a senior psychiatrist at Chiba University Hospital using the Mini International Neuropsychiatric Interview [19,20] and a primary diagnosis of bulimia nervosa or binge-eating disorder according to the *Diagnostic and Statistical Manual of Mental Disorders, 5th Edition* (DSM-5) criteria, females aged 16 to 65 years; BMI >17.5 kg/m^2, and access to the internet at home. A comorbid diagnosis was accepted if it was clearly secondary to bulimia nervosa and binge-eating disorder. Exclusion criteria involved those who were expected to interrupt CBT due to the following comorbidities: organic brain damage, dementia, psychotic disorders, bipolar disorder, current high risk of suicide, substance abuse or dependence, antisocial behaviors, and critical physical disease. In principle, no drug changes were allowed during the trial.

Cognitive Behavioral Therapists and Quality Control

A total of 16 one-on-one sessions were conducted between therapists and patients. The therapists were three clinical psychologists and one nurse; all therapists were certified public psychologists. All therapists had completed a CBT training course (ie, Chiba–Improving Access to Psychological Therapies project) [21].

Primary Outcome

The primary outcome was the change from week 1 (baseline) to week 16 (postintervention) in frequency of objective binge and purging episodes assessed by the Eating Disorder Examination Edition 16.0D (EDE 16D) [22,23]. It contains subscales that reflect the frequency of episodes of a particular behavior, frequency of days on which the behavior occurred, and psychopathologic severity of the eating disorder. In this study, we used only the frequency of episodes of behavior, which is an important feature of eating disorders. The frequency of objective binge episodes over 28 days, frequency of purging episodes (vomiting, laxative abuse, and diuretic abuse), and frequency of combined objective binge and purging episodes were calculated. Each subject was assessed at week 1 (baseline) and week 16 (postintervention).

Secondary Outcomes

Each patient was also assessed using the Eating Disorders Examination Questionnaire (EDE-Q) [23,24], Bulimic Investigatory Test, Edinburgh (BITE) [25,26], BMI, Patient Health Questionnaire (PHQ-9) [27,28], Generalized Anxiety Disorder scale (GAD-7) [29,30], EuroQol-5 Dimensions (EQ-5D-5L) [31,32], Working Alliance Inventory-Short Form (WAI-SF) [33], and Motivational Ruler [34] at weeks 1 (baseline), 8 (midintervention), and 16 (postintervention).

Psychopathologic Severity of Bulimia Nervosa or Binge-Eating Disorder

The psychopathologic severity of bulimia nervosa or binge-eating disorder was assessed using the EDE-Q [23,24], which is a self-contained, 28-item questionnaire derived from the EDE. The EDE-Q is scored on a 7-point Likert scale (0 to 6) on which a score of ≥4 indicates a clinical range. The global score on the EDE-Q is the sum of the 4 subscale scores (restraint, eating concern, shape concern, and weight concern) divided by 4. The presence and severity of symptoms of bulimia nervosa and cognitive and emotional signs and symptoms associated with hyperphagia were assessed using the BITE [25,26]. The BITE is a self-administered 33-item questionnaire that consists of 30 symptom subscales (eg, “Would you say that food dominated your life?”) and 3 severity subscales (“I don't eat all day,” “how to lose weight,” and “frequency of overeating”). Symptom subscale items are organized in yes/no format, and severity subscale items are organized in a Likert-type response format (5-point or 7-point, depending on the item). The symptom subscale has a minimum score of 0 and a maximum score of 30. However, the maximum severity subscale score is 39. The symptom subscale has 20 or more points ranging from “altered pattern of behavior” to “high probability of bulimia nervosa.” A severity subscale score of ≥5 indicates a clinically significant eating disorder.

Depression and Generalized Anxiety

The presence and severity of symptoms of depression experienced in the previous 2 weeks was evaluated using the PHQ-9 [27,28]. The self-administered questionnaire items are scored on a 4-point Likert-scale (0=not at all, 1=on several days, 2=half or more of the days, 3=almost every day). The minimum score is 0 and the maximum score is 27 (0 to 4 indicates no symptoms, 5 to 9 indicates mild symptoms, 10 to 14 indicates moderate symptoms, 15 to 19 indicates moderate to severe symptoms, and 20 to 27 indicates severe symptoms). The cutoff score for clinically significant depressive symptoms is 10. The presence and severity of generalized anxiety disorder was assessed using the GAD-7 [29,30], a self-administered questionnaire that assesses the severity of generalized anxiety disorder in the previous 2 weeks on a 4-point Likert scale (0=not at all, 1=one episode, 2=on half or more days, 3=almost every day). The minimum score is 0 and maximum score is 21 (0 to 4 indicates no symptoms, 5 to 9 indicates mild symptoms, 10 to 14 indicates moderate symptoms, and 15 to 21 indicates severe symptoms). The cutoff score for clinically significant symptoms of anxiety is 10.

Quality of Life

Quality of life was measured using the EQ-5D-5L questionnaire [31,32]. The EQ-5D-5L is self-administered with items that are scored from 0 (death) to 1 (in good health).

Therapeutic Relationship and Motivation

This study was the first to use ICBT via videoconference to provide therapist support for Japanese patients with bulimia nervosa and binge-eating disorder. ICBT via videoconferencing used the following scales to determine whether motivation and treatment compliance could be established. The strength of the therapeutic alliance was assessed using the WAI-SF [33]. This self-administered questionnaire consists of 12 items that measure 3 factors (ie, agreement on the task of treatment, agreement on the goal of treatment, and a bond between the therapist and patient). The WAI-SF is scored on a 7-point Likert scale; the minimum score is 12, and the maximum is 84. Motivation for treatment was assessed using the Motivational Ruler [34]. This instrument is a self-administered questionnaire containing 2 items (“How important is it for you to change and recover from your eating disorder?” and “How confident are you in your ability to change and recover from your eating disorder?”) that are measured on a 10-point Likert scale. The minimum score is 2, and the maximum score is 20.

Interventions

Participants entered the Web conference room by clicking on the URL in the email sent by the therapist. The intervention was performed for 50 minutes once a week for 16 weeks. The protocol included the following modules: assessment; explanation of CBT, motivation and goal setting, education on bulimia nervosa and how to keep a dietary diary, psychological education and monitoring for overeating behavior, psychological education and monitoring of diet, reforming of dietary habits (mindful diet), psychological education and self-monitoring of compensatory behavior, stress management and relaxation, understanding perfectionism, interpersonal relationships and assertions, acquisition of skills to associate with emotions (emotional adjustment, mindfulness, tolerance of distress, alexithymia), attention to body image, and prevention of recurrence. These modules are a development of those used in our previous study of face-to-face CBT for patients with bulimia nervosa in Japan [8,35].

Visual Aids

The use of visual aids promotes the learning process by enhancing motivation and understanding of complex concepts [36]. To deepen understanding, therapists presented a visual aid that consisted of several slides containing the key concepts of CBT using the screen-sharing feature of the videoconferencing software. After each session, study participants were emailed a set of password-protected homework slides.

Hardware

The therapists used a Surface Pro 2 computer (Microsoft Corp), which is a 2-in-1 detachable system that runs on Windows 10 Pro (Microsoft Corp). This computer has a display size of 10.6 inches and a resolution of 1920×1080 pixels.

Videoconferencing Software

A total of 3 licenses for videoconferencing software (Cisco Webex, Cisco Systems) were used. This system has International Organization for Standardization 27001 certification (for handling information security) and Statement on Standards for Attestation Engagements Number 16–compliant certification issued by a third party. Cisco Webex's use of a switching network with 128-bit Secure Sockets Layer encryption and public key infrastructure is considered to have resolved the problem of security [37]. We determined that Cisco Webex can be trusted for the purposes of this research because the software is stable and secure and protects personal information well.

Adverse Events

At the end of each session, therapists asked patients about their overall physical and mental state of health and instructed them to report any adverse events experienced after the intervention by email. No mental or physical adverse events were reported.

Statistical Analysis

Statistical analysis and reporting of this trial were conducted in accordance with the Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online TeleHealth (CONSORT-EHEALTH) guidelines [38]. The primary analysis compared differences in objective binge and purging on the EDE 16D between baseline and 16 weeks for all patients with bulimia nervosa or binge-eating disorder using the paired *t* test, and 95% confidence intervals were calculated. Analysis of secondary outcomes was performed in the same way. We also calculated the effect size of treatment using Cohen *d*, calculated as the mean difference after treatment divided by the pooled standard deviation. A Cohen *d*>0.20 was used as the criterion for a small effect, a value >0.50 as a medium effect, and >0.80 as a large effect [39]. All *P* values were 2-sided and those <.05 were considered statistically significant. All statistical analyses were performed with SAS software version 9.4 (SAS Institute Inc).

Results

Recruitment

Figure 1 shows the flow of participants through the study. Ten patients applied to participate in the study via our website. After screening by email and telephone, 2 patients failed to meet the inclusion criteria and were excluded (BMI <17.5 kg/m^2); 8 patients attended a baseline face-to-face assessment and were enrolled in the study. After enrollment, 1 patient declined to participate, and 2 patients dropped out during the intervention (1 refused to perform the homework; 1 could not combine the task assigned with her daily work). The analysis included the data of these 2 patients.

Demographic and Clinical Characteristics

The study population comprised 7 women of mean age 31.9 (SD 7.9; range 21-43) years. Their demographic and clinical characteristics are shown in Table 1. Five participants continued to receive pharmacotherapy during the trial (triazolam, n=1; flunitrazepam, n=1; clonazepam, n=1; sertraline, n=1; sodium valproate, n=1; quetiapine fumarate, n=1; aripiprazole, n=1; and clotiazepam, n=1).

Figure 1. Flow of participants.

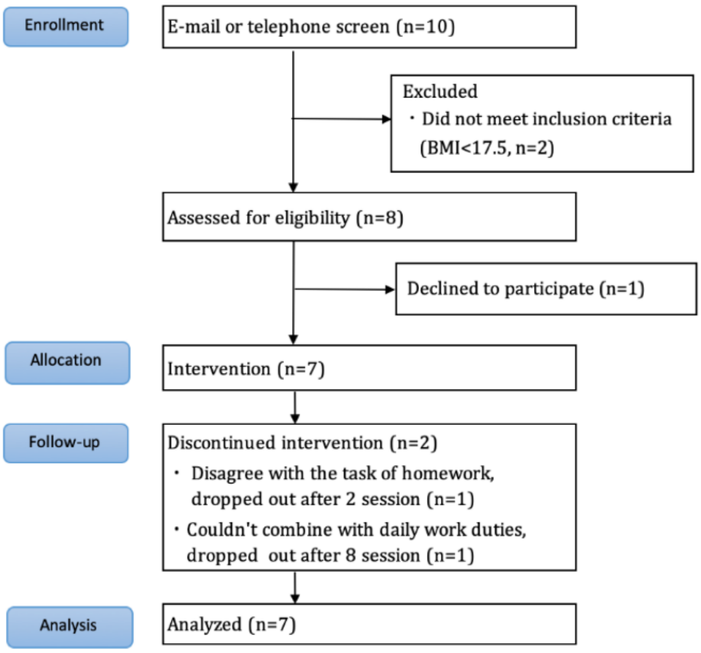


Table 1. Clinical and demographic characteristics of the study participants (n=7).

Characteristic	Participant							Value
	1	2	3	4	5	6	7	
Sex	Female	Female	Female	Female	Female	Female	Female	—
Age, mean (SD)	26	43	31	29	43	24	27	31.9 (7.9)
Education in years, mean (SD)	14	15	17	11	12	12	16	14.1 (2.4)
Duration of eating disorder in years, mean (SD)	12	17	17	12	19	5	2	12.0 (6.4)
BMI (kg/m^2), mean (SD)	17.8	22.2	20.4	18.4	36.0	20.0	30.0	23.5 (6.8)
Diagnosis	Bulimia nervosa	Bulimia nervosa	Bulimia nervosa	Bulimia nervosa	Binge-eating disorder	Bulimia nervosa	Binge-eating disorder	—
Comorbidity, n (%)								
Major depressive disorder	—	—	—	X	X	—	—	2 (29)
Dysthymia	—	—	—	X	—	—	X	2 (29)
Agoraphobia	—	—	—	—	X	—	—	1 (14)
Social anxiety disorder	X	—	—	—	—	—	X	2 (29)
Obsessive-compulsive disorder	—	—	X	—	—	—	—	1 (14)
Posttraumatic stress disorder	—	—	X	—	—	X	—	1 (14)
Generalized anxiety disorder	—	—	—	X	—	X	—	2 (29)
Schizophrenia	—	—	—	—	—	—	X	1 (14)
Past history, n (%)								
Major depressive disorder	—	—	—	—	X	—	—	1 (14)
Manic episode	—	—	—	—	X	—	X	2 (29)
Autism spectrum quotient, mean (SD)	28	13	15	33	33	19	17	22.6 (7.9)
Estimated IQ by JART[a], mean (SD)	106	106	98	89	94	100	106	99.9 (6.2)
Psychotropic drug, n (%)	—	—	X	X	X	X	X	5 (71)
Completer, n (%)	X	X	X	X	X	—	—	5 (71)

[a] JART: Japanese Adult Reading Test.

Primary Outcome

Figure 2 shows the change in the combined objective binge and purging episodes on the EDE 16D, and Figure 3 shows the detailed transition of symptoms for each patient. Although the mean combined episodes improved from 47.0 to 13.6 (71% reduction) and showed a medium effect size (Cohen *d*=–0.76), there was no significant reduction in the combined objective binge and purging episodes (EDE 16D −41; 95% CI −2.09 to 0.58; *P*=.17; Table 2). The frequency of objective binge episodes improved from 23.6 to 6.8 (71% reduction) and showed a large effect size (Cohen *d*=–0.82); no significant reduction in objective binge episodes (EDE 16D –20.60; 95% CI –2.18 to 0.53; *P*=.14). The frequency of purging episodes improved from 23.4 to 6.8 (71% reduction) and showed a medium effect size (Cohen *d*=–0.70); no significant reduction in the frequency of purging episodes (EDE 16D –20.40; 95% CI 2.01 to 0.62; *P*=.19). Additionally, abstinence was defined as none of these behaviors reported in the previous 28 days [18]. The abstinence rate from objective binge-eating, purging, and combined objective binge and purging was 43% (3/7).

Secondary Outcomes

Table 3 shows the change in secondary outcomes during the study. There were no significant changes in results.

Figure 2. Changes in frequency of objective binge and purging episodes during treatment. Bars indicate 1 standard deviation.

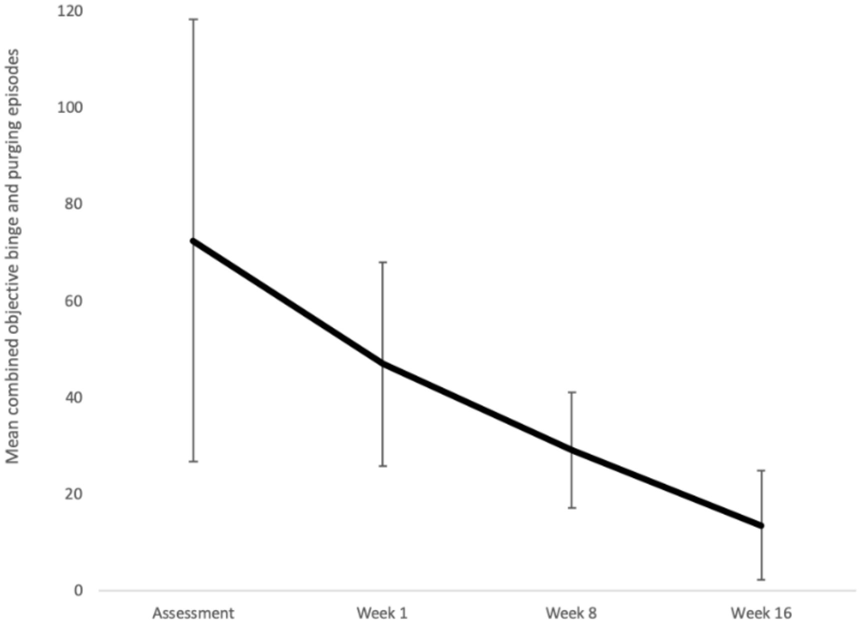


Figure 3. Change in number of episodes during treatment.

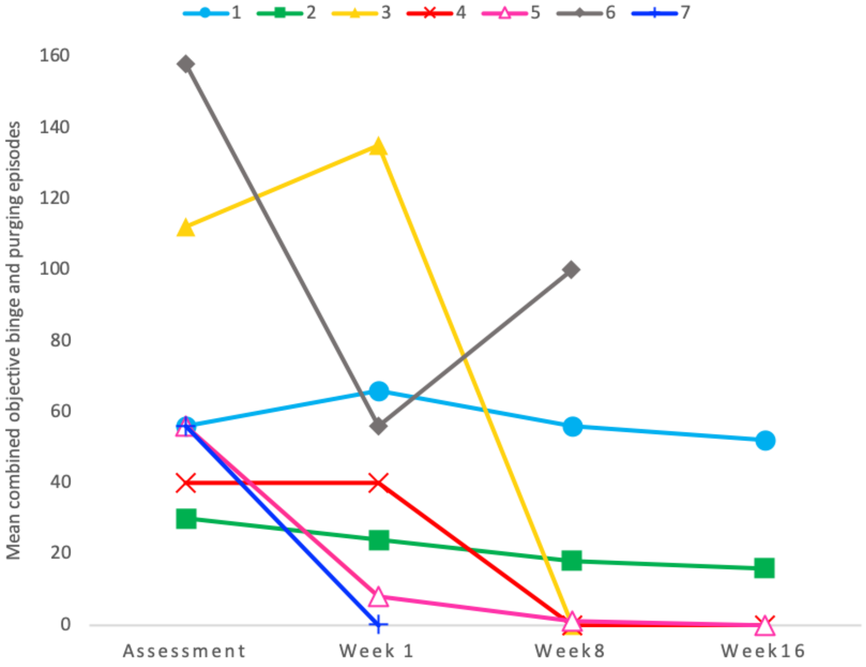


Table 2. Changes in the frequency of objective binge and purging episodes from pre- to post–internet-based cognitive behavioral therapy.

Episode and assessment time point	n	Mean (SD)	Cohen *d* (n=5)	95% CI	*P* value
Combined objective binge eating and purging episodes	—	—	–0.76	–2.09 to 0.58	.17
Assessment	7	72.6 (45.7)			
Week 1	7	47.0 (45.6)	—	—	—
Week 8	7	29.2 (40.9)			
Week 16	5	13.6 (22.6)	—	—	—
Objective binge eating episodes	—	—	–0.82	–2.18 to 0.53	.14
Assessment	7	32.6 (15.8)			
Week 1	7	29.2 (40.9)	—	—	—
Week 8	6	9.7 (11.9)			
Week 16	5	6.8 (11.3)	—	—	—
Purging episodes	—	—	–0.70	–2.01 to 0.62	.19
Assessment	7	40.0 (31.5)			
Week 1	7	13.6 (22.6)	—	—	—
Week 8	6	19.5 (31.6)			
Week 16	5	6.8 (11.3)	—	—	—

Table 3. Changes in psychological test scores from pre- and post–internet-based cognitive behavioral therapy.

Test name and assessment time point	n	Mean (SD)	Cohen *d* (n=5)	95% CI	*P* value
EDE-Q[a]					
Global score	—	—	**–0.48**	**–1.158 to 0.204**	**.35**
Assessment	7	3.2 (1.7)	—[b]	—	—
Week 1	7	2.2 (1.3)	—	—	—
Week 8	6	1.8 (1.3)	—	—	—
Week 16	5	1.3 (1.1)	—	—	—
Restraint subscale	—	—	**–0.33**	**–1.401 to 0.741**	**.50**
Assessment	7	3.1 (1.9)	—	—	—
Week 1	7	2.3 (1.9)	—	—	—
Week 8	6	1.6 (1.4)	—	—	—
Week 16	5	1.4 (0.9)	—	—	—
Eating concern subscale	—	—	**–1.07**	**–1.957 to –0.179**	**.08**
Assessment	7	3.3 (1.8)	—	—	—
Week 1	7	1.9 (1.7)	—	—	—
Week 8	6	1.2 (1.4)	—	—	—
Week 16	5	0.7 (0.2)	—	—	—
Shape concern subscale	—	—	**0.01**	**–0.273 to 0.297**	**.98**
Assessment	7	3.4 (1.9)	—	—	—
Week 1	7	2.5 (2.0)	—	—	—
Week 8	6	2.3 (2.0)	—	—	—
Week 16	5	1.6 (1.8)	—	—	—
Weight concern subscale	—	—	**0.26**	**0.044 to 0.484**	**.59**
Assessment	7	3.0 (2.0)	—	—	—
Week 1	7	2.2 (2.0)	—	—	—
Week 8	6	2.0 (1.6)	—	—	—
Week 16	5	1.5 (1.9)	—	—	—
BITE[c]					
Severity	—	—	**–1.04**	**–2.349 to 0.262**	**.08**
Assessment	7	9.6 (2.8)	—	—	—
Week 1	7	8.0 (4.0)	—	—	—
Week 8	6	7.8 (7.7)	—	—	—
Week 16	5	4.6 (3.4)	—	—	—
Symptoms	—	—	**–1.08**	**–2.069 to –0.087**	**.07**
Assessment	7	21.4 (4.1)	—	—	—
Week 1	7	17.0 (3.9)	—	—	—
Week 8	6	13.5 (4.6)	—	—	—
Week 16	5	12.0 (6.6)	—	—	—
BMI	—	—	**0.40**	**0.096 to 0.708**	**.42**
Assessment	7	23.5 (6.8)	—	—	—
Week 1	7	22.8 (6.0)	—	—	—
Week 8	6	22.4 (6.2)	—	—	—

Test name and assessment time point	n	Mean (SD)	Cohen *d* (n=5)	95% CI	*P* value
Week 16	5	23.1 (6.9)	—	—	—
PHQ-9[d]	—	—	**–0.35**	**–1.278 to 0.576**	**.48**
Assessment	7	14.0 (6.5)	—	—	—
Week 1	7	9.3 (3.8)	—	—	—
Week 8	6	8.2 (2.7)	—	—	—
Week 16	5	8.4 (4.8)	—	—	—
GAD-7[e]	—	—	**–0.52**	**–1.586 to 0.550**	**.31**
Assessment	7	11.1 (5.7)	—	—	—
Week 1	7	7.6 (5.1)	—	—	—
Week 8	6	8.3 (5.6)	—	—	—
Week 16	5	5.8 (1.5)	—	—	—
EQ-5D[f]	—	—	**0.46**	**–0.660 to 1.577**	**.36**
Assessment	7	0.7 (0.1)	—	—	—
Week 1	7	0.7 (0.2)	—	—	—
Week 8	6	0.8 (0.2)	—	—	—
Week 16	5	0.7 (0.1)	—	—	—
WAI-SF[g]	—	—	**0.13**	**–0.852 to 1.118**	**.78**
Assessment	7	—	—	—	—
Week 1	7	65.6 (8.4)	—	—	—
Week 8	6	63.7 (10.5)	—	—	—
Week 16	5	66.0 (4.4)	—	—	—
Motivational ruler	—	—	**0.00**	**–0.684 to 0.684**	**>.99**
Assessment	7	15.4 (3.6)	—	—	—
Week 1	7	16.4 (3.5)	—	—	—
Week 8	6	18.3 (1.9)	—	—	—
Week 16	5	16.0 (2.7)	—	—	—

[a] EDE-Q: Eating Disorder Examination Questionnaire.

[b] Not applicable.

[c] BITE: Bulimic Investigatory Test, Edinburgh.

[d] PHQ-9: 9-item Patient Health Questionnaire.

[e] GAD-7: 7-item Generalized Anxiety Disorder scale.

[f] EQ-5D: EurolQol–5 Dimension.

[g] WAI-SF: Working Alliance Inventory-Short Form.

Discussion

Principal Findings

This study examined the feasibility of videoconference-delivered CBT in patients with bulimia nervosa and binge-eating disorder. Interventions based on CBT for bulimia nervosa and binge-eating disorder were conducted, and symptom improvement and acceptance by patients were determined before and after the intervention. Our results suggest that ICBT via videoconference for treating bulimia nervosa and binge-eating disorder is feasible. This is the second study in to conduct ICBT via videoconference that we know of and the first to apply this treatment strategy in Japanese patients with bulimia nervosa and binge-eating disorder. Five out of 7 participants completed treatment, for a completion rate of 71%. The mean combined objective binge and purging episodes and mean frequency of objective binge episodes and purging episodes on the EDE 16D for the primary outcome decreased: combined objective binge and purging episodes were reduced from 47.0 to 13.6 (71% reduction) with a moderate Cohen *d* of –0.76, objective binge episodes were reduced from 23.6 to 6.8 (71% reduction) with a large Cohen *d* of –0.82, and purging episodes were reduced from 23.4 to 6.8 (71% reduction) with a moderate Cohen *d* of –0.70, but no change was statistically significant. The abstinence rate was 43% (3/7). There were no significant treatment-related changes in any of the secondary outcomes. No adverse events were reported. Hence, our findings revealed that ICBT via videoconference for patients with bulimia nervosa and binge-eating disorder is feasible.

Comparison With Previous Work

The results of this research are consistent with a previous RCT in the United States that used ICBT via videoconference and reported reductions in combined objective binge and purging episodes, objective binge episodes, and purging episodes (69%, 68%, and 70%, respectively) [18]. Our previous study with face-to-face CBT reported reduced combined objective binge and purging episodes, objective binge episodes, and purging episodes (51%, 50%, and 52%, respectively) [8]. This study has resulted in changes similar to our previous study. Furthermore, of the 5 people who completed treatment in this study, the abstinence rate was 60%. The abstinence rate among completers of ICBT via videoconference for bulimia nervosa and EDNOS was 37% [18]; the abstinence rate among face-to-face CBT for bulimia nervosa in our previous study was 40% [8]. Our results are better than previous studies about abstinence rate. Although previous studies did not report on effect size of combined objective binge and purging episodes [14,18], ICBT via videoconference had a moderate Cohen *d* of –0.66 for binging and a large Cohen *d* of –0.92 for purging at posttreatment, in our calculation [18]. According to a systematic review of ICBT, the range of Hedges effect sizes ranged from 0.75 to 1.05 for binge episodes and from 0.41 to 0.77 for purging episodes [14], which is similar to our findings. Furthermore, the results of our findings are similar to those of our previous study on CBT as a self-help strategy in Japanese patients with bulimia nervosa, in which the effect size of Cohen *d* for objective binge days (EDE 16D) was large, at 0.88, and for purging days was medium, at 0.67 [8]. Therefore, we believe that ICBT via videoconference is also feasible for Japanese patients with an eating disorder such as bulimia nervosa or binge-eating disorder.

First, for eating concern, a subscale of the EDE-Q, the effect size was large, although no significant difference was found in any of the secondary outcomes. Similar large effect sizes were observed for BITE severity and symptoms, which have been reported in previous studies [9,22]. Further, the BITE severity score (mean 4.6 [SD 3.4]) and symptoms score (mean 12.0 [SD 6.6]) were below the cutoff points of 5 and 20, respectively. The eating concern subscale assesses preoccupation with food, eating, or calories; fear of losing control over eating; eating in secret; social eating; and guilt about eating [26]. BITE, which is also related to the symptoms of overeating, had a similarly large effect size. Therefore, ICBT via videoconference may be promising for the treatment of overeating behavior and symptoms in patients with an eating disorder. However, there were no significant reductions in the restraint, shape concern, and weight concern subscales of the EDE-Q after treatment, and the effect size of Cohen *d* was small. A previous RCT in the United States using ICBT via videoconference reported a large effect size of Cohen *d* for these subscales of EDE-Q, in our calculation [18]. Similarly, our previous study on CBT as a self-help strategy in Japanese patients with bulimia nervosa reported a large effect size of Cohen *d* for these subscales of EDE-Q [8]. The effect of ICBT via videoconference on restraint, shape concern, and weight concern subscales of the EDE-Q may be relatively small in comparison with previous studies. Hence, there was no effect on dissatisfaction with body shape, aspiration for weight loss, or cognition to suppress food intake.

Second, the PHQ-9, which was used to evaluate symptoms of depression, had a small effect size, while the GAD-7, which was used to evaluate symptoms of anxiety, had a medium effect size. The mean GAD-7 score after treatment was below the cutoff value and decreased to a mild level. However, in our previous study of face-to-face CBT using the same treatment protocol, the effect sizes for PHQ-9 and GAD-7 were large [8]. The effect of ICBT via videoconference on symptoms of depression and anxiety associated with bulimia nervosa and binge-eating disorder may be relatively small in comparison with face-to-face CBT.

Third, the WAI-SF score in the range of 65.60 to 66.00 between pretreatment and posttreatment indicates that the patient's relationship with the therapist was at a consistently high level during the intervention. In a previous study in the United States that analyzed WAI scores in patients with bulimia nervosa or EDNOS who received ICBT via videoconference, there was no significant difference in WAI scores between those who received face-to-face CBT and those who received ICBT via videoconference; the WAI scores for ICBT via videoconference indicated that this treatment was also highly rated [40]. Furthermore, motivation ruler scores were also consistently high (16.4 to 16.0) during the intervention. This is consistent with the result of cognitive remediation therapy for anorexia nervosa (motivation ruler scores 14.3 to 16.1) [41]. Even if the therapist and patient did not participate in a session in the same room, it was possible to maintain an adequately high level of motivation and a satisfactory therapeutic alliance via videoconference. We confirmed for the first time that therapeutic alliance between the patient and therapist via videoconference can be sufficiently developed among Japanese patients with bulimia nervosa and binge-eating disorder. Hence, it is feasible to use ICBT via videoconference for patients with eating disorder who live in remote areas without them physically visiting a hospital.

Finally, the dropout rate was 29% (2/7) in this study. An RCT of ICBT via videoconference for patients with bulimia nervosa and EDNOS reported a dropout rate of 34% (21/64) [18]. A poor treatment alliance and complications arising from comorbid mental illness such as depression may affect the dropout rate or treatment response [42,43]. The motivation score in patients who dropped out was relatively high (participant 6 had a score of 15 out of 20; participant 7 had a score of 20 out of 20), suggesting that motivation is not a reason for discontinuing CBT. According to previous research, the likelihood of discontinuation increases in subjects with less education, more novelty-seeking behavior, previous experience with CBT, and a mismatch between the preferred treatment and that assigned [44]. According to a meta-analysis of dropout rates in studies that included face-to-face CBT in patients with eating disorders, the diagnostic entity, definition of dropout, symptom severity at baseline, quality of the study, and patient age had no effect on the likelihood of discontinuation of CBT; the estimated overall dropout rate was 24% [45], which is similar to the study we are reporting. In our previous face-to-face CBT study, the treatment dropout rate was very low at 8% (2/25) [8]. It is interesting to note that despite the ICBT via videoconference being conducted from the same treatment manual, the study we are reporting now had a dropout rate of 29%. Although this appears to be a factor in ICBT, our previous study of patients receiving ICBT via videoconference for anxiety and obsessive-compulsive disorder reported a dropout rate of only 3% (1/30) [46], suggesting that the likelihood of discontinuation may be disease-specific rather than treatment-related. The dropout rate in patients receiving ICBT via videoconference for an eating disorder may be higher than that in their counterparts receiving face-to-face CBT. The patients who dropped out of this study (one with bulimia nervosa and one with binge-eating disorder) had comorbidities that required treatment with oral psychotropic medication in both cases. One patient had schizophrenia, dysthymia, and social anxiety disorder, and the other had posttraumatic stress disorder and generalized anxiety disorder. Eating disorders are often combined with significant impairment of social function and other psychiatric disorders including depression, anxiety, obsessive-compulsive disorder, and substance abuse disorder [47-49]. The distribution of such complications in the patients in this study can be considered representative of patients with eating disorders.

No adverse events were reported, suggesting that ICBT via videoconference is safe for patients with bulimia nervosa, as per a previous study [18]. Therefore, ICBT via video conference can be safely performed in Japanese patients with bulimia nervosa and binge-eating disorder.

Limitations

There are some limitations in this study. The first is the small sample size, which meant that we could not determine whether our findings are due to a type II (β) error. G-power analysis was used to calculate the effect size, as described previously [9], which estimated the requirement of 15 samples. The second limitation is that the study was designed as a single-arm trial with no control group. To obtain statistically robust results, a clinical RCT is required that will estimate the sample size from the effect size of this study and will investigate the effectiveness of ICBT via videoconference to treat patients with bulimia nervosa and binge-eating disorder. We are presently planning such a trial in patients with eating disorders in Japan [UMIN000036825].

Conclusions

Our results suggest that ICBT via videoconference for treating bulimia nervosa and binge-eating disorder is feasible. To our knowledge, this study is the world’s second clinical study of ICBT via videoconference and is the first to apply this treatment strategy in Japanese patients with bulimia nervosa and binge-eating disorder. The frequency of objective binge and purging episodes on the EDE 16D was reduced after treatment (71%). Hence, ICBT via videoconference has the potential to be an effective treatment in Japanese patients with bulimia nervosa and binge-eating disorder. The lack of any adverse events suggests that ICBT via videoconference is safe for bulimia nervosa and binge-eating disorder. RCTs are now needed to examine the effectiveness of this treatment strategy.

Acknowledgments

We are very grateful for the support of our research nurses, Hiroe Ota and Yohko Suwabe. We also thank Dr Hideki Hanaoka, Mr Takumi Nakada, Ms Makiko Matsui, Ms Mayumi Ogawa, Ms Atsumi Yamazaki, and the staff at the Clinical Research Center of Chiba University Hospital for ensuring adherence to the research plan. This research was supported by the staff of the Department of Psychiatry, Chiba University Hospital. The authors express their sincere appreciation to everyone involved, in particular the study participants, for their cooperation. This work was supported by the Japan Society for the Promotion of Science Grants-in-Aid for Scientific Research (grant numbers 18K17313 and 19J00227) and Brain/MINDS Beyond program from the Japan Agency for Medical Research and Development (grant number 19dm0307104). The funding source had no role in the design or conduct of the study.

Authors' Contributions

SH designed the study and drafted the manuscript. NN developed the manual for bulimia nervosa, conducted the cognitive behavioral therapy, and managed the study. KM built the remote intervention system in our institution, created the digital questionnaire, and developed the manuscript. CS built and maintained the internet environment in our institution. HI, KO, and MT conducted the cognitive behavioral therapy. RS developed the manual for bulimia nervosa. YK analyzed the data. YH reviewed the manuscript. ES supervised the study. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, 5th Edition. Washington: American Psychiatric Association Publishing; 2013.
2. Chisuwa N, O'Dea JA. Body image and eating disorders amongst Japanese adolescents. A review of the literature. Appetite 2010 Feb;54(1):5-15. [doi: 10.1016/j.appet.2009.11.008] [Medline: 19941921]

3. Hofmann SG, Asnaani A, Vonk IJJ, Sawyer AT, Fang A. The efficacy of cognitive behavioral therapy: a review of meta-analyses. Cognit Ther Res 2012 Oct 1;36(5):427-440 [FREE Full text] [doi: 10.1007/s10608-012-9476-1] [Medline: 23459093]
4. Linardon J. Meta-analysis of the effects of cognitive-behavioral therapy on the core eating disorder maintaining mechanisms: implications for mechanisms of therapeutic change. Cogn Behav Ther 2018 Mar;47(2):107-125. [doi: 10.1080/16506073.2018.1427785] [Medline: 29378481]
5. Linardon J, Wade T, de la Piedad Garcia X, Brennan L. Psychotherapy for bulimia nervosa on symptoms of depression: a meta-analysis of randomized controlled trials. Int J Eat Disord 2017 Oct;50(10):1124-1136. [doi: 10.1002/eat.22763] [Medline: 28804915]
6. Vocks S, Tuschen-Caffier B, Pietrowsky R, Rustenbach SJ, Kersting A, Herpertz S. Meta-analysis of the effectiveness of psychological and pharmacological treatments for binge eating disorder. Int J Eat Disord 2010 Apr;43(3):205-217. [doi: 10.1002/eat.20696] [Medline: 19402028]
7. Linardon J, Wade TD, de la Piedad Garcia X, Brennan L. The efficacy of cognitive-behavioral therapy for eating disorders: a systematic review and meta-analysis. J Consult Clin Psychol 2017 Nov;85(11):1080-1094. [doi: 10.1037/ccp0000245] [Medline: 29083223]
8. Setsu R, Asano K, Numata N, Tanaka M, Ibuki H, Yamamoto T, et al. A single-arm pilot study of guided self-help treatment based cognitive behavioral therapy for bulimia nervosa in Japanese clinical settings. BMC Res Notes 2018 Apr 25;11(1):257 [FREE Full text] [doi: 10.1186/s13104-018-3373-y] [Medline: 29695260]
9. Carrard I, Rouget P, Fernández-Aranda F, Volkart A, Damoiseau M, Lam T. Evaluation and deployment of evidence based patient self-management support program for Bulimia Nervosa. Int J Med Inform 2006 Jan;75(1):101-109. [doi: 10.1016/j.ijmedinf.2005.07.031] [Medline: 16115793]
10. Ljotsson B, Lundin C, Mitsell K, Carlbring P, Ramklint M, Ghaderi A. Remote treatment of bulimia nervosa and binge eating disorder: a randomized trial of Internet-assisted cognitive behavioural therapy. Behav Res Ther 2007 Apr;45(4):649-661. [doi: 10.1016/j.brat.2006.06.010] [Medline: 16899213]
11. Strandskov SW, Ghaderi A, Andersson H, Parmskog N, Hjort E, Wärn AS, et al. Effects of tailored and ACT-influenced internet-based CBT for eating disorders and the relation between knowledge acquisition and outcome: a randomized controlled trial. Behav Ther 2017 Sep;48(5):624-637. [doi: 10.1016/j.beth.2017.02.002] [Medline: 28711113]
12. Zerwas SC, Watson HJ, Hofmeier SM, Levine MD, Hamer RM, Crosby RD, et al. CBT4BN: a randomized controlled trial of online chat and face-to-face group therapy for bulimia nervosa. Psychother Psychosom 2017;86(1):47-53 [FREE Full text] [doi: 10.1159/000449025] [Medline: 27883997]
13. Jacobi C, Beintner I, Fittig E, Trockel M, Braks K, Schade-Brittinger C, et al. Web-based aftercare for women with bulimia nervosa following inpatient treatment: randomized controlled efficacy trial. J Med Internet Res 2017 Sep 22;19(9):e321 [FREE Full text] [doi: 10.2196/jmir.7668] [Medline: 28939544]
14. Dölemeyer R, Tietjen A, Kersting A, Wagner B. Internet-based interventions for eating disorders in adults: a systematic review. BMC Psychiatry 2013 Aug 06;13:207 [FREE Full text] [doi: 10.1186/1471-244X-13-207] [Medline: 23919625]
15. Schlegl S, Bürger C, Schmidt L, Herbst N, Voderholzer U. The potential of technology-based psychological interventions for anorexia and bulimia nervosa: a systematic review and recommendations for future research. J Med Internet Res 2015 Mar 31;17(3):e85 [FREE Full text] [doi: 10.2196/jmir.3554] [Medline: 25840591]
16. Bakke B, Mitchell J, Wonderlich S, Erickson R. Administering cognitive-behavioral therapy for bulimia nervosa via telemedicine in rural settings. Int J Eat Disord 2001 Dec;30(4):454-457. [Medline: 11746307]
17. Simpson S, Knox J, Mitchell D, Ferguson J, Brebner J, Brebner E. A multidisciplinary approach to the treatment of eating disorders via videoconferencing in north-east Scotland. J Telemed Telecare 2003;9 Suppl 1:S37-S38. [doi: 10.1258/135763303322196286] [Medline: 12952717]
18. Mitchell JE, Crosby RD, Wonderlich SA, Crow S, Lancaster K, Simonich H, et al. A randomized trial comparing the efficacy of cognitive-behavioral therapy for bulimia nervosa delivered via telemedicine versus face-to-face. Behav Res Ther 2008 May;46(5):581-592 [FREE Full text] [doi: 10.1016/j.brat.2008.02.004] [Medline: 18374304]
19. Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, et al. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. J Clin Psychiatry 1998;59 Suppl 20:22-33. [Medline: 9881538]
20. Muramatsu K, Miyaoka H, Kamijima K, Muramatsu Y, Yoshida M, Otsubo T, et al. The patient health questionnaire, Japanese version: validity according to the mini-international neuropsychiatric interview-plus. Psychol Rep 2007 Dec;101(3 Pt 1):952-960. [doi: 10.2466/pr0.101.3.952-960] [Medline: 18232454]
21. Kobori O, Nakazato M, Yoshinaga N, Shiraishi T, Takaoka K, Nakagawa A, et al. Transporting cognitive behavioral therapy (CBT) and the improving access to psychological therapies (IAPT) project to Japan: preliminary observations and service evaluation in Chiba. J Ment Health Train Educ Pract 2014 Sep 02;9(3):155-166. [doi: 10.1108/JMHTEP-10-2013-0033]
22. Fairburn C, Cooper Z, O'Connor M. Eating disorder examination. In: Fairburn CG, editor. Cognitive Behavior Therapy and Eating Disorders. New York: Guilford Press; 2008.

23. Tachi T, Murakami K, Washizuka T, Ikuta N, Nishizono M, Miyake Y. Application of the eating disorder examination (EDE) to Japanese patients with eating disorders: reliability and validity of the Japanese version of EDE. Jpn J Psychosom Med 2005;45(10):792. [doi: 10.15064/jjpm.45.10_785]
24. Fairburn CG, Beglin SJ. Assessment of eating disorders: interview or self-report questionnaire? Int J Eat Disord 1994 Dec;16(4):363-370. [Medline: 7866415]
25. Henderson M, Freeman CP. A self-rating scale for bulimia. The "BITE". Br J Psychiatry 1987 Jan;150:18-24. [doi: 10.1192/bjp.150.1.18] [Medline: 3651670]
26. Nakai Y, Hamagaki S, Takagi R. [The validity of bulimic investigatory test of Edinburgh (BITE) and the survey of bulimia nervosa]. Clin Psychiatry 1998;40:711-716.
27. Spitzer RL, Kroenke K, Williams JB. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. JAMA 1999 Nov 10;282(18):1737-1744. [Medline: 10568646]
28. Muramatsu K, Miyaoka H, Kamijima K, Muramatsu Y, Yoshida M, Otsubo T, et al. The patient health questionnaire, Japanese version: validity according to the mini-international neuropsychiatric interview-plus. Psychol Rep 2007 Dec;101(3 Pt 1):952-960. [doi: 10.2466/pr0.101.3.952-960] [Medline: 18232454]
29. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006 May 22;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
30. Muramatsu K. Patient Health Questionnaire (PHQ-9, PHQ-15) Japanese version and Generalized Anxiety Disorder-7 Japanese version up to date. Stud Clin Psychol 2014;7:35-39.
31. Tsuchiya A, Ikeda S, Ikegami N, Nishimura S, Sakai I, Fukuda T, et al. Estimating an EQ-5D population value set: the case of Japan. Health Econ 2002 Jun;11(4):341-353. [doi: 10.1002/hec.673] [Medline: 12007165]
32. EuroQol Group. EuroQol—a new facility for the measurement of health-related quality of life. Health Policy 1990 Dec;16(3):199-208. [Medline: 10109801]
33. Tracey TJ, Kokotovic AM. Factor structure of the Working Alliance Inventory. Psychol Assess 1989 Sep;1(3):207-210. [doi: 10.1037/1040-3590.1.3.207]
34. Miller W, Rollnick S. Motivational Interviewing: Preparing People for Change. 2nd Edition. New York: Guilford Press; 2002.
35. Schmidt U, Treasure J, Alexander J. Getting better bit(e) by bit(e): a survival kit for sufferers of bulimia nervosa and binge eating disorders. London: Routledge; 1993.
36. Shabiralyani G, Hasan K, Hamad N, Iqbal N. Impact of visual aids in enhancing the learning process case research: district Dera Ghazi Khan. J Educ Pract 2015;6(19):226-233.
37. Ministry of Health, Labour, and Welfare. Guideline on Safety Management of Medical Information Systems, version 4. 2016. URL: https://www.mhlw.go.jp/stf/shingi2/0000119588.html [accessed 2018-12-02]
38. Eysenbach G. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. J Med Internet Res 2011;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]
39. Cohen J. Statistical Power Analysis for the Behavioral Sciences. 2nd Edition. Hillsdale: Lawrence Erlbaum Associates; 1988.
40. Ertelt TW, Crosby RD, Marino JM, Mitchell JE, Lancaster K, Crow SJ. Therapeutic factors affecting the cognitive behavioral treatment of bulimia nervosa via telemedicine versus face-to-face delivery. Int J Eat Disord 2011 Dec;44(8):687-691. [doi: 10.1002/eat.20874] [Medline: 22072405]
41. Kuge R, Lang K, Yokota A, Kodama S, Morino Y, Nakazato M, et al. Group cognitive remediation therapy for younger adolescents with anorexia nervosa: a feasibility study in a Japanese sample. BMC Res Notes 2017 Jul 25;10(1):317 [FREE Full text] [doi: 10.1186/s13104-017-2642-5] [Medline: 28743295]
42. Fernandez E, Salem D, Swift JK, Ramtahal N. Meta-analysis of dropout from cognitive behavioral therapy: magnitude, timing, and moderators. J Consult Clin Psychol 2015 Dec;83(6):1108-1122. [doi: 10.1037/ccp0000044] [Medline: 26302248]
43. 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]
44. Watson HJ, Levine MD, Zerwas SC, Hamer RM, Crosby RD, Sprecher CS, et al. Predictors of dropout in face-to-face and internet-based cognitive-behavioral therapy for bulimia nervosa in a randomized controlled trial. Int J Eat Disord 2017 May;50(5):569-577 [FREE Full text] [doi: 10.1002/eat.22644] [Medline: 27862108]
45. Linardon J, Hindle A, Brennan L. Dropout from cognitive-behavioral therapy for eating disorders: a meta-analysis of randomized, controlled trials. Int J Eat Disord 2018 May;51(5):381-391. [doi: 10.1002/eat.22850] [Medline: 29493805]
46. Matsumoto K, Sutoh C, Asano K, Seki Y, Urao Y, Yokoo M, et al. Internet-based cognitive behavioral therapy with real-time therapist support via videoconference for patients with obsessive-compulsive disorder, panic disorder, and social anxiety disorder: pilot single-arm trial. J Med Internet Res 2018 Dec 17;20(12):e12091 [FREE Full text] [doi: 10.2196/12091] [Medline: 30559094]
47. Flament MF, Godart NT, Fermanian J, Jeammet P. Predictive factors of social disability in patients with eating disorders. Eat Weight Disord 2001 Jun;6(2):99-106. [Medline: 11460850]
48. Halmi KA, Sunday SR. Temporal patterns of hunger and fullness ratings and related cognitions in anorexia and bulimia. Appetite 1991 Jun;16(3):219-237. [Medline: 1883249]

49. Kaye WH, Bulik CM, Thornton L, Barbarich N, Masters K. Comorbidity of anxiety disorders with anorexia and bulimia nervosa. Am J Psychiatry 2004 Dec;161(12):2215-2221. [doi: 10.1176/appi.ajp.161.12.2215] [Medline: 15569892]

Abbreviations

BITE: Bulimic Investigatory Test, Edinburgh
CBT: cognitive behavioral therapy
CONSORT-EHEALTH: Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and Online TeleHealth
DSM-5: Diagnostic and Statistical Manual of Mental Disorders, 5th Edition
EDE 16D: Eating Disorder Examination Edition 16.0D
EDE-Q: Eating Disorders Examination Questionnaire
EDNOS: eating disorders not otherwise specified
EQ-5D-5L: EuroQol–5 Dimensions
GAD-7: 7-item Generalized Anxiety Disorder scale
ICBT: internet-based cognitive behavioral therapy
PHQ-9: 9-item Patient Heallth Questionnaire
RCT: randomized controlled trials
WAI-SF: Working Alliance Inventory–Short Form

Edited by G Eysenbach; submitted 12.08.19; peer-reviewed by S Smith, K Mathiasen, H Wang; comments to author 02.09.19; revised version received 18.09.19; accepted 01.10.19; published 23.10.19.

Please cite as:
Hamatani S, Numata N, Matsumoto K, Sutoh C, Ibuki H, Oshiro K, Tanaka M, Setsu R, Kawasaki Y, Hirano Y, Shimizu E
Internet-Based Cognitive Behavioral Therapy via Videoconference for Patients With Bulimia Nervosa and Binge-Eating Disorder: Pilot Prospective Single-Arm Feasibility Trial
JMIR Form Res 2019;3(4):e15738
URL: http://formative.jmir.org/2019/4/e15738/
doi:10.2196/15738
PMID:31647472

Original Paper

A Promising Food-Coaching Intervention Program to Achieve Optimal Gestational Weight Gain in Overweight and Obese Pregnant Women: Pilot Randomized Controlled Trial of a Smartphone App

Ling-Jun Li[1], MD, PhD; Izzuddin M Aris[2], PhD; Wee Meng Han[3], MD; Kok Hian Tan[1], MD

[1]Department of Obstetrics & Gynecology, KK Women's and Children's Hospital, Singapore, Singapore

[2]Division of Chronic Disease Research Across the Lifecourse, Department of Population Medicine, Harvard Medical School and Harvard Pilgrim Health Care Institute, Boston, MA, United States

[3]Department of Nutrition, KK Women's and Children's Hospital, Singapore, Singapore

Corresponding Author:
Ling-Jun Li, MD, PhD
Department of Obstetrics & Gynecology
KK Women's and Children's Hospital
100 Bukit Timah Road, Children's Tower Level 3
Singapore, 229899
Singapore
Phone: 65 63941099
Email: queenie.li.l.j@gmail.com

Abstract

Background: Traditional dietary recommendations for achieving optimal gestational weight gain are ineffective for pregnant women due to the lack of real-time communication and tedious consultation processes.

Objective: In this pilot study, we aimed to determine the feasibility of a novel food-coaching smartphone app for controlling gestational weight gain and macronutrient intake among overweight and obese pregnant women.

Methods: We designed a randomized controlled trial and recruited 30 overweight and obese pregnant women (1:1 ratio) during 18-20 weeks of gestation and followed them up after 4 and 8 weeks, respectively. Both groups received standard pregnancy dietary orientation at recruitment, while the intervention group received 8 weeks of real-time food coaching via a smartphone app. This food-coaching smartphone app (Glycoleap, Holmusk, Singapore) aimed to improve care and outcomes for people with diabetes. Pregnant women using this app were able to upload food images (eg, a picture of a meal, a drink, or a dessert) and received real-time and detailed food-coaching comments and guidance provided by professional dietitians during the day (8 AM to 8 PM). We recorded detailed characteristics during recruitment and examined anthropometry at all visits. We compared the mean differences of the 8-week gestational weight gain and macronutrient intake between the two groups.

Results: Upon study completion, three subjects dropped out from the intervention, and one gave birth prematurely in the control group. The acceptance rate of the smartphone app was 90%. More participants achieved optimal gestational weight gain per week in the intervention group (8/12, 67%) than in the control group (5/14, 36%). After the 8-week intervention, women in the intervention group appeared to have lower gestational weight gain (mean difference=–0.08 kg; 95% CI –1.80 to 1.63) and cholesterol intake (mean difference=–31.73 mg; 95% CI –102.91 to 39.45) than those in the control group.

Conclusions: Our findings showed that this food-coaching smartphone app is feasible and favorable for weight gain control and cholesterol intake control among overweight and obese pregnant women. Although our results were not significant (perhaps, attributed to the small sample size), it provided proof of concept for the feasibility of applying such technology in future randomized controlled trials with a larger sample size, an earlier intervention onset, and a longer follow-up for overweight and obese pregnant women.

(JMIR Form Res 2019;3(4):e13013) doi:10.2196/13013

KEYWORDS

overweight, obesity, pregnant women; gestational weight gain; food diary; randomized controlled trial; smartphone app; food coaching; dietary recommendation; feasibility

Introduction

Overweight and obese pregnant women are often at an increased risk of a series of maternal and offspring adverse outcomes [1-3]. In developed countries, overweight and obese women accounted for 30%-50% of fertile women [4-6]. Therefore, the high prevalence of overweight and obesity among pregnant women signifies a substantial burden to public health welfare worldwide. Current clinical care management of overweight and obese women during pregnancy consists of information on healthy eating practices according to standard dietary guidelines both worldwide [7-10] and in Singapore [11,12], in terms of restricting daily energy intake; balancing the proportion of complex carbohydrate (33%-40%), protein (20%), and fat (40%); and lowering cholesterol intake However, the reported effects were equivocal due to limitations such as low compliance, delayed feedback, late initiation, and inefficient delivery of intervention [13,14].

In recent years, the use of mobile technology (such as smartphone apps) for patient care has been increasing. The feasibility and efficacy of such technology has been proven for weight management among pediatric obese patients [15] and for glucose control in mothers with gestational diabetes mellitus during pregnancy [16]. Furthermore, a systematic review summarized 12 studies using phone-based reporting interventions such as video call, phone calls, short messaging service (SMS) and smartphone apps and showed consistent evidence that although such approaches could help pregnant women control their gestational weight gain, they were not effective in preventing other pregnancy outcomes such as gestational diabetes mellitus [17]. Another systematic review of four randomized controlled trials (RCTs) did not draw any firm conclusions on the effects of mobile app interventions during pregnancy on maternal knowledge, behavior change, and perinatal health outcomes due to heterogeneity of interventions, comparators, and outcome measures across all RCTs [18]. In addition, based on the intervention evaluation alone, none of these mobile technologies provide a real-time communication between end users and medical workers or dietitians, thus likely affecting the compliance and efficacy of the phone-based intervention. Given the high prevalence of smartphone app usage among pregnant women, more rigorous studies are needed to optimize the study design and implementation of these technologies in improving maternal health outcomes. In this pilot randomized controlled trial study, we tested a food-coaching intervention program delivered through a smartphone app in overweight and obese pregnant subjects without gestational diabetes mellitus and examined its feasibility, acceptance, and preliminary utility during an 8-week follow-up in second trimester.

Methods

Study Design and Population

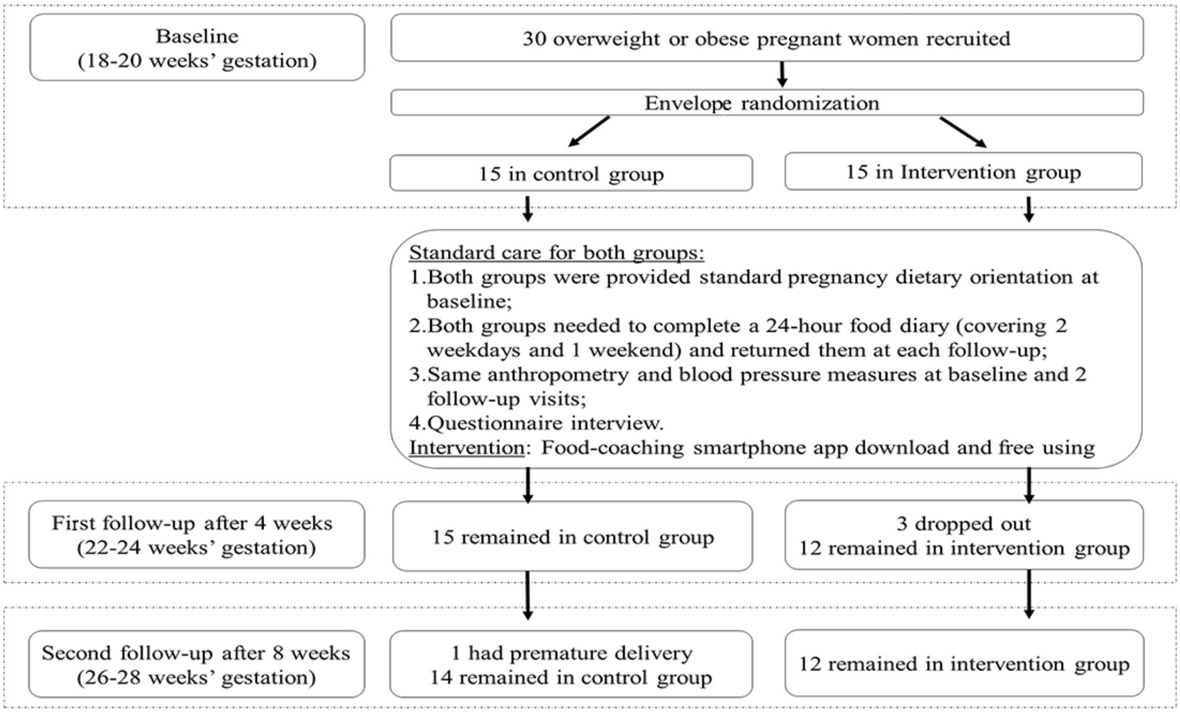
We conducted a prospective, two-arm, unblinded RCT in a subsidized clinic within a tertiary government hospital in Singapore (KK Women’s and Children’s Hospital [KKH]) between March and July 2018. We recruited pregnant women if they were Singapore citizens or permanent residents, overweight or obese (ie, prepregnancy or booking body mass index≥25 kg/m^2), age≥21 years, between 18 and 20 weeks’ gestation at the time of recruitment, planning to deliver in KKH, capable of reading and writing in English, and able to download and use the smartphone app. We excluded women with special dietary restrictions due to medical conditions such as type 1 or type 2 diabetes, gestational diabetes mellitus, hypertension, and chronic kidney disease.

We conducted the study according to the tenets of the Declaration of Helsinki and obtained approval by the SingHealth Centralized Institutional Review Board and the National Health Group’s Domain Specific Review Board. We obtained written informed consent from all pregnant women at baseline recruitment. As the primary aim of this study is to test the feasibility of the smartphone app in the pilot phase, we only obtained the local institutional review board approval (CIRB 2017/2132), but did not register for an online RCT number.

Randomization and Procedures

Research coordinators randomly assigned pregnant women to the intervention (food-coaching smartphone app) or control group (standard pregnancy dietary orientation) in a 1:1 allocation ratio, using the envelope randomization method (Figure 1).

Figure 1. Study flow of the randomized controlled trial.



Standard Pregnancy Dietary Orientation

The same research coordinators administered the standard dietary orientation to women in the intervention and control groups using standardized materials, including basic nutrition principles of choosing a wide variety of nutrient-dense foods, limiting intake of high-fat and high-sugar food and beverages, and guiding food portion sizes, with reference to the local gestational dietary guidelines by the Health Promotion Board [19]. We informed participants in both groups about the recommended gestational weight gain, together with some simple physical activity pointers.

A Food-Coaching Smartphone App

The intervention group participants could download and use the food-coaching smartphone app (Glycoleap, Holmusk, Singapore) for free for up to 8 weeks during the RCT. The app aimed to improve care and outcomes of people with type 2 diabetes in terms of diet control. Pregnant women using this app were able to upload food images and received real-time and detailed food-coaching comments and guidance from professional dietitians during the day (8 AM to 8 PM). According to the local pregnancy guideline [20], food dietitians rated the food image from 1 to 5 (1=the least recommended score, 5=the most recommended scored) and provided feedback in terms of degree and balance of the food items and composition (Multimedia Appendix 1). Although this smartphone app was available on both Apple and Google store platforms, we did not think there was a high chance for the controls to obtain access to this app for two reasons: (1) This smartphone app requires an in-store purchase of up to SG $30 (US $22) per month, which is not subsidized by any local Singaporean medical insurance. All participants were recruited from the government tertiary hospital's subsidized clinic; therefore, the chances for them to pay from their own pocket to afford additional medical care are low. (2) Only pregnant women randomized into the intervention group were informed of the name and company of the smartphone app, while those in the control group were not provided any information regarding the app. In addition, subjects in the intervention and control groups did not know of each other.

We recruited all subjects at baseline (18-20 weeks' gestation) and followed them up after 4 weeks (22-24 weeks' gestation) and 8 weeks (26-28 weeks' gestation).

Outcomes

Based on the usage of the food-coaching smartphone app, we assessed the feasibility of the smartphone app by collecting participant feedback in the intervention group, using an evaluation form at the 8-week follow-up visit (Multimedia Appendix 2). In addition, we assessed the compliance in the intervention group via the times of log in for each participant in the first half and second half of the follow-up. We calculated the proportion of participants with optimal second trimester gestational weight gain based on the Institute of Medicine guidelines for overweight (0.23-0.33 kg/week) and obese women (0.17-0.27 kg/week) [9,21]. We assisted all pregnant subjects in quantifying their food and beverage intake. Such methods of assessing dietary data have been widely published [22-24]. We obtained data on 24-hour food recall at recruitment and collected self-administered 3-day food diary data (Multimedia Appendix 3) in both groups at two follow-ups. We analyzed the dietary records using an online nutrient analysis software [25], which was derived from locally available foods [20,26,27]. The Singapore Health Promotion Board online guidelines provided specific dietary instructions upon completion of the 3-day food diary and provided a picture compendium to each subject to facilitate identification of each type and quantity of food consumed. The compendium of food pictures consisted of photographs of individual food items coded to reflect the portion size of food items and standardize the estimated amounts eaten. The guidelines also used standard bowls, glasses, and spoons of varying sizes to estimate the volume of fluids or the amount of food consumed. Research coordinators entered each recorded food item into an online nutrient analysis database (Food Information and Nutrient Database system, Health Promotion Board, Singapore) based on a food composition database of locally available foods [25]. The daily food intake data were then summed up, and the total daily energy, macronutrients (carbohydrate, fiber, fat, and protein), calcium, cholesterol, and sodium intakes were averaged over each 3-day period and listed in kilocalories, grams, and percentages in the data summary form. The intake was compared between the intervention and control groups at the two follow-up visits.

Statistical Analysis

We applied the Fisher exact test and Student *t* test for categorical and continuous variables, respectively, to compare characteristics between the intervention and control groups. We used linear regression to examine the effect of the food-coaching smartphone app on weight gain control and macronutrient intake between two groups. Mean differences as an estimate referencing the control group were shown in linear regression. We performed statistical analysis using STATA (Version 14.0. STATA Corp, College Station, Texas), set a two-tailed *P* value for significance at .05, and provided the 95% CIs for all estimates.

Results

Among 30 pregnant women recruited at baseline with 1:1 allocation, 26 (12 in the intervention group and 14 in the control group) completed the 8-week RCT. The detailed RCT flow chart is shown in Figure 1. Based on the log-in records, the uptake of this smartphone app among our intervention subjects was up to 90% at the beginning (n=15) and 70% at the end of the study (n=12), respectively. Based on the 12 returned user evaluation forms, 75% (n=9) of the app users found that the app was easy to operate, more than 80% (n=10) thought the food-coaching guidance was acceptably fast, and 90% (n=11) were satisfied and reported that the app had somewhat or greatly improved their own diet. Therefore, they would recommend the smartphone app to family and friends. Furthermore, according to the backlog records from the smartphone app provider, we assessed compliance by calculating the average log-in frequency in the first and second half of the 8-week follow-up. Each of the 12 subjects in the intervention group logged into the smartphone app 8 times per week in the first 4 weeks, on an average, but this number reduced to 2 times per week in the subsequent 4 weeks.

Tables 1-3 show the baseline and follow-up characteristics between the intervention and control groups. More participants met the Institute of Medicine recommendation for optimal gestational weight gain per week in the intervention group (4-week follow-up: 7/12, 58%; 8-week follow-up: 8/12, 67%) than in the control group (4-week follow-up: 8/15, 53%; 8-week follow-up: 5/14, 36%).

Table 1. Comparison of baseline characteristics between the intervention and control groups (baseline recruitment at 18-20 weeks' gestation).

Characteristics	Intervention group (n=15)	Control group (n=15)	*P* value[a]
Age (years), mean (SD)	29.3 (4.4)	30.7 (5)	.45
Ethnicity, n (%)			.15
Chinese	2 (13)	6 (40)	
Malay	12 (80)	8 (53)	
Indian	1 (7)	0 (0)	
Others	0 (0)	1 (6)	
Smoking history (yes), n (%)	4 (27)	1 (6)	.33
Parity, n (%)			.60
0	7 (47)	5 (33)	
1	4 (27)	7 (47)	
≥2	4 (27)	3 (20)	
Maternal college degree (yes), n (%)	3 (20)	5 (33)	.68
Paternal college degree (yes), n (%)	1 (7)	4 (27)	.33
Household income ≥US $6000/month, n (%)	3 (20)	6 (40)	.43
Past history of pregnancy outcomes			.16
Gestational diabetes mellitus, n (%)	0 (0)	3 (20)	
Hypertensive disorders during pregnancy, n (%)	0 (0)	1 (7)	
Prepregnancy weight (kg), mean (SD)	78.5 (15.8)	73.9 (6.8)	.31
Prepregnancy body mass index (kg/m^2), mean (SD)	31.0 (4.7)	29.4 (2.2)	.27
Gestational age at recruitment (weeks), mean (SD)	19.2 (2.1)	17.6 (2.8)	.08
Baseline body mass index (kg/m^2), mean (SD)	34.2 (4.3)	31.3 (2.8)	.04
Baseline systolic blood pressure (mm Hg), mean (SD)	118.6 (8.9)	116.8 (9.4)	.59
Baseline diastolic blood pressure (mm Hg), mean (SD)	69.4 (4.6)	64.1 (7.9)	.03

[a] Student *t* test or Fisher exact test.

Table 2. Comparison of follow-up measures between the intervention and control groups (4-week follow-up at 22-24 weeks' gestation).

Clinical measures	Intervention group (n=12)	Control group (n=15)	*P* value[a]
Anthropometric and blood pressure measures, mean (SD)			
First follow-up body mass index (kg/m^2)	36.4 (4.9)	31.9 (2.6)	.003
First follow-up systolic blood pressure (mm Hg)	121.9 (6.8)	120.9 (6.4)	.73
First follow-up diastolic blood pressure (mm Hg)	70.9 (7.5)	66.5 (6.7)	.15
Weight gain from baseline (kg)	1.3 (1.5)	1.5 (1.6)	.83
Patients obtaining optimal weight gain per week, n (%)	7 (58)	8 (53)	.67
Dietary measures, mean (SD)			
Energy intake (kcal)	1370.5 (359.4)	1514.0 (362.6)	.32
Carbohydrate (g)	177.5 (55.3)	189.5 (49.3)	.56
Protein (g)	52.6 (15.0)	65.1 (17.5)	.06
Total fat (g)	49.4 (15.9)	54.3 (17.3)	.46
Cholesterol (mg)	182.0 (90.6)	246.9 (109.7)	.11
Calcium (g)	513.4 (320.8)	513.3 (286.9)	.99
Dietary fiber (g)	12.9 (4.3)	12.9 (3.5)	.95
Sodium (g)	2.4 (0.8)	2.8 (0.8)	.18

[a] Student *t* test, Fisher exact test, or Wilcoxon signed rank test.

Table 3. Comparison of follow-up measures between the intervention and control groups (8-week follow-up at 26-28 weeks' gestation).

Clinical measures	Intervention group (n=12)	Control group (n=14)	*P* value[a]
Anthropometric and blood pressure measures, mean (SD)			
Second follow-up body mass index (kg/m^2)	36.0 (4.6)	32.4 (2.8)	.02
Second follow-up systolic blood pressure (mm Hg)	120.2 (7.5)	117.3 (7.3)	.33
Second follow-up diastolic blood pressure (mm Hg)	69.6 (9.1)	67.4 (5.8)	.47
Weight gain from baseline (kg)	2.9 (1.9)	3.0 (2.24)	.92
Patients obtaining optimal weight gain per week, n (%)	8 (67)	5 (36)	.43
Dietary measures, mean (SD)			
Energy intake (kcal)	1370.5 (359.4)	1514.0 (362.6)	.44
Carbohydrate (g)	1815.0)	170.6 (49.4)	.61
Protein (g)	56.7 (19.8)	58.6 (15.8)	.79
Total fat (g)	54.3 (20.7)	43.5 (17.5)	.15
Cholesterol (mg)	196.2 (89.3)	220.6 (87.9)	.48
Calcium (g)	461.0 (246.0)	489.1 (222.9)	.76
Dietary fiber (g)	13.0 (5.3)	12.9 (4.5)	.99
Sodium (g)	2.2 (0.7)	2.4 (0.8)	.55

[a] Student *t* test, Fisher exact test, or Wilcoxon signed rank test.

Although not significant, we found a trend among women in the intervention group who tended to have less weight gain than those in the control group at the 4-week follow-up (mean difference=–0.15 kg; 95% CI –1.51 to 1.21) and 8-week follow-up (mean difference=–0.08 kg; –1.80 to 1.63; Table 4). In addition, women in the intervention group tended to consume less cholesterol than those in the control group at the 4-week follow-up (mean difference=–64.87 mg; 95% CI –146.04 to 16.31) and 8-week follow-up (mean difference=–31.73 mg; 95% CI –102.91 to 39.45).

Table 4. Weight gain from baseline and dietary intake between the intervention and control groups.

Dietary components	Mean difference between intervention and control (ref) groups			
	4-week follow-up (22-24 weeks' gestation)[a]		8-week follow-up (26-28 weeks' gestation)[b]	
	β (95% CI)	*P* value	β (95% CI)	*P* value
Weight gain from baseline (kg)	–0.15 (–1.51 to 1.21)	.83	–0.08 (–1.80 to 1.63)	.92
Energy intake (kcal)	–143.55 (–431.66 to 144.56)	.32	123.99 (–222.74 to 470.72)	.47
Carbo (g)	–12.05 (–53.54 to 29.44)	.56	9.35 (–33.62 to 52.31)	.66
Protein (g)	–12.58 (–25.69 to 0.53)	.06	–2.08 (–16.73 to 12.57)	.77
Total fat (g)	–4.91 (–18.24 to 8.42)	.46	11.11 (–4.60 to 26.81)	.16
Cholesterol (mg)	–64.87 (–146.04 to 16.31)	.11	–31.73 (–102.91 to 39.45)	.37
Calcium (g)	0.04 (–241.08 to 241.16)	>.99	–19.16 (–211.31 to 172.99)	.84
Dietary fiber (g)	0.10 (–2.97 to 3.17)	.95	0.20 (–3.82 to 4.22)	.92
Sodium (g)	–413.16 (–1032.09 to 205.76)	.18	–98.83 (–692.65 to 494.98)	.73

[a] Intervention group: n=12; control group: n=15.

[b] Intervention group: n=12; control group: n=14.

Discussion

In this in a small sample RCT, we used a smartphone app to guide overweight and obese pregnant women to eat heathier in order to obtain optimal weight gain. We noted a high uptake of the smartphone app, as described above, and found evidence that pregnant women in the intervention group were more likely to have optimal gestational weight gain and consume less cholesterol compared with women in the control group. Although our pilot results were not significant, it provided proof of concept for the feasibility of applying such technology in future RCTs with a larger sample size, an earlier intervention onset, and a longer follow-up for overweight and obese pregnant women.

Although the dietary intervention has been proven to be effective in terms of weight gain control among overweight and obese pregnant women [28,29], there are huge variations in the delivery of dietary recommendation such as timing of intervention initiation, intensity of intervention, and feedback availability [13,28-31]. In a recently published systematic review summarizing 12 phone-based intervention studies on gestational weight gain control outcomes, most of the studies used phone calls or SMS to provide weight control guidelines, encouraged physical activity, and provided educational information about healthy nutrition. Several studies suggested that telephone communication is one of the most cost-effective tools to keep track of pregnant women's health [32]. However, phone-based interventions are typically initiated by health care providers and might not effectively motivate pregnant women to self-manage their behaviors. Given the high prevalence of smartphone app usage among pregnant women to improve their healthy behaviors, no firm conclusion has been drawn in terms of improving perinatal outcomes [18]. One of the major drawbacks is the lack of rigorous studies examining self-managing and self-regulatory behaviors in pregnant women. Our food-coaching app is more flexible and interactive and has greater variety of communication modalities, thus overcoming all aforementioned limitations in current technology. Given the high smartphone usage rate (up to 80%) among Singaporeans [33] and high prevalence (up to 30%) of overweight and obesity among pregnant women in tertiary hospital settings [4], use of a mobile app to promote healthy dietary intake is feasible and likely a promising intervention strategy. Our findings showed feasibility and acceptability of such a food-coaching smartphone app. For example, 75% of the app users found that the app was easy to operate, and more than 80% thought the food-coaching guidance was acceptably fast. Furthermore, 90% reported that the app had somewhat or greatly improved their own diet.

Interestingly, we did observe a trend of lower energy, carbohydrate, protein, total fat, and cholesterol intake in the first 4 weeks of the intervention, while most of the macronutrients did not maintain the same trend at 8 weeks of follow-up, except for cholesterol. This may be because pregnant subjects were more inclined to remember food items that are high in cholesterol (ie, saturated fat, red meat, full-fat dairy products) and were able to make an effort to avoid such food items compared with others. However, further studies are needed to verify our findings, as such findings with a small sample might be biased. Although the effect estimates are in the desired direction and supported the proof of concept as the primary focus in this pilot RCT, we still need to further verify the utility of such a smartphone app in a larger targeted population with an earlier phase intervention during pregnancy and a longer follow-up.

The strength of our study included a prospective RCT study design and the use of standard protocols in anthropometric measures and dietary assessments. However, our study has significant limitations regarding the loss to follow-up and a small sample for analysis.

In conclusion, our study provides proof of concept that smartphone technology is feasible and acceptable in clinical dietary guidance among overweight and obese pregnant women. In the future, we will adopt this food-coaching smartphone app and test its utility in a larger setting with a targeted population, earlier intervention, and longer follow-up throughout pregnancy.

Acknowledgments

This study is supported by Singapore National Medical Research Council Centre Grant (NMRC/CG/C008A/2017_KKH). L-JL is funded by Singapore National Medical Research Council Transition Award (NMRC TA/0027/2014) and (NMRC/CG/C008A/2017_KKH).

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 and the study was considered formative. However, readers are advised to carefully assess the validity of any potential explicit or implicit claims related to primary outcomes or effectiveness.

Multimedia Appendix 1
The electronic interface layout of the food-coaching smartphone app (Glycoleap).
[PDF File (Adobe PDF File), 50 KB - formative_v3i4e13013_app1.pdf]

Multimedia Appendix 2
Self-evaluation of acceptability on the GlycoLeap application during pregnancy.
[PDF File (Adobe PDF File), 161 KB - formative_v3i4e13013_app2.pdf]

Multimedia Appendix 3
Food diary form.
[PDF File (Adobe PDF File), 60 KB - formative_v3i4e13013_app3.pdf]

Multimedia Appendix 4
CONSORT‐EHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 2801 KB - formative_v3i4e13013_app4.pdf]

References

1. Torloni MR, Betrán AP, Horta BL, Nakamura MU, Atallah AN, Moron AF, et al. Prepregnancy BMI and the risk of gestational diabetes: a systematic review of the literature with meta-analysis. Obes Rev 2009 Mar;10(2):194-203. [doi: 10.1111/j.1467-789X.2008.00541.x] [Medline: 19055539]
2. Jolly MC, Sebire NJ, Harris JP, Regan L, Robinson S. Risk factors for macrosomia and its clinical consequences: a study of 350,311 pregnancies. Eur J Obstet Gynecol Reprod Biol 2003 Nov 10;111(1):9-14. [Medline: 14557004]
3. Yu Z, Han S, Zhu J, Sun X, Ji C, Guo X. Pre-pregnancy body mass index in relation to infant birth weight and offspring overweight/obesity: a systematic review and meta-analysis. PLoS One 2013;8(4):e61627 [FREE Full text] [doi: 10.1371/journal.pone.0061627] [Medline: 23613888]
4. Chong Y, Cai S, Lin H, Soh SE, Lee Y, Leow MK, GUSTO study group. Ethnic differences translate to inadequacy of high-risk screening for gestational diabetes mellitus in an Asian population: a cohort study. BMC Pregnancy Childbirth 2014 Oct 02;14:345 [FREE Full text] [doi: 10.1186/1471-2393-14-345] [Medline: 25273851]
5. Devlieger R, Benhalima K, Damm P, Van Assche A, Mathieu C, Mahmood T, et al. Maternal obesity in Europe: where do we stand and how to move forward? A scientific paper commissioned by the European Board and College of Obstetrics and Gynaecology (EBCOG). Eur J Obstet Gynecol Reprod Biol 2016 Jun;201:203-208. [doi: 10.1016/j.ejogrb.2016.04.005] [Medline: 27160501]
6. Mitchell S, Shaw D. The worldwide epidemic of female obesity. Best Pract Res Clin Obstet Gynaecol 2015 Apr;29(3):289-299. [doi: 10.1016/j.bpobgyn.2014.10.002] [Medline: 25487257]
7. American College of Obstetricians and Gynecologists. Challenges for Overweight and Obese Women URL: https://www.acog.org/Clinical-Guidance-and-Publications/Committee-Opinions/Committee-on-Health-Care-for-Underserved-Women/Challenges-for-Overweight-and-Obese-Women [accessed 2018-11-06]
8. Muktabhant B, Lawrie TA, Lumbiganon P, Laopaiboon M. Diet or exercise, or both, for preventing excessive weight gain in pregnancy. Cochrane Database Syst Rev 2015 Jun 15(6):CD007145. [doi: 10.1002/14651858.CD007145.pub3] [Medline: 26068707]
9. American College of Obstetricians and Gynecologists. ACOG Committee opinion no. 548: weight gain during pregnancy. Obstet Gynecol 2013 Jan;121(1):210-212. [doi: 10.1097/01.aog.0000425668.87506.4c] [Medline: 23262962]

10. National Institute for Health and Care Excellence. 2010 Jul. Preparing for pregnancy: women with a BMI of 30 or more URL: https://www.nice.org.uk/guidance/ph27/chapter/1-Recommendations#recommendation-1-preparing-for-pregnancy-women-with-a-bmi-of-30-or-more [accessed 2018-11-06]
11. Lee YS, Biddle S, Chan MF, Cheng A, Cheong M, Chong YS, et al. Health Promotion Board-Ministry of Health Clinical Practice Guidelines: Obesity. Singapore Med J 2016 Jun;57(6):292-300 [FREE Full text] [doi: 10.11622/smedj.2016103] [Medline: 27353244]
12. Nutrition During Pregnancy—Eating Right for Two. HealthHub URL: https://www.healthhub.sg/live-healthy/928/pregnancy-nutrition-during-pregnancy-eating-right-for-two [accessed 2018-11-06]
13. Flynn AC, Dalrymple K, Barr S, Poston L, Goff LM, Rogozińska E, i-WIP (International Weight Management in Pregnancy) Collaborative Group. Dietary interventions in overweight and obese pregnant women: a systematic review of the content, delivery, and outcomes of randomized controlled trials. Nutr Rev 2016 May;74(5):312-328. [doi: 10.1093/nutrit/nuw005] [Medline: 27083868]
14. Bennett CJ, Walker RE, Blumfield ML, Gwini S, Ma J, Wang F, et al. Interventions designed to reduce excessive gestational weight gain can reduce the incidence of gestational diabetes mellitus: A systematic review and meta-analysis of randomised controlled trials. Diabetes Res Clin Pract 2018 Jul;141:69-79. [doi: 10.1016/j.diabres.2018.04.010] [Medline: 29698713]
15. Chaplais E, Naughton G, Thivel D, Courteix D, Greene D. Smartphone Interventions for Weight Treatment and Behavioral Change in Pediatric Obesity: A Systematic Review. Telemed J E Health 2015 Oct;21(10):822-830. [doi: 10.1089/tmj.2014.0197] [Medline: 26290954]
16. Borgen I, Garnweidner-Holme LM, Jacobsen AF, Bjerkan K, Fayyad S, Joranger P, et al. Smartphone application for women with gestational diabetes mellitus: a study protocol for a multicentre randomised controlled trial. BMJ Open 2017 Mar 27;7(3):e013117 [FREE Full text] [doi: 10.1136/bmjopen-2016-013117] [Medline: 28348183]
17. Farzandipour M, Nabovati E, Anvari S, Vahedpoor Z, Sharif R. Phone-based interventions to control gestational weight gain: a systematic review on features and effects. Inform Health Soc Care 2018 Nov 27:1-16. [doi: 10.1080/17538157.2018.1540421] [Medline: 30479185]
18. Daly LM, Horey D, Middleton PF, Boyle FM, Flenady V. The Effect of Mobile App Interventions on Influencing Healthy Maternal Behavior and Improving Perinatal Health Outcomes: Systematic Review. JMIR Mhealth Uhealth 2018 Aug 09;6(8):e10012 [FREE Full text] [doi: 10.2196/10012] [Medline: 30093368]
19. Serné EH, de Jongh RT, Eringa EC, IJzerman RG, Stehouwer CDA. Microvascular dysfunction: a potential pathophysiological role in the metabolic syndrome. Hypertension 2007 Jul;50(1):204-211. [doi: 10.1161/HYPERTENSIONAHA.107.089680] [Medline: 17470716]
20. My Healthy Plate. 2010. Health Promotion Board URL: https://www.ntu.edu.sg/students/undergraduate/studentservices/healthandcounselling/documents/hpb_myhealthyplate_factsheet_fa(hires).pdf [accessed 2018-11-06]
21. Gilmore LA, Redman LM. Weight gain in pregnancy and application of the 2009 IOM guidelines: toward a uniform approach. Obesity (Silver Spring) 2015 Mar;23(3):507-511 [FREE Full text] [doi: 10.1002/oby.20951] [Medline: 25521748]
22. Crawford PB, Obarzanek E, Morrison J, Sabry ZI. Comparative advantage of 3-day food records over 24-hour recall and 5-day food frequency validated by observation of 9- and 10-year-old girls. J Am Diet Assoc 1994 Jun;94(6):626-630. [Medline: 8195550]
23. De Keyzer W, Huybrechts I, De Vriendt V, Vandevijvere S, Slimani N, Van Oyen H, et al. Repeated 24-hour recalls versus dietary records for estimating nutrient intakes in a national food consumption survey. Food Nutr Res 2011;55 [FREE Full text] [doi: 10.3402/fnr.v55i0.7307] [Medline: 22084625]
24. Ma Y, Olendzki BC, Pagoto SL, Hurley TG, Magner RP, Ockene IS, et al. Number of 24-hour diet recalls needed to estimate energy intake. Ann Epidemiol 2009 Aug;19(8):553-559 [FREE Full text] [doi: 10.1016/j.annepidem.2009.04.010] [Medline: 19576535]
25. Generate Nutrition Information Panel (NIP) Online. Health Promotion Board URL: https://focos.hpb.gov.sg/eservices/NIP/Add_Product.aspx [accessed 2018-11-06]
26. Chong MF, Chia A, Colega M, Tint M, Aris IM, Chong Y, GUSTO Study Group. Maternal Protein Intake during Pregnancy Is Not Associated with Offspring Birth Weight in a Multiethnic Asian Population. J Nutr 2015 Jun;145(6):1303-1310. [doi: 10.3945/jn.114.205948] [Medline: 25948786]
27. HealthHub. 2015. Healthy start for your pregnancy URL: https://www.healthhub.sg/live-healthy/840/healthy-start-for-your-pregnancy [accessed 2018-11-06]
28. Oteng-Ntim E, Varma R, Croker H, Poston L, Doyle P. Lifestyle interventions for overweight and obese pregnant women to improve pregnancy outcome: systematic review and meta-analysis. BMC Med 2012 May 10;10:47 [FREE Full text] [doi: 10.1186/1741-7015-10-47] [Medline: 22574949]
29. Thangaratinam S, Rogozinska E, Jolly K, Glinkowski S, Roseboom T, Tomlinson JW, et al. Effects of interventions in pregnancy on maternal weight and obstetric outcomes: meta-analysis of randomised evidence. BMJ 2012 May 16;344:e2088 [FREE Full text] [Medline: 22596383]
30. Wolff S, Legarth J, Vangsgaard K, Toubro S, Astrup A. A randomized trial of the effects of dietary counseling on gestational weight gain and glucose metabolism in obese pregnant women. Int J Obes (Lond) 2008 Mar;32(3):495-501. [doi: 10.1038/sj.ijo.0803710] [Medline: 18227847]

31. Thornton YS, Smarkola C, Kopacz SM, Ishoof SB. Perinatal outcomes in nutritionally monitored obese pregnant women: a randomized clinical trial. J Natl Med Assoc 2009 Jun;101(6):569-577. [Medline: 19585925]
32. Chan KL, Chen M. Effects of Social Media and Mobile Health Apps on Pregnancy Care: Meta-Analysis. JMIR Mhealth Uhealth 2019 Jan 30;7(1):e11836 [FREE Full text] [doi: 10.2196/11836] [Medline: 30698533]
33. Digital influence lab. Singapore Digital Marketing Statistics - 2015 URL: https://digitalinfluencelab.com/singapore-digital-marketing-stats/ [accessed 2018-11-06]

Abbreviations

KKH: KK Women’s and Children’s Hospital
RCT: randomized controlled trial
SMS: short messaging service

Edited by G Eysenbach; submitted 03.12.18; peer-reviewed by L Garnweidner-Holme, J Most, R Ciptaningtyas; comments to author 17.04.19; revised version received 22.05.19; accepted 21.07.19; published 24.10.19.

Please cite as:
Li LJ, Aris IM, Han WM, Tan KH
A Promising Food-Coaching Intervention Program to Achieve Optimal Gestational Weight Gain in Overweight and Obese Pregnant Women: Pilot Randomized Controlled Trial of a Smartphone App
JMIR Form Res 2019;3(4):e13013
URL: http://formative.jmir.org/2019/4/e13013/
doi:10.2196/13013
PMID:31651407

Original Paper

Accuracy of a Chatbot (Ada) in the Diagnosis of Mental Disorders: Comparative Case Study With Lay and Expert Users

Stefanie Maria Jungmann[1], PhD; Timo Klan[1], PhD; Sebastian Kuhn[2], MD; Florian Jungmann[2], MD

[1]Department of Psychology, Johannes Gutenberg-University Mainz, Mainz, Germany
[2]University Medical Center, Johannes Gutenberg-University Mainz, Mainz, Germany

Corresponding Author:
Stefanie Maria Jungmann, PhD
Department of Psychology, Johannes Gutenberg-University Mainz
Wallstraße 3
Mainz, 55122
Germany
Phone: 49 61313939201
Email: jungmann@uni-mainz.de

Abstract

Background: Health apps for the screening and diagnosis of mental disorders have emerged in recent years on various levels (eg, patients, practitioners, and public health system). However, the diagnostic quality of these apps has not been (sufficiently) tested so far.

Objective: The objective of this pilot study was to investigate the diagnostic quality of a health app for a broad spectrum of mental disorders and its dependency on expert knowledge.

Methods: Two psychotherapists, two psychology students, and two laypersons each read 20 case vignettes with a broad spectrum of mental disorders. They used a health app (Ada—Your Health Guide) to get a diagnosis by entering the symptoms. Interrater reliabilities were computed between the diagnoses of the case vignettes and the results of the app for each user group.

Results: Overall, there was a moderate diagnostic agreement (kappa=0.64) between the results of the app and the case vignettes for mental disorders in adulthood and a low diagnostic agreement (kappa=0.40) for mental disorders in childhood and adolescence. When psychotherapists applied the app, there was a good diagnostic agreement (kappa=0.78) regarding mental disorders in adulthood. The diagnostic agreement was moderate (kappa=0.55/0.60) for students and laypersons. For mental disorders in childhood and adolescence, a moderate diagnostic quality was found when psychotherapists (kappa=0.53) and students (kappa=0.41) used the app, whereas the quality was low for laypersons (kappa=0.29). On average, the app required 34 questions to be answered and 7 min to complete.

Conclusions: The health app investigated here can represent an efficient diagnostic screening or help function for mental disorders in adulthood and has the potential to support especially diagnosticians in their work in various ways. The results of this pilot study provide a first indication that the diagnostic accuracy is user dependent and improvements in the app are needed especially for mental disorders in childhood and adolescence.

(JMIR Form Res 2019;3(4):e13863) doi:10.2196/13863

KEYWORDS

artificial intelligence; eHealth; mental disorders; mHealth; screening; (mobile) app; diagnostic

Introduction

Background

Digital media have become enormously important in the health sector. Up to 80% of the internet users inform themselves on the Web about health [1], and about 60% of patients search for their symptoms on the internet before or after a visit to the doctor [2]. Experts estimate that there are over 380,000 health-related mobile apps worldwide [3].

Health apps play an important role not only in physical diseases but also particularly in mental health conditions and disorders [4-6]. For mental disorders, access to professional diagnosis and treatment is often difficult and delayed (eg, long waits and concerns about psychotherapy). In addition, there is considerable uncertainty in the population about the significance of the symptoms (eg, at what point feelings and behaviors are pathological). The advantage of health apps is low-threshold, locally and temporally flexible, and cost-efficient access [4]. The services are independent of the medical care situation, can be individually adapted and integrated into everyday life, and increase the self-help potential [5,7]. A systematic literature review showed that especially people who felt stigmatized by their problem or ashamed of it (eg, encopresis and eating disorders) use electronic mental health (e–mental health) [8]. Digital media can also be highly relevant for certain target groups. In mental disorders in childhood, for example, there are more possibilities for nonverbal recording of symptoms, and parents can be supported in coping with problems in everyday life [9]. As young people use new media every day (97% daily internet consumption), and mental health problems at this age are usually experienced as stigmatizing and shameful, the youth are considered particularly accessible to health apps [10]. In addition, health apps are promising for patients with chronic or recurrent phases of illness, which are particularly common in mental disorders.

In recent years, an enormous number of health apps have been developed for mental health conditions and disorders, the number of which is now hardly manageable. The proportion of health apps for mental health is about 29% of all health apps worldwide [11]. Health apps for mental health cover various areas of health promotion, prevention, screening and diagnostics, management, treatment, and aftercare [12]. These apps are usually aimed at consumers, that is, people suffering from symptoms. Recent developments also target professionals and, more recently, the public health care system (eg, pilot function and screening) [13,14].

Given the large number of health apps, the problem arises that they are used extensively but are usually not (sufficiently) evaluated and tested. Several reviews [15-18] have found that health apps for mental health have rarely been tested for their usefulness and effectiveness and often have ethical and legal shortcomings (eg, data privacy and safety). For example, Wisniewski et al [15] found that 15% to 45% of studied apps for anxiety, depression, and schizophrenia made medical claims, although these were rarely evidence-based, and no apps had Food and Drug Administration marketing approval. In addition, only 50% to 85% included a privacy policy [15]. Even if the apps have many benefits as described above, health-related internet use can also have negative or harmful effects on one’s emotional state and health behavior, as research shows, for example, on the phenomenon of cyberchondria [19]. Cyberchondria refers to an excessive health-related internet search resulting in an increase in emotional distress and health anxiety (eg, because of ambiguous information or serious disease [19]).

There is a particularly great need for research into apps for the screening or diagnosis of mental disorders [5]. This gap in research contrasts with the importance that diagnostic or screening tools can have, for example, in assigning patients to appropriate medical disciplines and practitioners.

Health Apps for Screening and Diagnosis of Mental Disorders

Regarding diagnostics using e–mental health, a distinction is made between the collection of objective and subjective data [5,20]. Objective data (mostly psychophysiological measures or behavioral activity) are recorded via sensors in or connected to the mobile phone or so-called wearables. For example, Valenza et al [21] showed that heart rate variability predicted mood swings in patients with a bipolar spectrum disorder. So far, there are few empirical findings on the use of wearables in mental disorders; only about 1.5% of studies on wearables deal with mental health [22]. A recent systematic review showed that objective data were promising in predicting moods and mood changes, but much more empirical evidence was needed to reliably evaluate potentials and risks [20].

There are countless health apps that assess subjective data, such as apps used for assessments (eg, Web-based questionnaires) or tracking (eg, monitoring mood or medication via diaries) [23]. Regarding self-report instruments that were adapted into a mobile phone app, there are few evaluated Web-based questionnaires on depression and posttraumatic stress disorder that showed a psychometric quality comparable with the paper-pencil version [23-25]. Some apps, such as Moodpath [26], include questions based on the operationalized diagnostic criteria of the International Classification of Diseases (ICD), tenth revision [27]. In Moodpath, users are asked different questions 3 times a day for 14 days according to the diagnostic criteria for depressive disorders. On the basis of the indicated symptom patterns, an algorithm determines possible depression (screening) and makes an assessment of severity. The results of diagnostic apps are often based on algorithms or artificial intelligence (AI), which means that computers can simulate complex human cognitions and actions.

Regarding mental tracking, a few apps on mood and affective disorders have been empirically investigated. For example, Hung et al [28] found in patients with depression that daily data on depression, anxiety, and sleep quality in a mobile phone app were significantly related to clinician-administered depression assessment at baseline. For bipolar affective disorder, a mobile phone app identified lower physical (location changes recorded via global positioning system) and social (outgoing messages) activities as significant predictors for increased depressive symptoms and lower physical but increased social activity for increased manic symptoms [29].

In contrast to apps for physical diseases (eg, Ada—Your Health Guide [30] and IBM Watson Health [31]), apps for mental health focus almost exclusively on a single symptom or single mental disorder, rather than on a broader spectrum. However, especially for the purpose of screening, it seems interesting and necessary at all 3 levels (eg, individual, practitioner, and public health system) that a single app asks for a variety of symptoms and mental disorders and provides information about the range of psychopathology. Only a few apps for mental health, such as WhatsMyM3 [32] (anxiety, depression, bipolar affective disorder, and posttraumatic stress) and T2 Mood Tracker [33] (anxiety, depression, head injury, and posttraumatic stress), assess multiple mental health conditions. However, these are usually limited to anxiety-depressive symptoms and have so far been little evaluated [23]. Therefore, in this study we used a medicine app that covers a wide range of physical and mental health conditions.

The aim of this pilot study was to test for the first time the diagnostic agreement of a medicine app and case vignettes over a broad spectrum of mental disorders. We expected at least moderate diagnostic agreement (ie, interrater reliability Cohen kappa≥0.41; hypothesis 1). As health apps are used both as a self-assessment at the consumer level and a diagnostic support system by experts and practitioners [34,35], we examined the diagnostic quality, depending on the user's level of expert knowledge (ie, 3 user groups: psychotherapists, psychology students, and laypersons). Given the less advanced state of development of diagnostic health apps for mental health than for physical diseases [5,36], we hypothesized that diagnostic accuracy for mental disorders is dependent on expert knowledge (eg, symptom checker includes fewer psychiatric terms, and alternative terms need to be entered; hypothesis 2).

Methods

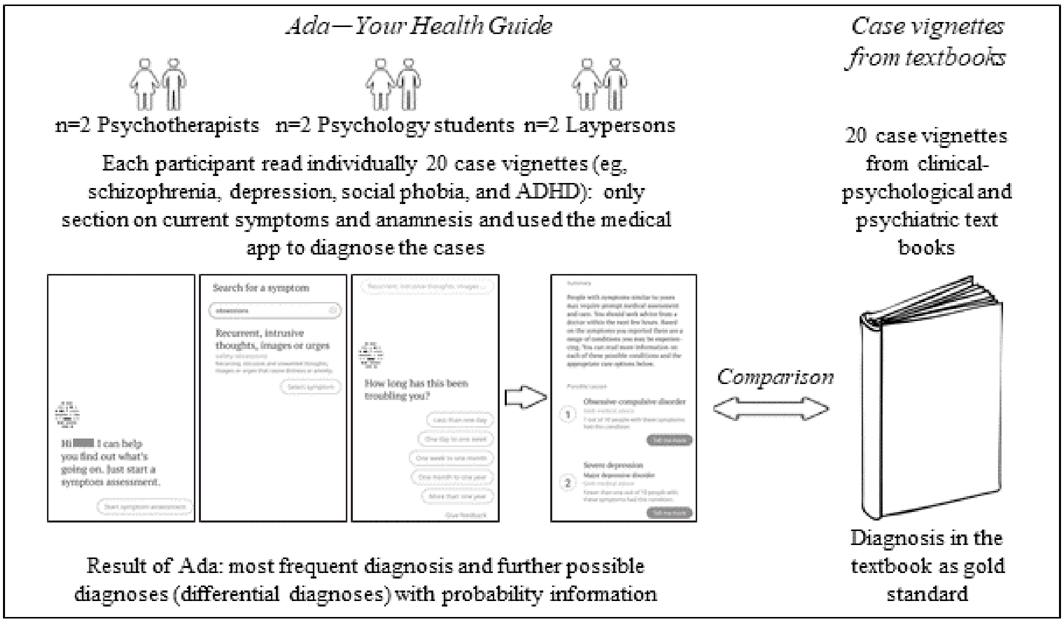
Design

A health app (Ada—Your Health Guide [30]) was used to diagnose 20 case vignettes from well-known textbooks of psychiatry and clinical psychology [37-40] by 3 groups: psychological psychotherapists, psychology students, and persons from the general population without previous professional knowledge of mental disorders (laypersons). Figure 1 illustrates the design and method.

Participants

Table 1 shows the sociodemographic characteristics of the participants.

Figure 1. Method and procedure of the study. ADHD: attention-deficit hyperactivity disorder.



Each participant read individually 20 case vignettes (eg, schizophrenia, depression, social phobia, and ADHD): only section on current symptoms and anamnesis and used the medical app to diagnose the cases

20 case vignettes from clinical-psychological and psychiatric text books

Hi [redacted] I can help you find out what's going on. Just start a symptom assessment.

How long has this been troubling you?

Result of Ada: most frequent diagnosis and further possible diagnoses (differential diagnoses) with probability information

Diagnosis in the textbook as gold standard

Table 1. Sociodemographic characteristics of the participants, subdivided into psychotherapists, psychology students, and laypersons.

Characteristics	Psychotherapists (n=2)	Psychology students (n=2)	Laypersons (n=2)	Statistics	
				P value	Effect size
Age (years), mean SD	40.0 (11.3)	22.0 (4.2)	40.0 (8.5)	.19	η^2_p=0.67
Sex, female, n (%)	1 (50)	1 (50)	1 (50)	>.99	Φ=0
Occupation, mean (SD) or with or without vocational training	14.0 (7.1) years professional experience[a]	4.0 (4.2) number of semesters	n=1 data scientist; n=1 without vocational training	—[b]	—[b]

[a] One of both is additionally a child and adolescent psychotherapist (first author).

[b] No comparative statistics possible due to different occupations.

Instruments

Health App for Diagnosis

Ada—Your Health Guide [30] is a Conformité Européenne–certified (ensures safe products within the European economic area) health app for the screening and diagnostic support of health conditions, primarily for physical diseases but increasingly also for mental health conditions and disorders. This app can be used both at the consumer level as a self-assessment app and by experts and practitioners as a diagnostic decision support system [35]. On the basis of AI, the chatbot asks for existing complaints adaptively and analogously to a medical or psychotherapeutic anamnesis interview. The Ada chatbot [30] is based on a medical database with constantly updated research findings. As a result, the diagnosis is determined that best matches the pattern of symptoms entered. The user is given a probability of a possible diagnosis and, as differential diagnoses, other less probable diagnoses (eg, 8/10 people with the symptoms described suffer from a depressive disorder). Patients or relatives (eg, parents and caregivers) receive an assessment of the urgency of seeking medical advice. Ada—Your Health Guide [35] was selected to investigate the above research questions for the following reasons (also in comparison with other symptom checkers [41,42]): (1) in addition to somatic symptoms, the app considers a wide spectrum of mental health conditions; (2) the app provides probabilities of possible and differential diagnoses (indications of comorbidities); (3) the app is widespread (>5 million users in >130 countries), publicly available, and free; (4) available in different languages (including English and German), and (5) in comparison with other symptom checkers (eg, Your.MD [43] and Babylon Health [44]), Ada provided more accurate diagnoses [42,45].

Knowledge of Mental Disorders

The user's knowledge of mental disorders was assessed on a 5-point Likert scale (1=not at all to 5=very good).

Procedure

After the informed consent, participants were instructed to carefully read the case vignette and then use the app to determine a diagnosis. A total of 20 case vignettes from psychiatry and clinical psychology textbooks were used, with 12 cases from adulthood [37,39] and 8 cases from childhood and adolescence [38,40]. All participants worked on one case after the other. The case vignettes were selected in such a way that a broad spectrum of mental disorders could be examined (see a list of mental disorders in Multimedia Appendix 1). The case vignettes included the initial symptoms before treatment (reason to seek treatment) and the anamnestic information, without naming or citing the diagnosis. The participants worked on the case with the health app on a tablet. The study duration was 3 to 6 hours per participant, divided into 2 to 3 individual sessions (most of the time was spent reading the 20 case vignettes). The participants (except the psychotherapists) received financial compensation (€10/hour) or a course credit (students).

Data Analysis

The main outcome was the agreement between the main diagnosis of the case vignette in the textbook and the result given by the app (the most probable diagnosis). Consistent labeling of the mental disorders was considered when assessing agreement. As an exception, the terms abuse and addiction were judged to agree, as the app did not distinguish between abuse and addiction to our knowledge. The diagnoses were compared at the level of 4-digit codes in the ICD (eg, anxiety disorders such as social anxiety and agoraphobia or personality disorders such as borderline personality disorder). If the 4th digit represents a more detailed specification (eg, obsessive-compulsive disorder: predominantly obsessive-compulsive behavior and thoughts or severity of the depressive episode), the 3-digit code match was counted (for the following disorders: depressive disorder, bipolar affective disorder, obsessive-compulsive disorder, conduct disorder, or schizophrenia). To consider the function and purpose of the screening and diagnostic app (eg, further diagnostic procedures required), no distinction was made between the subtypes of dementia (eg, Alzheimer and vascular dementia) and that of urinary incontinence (eg, stress incontinence and enuresis diurnal or nocturnal). The list of diagnoses in the textbooks and the results from the app can be found in Multimedia Appendix 1. The statistical outcomes were calculated as the percentage of agreement and the Cohen kappa coefficient (interrater reliability) for controlling random agreements. According to Landis and Koch [46], kappa values between 0.41 and 0.60 can be rated as moderate, between 0.61 and 0.80 as good, and >0.81 as very good. The agreement was checked if the secondary or differential diagnosis given by the app was also included (eg, bipolar disorder in the textbook but as a differential diagnosis in the app). Cohen *d* was calculated as effect size for group differences and partial eta-square for variance analyses. All statistical analyses were conducted using SPSS version 23 (IBM SPSS) [47], with an alpha level of .05. Following the study by Field [48], the Ryan, Einot, Gabriel, and Welsch Q procedure was used in post hoc tests to control the alpha error (same sample size; the Gabriel procedure was used when the sample sizes were different).

Results

Knowledge of Mental Disorders

Self-rated knowledge of mental disorders varied significantly depending on the group (ie, psychotherapists, students, or laypersons)—$F_{2,3}$=18.50; *P*=.02; partial eta-square=0.93. Post hoc analyses indicated that laypersons (mean 1.50, SD 0.71) reported significantly lower knowledge than students (mean 3.50, SD 0.71; *P*=.04) and psychotherapists (mean 5.00, SD 0; *P*=.01), with the last 2 groups having a marginally significant difference from each other (*P*=.08).

Percentage Agreement and Interrater Reliability

For mental disorders in adulthood, we found for the 72 case records (6 users×12 mental disorders), a percentage agreement of 68% and an interrater reliability according to Cohen kappa 0.64 between the textbook diagnosis and the result produced by the app. Taking into account the differential diagnoses, we found a percentage agreement of 85% and Cohen kappa 0.82. For mental disorders in childhood and adolescence, 48 case records (6 users×8 mental disorders) showed a percentage agreement of 42% (including differential diagnoses: 56%) and a Cohen kappa 0.40 (including differential diagnoses: kappa=0.52).

Table 2 shows the mean number (n), percentage (%), and Cohen kappa coefficients, differentiated among the 3 different user groups (ie, psychotherapists, students, and laypersons).

For mental disorders in adulthood, the Cohen kappa values were 0.78 (95% CI 0.60-0.95) for psychotherapists, 0.55 (95% CI 0.35-0.76) for students, and 0.60 (95% CI 0.39-0.80) for laypersons. Regarding case vignettes from childhood and adolescence, Cohen kappa values were numerically higher for psychotherapists (kappa=0.53, 95% CI 0.28-0.77) than for students (kappa=0.41, 95% CI 0.18-0.63) and laypersons (kappa=0.29, 95% CI 0.08-0.49).

Multimedia Appendix 1 lists the 20 mental disorders of the case vignettes as well as the main diagnoses in Ada Health and examples of differential diagnoses. The app mostly identified the main diagnosis (67% [8/12] of cases for adulthood and 44% [3.5/8] of cases for childhood and adolescence); it reported the differential diagnoses in an additional 17% (2/12) of cases for adulthood and 13% (1/8) of cases for childhood and adolescence. If the differential diagnoses are included, all diagnoses except undifferentiated somatization disorder, separation anxiety, and selective mutism in childhood were correctly detected.

Table 2. Mean number, percentage, and Cohen kappa coefficients for agreement between the textbook diagnosis and the result from Ada Health.

Case reports	Main diagnosis in Ada Health						Additional consideration of differential diagnoses in Ada Health					
	Psychotherapists		Students		Laypersons		Psychotherapists		Students		Laypersons	
	n (%)	kappa	n (%)	kappa	n (%)	kappa	n (%)	kappa	n (%)	kappa	n (%)	kappa
Adulthood (n_{max}=12)	9.5 (79)	0.78	7 (58)	0.55	7.5 (63)	0.60	11 (92)	0.91	10.5 (88)	0.87	8.5 (71)	0.69
Childhood and adolescence (n_{max}=8)	4.5 (56)	0.53	3.5 (44)	0.41	2.5 (31)	0.29	4.5 (56)	0.59	5 (63)	0.52	4 (40)	0.45

Number of Questions and Duration

To find a solution, the app had to ask an average of 34 questions per case (mean 33.78, SD 8.73) about the type and duration of the symptoms. There was no significant difference between the groups ($F_{2,117}$=1.89; *P*=.16; partial eta-square=0.03). The average time to complete was 409 seconds (SD 141.23). The groups differed in the average time for completion ($F_{2,96}$=9.93; *P*<.001; partial eta-square=0.17). Psychotherapists (mean 457.28, SD 138.61) and students (mean 415.82, SD 143.11), who did not differ from each other (*P*=.40), showed a significantly longer time for completion than the laypersons (the time recorded for only 1 layperson; mean 299.45, SD 141.23; *P*<.001).

Discussion

Principal Findings

In this pilot study, we tested whether a health app (Ada—Your Health Guide [30]) could detect mental disorders in children, adolescents, and adults. A total of 3 groups of users (ie, psychotherapists, psychology students, and laypersons) used the app to diagnose 20 case vignettes. Across all users, the agreement between the textbook diagnoses and the app was moderate (kappa=0.64) for mental disorders in adulthood and low (kappa=0.40) for that in childhood and adolescence. Adding differential diagnoses, good (kappa=0.82) and moderate (kappa=0.52) values, respectively, were obtained for interrater reliability.

When psychotherapists applied the app, there was a good agreement (kappa=0.78) between the results of the app and the diagnoses in the textbook on mental disorders in adulthood. This value is comparable with interrater reliabilities between 2 psychologists for diagnoses assessed with structured clinical interviews (kappa=0.71 for Axis I disorders and kappa=0.84 for personality disorders [49]). The diagnostic agreement was moderate (kappa=0.55/0.60) when students and laypersons used the app. The addition of differential diagnoses showed a good to very good interrater reliability (kappa=0.69-0.91). In 17% of the cases, the app did not give the diagnosis as the main diagnosis but as a differential diagnosis. Although the app assessed a different diagnosis as more likely, the main diagnosis of the case report was considered in some cases as a differential diagnosis.

For mental disorders in childhood and adolescence, a moderate diagnostic quality was found when psychotherapists (kappa=0.53) and students (kappa=0.41) used the app, whereas the quality was low for laypersons (kappa=0.29). In contrast to mental disorders in adulthood, the addition of differential diagnoses improved the diagnostic quality in childhood and adolescence to a lesser extent.

Taken together, only for mental disorders in adulthood, and when psychotherapists used the app, did Ada—Your Health Guide show good diagnostic quality. The app can serve as an indication of a mental health problem in the range of moderate agreement (adult mental disorders: students and laypersons; child and adolescent mental disorders: psychotherapists, students). With an average app time of 7 min, the app can be an efficient tool for the initial evaluation and screening of mental health problems and disorders. So, this pilot study indicates that expert knowledge tends to lead to better diagnostic quality when using the health app.

When comparing mental disorders in adulthood and childhood and adolescence, the app shows deficits for mental disorders in children and adolescents. For example, the app could not detect separation anxiety in childhood or selective mutism in any operation. On the one hand, this may be because of deficits in the app, on the other, mental disorders in childhood and adolescence are more often characterized by less specific symptom descriptions—children and adolescents show fewer specific symptoms and, from a developmental perspective, more frequent temporary subclinical symptoms [50]. This may also have led to confusion with the concrete naming and focusing of symptoms in childhood and adolescence. Examples include case reports on attention-deficit hyperactivity disorder (ADHD) and separation anxiety. In the ADHD case vignette, fears are mentioned first (eg, *would see ghosts*) [40]. In the case of separation anxiety, the initial focus is on describing the problematic relationship of the parents. In both cases, the hallmarks of the disorders are reported later and relatively profoundly. In addition, the app [30] may not include relevant terms and psychopathological characteristics, such as school fear and selective mutism. There is a clear need to catch up here. Especially in the case of enuresis, the results generated by the app, such as *mixed incontinence* or *stress incontinence*, made it clear that these were primarily terms pertaining to adults. As Ada—Your Health Guide [30] is based on a medical database with updated research findings, these deficits in the detection of mental disorders can also be because research activity in children and adolescents is significantly lower than that in adults. In the case of disorders with somatic symptoms (eg, undifferentiated somatization disorder), the diagnosis was more difficult because of the delimitation of psychological and physical symptoms. The overall interrater reliability in this study is lower than in studies that use structured clinical interviews [49].

It is important to consider the aims of screening and diagnostic apps. Health apps (eg, Ada—Your Health Guide [30]) do not aim to replace doctors or psychotherapists. Psychopathological symptoms can only be adequately understood and classified by a detailed anamnesis, the consideration of the temporal course, and the correct assessment of inclusion and exclusion criteria. For example, a severe, recurrent depressive disorder or multiple comorbidities worsen prognosis and require treatment (eg, combined treatment with psychotropic drugs) different from more circumscribed cases, such as a mild and single depressive episode. To our knowledge, there is currently no diagnostic app that captures this complexity (especially several comorbidities). Furthermore, the benefits of personal interaction should not be underestimated, as some behavioral abnormalities become apparent especially in direct contact (eg, hyperactivity or personality disorders), and unintended or intentional bias tendencies (eg, social desirability) can be more easily identified. Therefore, we consider the clarification of problems and diagnostics by experts to be of immense importance. The evaluated diagnostic apps should rather be regarded as low-cost, low-threshold, and time-efficient support in the diagnosis of mental disorders in adulthood [5]. There is great potential for the application of AI-supported diagnostics at the level of the consumer or patient, the experts, and the health care system, for example the following [14]:

- *Consumers and patients:* for example, screening of symptoms, combined with possible emotional relief for the affected person (eg, diagnosis as an explanation or treatment option) and a recommendation for action (eg, seeking medical advice).
- *Professionals:* for example, support in more efficient exploration and diagnosis (eg, bringing the result of the health app to the initial consultation), consideration and explanation of differential diagnoses, rapid reaction to significant symptoms (eg, suicidal intentions and alcohol consumption), and support in making indication decisions.
- *Macro/health care system:* for example, optimizing the assignment to treatment providers or treatment settings, supporting employees of other occupational groups in the health care system.

Limitations and Research Perspectives

In this study, the health app was only tested on case vignettes, and the user groups had a very small sample size. This limits the transferability of our results to everyday practice (low ecological validity). In addition, in the case of small samples, the performance of individual and outlier values plays a major role [51]. A recent study examined another symptom checker (Babylon Health [44]) that had comparable methodological limitations (case vignettes and small sample [52]). In contrast to this study, we investigated mental disorders for which the apps have so far been little developed, requiring a first pilot study. In addition, we focused on the question of whether the diagnostic quality is dependent on expert knowledge and examined the quality when experts, students, and laypersons used the app.

A next step will be to investigate the diagnostic accuracy of health apps for mental disorders in a direct interaction of practitioner and patient and with a larger sample. Depending on the research question, the design has to be differentiated. If the diagnostic quality is of interest, the agreement of the results of the app applied by the end user or patient could be compared with the current gold standard for the diagnosis of mental disorders, that is, structured or standardized interviews (eg, Diagnostic Interview for Mental Disorders [43]). If investigating the question of how well the health app can support clinicians in diagnosing mental disorders, the comparison of the clinical diagnosis with and without an additional health app should be examined. It should also be noted that the present design could not determine a match for *no diagnosis present* as the case vignettes always included a diagnosis. In a future naturalistic study with patients, this limitation would be removed.

Health apps are considered to be a support system rather than a substitute for doctors and psychotherapists, both by development companies and by doctors [53] and psychotherapists [5,54]. For example, a recent study [54] interviewed 720 general practitioners about future digitization in the health care system. Of them, 68% considered it unlikely that doctors would ever be replaced for diagnostic tasks. Previous findings on the appropriateness of the recommendation for further treatment vary between 33% [41] and 81% [55] agreement regarding the triage performance of the app and doctors or nurses, depending on, for example, the app used, the urgency of the treatment, and the judging person (doctor or nurse).

Combined with future research to test diagnostic accuracy, it would also be interesting to compare the extent to which differences exist when patients do the input themselves. As already mentioned, there is a clear need to catch up in the field of diagnostics in childhood and adolescence using the app tested here. Parents are often uncertain about the significance of existing symptoms, behavioral abnormalities, or developmental deficits. Even if electronic health systems are to be understood as diagnostic indications or screenings and not something that can replace a doctor or psychotherapist, such a system can provide parents with relevant information and initial instructions for action.

As the app is used particularly at the consumer level, and our pilot study indicated that diagnostic quality was lower among users from the general population and students, an important research perspective is to examine in which areas the weaknesses and deficits lie with nonprofessionals and how these can be addressed in further development. Such development could also be valuable, for example, for use in regions or countries with limited medical and psychotherapeutic care. The professional level would also benefit from a higher reliability of AI-supported diagnosis of mental disorders in childhood and adolescence. The fact that a patient is referred to an appropriate medical or psychotherapeutic specialty, for example, has relevant effects on the patient and the physician and can have considerable health economic effects.

As health apps collect and process highly sensitive health data, data security is of immense importance. Frequent shortcomings of current health apps are inadequate information about the nature and purpose of further processing of the data, missing or excessively complex data privacy statements, and comparatively easy access and manipulation by third parties [6,18,56]. Health apps should increasingly be certified based on defined catalogues of criteria and provided with a seal of quality, although this has rarely been done to date [57]. Overall, challenges remain to improve data security and the standardization of quality assurance, in particular, transparency for users, data protection control, and the handling of big data [14,36,57].

Conclusions

Health-related apps are also widely used for mental health conditions and disorders (in the general population and increasingly by practitioners and the public health system), but little is known about the diagnostic quality of health apps for mental disorders. This pilot study found that the diagnostic agreement between the health app and the diagnosis of the case vignettes for mental disorders was overall low to moderate. The diagnostic quality was shown to be dependent on the user and the type of mental disorder. Only when psychotherapists used the app for mental disorders in adulthood, good diagnostic agreements were found. Therefore, the health app should be used with caution in the general population and should be considered as a first indication of possible mental health conditions. In particular, improvements in the app with regard to mental disorders in childhood and adolescence and further research are needed.

Acknowledgments

The authors would like to thank Louisa Wagner for her support in the recruitment of participants and data collection.

Authors' Contributions

SMJ, SK, and FJ were involved in the study concept and design. SMJ and TK used the app as psychotherapists and were involved in statistical analysis. SMJ, TK, SK, and FJ interpreted and discussed the results. SJ was the main author of the first version of the paper; TK, SK, and FJ completed it, and all authors agreed to the final version.

Conflicts of Interest

None declared.

Multimedia Appendix 1
List of mental disorders.
[PDF File (Adobe PDF File), 96 KB - formative_v3i4e13863_app1.pdf]

References

1. Muse K, McManus F, Leung C, Meghreblian B, Williams JM. Cyberchondriasis: fact or fiction? A preliminary examination of the relationship between health anxiety and searching for health information on the Internet. J Anxiety Disord 2012 Jan;26(1):189-196. [doi: 10.1016/j.janxdis.2011.11.005] [Medline: 22137465]
2. Bertelsmann Stiftung. Gesundheitsinfos Wer suchet, der findet – Patienten mit Dr. Google zufrieden. [Health information Who searches, finds - patients are satisfied with Dr. Google.] URL: https://www.bertelsmann-stiftung.de/fileadmin/files/BSt/Publikationen/GrauePublikationen/VV_SpotGes_Gesundheitsinfos_final.pdf [accessed 2019-02-11] [WebCite Cache ID 766oJS9BN]
3. Lucht M, Boeker M, Krame U. University of Freiburg. 2015. Gesundheits-und Versorgungs-Apps Hintergründe zu deren Entwicklung und Einsatz. [Health and care apps: Backgrounds to their development and use.] URL: https://www.uniklinik-freiburg.de/fileadmin/mediapool/09_zentren/studienzentrum/pdf/Studien/150331_TK-Gesamtbericht_Gesundheits-und_Versorgungs-Apps.pdf [accessed 2019-02-11] [WebCite Cache ID 766lwArcr]
4. Bakker D, Kazantzis N, Rickwood D, Rickard N. Mental health smartphone apps: review and evidence-based recommendations for future developments. JMIR Ment Health 2016 Mar 1;3(1):e7 [FREE Full text] [doi: 10.2196/mental.4984] [Medline: 26932350]

5. Lüttke S, Hautzinger M, Fuhr K. E-Health in Diagnostik und Therapie psychischer Störungen: Werden Therapeuten bald überflüssig? [E-Health in diagnosis and therapy of mental disorders: Will therapists soon become superfluous?]. Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz 2018 Mar;61(3):263-270. [doi: 10.1007/s00103-017-2684-9] [Medline: 29318339]
6. Magee JC, Adut S, Brazill K, Warnick S. Mobile app tools for identifying and managing mental health disorders in primary care. Curr Treat Options Psychiatry 2018 Sep;5(3):345-362 [FREE Full text] [doi: 10.1007/s40501-018-0154-0] [Medline: 30397577]
7. Löcherer R, Apolináro-Hagen J. Efficacy and acceptance of webbased programs for self-help for facilitation of psycological health and handling of stress during academic studies. Wirksamkeit und Akzeptanz von webbasierten Selbsthilfeprogrammen zur Förderung psychischer Gesundheit und zur Stressbewältigung im Studium: Ein Scoping-Review der aktuellen Forschungsliteratur. [Efficacy and acceptance of webbased programs for self-help for facilitation of psycological health and handling of stress during academic studies.] URL: https://www.e-beratungsjournal.net/ausgabe_0117/Loecherer_Apolinario-Hagen.pdf [accessed 2019-02-11] [WebCite Cache ID 766mIUNWj]
8. Griffiths F, Lindenmeyer A, Powell J, Lowe P, Thorogood M. Why are health care interventions delivered over the internet? A systematic review of the published literature. J Med Internet Res 2006 Jun 23;8(2):e10 [FREE Full text] [doi: 10.2196/jmir.8.2.e10] [Medline: 16867965]
9. Döpfner M, Schürmann S, Lehmkuhl G. Wackelpeter und Trotzkopf. Hilfen für Eltern bei ADHS-Symptomen, hyperkinetischem und oppositionellem Verhalten. Mit Online-Materialien. 4., überarb. Aufl. [Jello-Pete and stubbornness: Help for Parents with ADHS-symptoms, hyperkinetic and opposing behavior. With online materials. 4th, revised edition.]. Weinheim, Basel: Beltz; 2011.
10. D'Alfonso S, Santesteban-Echarri O, Rice S, Wadley G, Lederman R, Miles C, et al. Artificial intelligence-assisted online social therapy for youth mental health. Front Psychol 2017;8:796 [FREE Full text] [doi: 10.3389/fpsyg.2017.00796] [Medline: 28626431]
11. Anthes E. Mental health: there's an app for that. Nature 2016 Apr 7;532(7597):20-23. [doi: 10.1038/532020a] [Medline: 27078548]
12. 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]
13. de Rosis S, Nuti S. Public strategies for improving eHealth integration and long-term sustainability in public health care systems: Findings from an Italian case study. Int J Health Plann Manage 2018 Jan;33(1):e131-e152 [FREE Full text] [doi: 10.1002/hpm.2443] [Medline: 28791771]
14. Kuhn S, MME, Jungmann S, Jungmann F. Deutsches Ärzteblatt International. Künstliche Intelligenz für Ärzte und Patienten: „Googeln" war gestern. [Artificial intelligence for doctors and patients: 'Googling' was yesterday.] URL: https://www.aerzteblatt.de/archiv/198854 [accessed 2019-02-11] [WebCite Cache ID 766okKhVs]
15. Wisniewski H, Liu G, Henson P, Vaidyam A, Hajratalli NK, Onnela J, et al. Understanding the quality, effectiveness and attributes of top-rated smartphone health apps. Evid Based Ment Health 2019 Feb;22(1):4-9. [doi: 10.1136/ebmental-2018-300069] [Medline: 30635262]
16. Torous J, Roberts LW. The ethical use of mobile health technology in clinical psychiatry. J Nerv Ment Dis 2017 Jan;205(1):4-8. [doi: 10.1097/NMD.0000000000000596] [Medline: 28005647]
17. Radovic A, Vona PL, Santostefano AM, Ciaravino S, Miller E, Stein BD. Smartphone applications for mental health. Cyberpsychol Behav Soc Netw 2016 Jul;19(7):465-470 [FREE Full text] [doi: 10.1089/cyber.2015.0619] [Medline: 27428034]
18. Grist R, Porter J, Stallard P. Mental health mobile apps for preadolescents and adolescents: a systematic review. J Med Internet Res 2017 May 25;19(5):e176 [FREE Full text] [doi: 10.2196/jmir.7332] [Medline: 28546138]
19. McMullan RD, Berle D, Arnáez S, Starcevic V. The relationships between health anxiety, online health information seeking, and cyberchondria: systematic review and meta-analysis. J Affect Disord 2019 Feb 15;245:270-278. [doi: 10.1016/j.jad.2018.11.037] [Medline: 30419526]
20. Dogan E, Sander C, Wagner X, Hegerl U, Kohls E. Smartphone-based monitoring of objective and subjective data in affective disorders: where are we and where are we going? Systematic review. J Med Internet Res 2017 Jul 24;19(7):e262 [FREE Full text] [doi: 10.2196/jmir.7006] [Medline: 28739561]
21. Valenza G, Nardelli M, Lanata' A, Gentili C, Bertschy G, Kosel M, et al. Predicting mood changes in bipolar disorder through heartbeat nonlinear dynamics. IEEE J Biomed Health Inform 2016 Jul;20(4):1034-1043. [doi: 10.1109/JBHI.2016.2554546] [Medline: 28113920]
22. Baig MM, GholamHosseini H, Moqeem AA, Mirza F, Lindén M. A systematic review of wearable patient monitoring systems - current challenges and opportunities for clinical adoption. J Med Syst 2017 Jul;41(7):115. [doi: 10.1007/s10916-017-0760-1] [Medline: 28631139]
23. van Ameringen M, Turna J, Khalesi Z, Pullia K, Patterson B. There is an app for that! The current state of mobile applications (apps) for DSM-5 obsessive-compulsive disorder, posttraumatic stress disorder, anxiety and mood disorders. Depress Anxiety 2017 Jun;34(6):526-539. [doi: 10.1002/da.22657] [Medline: 28569409]

24. Fann JR, Berry DL, Wolpin S, Austin-Seymour M, Bush N, Halpenny B, et al. Depression screening using the Patient Health Questionnaire-9 administered on a touch screen computer. Psychooncology 2009 Jan;18(1):14-22 [FREE Full text] [doi: 10.1002/pon.1368] [Medline: 18457335]
25. Bush NE, Skopp N, Smolenski D, Crumpton R, Fairall J. Behavioral screening measures delivered with a smartphone app: psychometric properties and user preference. J Nerv Ment Dis 2013 Nov;201(11):991-995. [doi: 10.1097/NMD.0000000000000039] [Medline: 24177488]
26. Goering M, Frauendorf F. Moodpath App. 2019. URL: https://mymoodpath.com/en/ [accessed 2019-02-11] [WebCite Cache ID 766oLDIU2]
27. World Health Organization. The ICD-10 Classification of Mental and Behavioural Disorders: Clinical Descriptions and Diagnosis Guidelines. Geneva: World Health Organization; 1992.
28. Hung S, Li M, Chen Y, Chiang J, Chen Y, Hung GC. Smartphone-based ecological momentary assessment for Chinese patients with depression: an exploratory study in Taiwan. Asian J Psychiatr 2016 Oct;23:131-136. [doi: 10.1016/j.ajp.2016.08.003] [Medline: 27969071]
29. Beiwinkel T, Kindermann S, Maier A, Kerl C, Moock J, Barbian G, et al. Using smartphones to monitor bipolar disorder symptoms: a pilot study. JMIR Ment Health 2016 Jan 6;3(1):e2 [FREE Full text] [doi: 10.2196/mental.4560] [Medline: 26740354]
30. Ada: Your health companion. URL: https://ada.com/ [accessed 2019-02-11] [WebCite Cache ID 766o3R2hN]
31. IBM. IBM Watson Health. URL: https://www.ibm.com/watson/health/ [accessed 2019-02-11] [WebCite Cache ID 766oiCB1k]
32. Hurowitz G, Post R. Whats My M3 - M3 Information. 2018. URL: https://whatsmym3.com/ [accessed 2019-02-11] [WebCite Cache ID 766oMNPro]
33. Military Health System. MHSRS 2019. URL: https://www.health.mil/Military-Health-Topics/Research-and-Innovation/MHSRS-2018 [accessed 2019-02-11] [WebCite Cache ID 766nvY0HA]
34. Albrecht U, Afshar K, Illiger K, Becker S, Hartz T, Breil B, et al. Expectancy, usage and acceptance by general practitioners and patients: exploratory results from a study in the German outpatient sector. Digit Health 2017;3:2055207617695135 [FREE Full text] [doi: 10.1177/2055207617695135] [Medline: 29942582]
35. Hoffmann H. ITU: Committed to connecting the world. Ada Health Our Approach to Assess Ada's Diagnostic Performance. URL: https://www.itu.int/en/ITU-T/Workshops-and-Seminars/20180925/Documents/3_Henry%20Hoffmann.pdf [accessed 2019-05-10]
36. Millenson ML, Baldwin JL, Zipperer L, Singh H. Beyond Dr. Google: the evidence on consumer-facing digital tools for diagnosis. Diagnosis (Berl) 2018 Sep 25;5(3):95-105. [doi: 10.1515/dx-2018-0009] [Medline: 30032130]
37. Stieglitz R, Baumann U, Perrez M. Fallbuch zur klinischen Psychologie und Psychotherapie. [Case Book on Clinical Psychology and Psychotherapy.]. Bern: Huber; 2007.
38. Petermann F. Fallbuch der klinischen Kinderpsychologie. [Casebook of Clinical Child Psychology.]. Göttingen: Hogrefe; 2009.
39. Freyberger H, Dilling H. Fallbuch Psychiatrie: Kasuistiken zum Kapitel V(F) der ICD-10. 2 Auflage. [Case Book Psychiatry: Casuistics on Chapter V(F) of the ICD-10. 2nd edition.] . Bern: Huber; 2014.
40. Poustka F, van Goor-Lambo G. Fallbuch Kinder- und Jugendpsychiatrie. Erfassung und Bewertung belastender Lebensumstände von Kindern nach Kapitel V (F) der ICD-10. [Case Book Child and Adolescent Psychiatry. Recording and evaluation of stressful life circumstances of children according to Chapter V (F) of the ICD-10.]. Göttingen: Hogrefe; 2008.
41. Semigran HL, Linder JA, Gidengil C, Mehrotra A. Evaluation of symptom checkers for self diagnosis and triage: audit study. Br Med J 2015 Jul 8;351:h3480 [FREE Full text] [doi: 10.1136/bmj.h3480] [Medline: 26157077]
42. WIRED. Can You Really Trust the Medical Apps on Your Phone? URL: https://www.wired.co.uk/article/health-apps-test-ada-yourmd-babylon-accuracy [accessed 2019-05-10]
43. Your.MD - Health Guide and Symptom Checker. URL: https://www.your.md/ [accessed 2019-05-10]
44. Babylon Health. URL: https://www.babylonhealth.com/ [accessed 2019-05-10]
45. Armstrong S. The apps attempting to transfer NHS 111 online. Br Med J 2018 Jan 15;360:k156. [doi: 10.1136/bmj.k156] [Medline: 29335297]
46. Landis JR, Koch GG. An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. Biometrics 1977 Jun;33(2):363-374. [doi: 10.2307/2529786] [Medline: 884196]
47. IBM. 2015. IBM SPSS Statistics. URL: https://www.ibm.com/in-en/products/spss-statistics [accessed 2019-10-03]
48. Field A. Discovering Statistics Using IBM SPSS Statistics. Fourth Edition. Thousand Oaks, California, United States: SAGE; 2013.
49. Lobbestael J, Leurgans M, Arntz A. Inter-rater reliability of the Structured Clinical Interview for DSM-IV Axis I Disorders (SCID I) and Axis II Disorders (SCID II). Clin Psychol Psychother 2011;18(1):75-79. [doi: 10.1002/cpp.693] [Medline: 20309842]

50. Merten EC, Cwik JC, Margraf J, Schneider S. Overdiagnosis of mental disorders in children and adolescents (in developed countries). Child Adolesc Psychiatry Ment Health 2017;11:5 [FREE Full text] [doi: 10.1186/s13034-016-0140-5] [Medline: 28105068]
51. Fraser H, Coiera E, Wong D. Safety of patient-facing digital symptom checkers. Lancet 2018 Nov 24;392(10161):2263-2264. [doi: 10.1016/S0140-6736(18)32819-8] [Medline: 30413281]
52. Razzaki S, Baker A, Perov Y, Middleton K, Baxter J, Mullarkey D, et al. arXiv. 2018. A Comparative Study of Artificial Intelligence and Human Doctors for the Purpose of Triage and Diagnosi. URL: https://marketing-assets.babylonhealth.com/press/BabylonJune2018Paper_Version1.4.2.pdf [accessed 2019-09-27]
53. WIRED. Stop Googling Your Symptoms – The Smartphone Doctor is Here to Help URL: https://www.wired.co.uk/article/ada-smartphone-doctor-nhs-gp-video-appointment [accessed 2019-05-10]
54. Blease C, Bernstein MH, Gaab J, Kaptchuk TJ, Kossowsky J, Mandl KD, et al. Computerization and the future of primary care: a survey of general practitioners in the UK. PLoS One 2018;13(12):e0207418 [FREE Full text] [doi: 10.1371/journal.pone.0207418] [Medline: 30540791]
55. Verzantvoort NC, Teunis T, Verheij TJ, van der Velden AW. Self-triage for acute primary care via a smartphone application: practical, safe and efficient? PLoS One 2018;13(6):e0199284 [FREE Full text] [doi: 10.1371/journal.pone.0199284] [Medline: 29944708]
56. Morera EP, Díez I, Garcia-Zapirain B, López-Coronado M, Arambarri J. Security recommendations for mHealth apps: elaboration of a developer's guide. J Med Syst 2016 Jun;40(6):152. [doi: 10.1007/s10916-016-0513-6] [Medline: 27147515]
57. Albrecht U, Hillebrand U, von Jan U. Relevance of trust marks and CE labels in German-language store descriptions of health apps: analysis. JMIR Mhealth Uhealth 2018 Apr 25;6(4):e10394 [FREE Full text] [doi: 10.2196/10394] [Medline: 29695374]

Abbreviations

ADHD: attention-deficit hyperactivity disorder
AI: artificial intelligence
e–mental health: electronic mental health
ICD: International Classification of Diseases

Edited by G Eysenbach; submitted 28.02.19; peer-reviewed by H Tunnell, G Wadley; comments to author 29.04.19; revised version received 26.05.19; accepted 31.08.19; published 29.10.19.

Please cite as:
Jungmann SM, Klan T, Kuhn S, Jungmann F
Accuracy of a Chatbot (Ada) in the Diagnosis of Mental Disorders: Comparative Case Study With Lay and Expert Users
JMIR Form Res 2019;3(4):e13863
URL: http://formative.jmir.org/2019/4/e13863/
doi:10.2196/13863
PMID:31663858

Original Paper

Virtual Reality as a Therapy Adjunct for Fear of Movement in Veterans With Chronic Pain: Single-Arm Feasibility Study

Christopher A Fowler[1], PhD; Lisa M Ballistrea[1], DPT; Kerry E Mazzone[2], DPT; Aaron M Martin[2], PhD; Howard Kaplan[3], MEd; Kevin E Kip[1,4], PhD; Katherine Ralston[1], BSc; Jennifer L Murphy[2,5], PhD; Sandra L Winkler[1], OTR/L, PhD

[1]Research and Development Service, James A Haley Veterans Hospital, Tampa, FL, United States
[2]Chronic Pain Rehabilitation Program, James A Haley Veterans Hospital, Tampa, FL, United States
[3]Advanced Visualization Center, Information Technology, University of South Florida, Tampa, FL, United States
[4]College of Public Health, University of South Florida, Tampa, FL, United States
[5]Department of Neurology, College of Medicine, University of South Florida, Tampa, FL, United States

Corresponding Author:
Christopher A Fowler, PhD
Research and Development Service
James A Haley Veterans Hospital
8900 Grand Oak Circle
Tampa, FL, 33705
United States
Phone: 1 813 558 3904
Email: christopher.fowler3@va.gov

Abstract

Background: Virtual reality (VR) has demonstrated efficacy for distraction from pain-related thoughts and exposure to feared movements. Little empirical VR research has focused on chronic pain management.

Objective: The purpose of this study was to examine the feasibility of VR as an adjunctive intervention for Veterans with chronic pain. We designed a hierarchy ranging from low-intensity pain distraction to high-intensity movement-based exposure for this purpose. VR apps were mapped onto the hierarchy.

Methods: Sixteen Veterans receiving inpatient chronic pain rehabilitation participated in daily VR sessions over a 3-week period. Trajectories across the distraction-to-exposure hierarchy and Veteran-reported intensity ratings were described and evaluated over time. Minimum clinically important differences (MCIDs), pre-post effect sizes, and 95% confidence intervals were examined for fear of movement using the Fear of Daily Activities Questionnaire (FDAQ) and Pain Outcomes Questionnaire-VA (POQ-VA; fear scale). This approach was applied to secondary outcomes: POQ-VA (pain intensity, interference, negative affect), Pain Catastrophizing Scale, and Patient-Specific Functioning Scale (PSFS). Session attendance, completion, and VR experiences were described.

Results: Ten of 14 Veterans (71%) who participated in three or more VR sessions completed the distraction-to-exposure hierarchy. Only three trajectories emerged more than once. Due to high completion rates, Veterans that completed the hierarchy could self-select nonhierarchy apps. Veterans rated all hierarchy levels (low, medium, high) near medium intensity. Self-selected activities were rated as high intensity. For kinesiophobia, six Veterans (38%) exceeded the MCID on the FDAQ and a small effect size improvement was observed (Cohen d=−0.35). The confidence interval (95% CI −0.71 to 0.01) indicated the possibility of a null effect. The POQ-VA fear scale yielded no effect (Cohen d=0.06, 95% CI −0.43 to 0.54). For secondary outcomes, Veterans exceeding MCID were calculated with complete data: pain intensity (1/15, 7%), pain catastrophizing (5/14, 36%), and patient-specific functioning (10/15, 67%). Effect sizes were large for patient-specific functioning (Cohen d=1.14, 95% CI 0.50-1.78), medium for mobility interference (Cohen d=−0.56, 95% CI −0.96 to −0.16), and small for pain intensity (Cohen d=−0.40, 95% CI −0.69 to −0.12) and catastrophizing (Cohen d=−0.41, 95% CI −0.79 to −0.02). No effects were observed for interference in daily activities (Cohen d=0.10, 95% CI −0.27 to 0.47) and negative affect (Cohen d=0.07, 95% CI −0.26 to 0.40). Veterans attended 85.2% (98/108) of VR sessions and completed 95% (93/96) of sessions attended. Twenty-minute sessions were rated as too short. No significant adverse events were reported.

Conclusions: Findings support the feasibility of VR as an adjunct for Veterans with chronic pain. However, the hierarchy will require modification, as evidenced by homogeneous intensity ratings. Veteran-selected activities presented the highest intensity ratings, largest outcome effect size (PSFS), and MCID. This highlights the important role of utilizing Veteran stakeholders in hierarchy modification, design of VR interventions, and outcome selection.

(JMIR Form Res 2019;3(4):e11266) doi:10.2196/11266

KEYWORDS

chronic pain; virtual reality; Veterans; pain management; rehabilitation; fear of movement; kinesiophobia; exposure therapy; distraction therapy

Introduction

Virtual Immersion

Virtual environments present an opportunity to safely and gradually expose Veterans with chronic pain to movements they avoid in the real world. Virtual reality (VR) describes “computer-generated simulations of three-dimensional objects or environments with seemingly real, direct, or physical user interaction” [1]. VR technologies use wearable devices to project a virtual environment and to track movements within it [2]. Sensory inputs (eg, visual, audio, tactile) give the user the illusion of “immersion” or being cognitively absorbed by a virtual environment [3,4]. Interaction with the immersive environment allows the user to co-create their experience giving them a subjective sense of “presence” in the virtual environment [3,5]. Immersion and presence can facilitate interventions using simulated 3D environments that set VR apps apart from in vivo behavioral treatments [6] and low-immersion 2D apps (eg, mobile phone) [7]. VR can serve as an adjunct to assist with the adoption of pain management skills in evidence-based interventions [4,8].

Virtual Reality for Pain Management

Virtual reality apps have demonstrated efficacy and feasibility for delivering pain management skills, including distraction and exposure therapies. A rapid review (20 studies, N=337) found short-term pain reduction (strong evidence) and analgesic effects (moderate evidence) [4]. A meta-analysis of controlled studies (16 studies, N=656) estimated a medium effect size pain reduction when using VR during medical procedures [9]. To date, VR research has prominently focused on acute, not chronic, pain [4,9].

Chronic Versus Acute Pain

Pain is defined as “an unpleasant sensory and emotional experience associated with actual or potential tissue damage or described in terms of such damage” [10]. Postinjury acute pain is a nociceptive physiological warning to limit or avoid certain behaviors to prevent further harm [11]. Acute pain lasts 3 months or less and helps facilitate physical healing [12]. In chronic pain, generalized hypersensitivity in the central nervous system results in overactive pain [13] and sensory [14] pathways. Chronic pain persists beyond the physical healing process (ie, >3 months) and serves no adaptive purpose [12]. Still, chronic pain may feel indistinguishable from acute pain and leads people to avoid movement, which negatively impacts functioning [13,15].

Fear of Movement

The fear-avoidance model of chronic pain posits a feedback loop with fear as a key component of avoidant behavior [16]. People with chronic pain mistakenly believe that pain sensations signal harm. Cognitive biases in pain processing (eg, behaviors that aggravate pain should be avoided to prevent reinjury) can result in kinesiophobia, or fear of movement, and subsequent pain avoidance as a means of self-protection [16,17]. Avoidance promotes a self-perpetuating cycle of physical deconditioning, negative affect, disability, and worse pain [17-19]. Interventions that operate on fear-avoidance principles (eg, graded physical therapy, cognitive behavioral therapy) aim to disrupt this cycle through gradual exposure to feared movements [15,20]. Reengagement in feared movements can modify pain interpretations (eg, pain signals harm), disrupt fear avoidance, and combat physical deconditioning [20-22]. This is often challenging because even safe movement can cause pain and emotional discomfort.

Pain Distraction

The most common VR intervention has been distraction therapy [4,9] not exposure therapy. Distraction is hypothesized to be a mechanism of action for VR in attenuating pain [23]. Distraction therapy is based on the assumption that people have finite cognitive resources for information processing [24]. Immersive VR consumes cognitive and attentional resources through sensory input, thereby limiting pain-processing capabilities [23,25]. A meta-analysis (14 studies, N=581) found large effect size reductions in acute and laboratory-induced pain when using VR distraction versus controls [26].

Distraction is important in the cognitive behavioral therapy for chronic pain protocol of Veterans Affairs (VA) [15]. A paucity of studies supports VR distraction for chronic pain. A within-group pilot study found a large effect size improvement in pain intensity following VR use [27]. A randomized crossover study found reduced pain intensity during VR compared with self-mediated distraction (eg, meditation, gaming), but not at posttest [28]. Still, VR provided superior distraction than self-mediated methods as evidenced by 56% less time thinking about pain [28]. Thus, VR distraction may be beneficial for chronic pain management.

Graded Exposure

People with chronic pain benefit from rehabilitation, not just immediate relief [8,22]. As such, passive pain distraction and more activating therapies, including exposure to feared movements, are used in comprehensive protocols [15]. Exposure therapy is important for chronic pain rehabilitation and compatible with VR [8,22]. A randomized controlled trial (RCT) assigned people with acute and chronic low back pain to receive 10 sessions of physical therapy with or without VR exposure (“virtual walking”) [29]. Participants who received VR before each session experienced medium-to-large improvements in kinesiophobia, walking distance, and disability. A feasibility study of VR dodgeball found increased spine flexion, no adverse events, high acceptability, and likelihood to recommend the game to others with chronic back pain [30]. One study found that 76% of participants indicated a preference for VR exposure therapy over in vivo exposure before randomization [31]. Only 3% indicated that they would refuse VR compared with 27% for in vivo. This evidence supports VR exposure as an efficacious and acceptable adjunct for chronic pain management.

Distraction-to-Exposure Hierarchy

This study was informed by the fear-avoidance model of chronic pain [16], assuming that gradual exposure to feared movements using VR can improve kinesiophobia and pain outcomes. Still, people with chronic pain experience sensory exacerbation which may be a VR contraindication. Passive distraction therapy apps may be useful to gradually integrate Veterans to VR use. Thus, we created a hierarchy of increasing intensity based on sensory integration theory [32]. Veterans began with low movement-intensive VR distraction apps and could gradually progress to medium and high movement-intensity apps (exposure). This helps ensure that VR does not over- or understimulate the user.

Gaps Addressed

This study addresses notable gaps in VR research. First, there is a dearth of chronic pain studies in the growing literature about VR for pain management [9,33]. Second, VR studies have examined individual applications of pain distraction and exposure. Despite the efficacy of both therapies and distraction inherent in VR, no identified studies used a hierarchy approach [23]. Third, we failed to identify any VR pain management studies conducted with Veterans. More research is needed given the need for alternative nonpharmacological treatments for Veterans [34], who experience chronic pain with greater prevalence than non-Veterans [35].

Specific Aims

This study examined the feasibility of VR as an adjunct for chronic pain management. Evidence obtained is intended to inform a future RCT that will test the efficacy of VR and aim to validate our current and future revised distraction-to-exposure hierarchy. Primary aims are to (1) describe and compare the Veteran trajectories and self-reported app intensity ratings over a 3-week treatment period on the distraction-to-exposure hierarchy, (2) estimate the proportion of Veterans experiencing minimum clinically important differences (MCIDs) and within-subject effect size and 95% confidence intervals (CIs) for fear of movement and secondary pain outcomes associated with the use of VR, and (3) pilot test this protocol to assess the feasibility of VR use to plan for a future RCT.

Methods

Study Setting

This study was conducted in the Chronic Pain Rehabilitation Program at the James A Haley Veterans Hospital in Tampa, FL. This program used a cognitive behavioral approach to target the biopsychosocial impact of chronic pain [36]. Veterans were referred to this unique 19-day residential treatment program from throughout the VA system. For a more comprehensive overview of this program, see Murphy and colleagues [36,37].

Participants and Recruitment

Each week, up to four Veterans entered the 19-day Chronic Pain Rehabilitation Program, and up to four graduated and were discharged. All Veterans (N=18) that entered the program over the 3-week study period were targeted for recruitment; one Veteran refused. Of the 17 Veterans that consented, one did not meet the inclusion criteria. The final sample consisted of 16 Veterans (89%). Inclusion criteria were (1) diagnosis of chronic pain syndrome (*International Classification of Diseases [ICD], Ninth Revision, Clinical Modification* code 338.4 and *ICD-10* code G89.4) and (2) negative screen for illicit substances and unprescribed opioids. Exclusion criteria were uncontrolled medical and psychological factors (ie, aggression, depression, psychosis, suicidality) that could interfere with rehabilitation. These criteria mirrored that of the chronic pain program [36,37].

Design

A hybrid type 1 implementation-effectiveness design [38] was used to collect clinical efficacy and preimplementation data for VR. Qualitative and quantitative preimplementation data were collected following each VR session (aims 1 and 3). This paper emphasizes quantitative outcomes. Fear of movement and pain outcomes were assessed using a within-participants pretest-posttest design (aim 2).

Intervention

The distraction-to-exposure hierarchy was built with input from chronic pain program clinical stakeholders [39]. The hierarchy started with low stimulation intensity and then moved to high movement intensity. Twelve commercially available VR apps, six per head-mounted display (HMD), were then chosen to fit intensity levels. Low-intensity distraction apps included mindfulness meditation [40] and visual imagery [41,42], which required minimal movement. Medium-intensity apps included virtual walking or swimming [43,44] and controlling aircraft or watercraft [45,46], which required head and neck movement. High-intensity apps were 3D painting [47,48] and music or rhythmic-based [49,50], which also required torso and upper extremity movement. Veterans alternated between two commercially available VR HMDs: Oculus Rift [51] and Samsung Oculus Gear VR [52]. Rift is an HMD with hand-tracking controllers, which is used with commercial gaming computers. Gear VR uses Samsung Galaxy Series mobile phones (S6 and above) to project virtual environments with sound. Both HMDs have been used for pain management research in medical settings [28,53,54]. Figure 1 shows our team members using each HMD.

Figure 1. Samsung Oculus Gear VR with supplemental hand controller (left) and Oculus Rift (right).

Primary Measures

Daily Rating Form

The eight-item daily rating form was created by the investigators to assess VR feasibility. Following each VR session, this form was used to track Veteran progress across the hierarchy (app selection), which HMD they used, and the number of sessions attempted and completed. Veterans reported their level of immersion using a single ordinal item adapted from Cole et al [55]: 1=aware and acknowledge the role of technology; 2=partially aware of technology, but perceived being inside a virtual environment; and 3=unaware of technology (complete virtual immersion). In addition, self-reported VR intensity (1=low, 2=medium, 3=high) and VR session length (1=too short, 2=just right, 3=too long) followed a similar approach [55]. Veterans could also provide feedback about their VR experience via three open-ended questions: likes, dislikes, and additional comments. This helped identify any facilitators (eg, HMD preferences), barriers (physical, psychological discomfort), and adverse events (eg, cybersickness, falls).

Fear of Movement

Kinesiophobia was assessed using two measures: the Pain Outcomes Questionnaire-VA (POQ-VA) [56] and Fear of Daily Activities Questionnaire (FDAQ) [57]. These scales had modest convergent validity (r=.29, P=.28) suggesting they could be examined separately. Both scales have demonstrated acceptable psychometric properties in chronic pain studies [56,57].

The POQ-VA [56] is a multidimensional instrument developed specifically for the Veteran population. The fear subscale measures kinesiophobia using two items (fear of reinjury, safe to exercise) on Likert-type scales ranging from 0 to 10 and summed, with higher scores indicating positive outcomes. No MCID standards were identified.

The 10-item FDAQ [57] was designed in accordance with the fear-avoidance model. It was used to assess common feared movements (eg, sitting, standing, lifting, walking). All items are measured on scales anchored by 0 (no fear) and 100 (maximal fear) and then averaged. The MCID for the FDAQ is a 12.90-point reduction from baseline [57].

Secondary Measures

Pain Outcomes Questionnaire-VA

Secondary outcomes were examined to identify promising outcomes for use with a future VR RCT [39]. Multiple secondary pain outcomes were collected using the POQ-VA [56]. These were interference with activities of daily living and mobility as well as negative affect. No MCID scores were identified. The POQ-VA also assesses pain intensity using the common pain Numeric Rating Scale [58] (0=no pain at all to 10=worst possible pain). The Numeric Rating Scale MCID is 2.10 points for moderate pain (baseline=4-6) and 2.80 points for severe pain (baseline ≥7) [59].

Pain Catastrophizing Scale

The 13-item Pain Catastrophizing Scale [60] was used to measure exaggerated beliefs about pain (eg, nothing I can do to reduce pain). Items are measured on a Likert-type scale anchored by 0 (not at all) and 4 (very) and summed with higher scores indicating maladaptive beliefs about pain. MCIDs of 38% or greater have been established for improved disability and pain intensity following pain rehabilitation [61].

Patient-Specific Functional Scale

The Patient-Specific Functional Scale (PSFS) [62] required Veterans to identify three activities that have been hindered because of their pain. These tasks were then rated 0 (unable to perform) to 10 (able to perform at prior level) and averaged.

The MCID for the PSFS is 1.30 to 2.29 points (small), 2.30 to 2.69 (medium), and 2.70 or higher (large) [63].

Procedures

Veterans were informed about the study during orientation to the pain program. Consenting procedures were performed in-person by the research team before their first physical therapy session. All study procedures were approved by the James A Haley VA Research and Development Committee and the University of South Florida (Tampa) Institutional Review Board (protocol: 00031503).

Veterans completed 20 minutes of VR during daily physical therapy sessions. During session 1, the hierarchy was described to the Veterans, and they began with low intensity guided meditation. Each session, they were asked which intensity VR APPS they would like to use that session. Following each session, research staff administered the daily rating form. Primary and secondary outcome measures were administered to Veterans at intake and discharge (approximately 3 weeks) to the chronic pain program to track improvements. The research staff retrieved these data from the VA's electronic medical record.

Statistical Analysis

To address aim 1 (describe and compare the Veteran trajectories and self-reported app intensity ratings on the distraction-to-exposure hierarchy), distributions of the Veteran-selected apps (proposed intensity range 1-3) were plotted across VR sessions. The frequency of Veteran trajectories toward completing the hierarchy were counted to identify common patterns. Veterans that completed less than three VR sessions were excluded from this analysis because the hierarchy could not be completed in two sessions. Veterans' median self-reported app intensity ratings were also calculated and plotted across the first nine VR sessions. Sessions 10 and 11 were excluded due to the low frequency of Veterans (N≤2) that attended more than nine sessions. Consistencies across proposed and self-reported VR app intensity were descriptively compared.

For aim 2 (estimate the proportion of Veterans experiencing MCIDs and within-subject effect size and 95% CI for fear of movement and secondary pain outcomes associated with the use of VR), changes in fear of movement and secondary outcomes were calculated and compared with established MCID scores for each respective measure excluding imputed missing values. The proportion of Veterans that exceeded MCID was calculated for each outcome. Within-participants Cohen *d* effect sizes ([post mean−pre mean]/SD difference) and 95% CI for fear of movement and secondary outcomes were calculated to examine the efficacy of VR [64,65]. Suspected outliers were assessed using multiple criteria. These included examining boxplots of these pre-to-post test change scores, Tukey fences [66], and clinical observation and consultation noted during the study. In the case of a suspected outlier, findings were reported with and without its inclusion.

For aim 3 (pilot test this protocol to assess the feasibility of VR use), compliance for the Veteran sample was calculated via the proportion of VR sessions attended (sessions attended/total scheduled sessions). Adherence was calculated using the proportion of completed 20-minute VR sessions (full sessions completed/sessions attended). Veteran ratings of the length of the session and their self-reported levels of immersion were plotted across sessions.

Availability of Data and Materials

The final deidentified datasets from this study (qualitative and quantitative) and the VR user manual will be made available by the corresponding author on reasonable request.

Results

Demographic Characteristics

Veterans ranged from 28 to 63 years, with a mean age of 49 (SD 12) years (Table 1). They were prominently male, and their racial/ethnic composition was primarily Caucasian or white. Median pain duration was 16.50 (IQR 14.62) years and baseline pain intensity was near the severe range (≥7) on the Numeric Rating Scale [38]. The primary pain location was low back; daily opioid use was low.

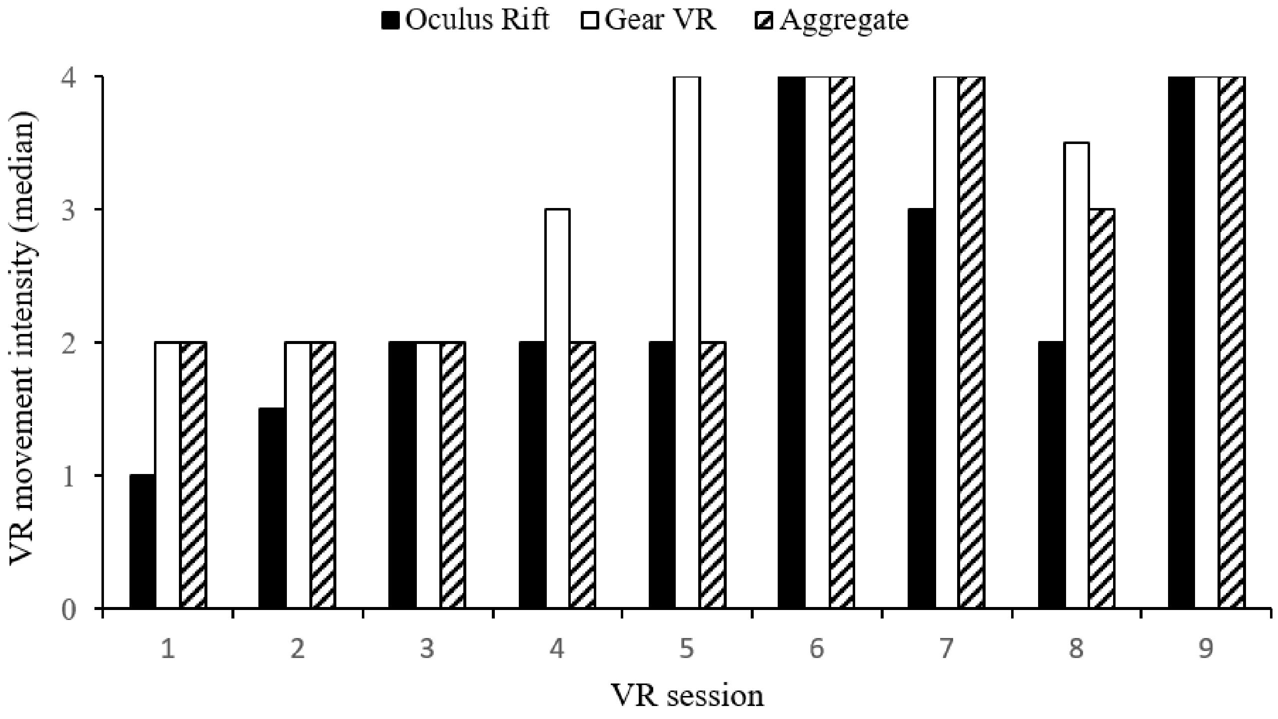
Table 1. Demographic characteristics for study sample (N=16).

Characteristic	Participants
Age (years), mean (SD)	48.88 (11.62)
Gender, n (%)	
Female	3 (19)
Male	13 (81)
Ethnicity, n (%)	
African American or black	4 (25)
Caucasian or white	8 (50)
Hispanic or Latino	2 (12)
Other	2 (13)
Pain location, n (%)	
Head	2 (13)
Low back	11 (69)
Other	3 (19)
Pain chronicity (years), median (IQR)	16.50 (14.62)
Virtual reality sessions, median (IQR)	7.50 (6.50)

Aim 1: Describe and Compare the Veteran Trajectories and Self-Reported App Intensity Ratings on the Distraction-to-Exposure Hierarchy

In total, 10 of 14 Veterans (71%) who participated in three or more VR sessions completed the hierarchy. Eleven different trajectories emerged among these 14 Veterans during completion of the hierarchy. Only three patterns (21%) occurred more than once, which indicated notable variability. Six of the initial eight Veterans (75%) reached the highest level of the hierarchy during the first week of testing. Because of this high frequency early in the study, Veterans that completed the hierarchy could self-select VR activities (eg, fishing, basketball free throws) in addition to hierarchy apps. Self-selected apps were assigned a fourth intensity level indicating they were beyond the hierarchy. Median progressions across the distraction-to-exposure hierarchy are presented in Figure 2.

Figure 2. Median Veteran progression across the distraction-to-exposure hierarchy. App movement intensity: 1=low, 2=medium, 3=high, 4=self-selected. VR: virtual reality.



Veterans rated the intensity of each VR session. Low- to medium-intensity ratings were reported for each of the first four sessions (median 2, IQR 1.5). Sessions five to nine were rated from medium to high intensity (median 3, IQR 1; see Figure 3). Comparisons between Veterans' self-reported intensity ratings and hierarchy movement intensity levels were examined. Veterans rated low movement intensity apps (ie, distraction) between low and medium intensity (median 1.5, IQR 1). Both medium- (median 2, IQR 1) and high-intensity (median 2, IQR 1) movement exposure apps were rated as medium intensity. Veteran-selected apps were rated as high intensity (median 3, IQR 1). Veteran intensity ratings across hierarchy levels are presented in Figure 4.

Figure 3. Veteran-reported intensities for virtual reality (VR) apps across sessions. Self-reported app intensity: 1=low, 2=medium, 3=high.

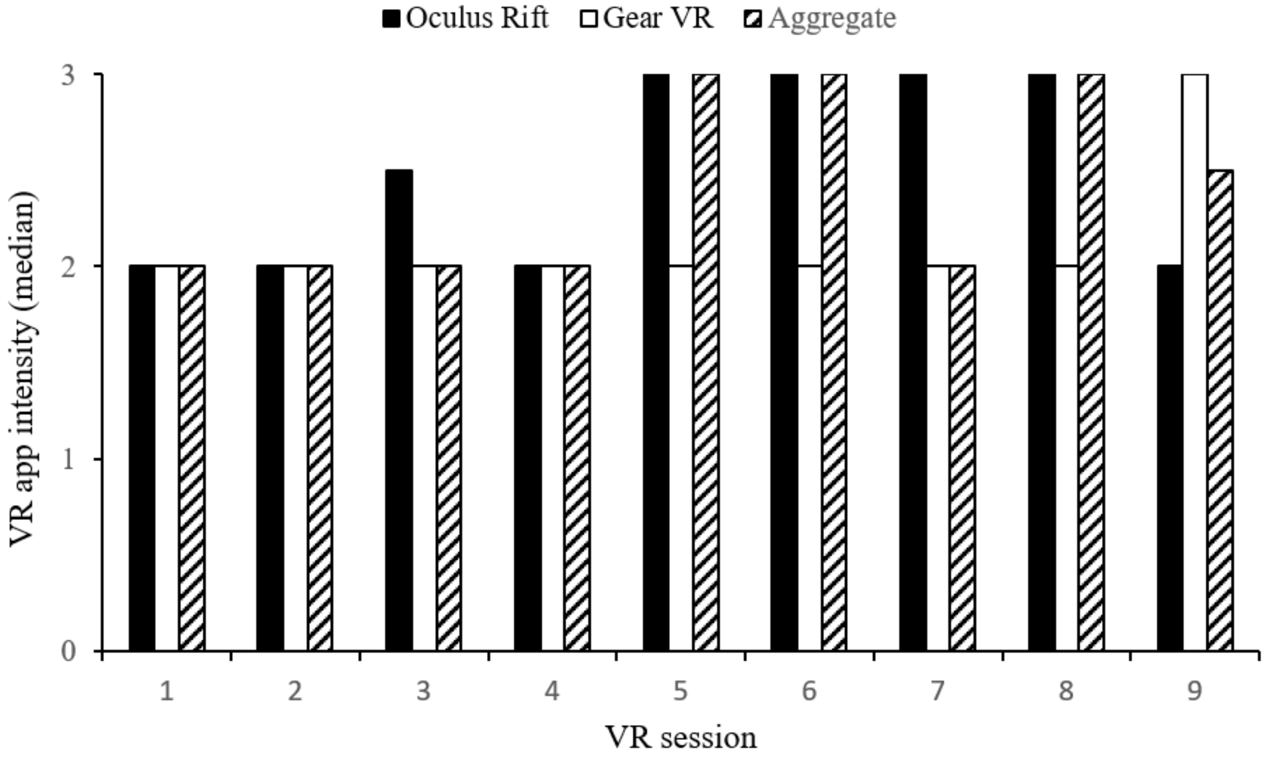
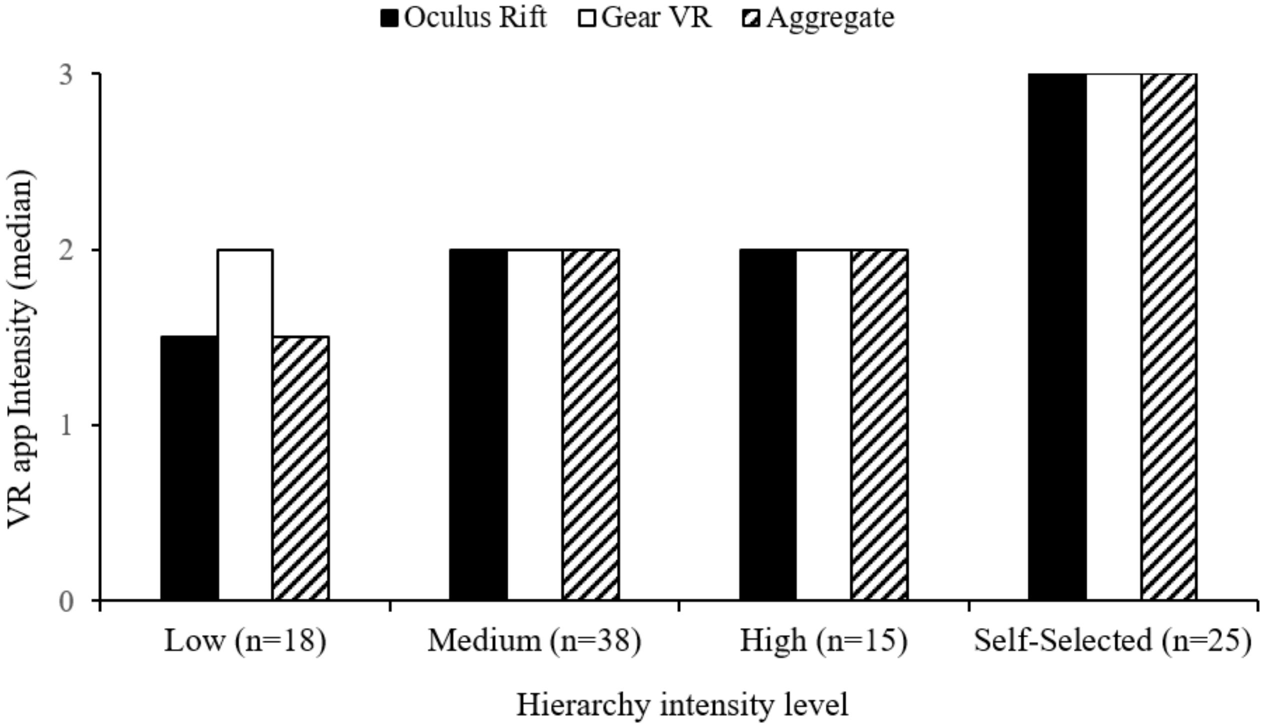


Figure 4. Veteran-reported intensities for virtual reality (VR) apps across movement intensity levels.



Aim 2: Estimate the Effect Size, 95% CI, and Proportion of Veterans Experiencing Minimum Clinically Important Difference for Fear of Movement and Pain Outcomes With Virtual Reality Use

For the primary outcome kinesiophobia, 10 of 16 Veterans (63%) had improved scores from baseline with six (38%) exceeding an MCID of 12.9 points or greater [58]. The observed effect size improvement was minimal (ie, Cohen *d*<0.20) on the FDAQ (Cohen *d*=−0.15, 95% CI −0.55 to 0.25). One Veteran was identified as a possible outlier. When this score was excluded, a small effect size reduction in fear of movement was observed (Cohen *d*=−0.35, 95% CI −0.71 to 0.01). However, there was notable variability in the confidence interval with the effect size ranging from no effect to a medium improvement suggesting a possible null finding.

Using the POQ-VA fear scale, there was little evidence of a reduction in kinesiophobia (Cohen *d*=−0.10, 95% CI −0.69 to 0.48). When a possible outlier score was excluded, the effect size changed directions but did not amount to a notable effect (Cohen *d*=0.06, 95% CI −0.43 to 0.54). Full sample scale scores for primary and secondary measures, effect sizes, 95% CIs, and MCID are presented in Table 2.

Table 2. Baseline and posttest scores for fear of movement and secondary outcomes (N=16).

Measure	Baseline	Posttest	Effect size (95% CI)	MCID[a], n (%)
Kinesiophobia, mean (SD)				
FDAQ[b]	59.20 (24.83)	56.45 (21.02)	−0.15 (−0.55, 0.25)	6 (38)
POQ-VA[c] Fear	12.44 (5.01)	11.91 (3.01)	−0.10 (−0.69, 0.48)	N/A[d]
Outliers excluded[e]				
FDAQ	60.28 (25.31)	54.75 (20.58)	−0.35 (−0.71, 0.01)	—[f]
POQ-VA Fear	12.31 (2.65)	12.07 (4.95)	0.06 (−0.47, 0.59)	—
Pain outcomes (POQ-VA), mean (SD)				
Interfere daily living	16.44 (12.36)	17.33 (12.03)	0.10 (−0.27, 0.47)	N/A
Interfere mobility	26.31 (9.58)	22.25 (10.82)	−0.56 (−0.96, −0.16)	N/A
Negative affect	28.56 (10.96)	29.09 (11.87)	0.07 (−0.26, 0.40)	N/A
Numeric Rating Scale, mean (SD)				
Pain intensity	6.88 (1.26)	6.38 (1.59)	−0.48 (−0.87, −0.10)	1 (7)
Outlier excluded	6.73 (1.16)	6.40 (1.64)	−0.40 (−0.69, −0.12)	—
PCS[g]	28.83 (10.39)	24.54 (15.45)	−0.41 (−0.79, −0.02)	5 (36)
PSFS[h]	3.60 (1.59)	5.98 (2.37)	1.14 (0.50, 1.78)	10 (67)

[a] MCID: minimum clinically important difference.

[b] FDAQ: Fear of Daily Activities Questionnaire.

[c] POQ-VA: Pain Outcomes Questionnaire-VA.

[d] Not applicable.

[e] N=15.

[f] MCID not recalculated.

[g] PCS: Pain Catastrophizing Scale.

[h] PSFS: Patient-Specific Functioning Scale.

Aim 2: Secondary Outcomes

The POQ-VA pain interference with mobility and activities of daily living as well as negative affect scales were examined. Veterans experienced a medium effect size improvement in interference with mobility (Cohen *d*=−0.56, 95% CI −0.96 to −0.16). There was variability with the effect size interval ranging from a small to large effect that was statistically significant (*P*<.05) because the interval did not contain the null value of zero. Conversely, Veterans experienced a slight exacerbation in interference with activities of daily living (Cohen *d*=0.10, 95% CI −0.27 to 0.47). A similar pattern was observed with negative affect (Cohen *d*=0.07, 95% CI −0.26 to 0.40). Intervals for the latter two findings suggested no effect.

Five of 15 Veterans (33%) with complete data had reduced pain intensity scores from baseline using the Numeric Rating Scale. However, only one Veteran (7%) met MCID of 2.10 points or greater for moderate pain intensity at baseline (rating=4-6) [59]. No Veterans with severe baseline pain (rating ≥7) exceeded the MCID [59]. The sample had a small to medium improvement in pain intensity (Cohen *d*=−0.49, 95% CI −0.87 to −0.11). The confidence interval suggested a minimal-to-large effect. When a possible outlier was excluded, a similar effect pattern remained (Cohen *d*=−0.40, 95% CI −0.69 to −0.12). These effects were statistically significant (*P*<.05) because the intervals did not contain the null value of zero.

Ten of 14 Veterans (71%) had reduced catastrophizing scores from baseline, and five (36%) exceeded the MCID of 38% or greater improvement [61]. Veterans experienced a small to medium effect size improvement in pain catastrophizing (Cohen *d*=−0.41, 95% CI −0.79 to −0.02). The interval ranged from a minimal to large and was statistically significant.

Finally, 14 of 15 Veterans (93%) reported improvements in patient-specific functional tasks that were previously hindered by their pain, as measured by the PSFS. Ten Veterans (67%) exceeded the MCID. Based on the scheme described by Abbott and Schmitt [63], Veterans MCID improvements were categorized as small (1.30-2.29; *n*=2), medium (2.30-2.69; *n*=2), and large (≥2.70; *n*=6). The observed effect size improvement in patient-specific functioning ranged from medium to large and was statistically significant (Cohen *d*=1.14, 95% CI 0.50-1.78, *P*<.001).

Aim 3: Pilot Test the Protocol to Assess the Feasibility of VR Use

Quantitative feasibility outcomes included levels of VR compliance, adherence, and session experiences. The compliance rate for this study (85.2%) was calculated via the number of VR sessions Veterans attended (n=98) divided by the total number of scheduled sessions (n=115). Of the 98 sessions attended, Veterans completed the full 20 minutes in 93 for an estimated adherence rate of 94.9%. The most common reason for missing (n=7) or shortened VR sessions (n=2) was physical therapy-related (eg, longer than expected sessions). Compliance (98/108, 90.7%) and adherence (93/96, 96.9%) rates were calculated accounting for physical therapy as the primary reason for missing VR sessions.

The VR experiences included Veteran-rated immersion, session length, and HMD preferences, and were plotted across sessions. Adverse events were also described. During sessions 1 to 3, the Veteran's median immersion ratings (median 2, IRQ 1) indicated that they were aware of using technology and were immersed in the virtual world. Sessions 4 to 9 saw ratings vary to include greater immersion in the virtual world with less technology awareness (range 2-3). Overall, median immersion ratings for the Oculus Rift and VR Gear HMDs were equal (median 2, IQR 1). Veterans typically rated Oculus Rift higher across the first five sessions and Gear VR higher across sessions 6 to 9 (Figure 5). Median ratings indicated that 20-minute sessions typically were "too short" (median 1, IQR 1) with no sessions rated lower than "just right." Overall ratings for each HMD were equal (median 1, IQR 1), although median ratings for Gear VR were occasionally lower (ie, too short) than Oculus Rift across sessions (Figure 6).

No significant adverse events, such as falls, occurred. Minor adverse events included cybersickness symptoms (dizziness: n=2 sessions; nausea: n=4 sessions) and one session being terminated early because the Veteran was "starting to feel the weight [Gear VR] on their neck." Interestingly, 2 of 16 Veterans (13%) purchased VR headsets during the study. Four additional Veterans (25%) requested VR purchasing information and were provided with a handout (eg, models, cost, reviews) on request.

Figure 5. Veteran-rated immersion across virtual reality (VR) sessions. Immersion rating scale: 1=using technology, 2=using technology and immersed in the virtual world, 3= completely immersed in the virtual world.

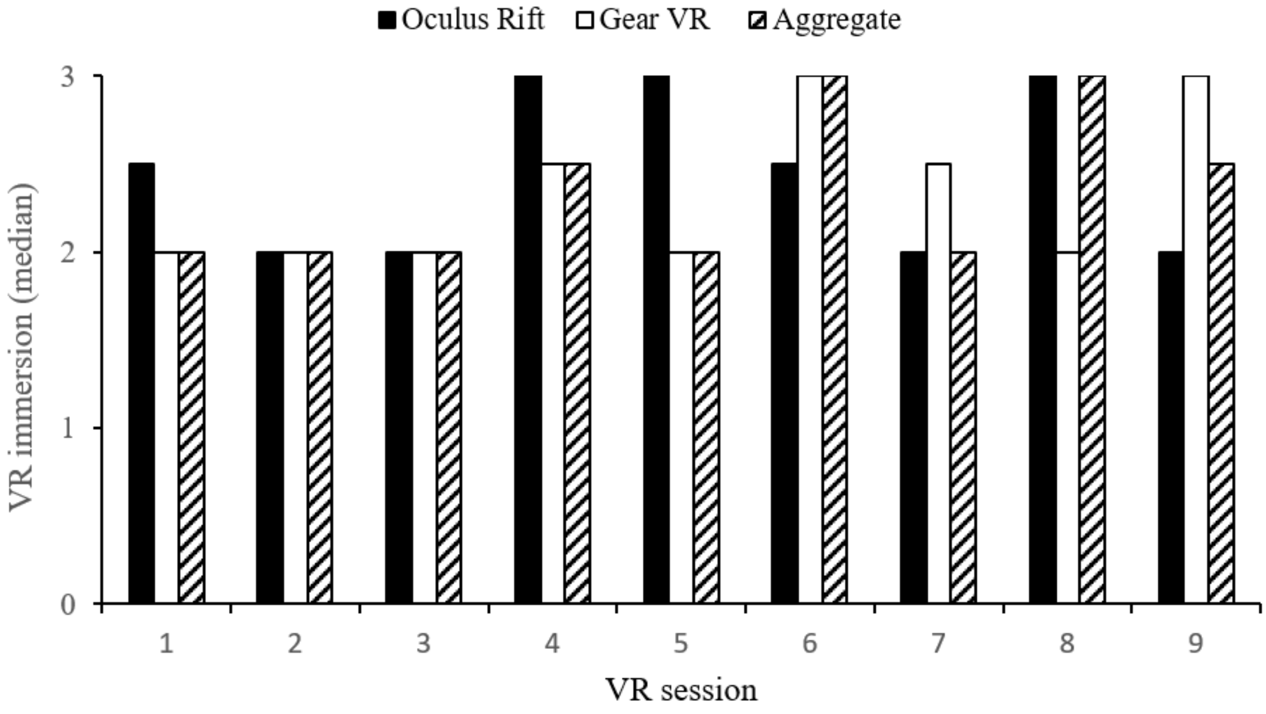
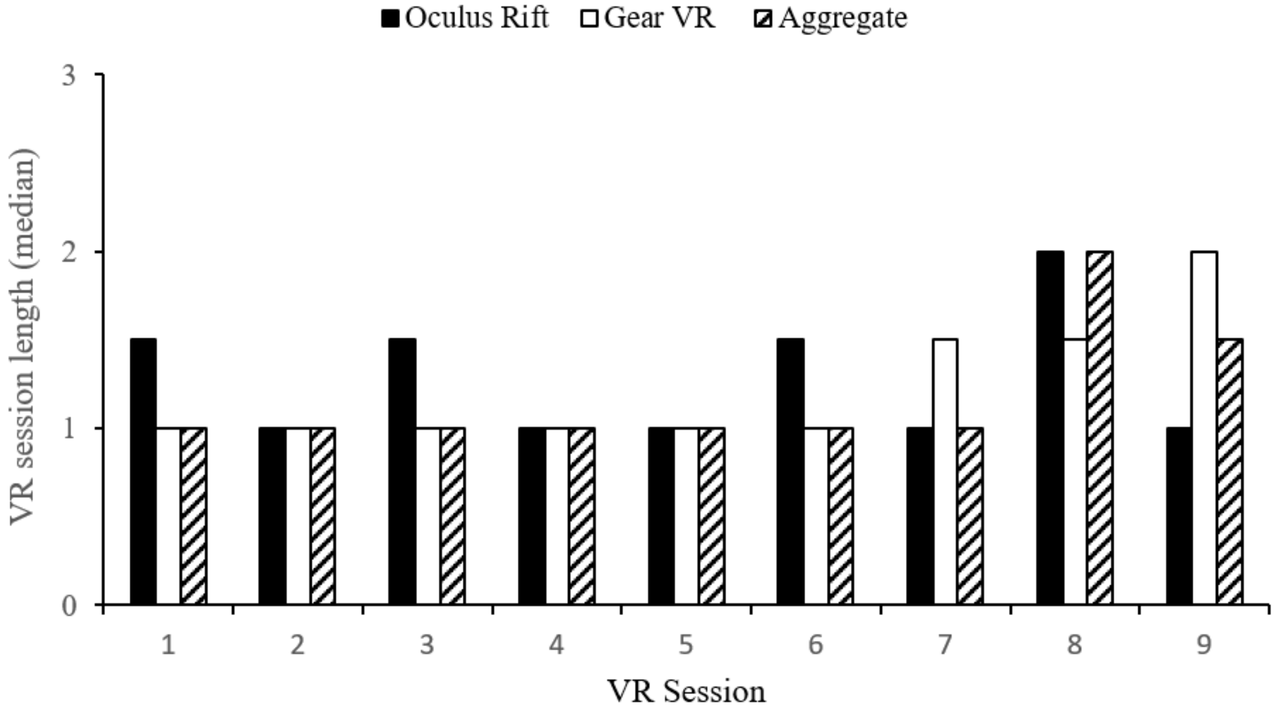


Figure 6. Veteran-rated virtual reality (VR) session length. Session length: 1=too short, 2=just right, 3=too long.



Discussion

Principal Findings

This study examined the feasibility of using a VR distraction-to-exposure hierarchy to improve fear of movement for Veterans with chronic pain. Descriptive findings suggest notable variability in Veteran trajectories across the hierarchy and that exposure apps requiring increased levels of movement were less challenging than anticipated. For the primary outcome (fear of movement), fear of daily activities, which measures specific movements, may be a more promising measure than the POQ-VA fear scale, which contains more general items (ie, exercise, reinjury). Promising pain outcomes for future research include interference with mobility, pain intensity, pain catastrophizing, and patient-specific functional activities. Veterans rated feeling immersed in VR even when they were cognizant of their technology use. They also rated their sessions as too short. Minimal adverse events were reported. This provided favorable evidence for the feasibility of VR for chronic pain management in a VA medical center. However, app selection, particularly for high movement intensity, will require modification before future research to align with the initial hierarchy.

Trajectories and App Intensity

The first aim of this study was to describe trajectories and self-reported app intensity ratings across the distraction-to-exposure hierarchy. Clinical consultation suggested that the majority of Veterans may not complete the hierarchy due to fear of movement. However, most Veterans reached or surpassed the highest movement level in six sessions. Differences between the Oculus Rift and Gear VR HMDs revealed most Veterans completed the hierarchy after only four sessions using Gear VR versus six sessions using the Oculus Rift. To help ensure Veterans did not lose interest in the study if they found the tasks to be too easy, the hierarchy was expanded to include Veteran-selected activities (eg, fishing, basketball free throws) once they completed the hierarchy.

To support our hierarchy, we anticipated that median self-reported app intensities would align with movement intensity on the same rating scale (1=low, 2=medium, 3=high). Self-reported app intensity ratings suggest that hierarchy apps did not calibrate as expected. Low movement distraction apps were rated more intensely than expected, particularly for Gear VR. Medium movement-intensity apps were rated as expected. Proposed high movement-intensity apps were rated similarly to medium intensity, which indicated that this was not a unique intensity level as anticipated. Similar Veteran ratings across proposed levels of the hierarchy indicate a possible range restriction in movement intensity. Median high intensity scores were not observed until Veterans engaged in self-selected activities. Despite high intensity ratings, post hoc integration of these apps into the distraction-to-exposure hierarchy was difficult due to their heterogeneity.

Fear of Movement and Pain Outcomes

The second aim was to estimate the effect size, 95% CI, and proportion of Veterans experiencing MCID for fear of movement and pain outcomes. These analyses were to inform design and sample size requirements for a future trial and selection of appropriate instruments for use with VR for chronic pain. The MCID for kinesiophobia indicated that 38% exceeded MCID on the FDAQ. After removal of an outlier, a small effect size improvement was observed for the primary outcome fear of movement using the FDAQ, but not the POQ-VA. However, there was a wide confidence interval associated with these effect sizes (and pain outcomes), which is common in small sample sizes [67].

For secondary pain measures, the MCID statistics, effect size, and confidence interval estimates were generally promising. For pain outcomes, 36% of Veterans exceeded MCID in catastrophizing, but only one met MCID for pain intensity. In total, 67% of Veterans experienced MCID for patient-specific functional outcomes, thus indicating the most promising outcome measure in this small study.

On the POQ-VA, improvements were observed in interference with mobility (medium) and pain intensity (small). The 95% CIs were statistically significant for these outcomes. No changes were observed for negative affect or interference with activities of daily living. There was a small effect size improvement for pain catastrophizing and the significant CI ranged from minimal to large. Similar to MCID, the most promising improvements were for patient-specific functional outcomes. A large effect size was observed in patient-specific functional outcomes with the confidence interval ranging from medium to large.

Feasibility of Virtual Reality

The final aim was to establish the feasibility of VR as a therapy adjunct. All Veterans completed the study. Session attendance was over 85% and increased to more than 90% after accounting for sessions missed for physical therapy. Adherence, as measured by session completion, was nearly 95% and approximately 97% when accounting for therapy. This adjustment was made because VR was proposed to complement evidence-based interventions, and we did not want to study participation to reduce therapy adherence. Immersion ratings indicated that Veterans simultaneously felt immersed in the virtual world and that they were using VR HMDs. Median immersion ratings increased in later sessions, which indicates that immersion increased over time and technology awareness was lessened. Immersion was more variable for Oculus Rift than Gear VR. Regarding VR dose, most 20-minute sessions were rated as too short. Ratings suggest that Veterans may prefer somewhat longer sessions for Gear VR than the more immersive Oculus Rift. Minimal adverse events, including cybersickness and the weight of the Gear VR aggravating pain, were reported. Finally, more than 30% of Veterans purchased VR HMDs or requested purchasing information during the study.

Clinical Implications

Aim 1

Considering the speed in which Veterans completed the hierarchy and their app intensity ratings, it is likely that the hierarchy was less incremental than designed. This would suggest that our approach to the hierarchy proved too conservative. Another possibility is that the intensity of their VR experience was shaped less by movement than their interest or engagement in the self-selected activities. The hierarchy will be modified before future use. The modification process will take a bottom-up approach to include Veteran stakeholders throughout this process of app selection, testing, and intensity rating [39]. Despite the movement heterogeneity in these Veteran-selected activities, these apps and their intensity ratings will be considered in the modification process. Having a structured app selection process for the hierarchy is important because of the goal to generalize to other VA hospitals and clinics. This is to ensure Veterans are not over- or underexposed to feared movements.

Aim 2

Effect sizes for kinesiophobia and pain outcomes were smaller than in previous research. It is possible that VR is more efficacious for acute than chronic pain. However, these effects may have been attenuated by the ceiling and range restriction effects in the hierarchy. Moreover, given the wide CI and the small sample, estimation of CI at a lower limit of 75% to 85% may be appropriate to complement descriptive information in VR feasibility work [68]. When considering the MCID statistics in accordance with observed effect sizes for fear of movement (FDAQ), interference with mobility, pain intensity, pain catastrophizing, and patient-specific functioning may be promising for future research. Interestingly, Veteran-selected activities hindered by pain had the largest effect sizes and proportion of MCID. As discussed previously, Veteran-selected activities produced the highest VR intensity ratings. These findings highlight the importance of considering user preferences in selecting meaningful outcomes in addition to the VR intervention itself.

Aim 3

Despite noted concerns with the hierarchy, this study further supports the feasibility of VR for pain management [30,69]. Specifically, the integration of VR technology itself was considered successful. Although VR was adjunct, Veterans both attended and completed sessions at high rates. Veterans typically rated VR sessions as too short, which was consistent with the success of a recent study that provided twelve 30-minute sessions using Oculus Rift [69]. It is notable that Veterans' ratings showed marginally lower preferences for 20-minute sessions using Oculus Rift. Given that immersion was also slightly higher for Oculus Rift than Gear VR, consideration for immersion level, HMD selection [70], and sensory demand [14,32] is important when considering session length with chronic pain populations.

Limitations

Consistent difficulties emerged when using Gear VR in this study. First, inconsistent hospital Wi-Fi hindered the importance of certain apps (eg, Guided Meditation). Second, Gear VR lacked a “kiosk mode” and apps would often time out and need to be restarted during sessions. This may have affected immersion. Third, Veterans indicated that environmental noise in the therapy gym was an issue, but only for Gear VR. This study used two Gear VR HMDs and one Oculus Rift. This approach was due to equipment availability and physical space considerations. The power supply on the gaming computer also shorted out during the first day of testing. Hence, only Gear VR was used during the first four testing days. These factors may have accelerated Veteran progress across the hierarchy. Use of a single HMD type—Oculus Rift—may be more beneficial for validation of VR hierarchies in busy medical settings.

Veterans used VR as an adjunct and were involved in on-going interdisciplinary pain management, which likely influenced Veteran improvement in treatment outcomes (see Murphy et al [36,37]). This may have also influenced Veterans' quick progressions across the hierarchy. Inclusion of a randomized control group is necessary to estimate the true added impact of VR versus treatment as usual.

This study had additional limitations, including a small sample size, which was limited by the allotted 3-week data collection period. This likely influenced the wide CI around the effect sizes [67]. Additionally, Veterans may not have been adequately challenged by the hierarchy. Finally, because this pilot study was unfunded, commercially available VR apps were used. A more optimal approach would be to develop or tailor VR apps to capture frequently avoided movements of varying intensity levels, such as safe strategies for bending or climbing stairs.

Conclusions

This study provided evidence that VR is feasible for chronic pain populations. In addition, this study expanded the knowledge base by demonstrating feasibility for VR as an adjunct for evidence-based chronic pain interventions in a medical setting. Implementation was more successful for VR technologies than the distraction-to-exposure hierarchy itself. Future research will focus on modification of this hierarchy to validate feared movements at varying intensity levels. This may be best accomplished using a bottom-up approach that includes Veterans in the intervention design and outcome selection processes.

Acknowledgments

The authors would like to thank Nicolle Angeli, PhD; Stacey Sandusky, PhD; and Megan Anderson, PTA, from the Chronic Pain Rehabilitation Program at the James A Haley Veterans Hospital for their assistance with participant recruitment, scheduling, and implementation strategies. Thank you to Spencer Mason from the Advanced Visualization Center at the University of South Florida for his assistance with VR technologies.

This material is the result of work supported with resources and the use of facilities at the James A Haley Veterans' Hospital. The contents of this manuscript do not reflect the views of the Department of Veterans Affairs of the United States Government.

Conflicts of Interest

JLM serves as the Chief Behavioral Health Officer of Karuna Labs. This is a recent appointment. JLM was not in this position during the conceptualization of this study, data collection, or drafting of this manuscript.

References

1. Dionisio JDN, Burns WGIII, Gilbert R. 3D Virtual worlds and the metaverse. ACM Comput Surv 2013 Jun 01;45(3):1-38. [doi: 10.1145/2480741.2480751]
2. Parkin S. MIT Technology Review. 2014 Apr 23. Oculus Rift: thirty years after virtual-reality goggles and immersive virtual worlds made their debut, the technology finally seems poised for widespread use URL: https://www.technologyreview.com/s/526531/oculus-rift/ [accessed 2019-10-21]
3. Steuer J. Defining virtual reality: dimensions determining telepresence. J Comm 1992 Feb 07;42(4):73-93. [doi: 10.1111/j.1460-2466.1992.tb00812.x]
4. Garrett B, Taverner T, Masinde W, Gromala D, Shaw C, Negraeff M. A rapid evidence assessment of immersive virtual reality as an adjunct therapy in acute pain management in clinical practice. Clin J Pain 2014 Dec;30(12):1089-1098. [doi: 10.1097/AJP.0000000000000064] [Medline: 24535053]
5. Cummings JJ, Bailenson JN. How immersive is enough? A meta-analysis of the effect of immersive technology on user presence. Media Psychology 2015 May 15;19(2):272-309. [doi: 10.1080/15213269.2015.1015740]
6. Freeman D, Reeve S, Robinson A, Ehlers A, Clark D, Spanlang B, et al. Virtual reality in the assessment, understanding, and treatment of mental health disorders. Psychol Med 2017 Oct;47(14):2393-2400 [FREE Full text] [doi: 10.1017/S003329171700040X] [Medline: 28325167]
7. Olmos-Raya E, Ferreira-Cavalcanti J, Contero M, Castellanos MC, Giglioli IAC, Alcañiz M. Mobile virtual reality as an educational platform: a pilot study on the impact of immersion and positive emotion induction in the learning process. EURASIA J Math Sci Tech Ed 2018;14(6):2045-2057. [doi: 10.29333/ejmste/85874]
8. Trost Z, Parsons TD. Beyond distraction: virtual reality graded exposure therapy as treatment for pain-related fear and disability in chronic pain. J Appl Biobehav Res 2014 Jun 10;19(2):106-126. [doi: 10.1111/jabr.12021]
9. Chan E, Foster S, Sambell R, Leong P. Clinical efficacy of virtual reality for acute procedural pain management: a systematic review and meta-analysis. PLoS One 2018;13(7):e0200987 [FREE Full text] [doi: 10.1371/journal.pone.0200987] [Medline: 30052655]
10. International Association for the Study of Pain, Task Force on Taxonomy. Part III: a current list with definitions and notes on usage. In: Merskey H, Bogduk N, editors. Classification Of Chronic Pain: Descriptions Of Chronic Pain Syndromes And Definitions Of Pain Terms. Seattle, WA: IASP Press; 1994:207-214.
11. Treede R, Rief W, Barke A, Aziz Q, Bennett MI, Benoliel R, et al. A classification of chronic pain for ICD-11. Pain 2015 Jun;156(6):1003-1007 [FREE Full text] [doi: 10.1097/j.pain.0000000000000160] [Medline: 25844555]
12. Grichnik KP, Ferrante FM. The difference between acute and chronic pain. Mt Sinai J Med 1991 May;58(3):217-220. [Medline: 1875958]

13. Nijs J, Paul van Wilgen C, Van Oosterwijck J, van Ittersum M, Meeus M. How to explain central sensitization to patients with 'unexplained' chronic musculoskeletal pain: practice guidelines. Man Ther 2011 Oct;16(5):413-418 [FREE Full text] [doi: 10.1016/j.math.2011.04.005] [Medline: 21632273]
14. Simons LE, Elman I, Borsook D. Psychological processing in chronic pain: a neural systems approach. Neurosci Biobehav Rev 2014 Feb;39:61-78 [FREE Full text] [doi: 10.1016/j.neubiorev.2013.12.006] [Medline: 24374383]
15. Murphy J, McKellar J, Raffa S, Clark M, Kerns R, Karlin B. Cognitive Behavioral Therapy for Chronic Pain Among Veterans: Therapist Manual. Washington, DC: US Department of Veterans Affairs; 2014.
16. Lethem J, Slade P, Troup J, Bentley G. Outline of a fear-avoidance model of exaggerated pain perception--I. Behav Res Ther 1983;21(4):401-408. [doi: 10.1016/0005-7967(83)90009-8] [Medline: 6626110]
17. Vlaeyen JW, Linton SJ. Fear-avoidance and its consequences in chronic musculoskeletal pain: a state of the art. Pain 2000 Apr;85(3):317-332. [doi: 10.1016/s0304-3959(99)00242-0] [Medline: 10781906]
18. Leeuw M, Goossens ME, Linton SJ, Crombez G, Boersma K, Vlaeyen JW. The fear-avoidance model of musculoskeletal pain: current state of scientific evidence. J Behav Med 2007 Feb;30(1):77-94. [doi: 10.1007/s10865-006-9085-0] [Medline: 17180640]
19. Linton S, Shaw W. Impact of psychological factors in the experience of pain. Phys Ther 2011 May;91(5):700-711. [doi: 10.2522/ptj.20100330] [Medline: 21451097]
20. George SZ, Zeppieri G. Physical therapy utilization of graded exposure for patients with low back pain. J Orthop Sports Phys Ther 2009 Jul;39(7):496-505. [doi: 10.2519/jospt.2009.2983] [Medline: 19574658]
21. Smeets RJ, Vlaeyen JW, Hidding A, Kester AD, van der Heijden GJ, van Geel AC, et al. Active rehabilitation for chronic low back pain: cognitive-behavioral, physical, or both? First direct post-treatment results from a randomized controlled trial [ISRCTN22714229]. BMC Musculoskelet Disord 2006 Jan 20;7:5 [FREE Full text] [doi: 10.1186/1471-2474-7-5] [Medline: 16426449]
22. Parsons TD, Trost Z. Virtual reality-graded exposure therapy as treatment for pain-related disability in chronic pain. In: Ma M, Jain LC, Anderson P, editors. Virtual, Augmented Reality And Serious Games For Healthcare 1 (Intelligent Systems Reference Library). Berlin, Heidelberg, DE: Springer; Apr 26, 2014:523-546.
23. Hoffman HG, Chambers GT, Meyer WJ, Arceneaux LL, Russell WJ, Seibel EJ, et al. Virtual reality as an adjunctive non-pharmacologic analgesic for acute burn pain during medical procedures. Ann Behav Med 2011 Apr;41(2):183-191 [FREE Full text] [doi: 10.1007/s12160-010-9248-7] [Medline: 21264690]
24. McCaul KD, Malott JM. Distraction and coping with pain. Psychol Bull 1984 May;95(3):516-533. [Medline: 6399756]
25. Hoffman HG, Meyer WJ, Ramirez M, Roberts L, Seibel EJ, Atzori B, et al. Feasibility of articulated arm mounted Oculus Rift Virtual Reality goggles for adjunctive pain control during occupational therapy in pediatric burn patients. Cyberpsychol Behav Soc Netw 2014 Jun;17(6):397-401 [FREE Full text] [doi: 10.1089/cyber.2014.0058] [Medline: 24892204]
26. Malloy KM, Milling LS. The effectiveness of virtual reality distraction for pain reduction: a systematic review. Clin Psychol Rev 2010 Dec;30(8):1011-1018. [doi: 10.1016/j.cpr.2010.07.001] [Medline: 20691523]
27. Wiederhold BK, Gao K, Sulea C, Wiederhold MD. Virtual reality as a distraction technique in chronic pain patients. Cyberpsychol Behav Soc Netw 2014 Jun;17(6):346-352 [FREE Full text] [doi: 10.1089/cyber.2014.0207] [Medline: 24892196]
28. Jin W, Choo A, Gromala D, Shaw C, Squire P. A virtual reality game for chronic pain management: a randomized, controlled clinical study. Stud Health Technol Inform 2016;220:154-160. [Medline: 27046570]
29. Yilmaz Yelvar GD, Çırak Y, Dalkılınç M, Parlak Demir Y, Guner Z, Boydak A. Is physiotherapy integrated virtual walking effective on pain, function, and kinesiophobia in patients with non-specific low-back pain? Randomised controlled trial. Eur Spine J 2017 Feb;26(2):538-545. [doi: 10.1007/s00586-016-4892-7] [Medline: 27981455]
30. Thomas JS, France CR, Applegate ME, Leitkam ST, Walkowski S. Feasibility and safety of a virtual reality dodgeball intervention for chronic low back pain: a randomized clinical trial. J Pain 2016 Dec;17(12):1302-1317 [FREE Full text] [doi: 10.1016/j.jpain.2016.08.011] [Medline: 27616607]
31. Garcia-Palacios A, Botella C, Hoffman H, Fabregat S. Comparing acceptance and refusal rates of virtual reality exposure vs. in vivo exposure by patients with specific phobias. Cyberpsychol Behav 2007 Oct;10(5):722-724. [doi: 10.1089/cpb.2007.9962] [Medline: 17927544]
32. Parham LD, Cohn ES, Spitzer S, Koomar JA, Miller LJ, Burke JP, et al. Fidelity in sensory integration intervention research. Am J Occup Ther 2007;61(2):216-227. [doi: 10.5014/ajot.61.2.216] [Medline: 17436844]
33. Garrett B, Taverner T, Gromala D, Tao G, Cordingley E, Sun C. Virtual reality clinical research: promises and challenges. JMIR Serious Games 2018 Oct 17;6(4):e10839 [FREE Full text] [doi: 10.2196/10839] [Medline: 30333096]
34. US Department of Veterans Affairs. 2017. VA/DoD Clinical Practice Guideline for Opioid Therapy for Chronic Pain URL: https://www.healthquality.va.gov/guidelines/Pain/cot/ [accessed 2018-04-19]
35. Dahlhamer J, Lucas J, Zelaya C, Nahin R, Mackey S, DeBar L, et al. Prevalence of chronic pain and high-impact chronic pain among adults-United States, 2016. MMWR Morb Mortal Wkly Rep 2018 Sep 14;67(36):1001-1006 [FREE Full text] [doi: 10.15585/mmwr.mm6736a2] [Medline: 30212442]
36. Murphy JL, Clark ME, Banou E. Opioid cessation and multidimensional outcomes after interdisciplinary chronic pain treatment. Clin J Pain 2013 Feb;29(2):109-117. [doi: 10.1097/AJP.0b013e3182579935] [Medline: 22751033]

37. Murphy JL, Phillips KM, Rafie S. Sex differences between Veterans participating in interdisciplinary chronic pain rehabilitation. J Rehabil Res Dev 2016;53(1):83-94 [FREE Full text] [doi: 10.1682/JRRD.2014.10.0250] [Medline: 27005932]
38. Krebs EE, Carey TS, Weinberger M. Accuracy of the pain numeric rating scale as a screening test in primary care. J Gen Intern Med 2007 Oct;22(10):1453-1458 [FREE Full text] [doi: 10.1007/s11606-007-0321-2] [Medline: 17668269]
39. 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]
40. Birckhead B, Khalil C, Liu X, Conovitz S, Rizzo A, Danovitch I, et al. Recommendations for methodology of virtual reality clinical trials in health care by an international working group: iterative study. JMIR Ment Health 2019 Jan 31;6(1):e11973 [FREE Full text] [doi: 10.2196/11973] [Medline: 30702436]
41. Guided Meditation VR. 2016. URL: https://guidedmeditationvr.com [accessed 2018-06-11] [WebCite Cache ID 7067ZzUXl]
42. nDreams. 2016. Perfect URL: http://www.ndreams.com/titles/perfectvr/ [accessed 2018-06-11] [WebCite Cache ID 7068TH8zl]
43. Now VR. Rest VR: Relax and Meditate. 2016. URL: http://www.relaxvr.co/ [accessed 2018-06-11] [WebCite Cache ID 706I8RXQm]
44. Picselica. 2016. Ocean Rift URL: http://www.picselica.com/?cat=4 [accessed 2018-06-11] [WebCite Cache ID 706IYMIf8]
45. Greener Games. 2017. Nature Treks URL: http://greenergames.net [accessed 2018-06-11] [WebCite Cache ID 7069aNcDw]
46. Immersive Entertainment Inc. 2016. Grand Canyon VR URL: https://www.immersiveentertainment.com/the-grand-canyon-vr-experience/ [accessed 2018-06-11] [WebCite Cache ID 706HP8KQ4]
47. Multiverse Entertainment. 2016. Reveries: Dream Flight URL: https://www.multiverseinc.com/dreamflight [accessed 2018-06-11] [WebCite Cache ID 706Ij5t5U]
48. Facebook. 2017. Paint VR URL: https://www.facebook.com/PAINTinVR/ [accessed 2011-06-11]
49. Google VR. 2017. Tilt Brush URL: https://www.tiltbrush.com [accessed 2018-06-11] [WebCite Cache ID 706HXsnr3]
50. Arrowiz. 2017. Beats Fever Paper URL: http://arrowiz.com/ [accessed 2018-06-11] [WebCite Cache ID 706IszBCR]
51. NiVision. The Show Must Go On. 2017. URL: https://nivisionag.wixsite.com/showmust-goon [accessed 2018-06-11] [WebCite Cache ID 706IEzvuJ]
52. Oculus VR. 2016. Oculus Rift URL: https://www.oculus.com/rift/ [accessed 2018-06-11] [WebCite Cache ID 706J17SJl]
53. Samsung Oculus VR. 2015. Samsung Oculus Gear VR URL: https://www.oculus.com/gear-vr/ [accessed 2018-06-11] [WebCite Cache ID 706J6M3PU]
54. Czub M, Piskorz J. Body movement reduces pain intensity in virtual reality–based analgesia. Int J Hum-Comput Int 2017 Dec 08;34(11):1045-1051. [doi: 10.1080/10447318.2017.1412144]
55. Tashjian VC, Mosadeghi S, Howard AR, Lopez M, Dupuy T, Reid M, et al. Virtual reality for management of pain in hospitalized patients: results of a controlled trial. JMIR Ment Health 2017 Mar 29;4(1):e9 [FREE Full text] [doi: 10.2196/mental.7387] [Medline: 28356241]
56. Cole J, Crowle S, Austwick G, Slater DH. Exploratory findings with virtual reality for phantom limb pain; from stump motion to agency and analgesia. Disabil Rehabil 2009;31(10):846-854. [doi: 10.1080/09638280802355197] [Medline: 19191061]
57. Clark ME, Gironda RJ, Young RW. Development and validation of the Pain Outcomes Questionnaire-VA. J Rehabil Res Dev 2003;40(5):381-395 [FREE Full text] [Medline: 15080223]
58. George SZ, Valencia C, Zeppieri G, Robinson ME. Development of a self-report measure of fearful activities for patients with low back pain: the fear of daily activities questionnaire. Phys Ther 2009 Sep;89(9):969-979 [FREE Full text] [doi: 10.2522/ptj.20090032] [Medline: 19608630]
59. Salaffi F, Stancati A, Silvestri CA, Ciapetti A, Grassi W. Minimal clinically important changes in chronic musculoskeletal pain intensity measured on a numerical rating scale. Eur J Pain 2004 Aug;8(4):283-291. [doi: 10.1016/j.ejpain.2003.09.004] [Medline: 15207508]
60. Sullivan MJ, Bishop SR, Pivik J. The Pain Catastrophizing Scale: development and validation. Psychol Assessment 1995;7(4):524-532. [doi: 10.1037/1040-3590.7.4.524]
61. Osman A, Barrios FX, Kopper BA, Hauptmann W, Jones J, O'Neill E. Factor structure, reliability, and validity of the Pain Catastrophizing Scale. J Behav Med 1997 Dec;20(6):589-605. [Medline: 9429990]
62. Stratford P. Assessing disability and change on individual patients: a report of a patient specific measure. Physiother Can 1995 Oct;47(4):258-263. [doi: 10.3138/ptc.47.4.258]
63. Abbott JH, Schmitt J. Minimum important differences for the patient-specific functional scale, 4 region-specific outcome measures, and the numeric pain rating scale. J Orthop Sports Phys Ther 2014 Aug;44(8):560-564. [doi: 10.2519/jospt.2014.5248] [Medline: 24828475]
64. Kadel R, Kip K. A SAS macro to compute effect size (Cohen's d) its confidence interval from raw survey data. 2012 Oct 01 Presented at: SAS Global Forum 2012; April 22-25, 2012; Orlando, FL.
65. Cohen J. Statistical Power Analysis For The Behavioral Sciences (2nd Edition). New York: Routledge; 1988.
66. Tukey JW. Exploratory Data Analysis. Boston, MA: Pearson; 1977.
67. Cohen J. The earth is round (p<.05). Am Psychol 1994;49(12):997-1003. [doi: 10.1037//0003-066x.49.12.997]

68. Lee EC, Whitehead AL, Jacques RM, Julious SA. The statistical interpretation of pilot trials: should significance thresholds be reconsidered? BMC Med Res Methodol 2014 Mar 20;14:41-49 [FREE Full text] [doi: 10.1186/1471-2288-14-41] [Medline: 24650044]
69. Garrett B, Taverner T, McDade P. Virtual reality as an adjunct home therapy in chronic pain management: an exploratory study. JMIR Med Inform 2017 May 11;5(2):e11 [FREE Full text] [doi: 10.2196/medinform.7271] [Medline: 28495661]
70. Tong X, Gromala D, Gupta D, Squire P. Usability comparisons of head-mounted vs. stereoscopic desktop displays in a virtual reality environment with pain patients. Stud Health Technol Inform 2016;220:424-431. [Medline: 27046617]

Abbreviations

CI: confidence interval
FDAQ: Fear of Daily Activities Questionnaire
HMD: head-mounted display
IQR: interquartile range
MCID: minimum clinically important difference
POQ-VA: Pain Outcomes Questionnaire-VA
PSFS: Patient-Specific Functional Scale
RCT: randomized controlled trial
VA: Veterans Affairs
VR: virtual reality

Edited by G Eysenbach; submitted 13.06.18; peer-reviewed by B Garrett, M Applegate; comments to author 10.07.18; revised version received 13.09.19; accepted 26.09.19; published 30.10.19.

Please cite as:
Fowler CA, Ballistrea LM, Mazzone KE, Martin AM, Kaplan H, Kip KE, Ralston K, Murphy JL, Winkler SL
Virtual Reality as a Therapy Adjunct for Fear of Movement in Veterans With Chronic Pain: Single-Arm Feasibility Study
JMIR Form Res 2019;3(4):e11266
URL: http://formative.jmir.org/2019/4/e11266/
doi:10.2196/11266
PMID:31670696

Original Paper

Web-Based Health Information Seeking Among Students at Kuwait University: Cross-Sectional Survey Study

Hasan Ashkanani[1], BMEDSC; Rabab Asery[1], BMEDSC; Fajer Bokubar[1], BMEDSC; Noor AlAli[1], BMEDSC; Shahad Mubarak[1], BMEDSC; Ali Buabbas[2], PhD; Abdullah Almajran[2], BSc, MSc

[1]Faculty of Medicine, Kuwait University, Jabriyah, Kuwait

[2]Department of Community Medicine, Faculty of Medicine, Kuwait University, Jabriyah, Kuwait

Corresponding Author:
Hasan Ashkanani, BMEDSC
Faculty of Medicine
Kuwait University
Block 4, Fourth Ring Road, Health Sciences Center
Jabriyah
Kuwait
Phone: 965 246 36559
Email: hasanashkanani@hotmail.com

Abstract

Background: Owing to the revolution in technology, the internet has become an important aspect of people’s lives. Modern technology is enabling people from diverse educational backgrounds to use the internet for several purposes, one of which is health information seeking. Recently, Web-based health information has become more popular among patients all over the world and among the general public.

Objective: This study aimed to investigate the use of Web-based health resources among undergraduate students from different faculties at Kuwait University.

Methods: The study employed a cross-sectional design with students selected from 8 faculties of Kuwait University, 4 faculties of Literature and 4 faculties of Science. Data were collected using structured questionnaires, and analysis was done using a chi-square test and binary logistic regression to determine the factors associated with seeking health information on the Web.

Results: The sample size obtained was 1132 with a response rate of 90.27% (1132/1254). Overall, the prevalence of students seeking Web-based health information was 92.66%. (1049/1132) The most significant factors associated with seeking health information on the Web were age, gender, faculty, year of study, primary source of internet, and level of experience with internet use. In total, 90.0% (325/361) of students who were aged older than 21 years used Web-based health information compared with 82.8% (275/332) of those who were aged 18 years. In addition, female students showed a higher prevalence (829/934, 88.8%) of Web-based health information seeking than males (210/270, 77.8%). Students who majored in faculties of Science were more likely to seek health information than those who majored in faculties of Literature. All the differences found in the study were statistically significant (P<.05).

Conclusions: The study concluded that many people use the internet for seeking health information. Sociodemographic factors have a significant association with Web-based health information seeking. Therefore, doctors must educate the public about the health information websites that can be trusted.

(JMIR Form Res 2019;3(4):e14327) doi:10.2196/14327

KEYWORDS

Kuwait; online; health care

Introduction

“The internet is a global network that enables computers and other communication gadgets to communicate directly and transparently” [1]. Owing to the revolution in technology, the internet has become an important aspect of people’s lives. Modern technology is enabling people from diverse educational backgrounds to use the internet for several purposes, one of which is health information seeking. People use the internet to search for several health-related topics, including facts about different diseases, diagnostic criteria, treatment methods, and side-effects of various medications.

Studies show that Web-based health information has become more popular among patients all over the world and among the general public. According to a study conducted in 2008, there was a 38% increase in the number of health-related websites in the Eastern Mediterranean Region [2]. Health information seekers can be defined as "internet users who search online for information about health topics whether they are acting as consumers, care givers, or e-patients" [3].

Previous studies concluded that most health-related internet searches are conducted to look for information related to a specific medical condition, both before and after seeking a consultation from a health care professional. This development was very prominent within the last 10 to 15 years. Concerns about information found on the Web include the quality of information found, whether the information would improve the patient's health, and whether the patient-doctor relationship will improve as patients become more informed or be damaged if doctors have difficulty when they recommend changes to patients with preconceived notions [4]. Doctors and the organizations that support them have to fully understand their role in the health revolution that was the result of Web-based health information seeking, both of which must improve their awareness of the patients and their needs. [5].

It is thought that the use of the internet by patients will raise their awareness about their conditions, whereas its use by doctors will assist their ability to make evidence-based decisions. A study conducted in the United States concluded that the challenge faced by the doctors is translating scientific advances into meaningful advice for clinical care [6].

As a source of information, the internet has several advantages over traditional sources. These advantages include easy access to information, frequently updated material, confidentiality, and interactivity, which leads to better understanding. Web-based resources of health information used to not be readily available for public consumption, as the internet was not known as a destination for medical information. Therefore, patients did not have any health information, and their roles were limited to passive roles that consisted of complying to the orders of their health care providers. However, with the rise of innumerable websites and apps that provide health information in a simplified manner, there was a shift in the paradigm, and in recent years, patients have been opting for a more active role in seeking health care. Using Web-based resources for health purposes empowers the patients to seek a more active role [4].

In spite of the various advantages of the internet, it can also be harmful when consumers use it without evaluation of the quality of the information sought. This can lead to misinterpretation of the information, affecting health behavior and outcomes [4,7].

A previous study conducted in 2003 had concluded that 61% of citizens of the United States who use the internet have searched for health information on the Web compared with 58.4% of citizens in Saudi Arabia [8,9]. A study conducted in Kuwait found that 93.2% of patients use the internet for one or more purposes and that 62.9% of these patients use it for seeking health-related information. In addition, it was concluded that gender, nationality, level of education, and using a computer at work were significant factors for obtaining health information on the Web [10]. This is consistent with another study conducted in Saudi Arabia, which showed that gender, education, employment, and income affect Web-based health information seeking [8].

Although there has been an increase in the use of the internet in health care, little research has been conducted to study the impact that the availability and integrity of health care information on the internet has on users. This study aimed to investigate the use of Web-based health resources among undergraduate students at Kuwait University.

The objective of the study was 3-fold: (1) to determine the prevalence of Web-based health information seeking among students at Kuwait University; (2) to determine the influence of the information found on the Web on the attitudes of students at Kuwait University; and (3) to identify the association between the sociodemographic factors of the students and their health information–seeking patterns.

Methods

Research Setting

Kuwait University is the first public university in the State of Kuwait and was established in 1966. Its vision is to provide world-class education, and it is committed to advancing, preserving, and disseminating knowledge, in addition to preparing educated, enlightened, and qualified human resources to realize the society's developmental needs. It consists of 17 faculties, divided into faculties of Literature and faculties of Science.

Study Design and Study Sample

This cross-sectional study was conducted among undergraduate students at Kuwait University across various faculties of the university. The study sample included 8 randomly selected faculties, stratified as 4 faculties of Literature and 4 faculties of Science. The selected faculties of Science were Medicine, Dentistry, Pharmacy, and Science, and the selected faculties of Literature were Law, Sharia and Islamic Studies, Education, and Business Administration (the total population of the 8 faculties was 19,359 students). A purposive convenient sampling procedure was used for students involved in the study, and the instrument used was a validated questionnaire. Students were approached in between classes and during their breaks, and participation was voluntary. The study included students aged between 18 and 25 years. Any student aged younger than 18 years or older than 25 years was excluded from the sample population. A total of 1253 questionnaires were collected during this period, and the response rate was 99.12% (1242/1253).

Ethical Approval

A Human Subject Form describing the ethical aspects of the study was approved by the Faculty of Medicine Ethical Committee, and an informed consent sheet was prepared, both ensuring protection of the participant's autonomy, beneficence, and justice. There were no risks to participation in this study, and benefits of the study were explained to each participant while obtaining their informed consent. Rejection rate as per faculty is provided in Multimedia Appendix 1.

The study protocol and data collecting instruments were reviewed and approved by the Health Sciences Ethical Committee for Student Research. In addition, the permission of the vice deans of each faculty and university in which the study was conducted was sought. The objectives of the study were explained to the students before acquiring their written consent. The anonymous questionnaires that were used assured privacy and confidentiality of the data collected. All students participated voluntarily and understood that they had the right to refuse to partake in the study at any time without any academic penalty (Multimedia Appendix 2).

Study Questionnaire and Pretesting

A validated questionnaire from a study conducted in Ghana was used to collect data from the students [1]. However, the questionnaire was modified to better suit the students at Kuwait University. The questionnaire was translated and distributed in both English and Arabic for convenience. It consisted of 31 questions divided into 7 sections (Multimedia Appendix 3). The first section consisted of questions regarding students' basic sociodemographic characteristics (age, gender, and nationality). The other sections were validated tools to assess students' (1) access and use of Web-based resources; (2) use of devices, apps, and platforms; (3) use of Web-based resources for health purposes; (4) usage of the information found on the Web; (5) rating of the legitimacy of health information sought using Web-based sources; and (6) attitudes and beliefs about the use of Web-based resources for finding health information.

The questionnaire was reviewed by 2 faculty members, and their feedback was incorporated during revision of the questionnaire. The questionnaire was pretested on 10 students similar to the target population, and both the English and Arabic versions were revised as required after pretesting. The questionnaires took on average 5 to 10 min to complete.

Data Collection

All the data were collected in a 2-week period (from March 29 to April 14, 2018). A total of 1253 questionnaires were collected during this period.

Statistical Analysis

Data were entered, cleaned, and analyzed using the Statistical Package for the Social Sciences version 25 (IBM). Univariate analysis was performed to calculate percentages, frequencies, means, and standard deviations. Significant associations between dependent and independent categorical variables were tested using the Pearson chi-square test. The binary logistic regression model was used to adjust the odds ratio for potential cofounding variables. Age, gender, and level of experience with internet use were the variables included as covariates in the model for adjustment as they were significantly associated with the score in the crude analysis. A *P* value of ≤.05 and 95% CI were considered to be the levels of significance.

Results

Table 1 presents the characteristics of the study participants. The variables are gender, age, faculty, nationality, year of study, frequency of Web use, main device used to access the internet, type of internet used on campus, primary source of internet access, and the level of experience with Web use. The study participants were 1132 students from 8 different faculties. The age of the participants varied from 18 to 25 years with an average of 20 years.

Most of the respondents were Kuwaitis (993/1132, 87.72%) and females (880/1132, 77.73%). Most of the students were from the Faculty of Business Administration (192/1132, 16.96%) with only few from the Faculty of Pharmacy (76/1132, 6.71%). Students in their first year of study were the most (274/1132, 24.20%), and students in their seventh year were the least (14/1132, 1.23%). The students most often accessed the Web daily (1063/1132, 93.90%) using smartphones as the main device (1089/1132, 96.20%). Mobile data was used the most on campus (726/1132, 64.13%), with only few students using the internet cable (3/1132, 0.26%). In addition, the primary source of internet use was a personal device and the least was the university computer laboratory. Most students were very experienced with Web use (623/1132, 55.03%), whereas few students rated themselves as not experienced (11/1132, 0.97%).

The sociodemographic status of the students is presented in Table 1. The mean age of the students was 20.7 years, and most of the students who participated in the study were Kuwaitis (993/1132, 87.72%) and females (880/1132, 77.73%). Approximately 16.96% (192/1132) of the students were from the Faculty of Business Administration, whereas 6.71% (76/1132) of the students were from the Faculty of Pharmacy. Almost 24.20% (274/1132) of the students in the study were freshmen. Most of the students used the Web daily (1063/1132, 93.90%). In addition, 96.20% (1089/1132) of the students used their smartphones as the main device while accessing the Web. Mobile data was the most used type of internet on campus (726/1132, 64.13%). Most of the students used their personal device as the primary source of internet (972/1132, 85.86%). Almost half of the students (1132/623, 55.03%) were very experienced with using the internet.

Table 1. Sociodemographic characteristics of the study sample (N=1132).

Variables	Values
Age (years), mean (SD)	20 (1.919)
Nationality, n (%)	
Kuwaiti	993 (87.72)
Non-Kuwaiti Arab	123 (10.86)
Non-Kuwaiti non-Arab	14 (1.23)
Others	2 (0.17)
Gender, n (%)	
Male	252 (22.26)
Female	880 (77.73)
Faculty, n (%)	
Medicine	190 (16.78)
Pharmacy	76 (6.71)
Dentistry	60 (5.30)
Science	170 (15.01)
Education	163 (14.39)
Business administration	192 (16.96)
Sharia and Islamic studies	173 (15.28)
Law	108 (9.54)
Year of study, n (%)	
First	274 (24.20)
Second	182 (16.07)
Third	268 (23.67)
Fourth	223 (19.69)
Fifth	117 (10.33)
Sixth	47 (4.15)
Seventh	14 (1.23)
How often do you use the Web (days), n (%)	
Daily	1063 (93.90)
1-3	23 (2.03)
4-7	44 (3.88)
Never	2 (0.17)
Main device used, n (%)	
Smartphone	1089 (96.20)
Laptop	23 (2.03)
Tablet	14 (1.23)
Desktop computer	6 (0.53)
Types of internet used on campus, n (%)	
Mobile data	726 (64.13)
Campus Wi-Fi	71 (6.27)
Internet cable	3 (0.26)
Personal Wi-Fi	332 (29.32)
Primary source of internet use, n (%)	

Variables	Values
University computer laboratories	29 (2.56)
University library	34 (3.00)
Home	97 (8.56)
Personal device	972 (85.86)
Level of experience with internet use, n (%)	
Very experienced	623 (55.03)
Fairly experienced	498 (43.99)
Not experienced	11 (0.97)

Table 2 shows significant associations between students who seek Web-based health information (n=1126) and age, gender, faculty, year of study, primary source of internet, and level of experience with *P* values <.05 with the exception of nationality, frequency of using the internet, main device used, type of internet used on campus, and internet use in years. Significant associations were identified as *P*<.05.

The association between the sociodemographic factors of the students and Web-based health information seeking is presented in Table 2. In total, 90.0% (325/361) of the older age group (more than 21 years of age) used the Web for health information seeking compared with 82.8% (275/332) of the relatively young age group (18-19 years of age). The association between age and Web-based health information seeking is significant (*P*=.006). However, there was an insignificant association between the students' nationality and seeking Web-based health information. Female students showed higher prevalence of health information seeking (829/934, 88.8%) than their male counterparts (210/270, 77.8%; *P*<.001).

The association between faculties and Web-based health information seeking was found to be statistically significant (*P*=.01). The highest prevalence of seeking health information on the Web was among students of the faculties of Pharmacy (79/83, 95.2%), Dentistry (63/68, 92.6%), Medicine (178/198, 89.9%), and Science (167/191, 87.4%). On the contrary, the faculties that had a lower prevalence were faculties of Sharia and Islamic studies (152/181, 84.0%), Business Administration (167/199, 83.9%), Law (98/118, 83.1%), and Education (137/169, 81.1%). Most of the students who used the Web for health information were from the sixth year (54/58, 93.1%), whereas first year students were the least to use the Web for retrieving health information (233/284, 82.0%; *P* value=.003).

In addition, the study showed no significant association between the frequency of Web use, main device used, and the type of internet used on campus with Web-based health information seeking. In addition, there was no significant association between internet use in years with Web-based health information seeking.

Most of the students (874/997, 87.7%) used personal devices as their primary source of internet, whereas the university library was the least used source (32/44, 72.7%). About 88.6% (576/650) of the students were very experienced with internet use, whereas 79% (15/19) were not experienced. Both associations between the primary source of internet and level of experience and Web-based health information seeking were found to be significant (*P* values=.005 and .008, respectively). After adjusting for odds ratio, nationality, faculty, year of study, frequency of internet use, main device used, type of internet used on campus, and primary source of internet were found insignificant, whereas age, gender, and level of experience were found to be significant.

Table 3 shows significant associations between students who use Web-based health information seeking and apps, search engines, YouTube, emails to doctors, and websites (*P*<.05). The students' use of devices, apps, and platforms to seek health information is demonstrated in Table 3. The study found that students use search engines, for example, Google, Bing, and Yahoo, the most to seek health information (954/1091, 87.44%). YouTube (882/1007, 87.58%) and apps for smartphones and tablets (93/104, 89.4%) were found to be the second and third most used to seek health information, respectively. The least used method to seek health information was email to doctors (282/312, 90.4%). Twitter was the most used (314/358, 87.7%) social media platform in terms of seeking health information, whereas Facebook was the least used (10/837, 1.1%) social media platform.

Table 2. Relationship between sociodemographics and Web-based health information seeking (N=1126).

Variables	Total, N	Web-based health information seeking, n (%[a])	*P* value[b]
Age (years)			
18-19	332	275 (82.8)	.006[c]
20-21	433	371 (85.7)	.006[c]
More than 21	361	325 (90.0)	.006[c]
Nationality			
Kuwaiti	1045	898 (85.93)	.77[d]
Non-Kuwaiti Arab	132	117 (88.6)	.77[d]
Non-Kuwaiti non-Arab	17	15 (88.2)	.77[d]
Others	2	2 (100.0)	.77[d]
Main device used			
Smartphone	1123	971 (85.77)	.69[d]
Laptop	38	32 (84.2)	.69[d]
Tablet	29	25 (86.2)	.69[d]
Desktop computer	12	9 (75.0)	.69[d]
Type of internet used on campus			
Mobile data	743	644 (86.7)	.75[d]
Campus Wi-Fi	91	76 (83.5)	.75[d]
Internet cable	3	3 (100.0)	.75[d]
Personal Wi-Fi	370	318 (85.9)	.75[d]
Primary source of internet			
University computer laboratory	39	30 (76.9)	.005[d]
University library	44	32 (72.7)	.005[d]
Home	126	104 (82.5)	.005[d]
Personal device	997	874 (87.7)	.005[d]
Gender			
Male	270	210 (77.8)	<.001[d]
Female	934	829 (88.8)	<.001[d]
Faculty			
Medicine	198	178 (89.9)	.01[d]
Pharmacy	83	79 (95.2)	.01[d]
Dentistry	68	63 (92.6)	.01[d]
Science	191	167 (87.4)	.01[d]
Education	169	137 (81.1)	.01[d]
Business administration	199	167 (83.9)	.01[d]
Sharia and Islamic studies	181	152 (84.0)	.01[d]
Law	118	98 (83.1)	.01[d]
How often do you use the internet			

Variables	Total, N	Web-based health information seeking, n (%[a])	*P* value[b]
Daily	1079	938 (86.93)	.1[d]
1-3 days	53	40 (75.5)	.1[d]
4-7 days	73	61 (83.6)	.1[d]
Never	1	1 (100.0)	.1[d]
Year of study			
First	284	233 (82.0)	.003[c]
Second	199	168 (84.4)	.003[c]
Third	277	241 (87.0)	.003[c]
Fourth	234	207 (88.5)	.003[c]
Fifth	127	112 (88.2)	.003[c]
Sixth	58	54 (93.1)	.003[c]
Seventh	25	23 (92.0)	.003[c]
Level of experience with internet use			
Very experienced	650	576 (88.6)	.008[c]
Fairly experienced	537	449 (83.6)	.008[c]
Not experienced	19	15 (79)	.008[c]
Internet use (years)			
Less than 1	167	161 (96.4)	.11[d]
1-2	221	217 (98.2)	.11[d]
3-4	252	251 (99.6)	.11[d]
More than 4	406	398 (98.0)	.11[d]

[a] Row percentages.

[b] Significance: *P*<.05.

[c] Chi-square test for trend.

[d] Pearson chi-square test.

Table 3. Relationship between internet access methods and Web-based health information seeking

Variables	All (N)	Web-based health information seeking, n (%)	*P* value
Apps for smartphones and tablets	1006	891 (88.56)	<.001
Search engines (eg, Google, Bing, and Yahoo)	1091	954 (87.44)	<.001
Social media	901	790 (87.7)	.008
The most used social media platform			
Twitter	358	314 (87.7)	.96
Instagram	304	266 (87.5)	.96
Facebook	11	10 (90.9)	.96
Snapchat	156	134 (85.9)	.96
Messaging apps	104	93 (89.4)	.96
YouTube	1007	882 (87.58)	.007
Emails to doctors	312	282 (90.4)	.03
Websites	881	785 (89.1)	<.001

Table 4 shows that students mostly use Web-based sources to get health information to read about causes and symptoms of an illness (468/1049, 44.61%), followed by finding information to decide if a visit to the doctor is needed (206/1049, 19.63%). On the contrary, Web-based resources are used least to contact doctors on the Web (49/1049, 4.67%). In addition, most students (684/1049, 65.20%) reported that they actually find health information through Web-based resources, compared with a few who did not (378/1049, 36.03%).

The study revealed that 44.61% (468/1049) of the students used the internet to read about the causes and the symptoms of an illness. Approximately 19.63% (206/1049) of the students seek health information on the Web to decide if they need to consult a doctor, and 18.01% (189/1049) of the students seek Web-based health information before their appointment. The study also showed that 14.58% (153/1049) used the internet to find information after an appointment with a health professional. On the contrary, only 4.67% (49/1049) of the students contact their doctors on the Web. In addition, 65.20% (684/1049) of the students reported finding the desired health information through Web-based resources.

Table 5 describes how students rate the importance of several factors when reading about health information on websites. The study found that the most important factor is the credibility of the information sought (746/1049, 71.11%), followed by the accuracy of the information found (674/1049, 64.25%), and how easy the website is to read (625/1049, 59.58%). On the contrary, the students rated the interactivity of a website as the least important factor (418/1049, 39.84%), followed by the appearance and layout of the website as the second least impactful factor (263/1049, 25.07%) to determine the importance when reading about health information on websites.

The study found that the most important factor for reading health information on the Web chosen by the students was the credibility of the information sought (746/1049, 71.11%). Other important factors include the accuracy of the information found (674/1049, 64.25%) and how easy the website is to read (625/1049, 59.58%). On the contrary, the students rated the interactivity of a website as the least important factor (292/1049 27.83%), followed by the appearance and layout of the website as the second less important factor (377/1049, 35.93%).

Table 6 shows the students' feelings after obtaining health information on the Web. The study found that the most felt emotion to the least was relief (357/1049, 34.03%), confusion (311/1049, 29.64%), stress and anxiety (244/1049, 23.26%), and curiosity (141/1049, 13.44%).

Table 4. The frequency of internet use to get health information on the Web (N=1049).

Variables	Values, n (%)			
	Always	Often	Occasionally	Never
Find information to decide if you need to consult a doctor	206 (19.63)	303 (28.88)	436 (41.56)	117 (11.15)
Find health information before a medical appointment	189 (18.01)	284 (27.07)	383 (36.51)	207 (19.73)
Find information after an appointment with a health professional	153 (14.58)	271 (25.83)	357 (34.03)	282 (26.88)
Contact your doctor on the Web	49 (4.67)	111 (10.58)	226 (21.54)	678 (64.63)
Read about causes and symptoms of an illness	468 (44.61)	342 (32.60)	209 (19.92)	41 (3.90)
What is the frequency rate of finding the health information through Web-based resources	209 (19.92)	475 (45.28)	361 (34.41)	17 (1.62)

Table 5. Rating the importance of several factors when reading about health information on websites.

Factor	Values, n (%)		
	Not important	Fairly important	Very important
Accuracy	78 (7.43)	295 (28.12)	674 (64.25)
Current Information	83 (7.91)	360 (34.31)	602 (57.38)
Credibility	49 (4.67)	248 (23.64)	746 (71.11)
Comprehensiveness	101 (9.62)	396 (37.75)	547 (52.14)
Ease of understanding	80 (7.62)	341 (32.50)	623 (59.38)
Readability	83 (7.91)	331 (31.55)	625 (59.58)
Confidentially	239 (22.78)	316 (30.12)	482 (45.94)
Interactivity	418.0 (39.84)	333 (31.74)	292 (27.83)
Appearances	263 (25.07)	405 (38.60)	377 (35.93)

Table 6. Students' feelings after obtaining health information on the Web.

Variables	Values, n (%)
Relieved	357 (34.03)
Stressed and anxious	244 (23.26)
Confused	311 (29.64)
Curious	141 (13.44)

Most of the students were relieved after obtaining health information on the Web (357/1049, 34.03%). In contrast, 29.64% (311/1049) of the students reported getting confused after searching health information on the Web. About 23.26% (244/1049) of the students reported feeling stressed and anxious, whereas 13.44% (141/1049) reported feeling curious after retrieving health information on the Web.

Table 7 describes whether the Web-based resources obtained increased the awareness of the students about the topic. Most of the students agreed (508/1049, 48.42%) with the fact that their awareness has increased after seeking health information on the Web.

Most of the students at Kuwait University agree (803, 76.54%) that seeking health information on the Web increased their awareness about the topic that they read. However, only 4.19% (44/1049) disagreed to the fact that seeking Web-based health information increases their awareness.

Table 8 describes how often obtaining health information on the Web has improved the students' health. Most of the students reported that their health improved occasionally (517/1049, 49.28%) after obtaining the health information on the Web.

Around 49.28% (517/1049) of the students reported that their personal health improved occasionally after using Web-based health information. Meanwhile, 11.91% (125/1049) of them reported that using Web-based health information always improved their personal health. Furthermore, the study showed that 8.48% (89/1049) of the students reported that their health never improved after obtaining health information on the Web.

Table 9 illustrates the binary logistic regression of significant sociodemographic factors associated with Web-based health information seeking. Students aged 21 years and above are 2.391 times more likely to seek health information on the Web than students aged 18 to 19 years. Age was significantly associated with Web-based health information seeking (*P*=.001). In addition, female students are 2.781 more likely to seek health information on the Web than male students. There was also a significant correlation between gender and Web-based health information seeking with a *P* value of <.001. Moreover, students who are not experienced with internet use are 0.528 times less likely to seek health information than students who are very experienced with internet use. This relationship between the level of experience with internet use and health information seeking was significant (*P*=.01). The table was adjusted for all significant variables in the chi-square table. Significant associations were identified as *P*<.05.

Students aged 21 and above are 2.391 times more likely to seek health information on the Web than students aged (18-19) In addition, female students are 2.781 more likely to seek health information on the Web than male students. Moreover, students who are not experienced with internet use are 0.528 times less likely to seek health information than students who are very experienced with internet use. The relationship between age (*P*=.001), gender (*P*<.001), and level of experience with internet use (*P*=.01) and health information seeking were significant.

Table 7. Increased awareness about the topic from resources obtained on the Web.

Variables	Values, n (%)
Strongly agree	295 (28.12)
Agree	508 (48.42)
Not sure	207 (19.73)
Disagree	34 (3.24)
Strongly disagree	10 (0.95)

Table 8. Frequency of improvement of personal health after seeking health information on the Web.

Variables	Values, n (%)
Always	125 (11.91)
Usually	325 (30.98)
Occasionally	517 (49.28)
Never	89 (8.48)

Table 9. Binary logistic regression of significant sociodemographic factors associated with Web-based health information seeking (n=1126).

Sociodemographic factors	Adjusted odds ratio (95% CI)	*P* value[a]
Age (years)		
18-19	Reference	.001
20-21	1.322 (0.884-1.977)	.001
More than 21	2.391 (1.496-3.821)	.001
Gender		
Male	Reference	<.001
Female	2.781 (1.912-4.044)	<.001
Level of experience		
Very experienced	Reference	.01
Fairly experienced	0.600 (0.421-0.855)	.01
Not experienced	0.528 (0.528-0.160)	.01

[a] Significance: *P*<.05.

An illustration of the percentage of barriers faced while using the Web for health information is presented in Multimedia Appendix 4. Approximately 46.7% of the students had difficulty in understanding medical terms; however, 7.3% of the students reported the absence of facilitators as the problem that they faced when seeking health information on the Web.

The percentage for what makes the Web preferable in seeking health information is shown in Multimedia Appendix 5. It was found that about 83.74% (948/1132) of the students search for health information on the Web out of interest and curiosity. On the contrary, 22.34% (253/1132) of the students use Web-based resources owing to a lack of medical insurance.

The percentage of students' usage of information found on the Web is illustrated in Multimedia Appendix 6. The percentage of those who used Web-based health information to make lifestyle changes was 68.63% (777/1132). Meanwhile, 22.08% (250/1132) of the students changed their medication without discussing it with their doctor after reading health information on the Web.

Discussion

This study found an overall high prevalence of Web-based health information seeking, with a variation that was noted according to gender, year of study, and faculty. Other variables were identified that affect the association, and the effect on attitudes was explored.

Prevalence of Web-Based Health Information Seeking Among Students at Kuwait University

The findings of the study revealed that most of the students 86.21% (976/1132) reported using Web-based resources to obtain health information, whereas 13.86% (157/1132) of the students have never used Web-based resources for health information. The same results were found in several studies conducted in Kuwait, Saudi Arabia, and Qatar that showed the prevalence to be 62.9%, 58.4%, and 71.1%, respectively [10-12]. In the United States, a study reported the high prevalence of using Web-based resources to obtain health information (61%) [9]. In contrast, a study conducted in Europe reported that the prevalence of using the Web to retrieve health information was 41.4%, 38.7%, and 33.5% in Denmark, the Netherlands, and Sweden, respectively. The same study also showed the prevalence to be 11.7%, 13.5%, 14%, and 15.3% in Greece, Spain, Portugal, and France, respectively [13]. The increased prevalence of using the Web to retrieve health information could be attributed to the fact that the Web is becoming more easily accessible and affordable to everyone, in addition to the availability of personal devices and campus Wi-Fi, which is provided to the public with minimal charge.

Association Between the Sociodemographic Factors of the Students and Their Health Information–Seeking Patterns

This study found that most of the Web-based health seekers were aged older than 21 years, whereas a minority of the seekers were within the younger demographic. A study conducted in Portland also found that young adults aged between 18 to 29 years use Web-based resources to seek health information [14]. In addition, it was found that people aged less than 30 years search health information on the Web in a study conducted in France [13]. With regard to the academic year, this study showed that students in their senior years sought health information more than students in their junior years. An increase in searching the Web for health information was noticed in the older age group (21 years and older) compared with the younger age group (18 to 19 years). These trends could be attributed to the increased independence as the students grow older, because they are more likely to act without consulting any family member about their condition. This necessitates prior knowledge about the health situation, which can be obtained through Web-based health information seeking.

In addition, the results of this study showed that non-Kuwaiti Arab students (n=117) sought health information on the Web more than Kuwaiti students (n=898). This has been contradicted in a study conducted in Kuwait, in which Kuwaiti participants (n=58) were more likely to use Web-based resources to obtain health information than non-Kuwaitis (n=71) [10]. The higher percentage of Web-based health information seeking among non-Kuwaitis could be due to the lack of medical insurance and due to the fact that a recent increase in health service expenses was applied to expatriates.

With regard to gender, this study shows a positive association that was significant. The results of several studies that were conducted in Egypt and France were consistent with this study in that there was a significant association between gender and health information seeking on the Web [7,13]. In addition, this study also showed that female students seek health information more than their male counterparts with an adjusted odds ratio of 2.781. This is supported by a study conducted in France, which showed that the odds of female students seeking health information on the Web are 1.64 times higher than the male students [13]. These results were consistent with a study conducted in Nigeria [15]. In contrast, the study that was conducted in Kuwait contradicted this result, indicating that Web-based health information seeking was higher in males (n=96) than in females (n=33) [10]. The difference in the proportion of health information seeking among females and males could be explained by the fact that women are more curious and concerned about their health, in addition to the increased stigma about female reproductive health in Kuwait.

The findings of this study show that students who majored in faculties of Science were more likely to seek health information than those who majored in faculties of Literature. These results were supported by a study done in Nigeria, which found that students in Science faculties use Web-based resources more than students in non-Science faculties [16]. In this study, we found that students in the Faculty of Pharmacy sought health information on the Web the most and students in the Faculty of Education were the least to seek health information on the Web. Additionally, the study conducted in Nigeria showed that the highest percentage of Web use for health seeking was among Health Sciences students, whereas the lowest proportion of Web use to seek health information was among students of the Faculty of Education [16]. This could be justified by the fact that medical, dentistry, and pharmacy students in particular are more likely to search for health information on the Web as it is related to their majors; therefore, Web-based resources are used to learn about specific health information for academic purposes, rather than being sought to improve health.

With regard to the main device used to obtain health information on the Web, this study found that most students used the smartphone, which can be explained by the unparalleled accessibility that mobile phones offer, which is access to Web-based resources. This is supported by a study conducted in Kuwait in 2017, which showed that 99.5% of university students reported owning or using a mobile phone [17].

In the study, it was found that the most frequently used type of internet on campus was mobile data (n=644) and the least was internet cable (n=3). This result is supported by the study in Ghana, which reported that the main type of internet used among students to seek Web-based health information was mobile data (n=268) and the least was local area network (n=33) [1].

This research also showed that the primary source of internet used by students to obtain health information was personal devices (n=874), compared with the study conducted in Ghana, which showed campus laboratories and Wi-Fi to be the primary source of internet used (n=268) [1]. These findings could be explained by the fact that young adults in Kuwait are up to date with the advancements in technology and so most of them own a mobile phone, which makes it more convenient to use mobile data.

The findings of this study showed that students who are more experienced with Web use (576/650, 88.6%) rely on Web-based resources to seek health information more than those who are fairly experienced (449/537, 83.6%) or inexperienced (15/19, 79%). In a study conducted in the United States in (2011-2014), it was found that throughout the 4 years of this study, participants who had a high level of Web experience were more likely to seek health information than those who had no level of Web experience [18]. This result was consistent with another study conducted on adolescents from the United Kingdom and the United States, which found that a significant number of participants who had previous Web experience were actively involved in Web-based health information retrieval [19]. This difference in proportion may be explained by the difficulties faced by inexperienced individuals when seeking health information on the Web, compared with those with prior experience with Web use.

Influence of the Information Found on the Web on the Attitudes of Students at Kuwait University

In this study, it was found that most of the students use the health information that they sought on the Web to make lifestyle changes and discuss this information with their doctors, and only a few of them used the health information to change their medication without discussing it with their doctors. Similar results were found in a study conducted in Ghana; it was found that most of the students used the health information to make lifestyle changes and a few of them used it to make, cancel, or change their appointments with their doctors [19]. In addition, a study that was conducted in India found that most of their respondents shared the health information with friends and family and that the major factors that affect Web-based health information seeking were behavioral and habitual changes [20].

The findings of this study also revealed that most of the students felt relieved after reading about health information on the Web (357/1049, 34.03%).

Moreover, most of the students agreed (508/1049, 48.42%) with the fact that their awareness was increased after seeking health information on the Web. In addition to that, it was found that most of the students reported that their health improved after reading health information on the Web (517/1049, 49.28%). This result was supported by the findings in a Ghana study that revealed that the students' health conditions improved after seeking health information on the Web [1]. This can be attributed to the fact that finding health information on the Web increases the individual's awareness about their health, which can lead to making better informed health decisions.

With regard to the barriers that the students face while searching for health information on the Web, the results showed that the dominant barriers were understanding medical terms and a lack of Web searching skills. This dominance of the difficulty in understanding medical terminology could be explained by the fact that most of the students have no knowledge of medical terminology and most of the websites used for health information seeking use medical terms. On the contrary, the lack of skill needed could be attributed to a lack of proper training in effective Web use for health-seeking purposes.

Finally, this study focused on the factors that made the Web preferable for seeking health information. Interest and curiosity were the dominant factors 83.74% (948/1132) that affected the students to consider Web-based health information preferable. This can be explained by the innumerable amount of health information accessible to students through these recourses. The least reported factor to make the Web a preferable source of health information is lack of medical insurance. This can be due to the fact that health care is provided to Kuwaiti people free of charge, and the sample mostly consists of Kuwaiti students.

Conclusions

The study concluded that there is a large prevalence of Web-based health information seeking among students at Kuwait University. In addition, sociodemographic factors, for example, age, gender, faculty, year of study, primary source of internet, and level of experience with the Web, were significant correlates to Web-based health information seeking among the students. The study also found that more than half of the students who seek health information on the Web made lifestyle changes or discussed the information found with their doctor, although this finding was not statistically significant, possibly because of chance. There is an increasing trend in Web-based health information seeking among college students worldwide. Therefore, people should become more aware about the quality of the information sought, and doctors should educate the people to increase that awareness.

Strengths

Few studies have been conducted in Kuwait to study the association between seeking health information on the Web and its influence on the population. Another strength is the appropriate sample size, covering 8 faculties from both Science and Literature. Moreover, the 8 faculties were randomly selected from the available faculties of Kuwait University. Nevertheless, some limitations were identified.

Limitations

The limitations of the study are as follows: (1) the fact that university students may not be totally representative of all health information seekers as it is a young population and not all age groups are included in the study, and therefore, the study does not include people from different educational backgrounds and age groups; (2) in addition, the study was conducted in Kuwait University, in which most of the students are Kuwaiti, unlike the actual population in Kuwait, which consists of a majority of non-Kuwaiti individuals; and (3) most of the study population consisted of females, which could have affected the results with regard to gender.

Recommendations

Future studies should be conducted to (1) assess the level of awareness of Web users when it comes to the reliability and integrity of websites that are used to seek health information, (2) increase the awareness of the users with regard to trusted websites, and (3) include cultural factors and their effects on Web-based health information–seeking behaviors of users, including students, and factors that affect Web-based health information seeking, such as having a chronic illness or having a family member affected by an illness.

Conflicts of Interest

None declared.

Multimedia Appendix 1
The rejection rate.
[PNG File , 196 KB - formative_v3i4e14327_app.png]

Multimedia Appendix 2
Informed consent form, human subjects form, and ethical portfolio.
[PDF File (Adobe PDF File), 260 KB - formative_v3i4e14327_app.pdf]

Multimedia Appendix 3
Data collection instruments (questionnaires).
[PDF File (Adobe PDF File), 186 KB - formative_v3i4e14327_app.pdf]

Multimedia Appendix 4
Percentages of barriers faced while using the internet for health information.
[PNG File , 92 KB - formative_v3i4e14327_app.png]

Multimedia Appendix 5

Percentage of what makes the internet preferable for seeking health information.
[PNG File , 113 KB - formative_v3i4e14327_app.png]

Multimedia Appendix 6
Percentage of students' usage of information found on the Web.
[PNG File , 122 KB - formative_v3i4e14327_app.png]

References

1. Asibey BO, Agyemang S, Dankwah AB. The internet use for health information seeking among Ghanaian university students: a cross-sectional study. Int J Telemed Appl 2017 Oct 31;2017:1756473 [FREE Full text] [doi: 10.1155/2017/1756473] [Medline: 29225620]
2. Al-Shorbaji N. E-health in the eastern Mediterranean region: a decade of challenges and achievements. East Mediterr Health J 2008(14 Suppl):S157-S173 [FREE Full text] [Medline: 19205616]
3. Higgins O, Sixsmith J, Barry MM, Domegan C. A Literature Review on Health Information-Seeking Behaviour on The Web: A Health Consumer and Health Professional Perspective. Stockholm, Sweden: European Center for Disease Prevention and Control; 2011.
4. Murray E, Lo B, Pollack L, Donelan K, Catania J, White M, et al. The impact of health information on the internet on the physician-patient relationship: patient perceptions. Arch Intern Med 2003 Jul 28;163(14):1727-1734. [doi: 10.1001/archinte.163.14.1727] [Medline: 12885689]
5. Ball MJ, Lillis J. E-health: transforming the physician/patient relationship. Int J Med Inform 2001 Apr;61(1):1-10. [doi: 10.1016/s1386-5056(00)00130-1] [Medline: 11248599]
6. Lenfant C, Taggart VS. National asthma education and prevention program: expert panel report 2. Guidelines for the diagnosis and management of asthma: translating research for clinicians and patients. Allergol Int 1999 Oct;48(4):223-230. [doi: 10.1046/j.1440-1592.1999.00146.x]
7. Ghweeba M, Lindenmeyer A, Shishi S, Abbas M, Waheed A, Amer S. What predicts online health information-seeking behavior among Egyptian adults? A cross-sectional study. J Med Internet Res 2017 Jun 22;19(6):e216 [FREE Full text] [doi: 10.2196/jmir.6855] [Medline: 28642216]
8. AlGhamdi KM, Moussa NA. Internet use by the public to search for health-related information. Int J Med Inform 2012 Jun;81(6):363-373. [doi: 10.1016/j.ijmedinf.2011.12.004] [Medline: 22217800]
9. Liang B, Scammon DL. Incidence of online health information search: a useful proxy for public health risk perception. J Med Internet Res 2013 Jun 17;15(6):e114 [FREE Full text] [doi: 10.2196/jmir.2401] [Medline: 23773974]
10. Alkhatlan HM, Rahman KF, Aljazzaf BH. Factors affecting seeking health-related information through the internet among patients in Kuwait. Alexandria Med J 2017;54(4):331-336. [doi: 10.1016/j.ajme.2017.05.008]
11. AlSaadi MM. Evaluation of internet use for health information by parents of asthmatic children attending pediatric clinics in Riyadh, Saudi Arabia. Ann Saudi Med 2012;32(6):630-636 [FREE Full text] [doi: 10.5144/0256-4947.2012.630] [Medline: 23396028]
12. Choudhury SM, Arora T, Alebbi S, Ahmed L, Aden A, Omar O, et al. How do Qataris source health information? PLoS One 2016;11(11):e0166250 [FREE Full text] [doi: 10.1371/journal.pone.0166250] [Medline: 27832112]
13. Spadaro R. European Commission. 2003. Eurobarometer 58.0: European Union Citizens and Sources of Information About Health. URL: https://ec.europa.eu/health/ph_information/documents/eb_58_en.pdf [accessed 2019-09-10]
14. Gowen LK. Online mental health information seeking in young adults with mental health challenges. J Technol Hum Serv 2013 Apr;31(2):97-111. [doi: 10.1080/15228835.2013.765533]
15. Oriogu CD, Subair R, Oriogu-Ogbuiyi DC. Use of internet health information resources and information seeking behaviour among health professionals in Federal Medical Center, Abuja. Libr Philos Pract 2017 Feb;2017(1):- [FREE Full text]
16. Awoleye MO, Siyanbola WO, Oladipupo OF. Adoption assessment of internet usage amongst undergraduates in Nigeria universities -a case study approach. J Technol Manag Innov 2008;3(1):- [FREE Full text]
17. Alzougoo B, AlMansour J. The Use of Smartphone for Learning Activities by University Students in Kuwait. In: Proceedings of Teaching and Education Conferences. 2017 Presented at: TEC'17; April 24-27, 2017; Venice, Italy URL: https://ideas.repec.org/p/sek/itepro/4907508.html [doi: 10.26484/2017/tjsss00217072]
18. Gray NJ, Klein JD, Noyce PR, Sesselberg TS, Cantrill JA. Health information-seeking behaviour in adolescence: the place of the internet. Soc Sci Med 2005 Apr;60(7):1467-1478. [doi: 10.1016/j.socscimed.2004.08.010] [Medline: 15652680]
19. McMullan M. Patients using the internet to obtain health information: how this affects the patient-health professional relationship. Patient Educ Couns 2006 Oct;63(1-2):24-28. [doi: 10.1016/j.pec.2005.10.006] [Medline: 16406474]
20. Shubha HS. Relationship between internet use and health orientation: a study among university students. J Commun Media Technol 2015;5:103-113 [FREE Full text]

Edited by G Strudwick; submitted 09.04.19; peer-reviewed by H Alshawaf, R Chan; comments to author 11.06.19; revised version received 13.06.19; accepted 15.06.19; published 31.10.19.

Please cite as:
Ashkanani H, Asery R, Bokubar F, AlAli N, Mubarak S, Buabbas A, Almajran A
Web-Based Health Information Seeking Among Students at Kuwait University: Cross-Sectional Survey Study
JMIR Form Res 2019;3(4):e14327
URL: http://formative.jmir.org/2019/4/e14327/
doi:10.2196/14327
PMID:31473592

Original Paper

A Mobile Phone App for the Provision of Personalized Food-Based Information in an Eating-Out Situation: Development and Initial Evaluation

Katherine Marie Appleton[1], PhD; Jeff Bray[2], PhD; Sarah Price[2], PhD; Gernot Liebchen[3], PhD; Nan Jiang[3], PhD; Ioannis Mavridis[4], PhD; Laure Saulais[5], PhD; Agnès Giboreau[5], PhD; Federico J A Perez-Cueto[6], PhD; Rebecca Coolen[7], MA; Manfred Ronge[7]; Heather Hartwell[2], PhD

[1]Research Centre for Behaviour Change, Department of Psychology, Faculty of Science and Technology, Bournemouth University, Poole, United Kingdom

[2]Faculty of Management, Bournemouth University, Poole, United Kingdom

[3]Faculty of Science and Technology, Bournemouth University, Poole, United Kingdom

[4]University of Macedonia, Thessaloniki, Greece

[5]Food and Hospitality Research Center, Institut Paul Bocuse, Ecully, France

[6]Department of Food Science, University of Copenhagen, Copenhagen, Denmark

[7]Ronge and Partner, Baden-bei-Wien, Austria

Corresponding Author:
Katherine Marie Appleton, PhD
Research Centre for Behaviour Change, Department of Psychology
Faculty of Science and Technology
Bournemouth University
Poole House
Fern Barrow
Poole, BH12 5BB
United Kingdom
Phone: 44 1202 965985
Email: k.appleton@bournemouth.ac.uk

Abstract

Background: Increasing pressure from governments, public health bodies, and consumers is driving a need for increased food-based information provision in eating-out situations. Meals eaten outside the home are known to be less healthy than meals eaten at home, and consumers can complain of poor information on the health impact and allergen content of meals eaten out.

Objective: This paper aimed to describe the development and early assessment of a mobile phone app that allows the provision of accurate personalized food-based information while considering individual characteristics (allergies, diet type, and preferences) to enable informed consumer choice when eating out.

Methods: An app was designed and developed to address these requirements using an agile approach. The developed app was then evaluated at 8 public engagement events using the System Usability Scale (SUS) questionnaire and qualitative feedback.

Results: Consideration of the literature and consultation with consumers revealed a need for information provision for consumers in the eating-out situation, including the ability to limit the information provided to that which was personally relevant or interesting. The app was designed to provide information to consumers on the dishes available in a workplace canteen and to allow consumers the freedom to personalize the app and choose the information that they received. Evaluation using the SUS questionnaire revealed positive responses to the app from a range of potential users, and qualitative comments demonstrated broad interest in its use.

Conclusions: This paper details the successful development and early assessment of a novel mobile phone app designed to provide food-based information in an eating-out situation in a personalized manner.

(JMIR Form Res 2019;3(4):e12966) doi:10.2196/12966

KEYWORDS

eating; eating behavior; food; diet; mhealth; mobile app; digitalhealth; smartphone

Introduction

Eating Out

Eating out (defined as eating food that is prepared by others and consumed out of the home in food establishments such as restaurants, cafes, canteens, and fast food outlets) is a growing trend [1-3]. Food consumed out of the home can represent a significant contribution to daily energy intake [4,5] and is often higher in calories, fat, and sugar, lower in fiber, and served in larger portions than food consumed in the home [4,5]. It has been suggested that both the increasing practice of eating out and the increasing energy provision in this situation contribute to the obesogenic environment [1,2,6-8], and positive associations between the increased prevalence of eating out and high body weight in recent years have been made [6,8]. Obesity and its related conditions constitute a major public health concern, and strategies for prevention are required [9]. Furthermore, most adults underestimate their calorie intake when eating out and are therefore unlikely to mitigate their own risk [1,2].

Food-Based Information

With an aim of reducing overeating, a regulation within Europe (EU No 1169/2011) makes it mandatory for all prepacked foods to display energy value and amounts of fat, saturates, carbohydrates, protein, sugars, and salt per 100 g or 100 ml food item [10]. Presently, only the United States has passed legislation requiring the provision of this type of labeling for out-of-home food provision [11], although voluntary labeling, particularly of energy content, is undertaken in various countries across the world [12]. Importantly, however, consumers are also asking for increased food-based information in the eating-out scenario [13-17], and the provision of information has been found to facilitate the adoption of healthier nutrition practices in this scenario [1-3,15,16,18,19]. Bates et al [1] and Burton et al [2] report lower repurchase intentions for unhealthy dishes following exposure to objective calorie and nutrient information. VanEpps et al [19] report fewer calories ordered following the provision of nutrition information, and Hammond et al [18] report fewer calories ordered and consumed following a nutrition-labeling intervention. Not all studies have reported benefits [20], but effects are typically small, and a lack of effects has been largely attributed to poor study size or methodology [20].

Food-based information provision in an eating-out situation, however, can result in menus or information boards that are described as cluttered and untidy, the consequences of which are that consumers feel overwhelmed and report that they would not use the information provided [13,16,17,21]. Consumers can report an unwillingness to search for information or high costs of this practice compared with benefits [1,22,23]. Food operators also acknowledge constraints to providing ingredient information on menus [15,16,21], including a reluctance to overload the menu with visual clutter, as well as a lack of knowledge on the part of the operator and a high perceived cost to the business, particularly where menus and dish specifications may change frequently [15,16,21].

Furthermore, different consumers can prefer or typically act on different types of information [2,16,17,22,24]. Yepes [24], for example, found menu calorie labeling was most valued by health conscious and older consumers and was effective in reducing calorie consumption in these individuals, whereas Burton et al [2] demonstrate impacts of calorie information on *low-health consciousness* consumers and impacts of calorie information and a color coding system for both *low-* and *high-health consciousness* consumers. Ellison et al [3] also found increased impacts of calorie information and a color-coding system compared with calorie-based information alone. The lack of effects of food information in some studies has been attributed to the type of information provided combined with the abilities of the consumer [20,25]. Different types of information may also be differentially effective in different settings. Bleich et al [20] and Dumanovsky et al [25] both report increased impacts of calorie information in fast food restaurants compared with sit-down type establishments. Food-based decisions are often made using simple heuristic processing methods [25], and decisions that are required under time pressure and high cognitive load can further be reliant on simple or automatic processing strategies, such as a calorie limit. Decisions that take place in a more relaxed setting, however, such as at home, can be more consciously taken and so more informed, for example, through the consideration of complex ingredient information [19,25].

Food-Based Information Provision

For decades, printed mediums have been the most common platform used for information provision in an eating-out scenario, but digital or electronic menus can deliver rich knowledge in an efficient manner [16,26] and may offer a neat solution to the existing constraints highlighted earlier [16,22,26]. Several studies have highlighted the potential that technology may hold in providing information [16,22], and a small number of mobile phone apps have been developed specifically for providing food-based information [27,28]. *SmartAPPetite* encourages people to eat local and healthy food [29]. *Tapingo* provides university students with canteen-based information and allows consumers to order food [30], and the *Smartmenu system* discussed by Pieskä et al [31] allows consumers to browse a menu, check additional information such as nutrient profile, and order items. The potential value of mobile phone apps for providing information is also increased by the rapidly growing numbers of mobile phones. Penetration rates of 68.4% in North America and 64.7% in Western Europe have been reported, with estimations of use by 32% of the global population [32], ranging recently in the top 50 markets between 11.2% in Ethiopia and 82.2% in the United Kingdom [33].

Increased information provision via digital platforms may also enable transparency and evidence of greater integrity for the food service operator [16,17,26,34]. Consumers with specific dietary needs are often limited in their choices not just by their personal constraints but also by a lack of information available from serving staff or a lack of trust in the information provided [22,26,35]. Lack of control, insufficient knowledge, and a lack of trust in any information provided can be key concerns when eating out [22,23,26,34,36]. Trust is an important component of health-based decision making [34-36], and catering operators that are open and transparent demonstrate commitment and trustworthiness to consumers [17,34]. Furthermore, even if the actual content is not always used, consumers can be reassured by the presence of such information [37,38]. Thus, food operators will also potentially benefit from increased information provision [16,21,26].

To enable healthy decision making in an eating-out situation, communication with consumers is clearly required, but any such communication should be carefully considered to ensure that it is well understood, suitable for each consumer, and suited to specific dishes and food operators [1-3,16,17,22,24,28]. Many existing apps use generic recipe data and so provide only generic information [27,28,39]. This generic information typically provides only estimations of nutrient content and is rarely sufficiently detailed to protect consumers with allergies or other very specific dietary requirements. For increased specificity, information should ideally be provided for every specific recipe from every specific manufacturer or food operator. For many consumers, more detailed and specific information is preferable, and more detailed information could increase trust and return business for food providers. For allergens, very specific information is required.

A Mobile Phone–Based Technological Solution

This study aimed to develop a mobile phone app to provide consumers with food-based information in an eating-out situation. The app was intended to provide information as required for current European legislation and as recommended by a number of public health bodies, also with the consideration of consumer desires. The study was undertaken for the workplace canteen situation. There is a growing acceptance that the food provided in a workplace canteen setting can have a significant impact on health [40], as this is a captive environment where the contribution of the meal served could constitute an important element of the overall diet owing to the frequency of use. It is estimated that most employees eat one or more meals per day while they are at work [41]. Promotion of healthy diets in the workplace will also have benefits not only for the individual but also for employers and society [40]. The workplace canteen also typically offers limited dishes that can be more easily, accurately prespecified than may be the case for a chef-led restaurant [21]. The study was undertaken in Europe with a focus on European consumers, but evaluations also extended the study beyond Europe.

Methods

Overview

The study was undertaken in 3 stages. First, the requirements for the app were elicited through consideration of current legislative and scientific literature and consultation with potential end users, and then prioritized with potential stakeholders using the Must have, Should have, Could have, Won’t have (MoSCoW) method [42]. Second, a mobile phone app was designed and developed using an agile approach. Finally, the developed app was evaluated at 8 public engagement events using the System Usability Scale (SUS) questionnaire [43,44] and qualitative feedback. Full ethical approval was granted from Bournemouth University Research Ethics Committee before commencement. The research complied with Directive 95/46/EC of the European Parliament on the Protection of Individuals and with Directive 2002/58/EC of the European Parliament concerning the processing of personal data and the protection of privacy.

Stage 1: Defining the App Requirements

Eliciting Government Requirements and Recommendations

Current legislation and recommendations for Europe were obtained from the European Union and from relevant public health agencies in Europe.

Eliciting Consumer Requirements

Consumer requirements were gained from a search of the scientific literature and from consultations with consumers. The search of the scientific literature was undertaken using known articles and snowballing from these by also looking at cited and citing articles. Consumer consultations were undertaken using one mixed methods study comprising a qualitative and then quantitative component [45,46], and one qualitative study [47]. A formal review of existing apps was not undertaken as part of the development work. No comparable apps with the necessary specificity and flexibility of dish information existed as far as we were aware, and few apps are developed following (reported) formal consultation at an early stage with end users; thus, a review of existing apps was unlikely to be informative.

The consumer consultation work has been published in detail elsewhere [45-47]. Briefly, the mixed methods study [45,46] used 8 focus groups of canteen users (N=40), 2 groups in each of the 4 European countries (Denmark, France, Greece, and the United Kingdom), to elicit the criteria used for making food-based decisions in a canteen scenario and known formats of food-based information provision. The decision-related criteria were then used in a best-worst scaling questionnaire to ascertain the relative importance of these criteria for making food-based decisions in 452 employees (Denmark [N=100], France [N=100], Greece [N=100], the United Kingdom [N=152]), who had access to a canteen at their place of work. The full sample was largely composed of females (61.1%, 276/452), aged 20 to 29 years (51.3%, 232/452), who had completed some form of higher tertiary education (74.1%, 335/452), and who now worked full time (60.4%, 273/452) in occupations classified as technicians and associate professionals (74.1%, 335/452). The known formats of and preferences for food information were also considered in a second best-worst scaling questionnaire administered to the same individuals to ascertain the most preferred format for the provision of food information. The study was undertaken in several European countries to enhance the generalizability of the work. Focus groups were used in a small sample to ascertain relevant criteria of canteen use and consumer preferences for food information provision in depth, given a lack of existing data in this area. The questionnaire was then used to determine the relative importance of these criteria and information formats in a much larger sample. The best-worst scaling method requires respondents to choose their most preferred and least preferred option of several sets, allowing relative assessments of the criteria of interest without the use of absolute judgments that can differ between contexts and cultures.

The additional qualitative study [47] used 4 focus groups of canteen users from the United Kingdom (N=28) to confirm the information desired for food-based decisions in a canteen scenario and ascertain attitudes and opinions toward the use of information communication technology (ICT) in this context. Focus groups were again used in a small sample to ascertain attitudes and opinions in depth, given a lack of existing data in this area.

Prioritizing Requirements

Elicited requirements for the app were then prioritized using MoSCoW principles following full consideration by the research team and potential stakeholders to ensure wide use of the app and increased transferability. The MoSCoW method [42] is a technique used in software development to prioritize the importance of the delivery of all identified requirements. Requirements are categorized as *Must have*, *Should have*, *Could have*, and *Won't have*, based on importance, and then prioritized during the development process in this order. Requirements identified as *must have* are considered central to project success; those identified as *should have* are considered important, but not necessary; those identified as *could have* are considered desirable but not necessary; and those identified as *won't have* are considered least important [42]. The research team included academic researchers in eating behaviors, hospitality, and food service (KMA, JB, SP, FJAPC, and HH), academic researchers and developers in computing and app development (GL, NJ, and IM), caterers, food operators, and personnel from the catering industry (RC and MR) and researchers working within the food industry (LS and AG). Additional potential stakeholders from the food industry (caterers, food operators, and personnel from the food and hospitality industries) were also consulted. Each of the elicited requirements were discussed and considered for inclusion in the app.

Stage 2: Designing and Developing the App

The app was intended for use in a workplace canteen using a predetermined food menu offering a number of dishes and side dishes per day and was envisaged as a consumer-facing user interface attached to a back-end database. All aspects of the app from the user's perspective were carefully considered from visual esthetics of the user interface (eg, logo, color scheme, and picture placement) to the workflow of user tasks. A user journey map was first created to visualize the timeline of interactions with the potential app from the landing page. Wireframes of each app screen were then produced using Balsamiq (Balsamiq Studios). These wireframes focused on app screen layout and content structure and were organized to reflect the user journey map. These wireframes were then mapped to mockups showing the actual visual designs for each screen. An interactive prototype was created using InVision (InVision Technologies, Inc), and from this an Android app was developed using native Android Studio. Particular consideration was also required for the setup of the back-end database, as databases rely on precise definitions of entities stored, whereas natural languages do not have this level of precision and contain synonyms and terminology imprecision. For instance, the terms *dish*, *food item*, and *meal* needed clarification. The specification of ingredients, for example, milk, within ingredients, for example, cheese, within dishes, also added complexity.

The app was developed for Android mobile phones following Google's Material Design Guidelines and industrial best practices, with reference to the adapted Technology Acceptance Model (TAM) [48-50]. The adapted TAM proposes that technology usage is positively predicted by *perceived usefulness* ("the degree to which a person believes that using a particular system would enhance his or her job performance") [48], *perceived ease of use* ("the degree to which a person believes that using a particular system would be free of effort") [48], *perceived enjoyment* ("the extent to which the activity of using the [technology] is perceived to be enjoyable in its own right, apart from any performance consequences that may be anticipated") [49], and *perceived visual attractiveness* ("the degree to which a person believes that the [technology] is aesthetically pleasing to the eye") [50].

Throughout app development, several consultation meetings were held with all stakeholders to make sure the app would meet all demands and support further needs. A prototype app was developed to include all requirements identified as *must have* and *should have* using the MoSCoW framework. This prototype was then assessed and refined by the research team to create a second version.

Stage 3: Initial Evaluation of the App

Following development in 2016, version 2 of the prototype app was demonstrated to potential stakeholders and end users at several public engagement events in 2017, where assessments of the app were made using a questionnaire.

Public Engagement Events

Data were collected from adult participants at 8 global public research engagement events. These public engagement events were largely intended to promote research for the general public and included displays from a range of disciplines—that is, they were not held specifically to demonstrate or promote our app. The app was demonstrated, and data were collected at 4 public engagement events held in the United Kingdom, 1 event held in France, 1 event held in Denmark, 1 event held in Malaysia, and 1 event held in China. These events were chosen to represent the European consumers in the countries involved in the study, but also expanded consideration of the app to different markets. Demonstration at these events was intended to gain feedback and insights from a wide range of potential users of the app. These events are typically attended by members of the public who are educated and have an interest in science, education, and/or progress and growth. All attendees who were interested were permitted to try the app and provide data. There were no inclusion/exclusion criteria, to increase generalizability, but the information on the app was provided in English. Mock data on 4 mock dishes from 1 mock canteen were included in the app for demonstration purposes. Respondents were asked to view and manipulate the app for as long as they wished and then complete the evaluation questionnaire in paper form.

Evaluation

Assessments of the app were made using the SUS questionnaire [43,44]. The SUS consists of a 10-item questionnaire based on 5-point Likert scales (5=strongly agree to 1=strongly disagree) to assess usability. In total, 5 questions are positively phrased: “I think that I would like to use this system frequently,” “I thought the system was easy to use,” “I found the various functions in this system were well integrated,” “I would imagine that most people would learn to use this system very quickly,” and “I felt very confident using the system.” A total of 5 questions are negatively phrased: “I found the system unnecessarily complex,” “I think that I would need the support of a technical person to be able to use this system,” “I thought there was too much inconsistency in this system,” “I found the system very cumbersome to use,” and “I needed to learn a lot of things before I could get going with this system.” An additional question was also added to the end of the questionnaire “I believe the FoodSmart App will be useful to customers in a canteen setting to help them to get informed about the dishes offered.” There was also the opportunity for respondents to leave open-ended comments if desired. The SUS is recommended to assess usability, because it is technology-independent, short (and therefore easy to complete and analyze), and can provide a single score per person [51]. It was chosen for this study because it has been extensively used on a variety of products and systems, notably for measuring the functionality of apps. It is a well-validated instrument and supported by a large pool of comparison data.

Analysis

Questionnaire responses were entered into SPSS version 24. All questionnaires providing complete SUS data were used for analysis, although not all questionnaires included complete demographic data. The results of the tests for normality were acceptable, and the results of a confirmatory principal component analysis (Varimax rotation) on the responses from the 10 SUS questions were comparable with those conducted by other researchers [51,52]. Analysis of the raw item data gave 2 factors that explained 53.8% of the variance, consisting of (1) the 5 positively phrased items (explaining 41.4% of the variance) and (2) the 5 negatively phrased items (explaining a further 12.4% of the variance). Overall SUS scores based on the 10 SUS questions were calculated as described by Brooke [43]. Values of positively phrased items were reduced to a 0 to 4 scale by subtracting 1. Values of the negatively phrased items were subtracted from 5 to reverse their sense and to similarly reduce them to a 0 to 4 scale. The sums of the 10 items together were multiplied by 2.5 to provide individual SUS scores out of 100. SUS scores were analyzed using descriptive statistics and *t* tests. For significance testing, probability was considered at $P<.05$. The additional question given at the end of the questionnaire was considered as a separate individual question and analyzed using descriptive statistics only. Open-ended feedback comments were analyzed through open coding techniques and grouped into themes to provide an added richness of understanding.

Results

Stage 1: Defining the App Requirements

Government Requirements and Recommendations

European legislation requires the provision of information for all dishes served in an eating-out situation on 14 common allergens (celery, cereals, crustaceans, eggs, fish, lupin, milk, molluscs, mustard, nuts, peanuts, sesame, soybeans, and sulfur; EU regulation No 1169/2011). Recommendations at the European level also suggest the inclusion of information on calorie content and amounts of fat, saturates, carbohydrates, protein, sugars, and salt of a food per 100 g or 100 ml or per weight of portion served [10].

Consumer Requirements

The study by Price et al [45,46] revealed 8 criteria for food selection in a canteen setting: *Animal Welfare* (how an animal is coping with the conditions in which it lives, where an animal is in a good state of welfare if it is healthy, comfortable, well nourished, safe, able to express innate behavior, and is not suffering from unpleasant states); *Environmental Impact* (the effect the food production has on the environment); *Fair Trade* (Fair Trade aims to help producers to get a fair price for their products so as to reduce poverty, provide the ethical treatment of workers and farmers, and promote environmentally friendly and sustainable practices); *Naturalness* (the extent to which fresh ingredients are used, or that there is less use of processed foods containing additives and preservatives); *Nutrition* (the nutritional composition of the food); *Organic* (food produced in a way that respects natural life cycles, minimizes the human impact on the environment, and operates as naturally as possible); *Provenance* (where the food was grown/produced); and *Value for Money* (the ratio between the perceived quality of the dish and the price paid for it). Consumers recognized a variety of formats currently used to provide food-based information to consumers and expressed a desire for information that was easy to understand and easy to use, and preferences for only the information that was relevant to themselves, such as for their religious beliefs and their dietary preferences, eg, vegetarianism. Results from the questionnaire study on food-based decision making suggested that in all 4 European countries, the criteria of Nutrition, Naturalness, and Value for Money were those most valued by consumers, followed by criteria based on Animal Welfare, Organic foods, and Provenance, followed by those of Environmental Impact and Fair Trade [45]. Results from the questionnaire study on food-based information provision suggested that in all 4 European countries, the formats most preferred by consumers were traffic light labeling, information boxes, and quality assurance markings. Latent class cluster models also identified 5 clusters of consumers in relation to information use, described as *Heuristic Processors* (individuals who preferred easy-to-find-and-use information); *Brand Orientated* (individuals who were persuaded by brand authority); *Systematic Processors* (individuals who prefer more detailed information); *Independent Processors* (individuals who use a mix of heuristic and systematic processing); and *Tech-Savvy* (individuals who indicated a high preference for technology and interactive displays) [46]. Some differences between countries were also found [45,46].

The study by Bray et al [47] confirmed the importance for consumers of information on nutritional content (Nutrition), ingredients and allergens (Naturalness), and confirmed desires for information to be presented in a clear and simple manner, desires for different information by different consumers, and a desire to personalize the information that each person receives. Options to remove or override any personal preferences were also preferred, as opposed to limits to free choice. In relation to ICT, consumers currently used ICT for food-based reasons related to: *Marketing*, for example, discounts or loyalty schemes; *Increased convenience*, for example, Web-based booking and viewing menus in advance; and *Accessing additional information*, for example, identifying ingredients and customer reviews. Avenues for future ICT usage focused on the following: the provision of digital menus and increased information per dish; the consumer gaining control over and confidence in what they were eating; the value of personalized (relevant) information provision; and the value of additional information/free choice despite personalization. The idea of using a mobile phone app to provide personalized information was generally supported.

The literature in the field was also found to highlight similar opinions and concerns regarding food-based decision making [13,21,38,53] and supports similar conclusions regarding the provision of nutritional information in a retail setting [15,16,37] and in an eating-out setting [3,19,38,53]. Similar opinions toward technological solutions were also found [15,16,53].

Requirement Priorities

The priorities for the app based on MoSCoW principles [42] are presented in Table 1.

Table 1. MoSCoW (Must have, Should have, Could have, Won't have) requirements for the app.

MoSCoW	Requirements[a]
Must have	• Provide detailed and accurate dish information as supplied by the manufacturer, including ingredients and allergens; • Include nutrient information (calories, sugar, fat, saturated fat, and salt); • Include information allowing dietary classifications; • Include price per dish, allowing assessments of "value for money"; • Provide the information in an easily accessible format; • Enable quick information access, eg, via a QR (quick response) code; • Allow users to store personal preferences about dietary needs and requirements, for example, religion, vegetarian, and vegan; • Tailor menu presentation based on user profile; • Warn users for certain dishes based on user preferences, for example, allergens and religious dietary needs
Should have	• Adopt a traffic light type coding system for the nutritional information; • Provide additional detailed information if required; • Provide a calorie calculator allowing assessment of a whole meal composed of several dishes; • Allow users to set a desired calorie limit per dish; • Allow presentation of all dishes to retain free choice for the consumer while retaining a tailored presentation based on the user profile
Could have	• Provide information about ingredient provenance and organic nature; • Provide information about animal welfare, environmental impact, and fair trade nature of all ingredients; • Allow users to set favorite food region; • Allow users to set favorite dish or specific food items; • Enable recommendations based on user preferences; • Store previous purchase history; Enable recommendations based on previous consumption; • Provide warnings of over or excess consumption; • Provide personalized food messages for each user; • Allow sharing via social media; • Allow users to take photos of dishes/meals chosen; • Allow users to search for dishes; • Allow users to access menus in advance; Include functionality to preorder meals; • Include functionality to feedback dish choices to a canteen; • Include functionality to feedback comments/suggestions to a canteen;
Won't have	• Provide generic dish information; • Limit consumer choice; • Provide information on allergen traces; • Provide advertisements; • Support push notifications, for example, for special offers; • Include functionality to allow users to pay via the app; • Include functionality to feedback sales to a canteen

[a] Definitions: a) Dish: can be made up of several food items, for example, lasagna with side salad; b) Food item: something a consumer can buy, which has nutritional facts and can fit a food classification; c) Nutritional fact: a fact about the nutritional values of a food item (eg, salt level or sugar level); d) Food classification: information about food items in relation to dietary classifications such as vegetarian, vegan, kosher, or halal.

Stage 2: Designing and Developing the App

The app was developed as a consumer-facing user interface attached to a back-end canteen-based database. The back-end database held all required information per dish (ingredients, allergens, and nutritional composition), as supplied by caterers and food manufacturers. The app was designed such that caterers and food manufacturers would be given free and unlimited access to the database to upload information for as many dishes as they wished based on their own dish specifications and could update this information as often as they wished. Information in the database is stored per canteen.

The user interface was designed to allow consumers to view all information provided by caterers and food manufacturers, and to manipulate the information displayed if desired, through the selection of settings on the user interface that allowed consumers to input personal details and preferences. The information for each dish could be revealed by accessing a menu or by scanning a QR (quick response) code placed on a menu or dish label, thus allowing very quick access to all information if desired.

Particular consideration, throughout development, was given to the storage and security of personal data from users and dish and ingredient data from manufacturers; thus, all personal information is stored on the user's device, and all industry data are owned and managed by the operator.

Version 1 of the prototype included features to meet all requirements identified as *must have* and most of the features identified as *should have* using the MoSCoW method [42]. Full details of the user interface are given in Table 2. In addition to the features identified, version 1 of the app also included functionality for users to label dishes that they liked, but this functionality has not yet been linked to databases to allow feedback to caterers. Version 2 of the prototype app included all features identified as *must have* and *should have*, and so included all features of version 1, plus 3 additional features. First, the information provided per dish was no longer presented on a single screen but split over 3 screens (overview/description; nutritional information; and ingredients/allergens) to facilitate users to access only desired information and avoid information overload. Second, the app included a calorie calculator. This facility allowed consumers to select the dishes they intended to consume, and a value for total calorie content was automatically calculated and provided for the meal as a whole. This facility recognizes that individuals do not typically only consume single dishes. Finally, the app included a notification system allowing caterers to send messages to users, for example, on special offers and promotions, recommendations, or advice. This feature has also not yet been fully activated.

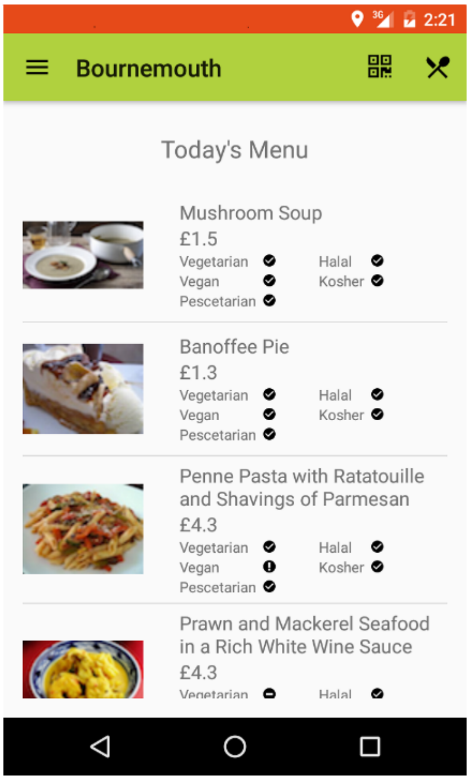
Table 2. Details of the prototype app per user interface screen.

Screen	Display	User actions
Basic (nonpersonalized) functionality		
Screen 1	Welcome and option for tutorial	User swipes to progress
Screen 2	Personalization screen	User has option to personalize the app (personalized functionality) or skip this (basic functionality)
Screen 3	Display of local canteens with available information based on Global Positioning System locator	User selects desired canteen and clicks option to see menu
Screen 4	Menu for the day is displayed pictorially, consisting of dish name, picture, price, and diet classification	Users can view all dishes available. Users can view full information per dish by clicking on any dish
Screen 5	Information (description, energy, portion size, ingredients; allergens; and nutritional content (gram per 100 g) of fat, saturated fat, carbohydrate, sugars, fiber, protein, salt, using the traffic light system) is displayed for the dish	User can view all information. User can also toggle a heart symbol to send feedback to the caterer that they like the dish. Activation with the QR (quick response) code (per dish) results in immediate arrival at Screen 5
Personalized functionality (optional)		
Screen 2	Personal preferences are available based on the following: Canteen selection (local canteens available); Diet type: vegetarian, vegan, pescatarian, halal, kosher; Allergens: celery, cereals, crustaceans, eggs, fish, lupin, milk, molluscs, mustard, nuts, peanuts, sesame, soybeans, sulfur; Dish calories	User selects preferences for canteen, diet type and allergens by moving a bar from "selected" to "not selected." These are saved automatically on the consumer's mobile phone and remain stored or can be updated at any time. The default selection is "not selected." Users can also set a desired maximum amount of calories per dish using a sliding scale and is provided with WHO (World Health Organization) recommendations for men and women
Screen 4	When personal preferences have been selected, the menu is provided such that preferred dishes are provided at the top, and less preferred dishes are provided at the bottom of the list. Dishes that do not fit the user profile based on diet type and allergens are provided grayed over	Users can view all dishes available. Users can view all information per dish by swiping across the dish

Details of the app, per screen, are given in Table 2. Images of the app screens are also displayed in Figures 1-3. The app is currently available for download from the Google Play Store. A video demonstration of both versions of the prototype app can be viewed in the dissemination section of the project website [54].

Figure 1. Screenshot of the FoodSMART app: screen 1.

Figure 2. Screenshots of the FoodSMART app: screen 4.



Today's Menu

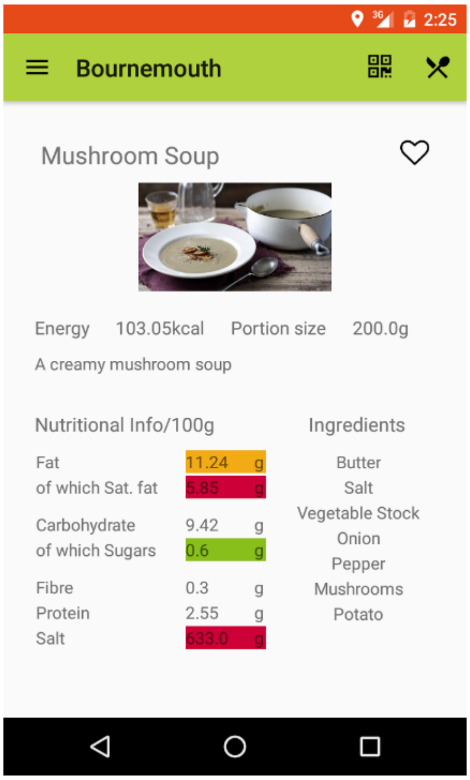
Mushroom Soup
£1.5
Vegetarian ✓ Halal ✓
Vegan ✓ Kosher ✓
Pescetarian ✓

Banoffee Pie
£1.3
Vegetarian ✓ Halal ✓
Vegan ✓ Kosher ✓
Pescetarian ✓

Penne Pasta with Ratatouille and Shavings of Parmesan
£4.3
Vegetarian ✓ Halal ✓
Vegan ! Kosher ✓
Pescetarian ✓

Prawn and Mackerel Seafood in a Rich White Wine Sauce
£4.3
Vegetarian ⊖ Halal ✓

Figure 3. Screenshots of the FoodSMART app: screen 5.



Mushroom Soup

Energy 103.05kcal Portion size 200.0g

A creamy mushroom soup

Nutritional Info/100g

Fat	11.24	g
of which Sat. fat	5.85	g
Carbohydrate	9.42	g
of which Sugars	0.6	g
Fibre	0.3	g
Protein	2.55	g
Salt	633.0	g

Ingredients

Butter
Salt
Vegetable Stock
Onion
Pepper
Mushrooms
Potato

Stage 3: Initial Evaluation of the App

A total of 233 respondents evaluated the prototype app and provided usable questionnaire data: 79 respondents from the UK events, 54 respondents from the French event, 18 respondents from the Danish event, 34 respondents from the Malaysian event, and 48 respondents from the Chinese event. Of these, 81 (34.7%, 81/233) were male, 143 (61.3%, 143/233) were female; 115 (49.3%, 115/233) were aged 20 to 29 years, 39 (16.7%, 39/233) were aged 30 to 39 years, 35 (15.0%, 35/233) were aged 40 to 49 years and 32 (13.7%, 32/233) were aged 50 years and older. Demographic details were incomplete from some respondents.

The mean overall SUS score was 68.3 (SD 15.4). Responses to individual questions are given in Table 3. Mean responses for all questions were positive relative to the midpoint of the scale and so represented a positive perception of the app (smallest t_{229}=26.40; P<.01).

Total SUS scores by gender, age, and event location are given in Table 4. Scores were comparable for males and females and no linear trend with age was found (largest t_{196}=0.81; P=.42). Differences were found between event locations, where higher total SUS scores were found in the United Kingdom and Denmark, compared with France and also with Malaysia and China ($F_{4,201}$=11.96; P<.01).

The following common themes emerged from the open-ended feedback:

- Good for informing healthy choices: Consumers liked the provision of various health and nutritional information in a manner that they found useful, for example, “Very useful for health conscious people, especially those who are busy, usually hard for them to buy healthy foods they wanted.”
- Particularly useful for those on restricted diets: Some participants highlighted a particular interest for those consumers who are health conscious or observing a restricted diet, for example, “Extremely useful and important app that would also be important to people with dietary restrictions.”
- Easy to use: Consumers offered support for the provision of information in a simple manner, and the use of the QR code, menus, and screen swipes to aid full use.
- Support for the personalization aspect and specific functions: The concept of using a mobile phone app to access personalized information on the nutrients, ingredients, and allergens was viewed positively by many participants. The calculator function was well received, and respondents felt it was easy to use, for example, “Very interesting and promising tool for personal use as well as for service providers—the calculator function is good.”
- Concerns around data accuracy: A small number of respondents expressed concerns around the possible accuracy of information provided through the app, not realizing the direct link between the caterer or food manufacturer and the information provided.
- A wish for improved functionality: Some consumers also suggested a desire for additional features, such as links with social media, or an option to provide feedback to suppliers about missing or questionable information.

Table 3. Mean and standard deviation of responses to all individual questions in the SUS questionnaire plus the additional question, for all consumers (N=233).

Question	Value, mean (SD)[a]
I think that I would like to use this system frequently	2.4 (1.1)
I found the system (was not) unnecessarily complex	2.8 (1.0)
I thought the system was easy to use	2.8 (0.9)
I think that I would (not) need the support of a technical person to be able to use this system	3.0 (1.1)
I found the various functions in this system were well integrated	2.7 (0.8)
I thought there was (not) too much inconsistency in this system	2.7 (0.9)
I would imagine that most people would learn to use this system very quickly	2.9 (1.0)
I found the system (not) very cumbersome to use	2.7 (1.0)
I felt very confident using the system	2.6 (1.0)
I did (not) need to learn a lot of things before I could get going with this system	2.7 (1.1)
I believe the FoodSmart App will be useful to customers in a canteen setting to help them to get informed about the dishes offered	3.1 (0.9)

[a] 0=strongly disagree and 4=strongly agree.

Table 4. Mean, SD, minimum, and maximum total System Usability Scale scores by gender, age, and event location for all consumers (N=233).

Demographic characteristic[a]	SUS[b] score, mean (SD)	Minimum-maximum
Gender		
Male (n=81)	69.8 (14.6)	32.5-100.0
Female (n=143)	67.9 (15.8)	27.5-100.0
Age (years)		
20-29 (n=115)	67.0 (13.6)	27.5-97.5
30-39 (n=39)	73.9 (14.2)	35.0-97.5
40-49 (n=35)	64.4 (19.7)	32.5-95.0
50+ (n=32)	71.9 (16.4)	45.0-100.0
Event location		
United Kingdom (n=65)	76.3 (13.7)	42.5-100.0
France (n=50)	67.8 (14.4)	32.5-95.0
Denmark (n=15)	73.0 (16.8)	45.0-97.5
Malaysia (n=30)	62.8 (9.6)	35.0-75.0
China (n=42)	58.6 (15.6)	27.5-87.5

[a] Numbers by gender, age, and event location do not equal 233 owing to incomplete demographic information in returned questionnaires from some respondents.

[b] System Usability Scale.

Discussion

This study aimed to develop and evaluate a mobile phone app to provide consumers with food-based information in a workplace canteen setting in a manner that also allowed them to limit and/or personalize the information they received if desired. Initial stages of the study ascertained the information that European consumers (legally) should and would like to receive about the food on offer and established that consumers would like to personalize this information and receive this information in a certain manner. A mobile phone app was subsequently developed to address these identified needs. Early assessments of the developed app suggest support for the app and use, if available.

Consumers were able to provide clear suggestions for information provision, which largely matched those provided by current legislation and previous research. App development resulted in a fully functioning prototype app, and the prototype was then evaluated positively. The SUS feedback gained highlights the strong usability and clarity of the developed app. Our app received a mean SUS score of 68.3. According to Bangor et al [51], scores in the range of 68 to 71 are typical of customer premise equipment (eg, phones and modems), graphical user interfaces, and interactive voice response phone systems. According to Kortum and Bangor [55], a score of 68.3 is about the value for Global Positioning System systems and slightly lower than that for digital video recorders. Furthermore, qualitative comments reaffirmed the potential ease of use and emphasized the value of information that was provided in a simple manner. Comparable SUS scores were also found between males and females and regardless of age, suggesting comparable appeal, but the app was found to be more positively evaluated in the United Kingdom and Denmark, compared with France and then compared with Malaysia and China. These location-based differences may warrant further examination and highlight a need for cultural considerations in app development and testing.

The study has confirmed that many consumers seek additional information when eating out and are keen to consider the constituents and nature of the food that they eat [13,15,16,53]. This study also confirms that the nature of the desired information can vary between consumers [1,2,13,16] and that consumers like the idea of personalized information [13]. The use of mobile phone technology facilitated development and was considered appropriate by consumers. The use of mobile phone technology to display detailed information on menus has been demonstrated previously to be successful [16,38], with particular focus on increased capabilities to navigate quickly and the provision of added value information that is integrated fundamentally and placed prominently [16].

Other studies have also indicated that consumers have greater trust in information provided through a technological solution rather than that gained from serving staff [16,17], and although a small number of respondents expressed concerns around the possible accuracy of information provided through our app, it may be that some additional information is required to inform consumers of the close links between the app and food providers, and the abilities for immediate and frequent updates where necessary. Some studies also report a need for more traditional forms of information provision to supplement technological information provision [53], but increased familiarity with technological solutions may reduce these concerns.

Components of the app were also designed to benefit caterers and food providers. The app is quick and easy to populate and does not take technological skill or experience. Difficulties were encountered during the development stage, however, owing to competing requirements and considerations from different stakeholders. Competition over requirements was experienced among stakeholders dependent on their intended end-use of the app. Difficulties also arose as a result of the differential use and consideration of similar terms and similar concerns between stakeholders, and the relative importance given to issues such as consumer privacy and data storage.

Further testing of our app is required. The app has been demonstrated as likely to be used and potentially useful, but further research is required in real canteen settings, both in terms of consumer acceptability and in terms of value to the consumer. Work is required specifically to demonstrate the value of the app for improving meal choice in a canteen scenario, for example, through the greater selection of more healthy meals, and for improving consumption, for example, through the consumption of less daily calories. Various research demonstrates that increased information may not necessarily be used or does not benefit all consumers [15,25,53], and some research even demonstrates less healthy consumption following the provision of nutritional information [56]. Considerable research demonstrates a strong distinction between the intention to perform a behavior and actually undertaking that behavior [57]. Behavioral tests of the benefit of the app are required. Preferably, these tests would be conducted in the form of randomized controlled trials to reduce potential bias, using behavioral outcomes based on food selection and consumption, and where calculations of benefits for health compared with relative costs would also be possible. We also accept that not all canteen users would use a mobile phone app, thus testing needs to be conducted in an appropriate volunteer sample.

Evaluation of the app from a service provider's perspective may also be of value. Consideration of additional criteria or additional functionality for food service operators such as feedback mechanisms could be of use. Food service providers are increasingly using digital means for informing and understanding consumers [15,16,58]. Some of the *could have* criteria that were not addressed during our development could be implemented. Facilities for consumer feedback, for example, via stores of purchases and purchase history may be beneficial.

Facilities for providing recommendations based on purchase storage and history may also be desirable for consumers. Using purchasing data from 39 burger restaurants, it has previously been found that software recommendations can change the mix of items purchased. The share of adult main course items requesting *no sauce* increased by 6.8%, the share of children's meals with apples instead of chips rose by 7.0% and the share of a breakfast sandwich without sausage increased by 3.8% [59]. Although these changes indicate only modest order refinement, the results suggest that targeted, adaptive food-based information could have behavioral potential. Facilities for social comparison may also be beneficial. Many apps related to social activities include a *share* option to allow others to view choices or allow comparisons between users or with an established norm. Notably, our app also currently does little to foster motivation for changes from the consumer. Engagement with the app requires consumers to be motivated to gain food-based information and then to act on this, and additional functionality, for example, feedback options for others to comment on food choices, through *likes*, may encourage and facilitate this motivation and so facilitate engagement with and action based on the app. Increased information provision may also be desirable to include an ability to flag any ingredient as an allergen. Allergies exist to more than the specified 14 commonly occurring allergens, and consumers have previously expressed a desire for all ingredients to be considered as potential allergens [26].

Benefits for health may also be enhanced by the inclusion of additional features not requested by the consumer. Considerable interest is currently focusing on the use of *nudges* to encourage consumers to select and consume more healthy dishes, or those that are more sustainably sourced [60-62]. The order in which items are displayed on the app may *nudge* individuals to select dishes toward the top of the list compared with those further down [62]. Tests of strategies such as these would be of interest.

Additionally, different versions of the app may be desirable, for example, through the use of different formats, different controls, or different setups. Tsai and Cheng [63], for example, identified 4 consumer clusters based on willingness to use technology (technology explorers, technology recipients, technology optimists, and technology laggards), and the app may be differentially viewed by these groups. Others also report the different use and preference of technology in different consumers [16,59].

Consideration of the non-European market in testing and development would also be of value. Food provision and the interests and needs of consumers can differ widely between countries and cultures, and further consideration of these differences is required. Although our app was developed in Europe and may be transferable to other Western cultures, it may not transfer to other more developing cultures where food provision is less regulated, for example, where the majority of takeaway food is sold by individual street vendors.

In conclusion, this study demonstrates the development and early positive evaluation of a prototype mobile phone app for the provision of food-based information in a canteen scenario in a manner that can be personalized. The study also confirms an interest by consumers in food-based information provision in eating-out scenarios, demonstrates a wish by consumers for abilities to personalize and limit the information provided, and demonstrates the value of a mobile phone app as a potential solution to current needs. Further research allowing further refinement of the app and demonstrating a health benefit from use of the app is required.

Acknowledgments

This study was funded by the European Commission, under grant agreement #643999: FoodSMART. The funder had no role in the conduct of the research, in the decision to publish, or in reviewing or approving the manuscript. The authors would like to thank Aris Koutsiamanis and Aimilia Tasidou, University of Thessaloniki, Greece, and Toby Osman, Bournemouth University, United Kingdom, for help with app development.

Conflicts of Interest

None declared.

References

1. Bates K, Burton S, Huggins K, Howlett E. Battling the bulge: menu board calorie legislation and its potential impact on meal repurchase intentions. J Consum Market 2011 Mar 22;28(2):104-113. [doi: 10.1108/07363761111115944]
2. Burton S, Creyer EH, Kees J, Huggins K. Attacking the obesity epidemic: the potential health benefits of providing nutrition information in restaurants. Am J Public Health 2006 Sep;96(9):1669-1675. [doi: 10.2105/AJPH.2004.054973] [Medline: 16873758]
3. Ellison B, Lusk JL, Davis D. Looking at the label and beyond: the effects of calorie labels, health consciousness, and demographics on caloric intake in restaurants. Int J Behav Nutr Phys Act 2013 Feb 8;10:21 [FREE Full text] [doi: 10.1186/1479-5868-10-21] [Medline: 23394433]
4. Orfanos P, Naska A, Trichopoulou A, Grioni S, Boer JM, van Bakel MM, et al. Eating out of home: energy, macro- and micronutrient intakes in 10 European countries. The European prospective investigation into cancer and nutrition. Eur J Clin Nutr 2009 Nov;63(Suppl 4):S239-S262. [doi: 10.1038/ejcn.2009.84] [Medline: 19888277]
5. Orfanos P, Naska A, Trichopoulos D, Slimani N, Ferrari P, van Bakel M, et al. Eating out of home and its correlates in 10 European countries. The European prospective investigation into cancer and nutrition (EPIC) study. Public Health Nutr 2007 Dec;10(12):1515-1525. [doi: 10.1017/S1368980007000171] [Medline: 17582244]
6. Bhutani S, Schoeller DA, Walsh MC, McWilliams C. Frequency of eating out at both fast-food and sit-down restaurants was associated with high body mass index in non-large metropolitan communities in Midwest. Am J Health Promot 2018 Jan;32(1):75-83 [FREE Full text] [doi: 10.1177/0890117116660772] [Medline: 27574335]
7. Edwards J, Engstrom K, Hartwell H. Overweight, obesity and the food service industry. Food Serv Technol 2005 Jun;5(2-4):85-94. [doi: 10.1111/j.1471-5740.2005.00115.x]
8. Ma Y, Bertone ER, Stanek EJ, Reed GW, Hebert JR, Cohen NL, et al. Association between eating patterns and obesity in a free-living US adult population. Am J Epidemiol 2003 Jul 1;158(1):85-92. [doi: 10.1093/aje/kwg117] [Medline: 12835290]
9. World Obesity Day. 2017. The First Unified World Obesity Day Will Take Place on 4th March 2020 URL: http://www.obesityday.worldobesity.org/ [accessed 2018-11-27] [WebCite Cache ID 74F9Uv7OY]
10. European Commission. 2015. Labelling and Nutrition URL: https://ec.europa.eu/food/safety/labelling_nutrition_en [accessed 2018-11-27] [WebCite Cache ID 74F9AY8z3]
11. US Food and Drug Administration. 2015. Overview of FDA Labeling Requirements for Restaurants, Similar Retail Food Establishments and Vending Machines URL: https://www.fda.gov/food/food-labeling-nutrition/overview-fda-labeling-requirements-restaurants-similar-retail-food-establishments-and-vending [accessed 2018-11-27] [WebCite Cache ID 74FA6QeKd]

12. Fernandes AC, Oliveira RC, Proença RP, Curioni CC, Rodrigues VM, Fiates GM. Influence of menu labeling on food choices in real-life settings: a systematic review. Nutr Rev 2016 Aug;74(8):534-548. [doi: 10.1093/nutrit/nuw013] [Medline: 27358442]
13. Banterle A, Cavaliere A, Ricci EC. Food labelled information: an empirical analysis of consumer preferences. Int J Food Syst Dyn 2012;3:156-170. [doi: 10.18461/ijfsd.v3i2.325]
14. Case DO, Given LM. Looking for Information: A Survey of Research on Information Seeking, Needs and Behavior. Fourth Edition. Bingley, United Kingdom: Emerald Group Publishing; 2016.
15. Din N, Zahari MS, Othman C, Abas R. Restaurant operator's receptiveness towards providing nutritional information on menu. Procedia Soc Behav Sci 2012;50:699-709. [doi: 10.1016/j.sbspro.2012.08.073]
16. Lessel P, Böhmer M, Kröner A, Krüger A. User Requirements and Design Guidelines for Digital Restaurant Menus. In: Proceedings of the 7th Nordic Conference on Human-Computer Interaction: Making Sense Through Design. 2012 Presented at: NordiCHI'12; October 14-17, 2012; Copenhagen, Denmark p. 524-533. [doi: 10.1145/2399016.2399096]
17. Price S, Hartwell H, Hemingway A, Chapleo C. Workplace foodservice; perception of quality and trust. Appetite 2016 Feb 1;97:169-175. [doi: 10.1016/j.appet.2015.12.006] [Medline: 26686582]
18. Hammond D, Lillico HG, Vanderlee L, White CM, Reid JL. The impact of nutrition labeling on menus: a naturalistic cohort study. Am J Health Behav 2015 Jul;39(4):540-548. [doi: 10.5993/AJHB.39.4.10] [Medline: 26018102]
19. VanEpps E, Downs J, Loewenstein G. Calorie label formats: using numeric and traffic light calorie labels to reduce lunch calories. J Public Policy Mark 2016;35:26-36 [FREE Full text]
20. Bleich SN, Economos CD, Spiker ML, Vercammen KA, VanEpps EM, Block JP, et al. A systematic review of calorie labeling and modified calorie labeling interventions: impact on consumer and restaurant behavior. Obesity (Silver Spring) 2017 Dec;25(12):2018-2044 [FREE Full text] [doi: 10.1002/oby.21940] [Medline: 29045080]
21. Price S, Bray J, Brown L. Enabling healthy food choices in the workplace: the canteen operators' perspective. Int J Workplace Health Manag 2017 Aug 7;10(4):318-331. [doi: 10.1108/IJWHM-12-2016-0087]
22. Lowe B, Fraser I, Souza-Monteiro DM. A change for the better? Digital health technologies and changing food consumption behaviors. Psychol Mark 2015 Apr 20;32(5):585-600. [doi: 10.1002/mar.20802]
23. Nocella G, Romano D, Stefani G. Consumers' attitudes, trust and willingness to pay for food information. Int J Consum Stud 2014 Jan 23;38(2):153-165. [doi: 10.1111/ijcs.12080]
24. Fotouhinia-Yepes M. Menu calorie labelling in a fine dining restaurant: will it make a difference? J Qual Assur Hospit Tourism 2013 Jul;14(3):281-293. [doi: 10.1080/1528008X.2013.802550]
25. Dumanovsky T, Huang CY, Nonas CA, Matte TD, Bassett MT, Silver LD. Changes in energy content of lunchtime purchases from fast food restaurants after introduction of calorie labelling: cross sectional customer surveys. Br Med J 2011 Jul 26;343:d4464 [FREE Full text] [doi: 10.1136/bmj.d4464] [Medline: 21791497]
26. Ogawa M, Tanaka A, Noda K, Kawai A, Amoroso DL. Research on food allergy information using smart mobile media devices to enhance communication at restaurants. Int E-Bus Res 2012;8(3):1-17. [doi: 10.4018/jebr.2012070101]
27. European Commission. 2016. Irst European Directory of Health Apps Recommended by Patients and Consumers URL: http://www.patient-view.com/uploads/6/5/7/9/6579846/ [accessed 2018-11-27] [WebCite Cache ID 74FAqueOA]
28. Flaherty SJ, McCarthy M, Collins A, McAuliffe F. Can existing mobile apps support healthier food purchasing behaviour? Content analysis of nutrition content, behaviour change theory and user quality integration. Public Health Nutr 2018 Feb;21(2):288-298. [doi: 10.1017/S1368980017002889] [Medline: 29081322]
29. SmartAPPetite: Buy Local, Eat Smart, Get Healthy. 2018. URL: http://www.smartappetite.ca [accessed 2018-11-27] [WebCite Cache ID 74FBPwU51]
30. Barfield J. Uatrav. 2014. Food Service App Allows Students to Order Meals With Smartphones, Avoid Long Lines URL: https://home.tapingo.com/ [accessed 2018-11-27] [WebCite Cache ID 74FAPqvRQ]
31. Pieskä S, Liuska M, Jauhiainen J, Auno A, Oy D. Intelligent Restaurant System SmartMenu. In: Proceedings of the 4th International Conference on Cognitive InfoCommunications. 2013 Presented at: CogInfoCom'13; December 2-5, 2013; Budapest, Hungary p. 2-5. [doi: 10.1109/CogInfoCom.2013.6719177]
32. Statista. 2018. Smartphone User Penetration as Percentage of Total Global Population From 2014 to 2021 URL: https://www.statista.com/statistics/203734/global-smartphone-penetration-per-capita-since-2005/ [accessed 2018-11-27] [WebCite Cache ID 74FBAlFN3]
33. Newzoo. 2019. Top 20 Countries/Markets by Smartphone Users and Penetration URL: https://newzoo.com/insights/rankings/top-50-countries-by-smartphone-penetration-and-users/ [accessed 2019-06-03]
34. Tonkin E, Wilson AM, Coveney J, Webb T, Meyer SB. Trust in and through labelling – a systematic review and critique. Br Food J 2015 Jan 5;117(1):318-338. [doi: 10.1108/BFJ-07-2014-0244]
35. Meyer SB, Coveney J, Henderson J, Ward PR, Taylor AW. Reconnecting Australian consumers and producers: identifying problems of distrust. Food Policy 2012 Dec;37(6):634-640. [doi: 10.1016/j.foodpol.2012.07.005]
36. Coveney J. Food and trust in Australia: building a picture. Public Health Nutr 2008 Mar;11(3):237-245. [doi: 10.1017/S1368980007000250] [Medline: 17672922]

37. Hodgkins C, Barnett J, Wasowicz-Kirylo G, Stysko-Kunkowska M, Gulcan Y, Kustepeli Y, et al. Understanding how consumers categorise nutritional labels: a consumer derived typology for front-of-pack nutrition labelling. Appetite 2012 Dec;59(3):806-817. [doi: 10.1016/j.appet.2012.08.014] [Medline: 22918174]
38. Yepes M. Mobile tablet menus attractiveness and impact of nutrition labeling formats on millennials? Food choices. Cornell Hosp Q 2014 Aug 28;56(1):58-67. [doi: 10.1177/1938965514546371]
39. Teixeira V, Voci SM, Mendes-Netto RS, da Silva DG. The relative validity of a food record using the smartphone application MyFitnessPal. Nutr Diet 2018 Apr;75(2):219-225. [doi: 10.1111/1747-0080.12401] [Medline: 29280547]
40. World Health Organisation. 2004. The Global Strategy on Diet, Physical Activity and Health (DPAS) URL: https://www.who.int/nmh/wha/59/dpas/en/ [accessed 2018-11-27] [WebCite Cache ID 74FB1Ku2e]
41. Lassen AD, Thorsen AV, Sommer HM, Fagt S, Trolle E, Biltoft-Jensen A, et al. Improving the diet of employees at blue-collar worksites: results from the 'Food at Work' intervention study. Public Health Nutr 2011 Jun;14(6):965-974. [doi: 10.1017/S1368980010003447] [Medline: 21208479]
42. Clegg D, Barker R. CASE Method Fast-track: A RAD Approach. Boston, Massachusetts, United States: Addison-Wesley; 1994.
43. Brooke J. SUS: a quick and dirty usability scale. In: Jordan PW, Thomas B, McClelland IL, Weerdmeester B, editors. Usability Evaluation In Industry. London: Taylor & Francis; 1996.
44. Brooke J. SUS: a retrospective. J Usability Stud 2013;8:29-40 [FREE Full text]
45. Price S, Viglia G, Hartwell H, Hemingway A, Chapleo C, Appleton K, et al. What are we eating? Consumer information requirement within a workplace canteen. Food Qual Prefer 2016 Oct;53:39-46 [FREE Full text] [doi: 10.1016/j.foodqual.2016.05.014]
46. Bray J, Hartwell H, Price S, Viglia G, Kapuściński G, Appleton K, et al. Food information presentation: consumer preferences when eating out. Br Food J 2019 Jun 19;121:1744-1762. [doi: 10.1108/BFJ-09-2018-0605]
47. Bray J, Hartwell H, Perez-Cueto F, Mavridis I. FoodSMART: Think Smart, Eat Smarter. In: Food Product Marketing Conference. 2018 Presented at: FPM'18; June 13-14, 2018; Bournemouth, UK.
48. 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]
49. Davis FD, Bagozzi RP, Warshaw PR. User acceptance of computer technology: a comparison of two theoretical models. Manag Sci 1989 Aug;35(8):982-1003. [doi: 10.1287/mnsc.35.8.982]
50. van der Heijden H. User acceptance of hedonic information systems. MIS Q 2004;28(4):695-704. [doi: 10.2307/25148660]
51. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the system usability scale. Int J Human Computer Interact 2008 Jul 30;24(6):574-594. [doi: 10.1080/10447310802205776]
52. Lewis JR, Sauro J. Revisiting the factor structure of the system usability scale. J Usability Stud 2017;12:183-192 [FREE Full text]
53. Cohen DA, Babey SH. Contextual influences on eating behaviours: heuristic processing and dietary choices. Obes Rev 2012 Sep;13(9):766-779 [FREE Full text] [doi: 10.1111/j.1467-789X.2012.01001.x] [Medline: 22551473]
54. Bournemouth University. 2019. FoodSMART URL: https://microsites.bournemouth.ac.uk/foodsmart/ [accessed 2019-09-20]
55. Kortum PT, Bangor A. Usability ratings for everyday products measured with the system usability scale. Int J Human Computer Interact 2013 Jan;29(2):67-76. [doi: 10.1080/10447318.2012.681221]
56. Appleton KM, Pidgeon HJ. 5-a-day fruit and vegetable food product labels: reduced fruit and vegetable consumption following an exaggerated compared to a modest label. BMC Public Health 2018 May 15;18(1):624 [FREE Full text] [doi: 10.1186/s12889-018-5528-0] [Medline: 29764392]
57. de Ridder D. Health psychology. In: Smelser NJ, Baltes PB, editors. International Encyclopedia of the Social & Behavioral Sciences. Second Edition. London: Pergamon; 2015:674-678.
58. Ruiz-Molina M, Gil-Saura I, Berenguer-Contrí G. Information and communication technology as a differentiation tool in restaurants. J Foodserv Bus Res 2014 Nov 17;17(5):410-428. [doi: 10.1080/15378020.2014.967639]
59. Bedard K, Kuhn P. Micro-marketing healthier choices: effects of personalized ordering suggestions on restaurant purchases. J Health Econ 2015 Jan;39:106-122 [FREE Full text] [doi: 10.1016/j.jhealeco.2014.10.006] [Medline: 25544398]
60. Friis R, Skov LR, Olsen A, Appleton KM, Saulais L, Dinnella C, et al. Comparison of three nudge interventions (priming, default option, and perceived variety) to promote vegetable consumption in a self-service buffet setting. PLoS One 2017;12(5):e0176028 [FREE Full text] [doi: 10.1371/journal.pone.0176028] [Medline: 28562678]
61. Saulais L, Massey C, Perez-Cueto FJ, Appleton KM, Dinnella C, Monteleone E, et al. When are 'Dish of the Day' nudges most effective to increase vegetable selection? Food Policy 2019 May;85:15-27. [doi: 10.1016/j.foodpol.2019.04.003]
62. Wilson AL, Buckley E, Buckley JD, Bogomolova S. Nudging healthier food and beverage choices through salience and priming. Evidence from a systematic review. Food Qual Prefer 2016 Jul;51:47-64. [doi: 10.1016/j.foodqual.2016.02.009]
63. Tsai MC, Cheng CC. Consumers' intention to use Mobile-Menu and destination clusters. Int J Mob Commun 2013;11(1):36-55. [doi: 10.1504/IJMC.2013.050994]

Abbreviations

ICT: information communication technology
MoSCoW: Must have, Should have, Could have, Won't have
QR: quick response
SUS: System Usability Scale

Edited by G Eysenbach; submitted 28.11.18; peer-reviewed by J Shirk, G Hendrie, J Nunez-Cordoba, W Han; comments to author 01.05.19; revised version received 23.06.19; accepted 19.08.19; published 04.11.19.

Please cite as:
Appleton KM, Bray J, Price S, Liebchen G, Jiang N, Mavridis I, Saulais L, Giboreau A, Perez-Cueto FJA, Coolen R, Ronge M, Hartwell H
A Mobile Phone App for the Provision of Personalized Food-Based Information in an Eating-Out Situation: Development and Initial Evaluation
JMIR Form Res 2019;3(4):e12966
URL: http://formative.jmir.org/2019/4/e12966/
doi:10.2196/12966
PMID:31682575

Original Paper

Formative Evaluation to Build an Online Parenting Skills and Youth Drug Prevention Program: Mixed Methods Study

Lawrence Matthew Scheier[1*], PhD; Karol L Kumpfer[2,3*], PhD; Jaynie Litster Brown[3*], BA; QingQing Hu[4*], MA

[1]LARS Research Institute, Scottsdale, AZ, United States
[2]University of Utah, Salt Lake City, UT, United States
[3]Strengthening Families Program, LLC, Salt Lake City, UT, United States
[4]Division of Public Health, University of Utah, Salt Lake City, UT, United States
[*]all authors contributed equally

Corresponding Author:
Lawrence Matthew Scheier, PhD
LARS Research Institute
15029 North Thompson Peak Blvd
Suite B111-443
Scottsdale, AZ, 85260
United States
Phone: 1 702 630 7584
Email: scheier@larsri.org

Abstract

Background: Family-based drug prevention programs that use group-based formats with trained facilitators, such as the Strengthening Families Program (SFP), are effective in preventing underage drinking and youth drug use. However, these programs are resource-intensive and have high costs and logistical demands. Tailoring them for Web-based delivery is more cost-effective and makes it easier to scale these programs for widespread dissemination. This requires the active involvement of all key stakeholders to determine content and delivery format.

Objective: The aim was to obtain consumer, agency stakeholder, and expert input into the design of a Web-based parenting skills training and youth drug prevention program.

Methods: We conducted 10 focus groups with 85 adults (range 4-10, average 8 per group), 20 stakeholder interviews with family services agency staff, and discussed critical design considerations with 10 prevention scientists and e-learning experts to determine the optimal program content and technology features for SFP Online. Focus group participants also answered survey questions on perceived barriers to use, desired navigational features, preferred course format, desired content, preferred reward structures, course length, interactive components, computer efficacy, and technology use. Descriptive statistics were used to examine consumer characteristics; linear regression was used to examine relations between SFP exposure and four continuous outcome measures, including desired program content, interactive technology, and concerns that may inhibit future use of SFP Online. Logistic regression was used as a binary measure of whether consumers desired fun games in the SFP Online program.

Results: Three broad thematic categories emerged from the qualitative interviews enumerating the importance of (1) lesson content, (2) logistics for program delivery, and (3) multimedia interactivity. Among the many significant relations, parents who viewed more SFP lessons reported more reasons to use an online program (beta=1.48, *P*=.03) and also wanted more interactivity (6 lessons: beta=3.72, *P*=.01; >6 lessons: beta=2.39, *P*=.01), parents with less interest in a mixed delivery format (class and online) reported fewer reasons to use the online program (beta=−3.93, *P*=.01), comfort using computers was negatively associated with concerns about the program (beta=−1.83, *P*=.01), having mobile phone access was related to fewer concerns about online programs (beta=−1.63, *P*=.02), willingness to view an online program using a mobile phone was positively associated with wanting more online components (beta=1.95, *P*=.02), and parents who wanted fun games wanted more interactivity (beta=2.28, *P*=.01).

Conclusions: Formative evaluation based on user-centered approaches can provide rich information that fuels development of an online program. The user-centered strategies in this study lay the foundation for improving SFP Online and provide a means to accommodate user interests and ensure the product serves as an effective prevention tool that is attractive to consumers, engaging, and can overcome some of the barriers to recruitment and retention that have previously affected program outcomes in family-based prevention.

(JMIR Form Res 2019;3(4):e14906) doi:10.2196/14906

KEYWORDS

formative evaluation; parenting skills; drug prevention; focus groups; key stakeholders; consumer preference survey; internet intervention

Introduction

Background

Youth alcohol and drug use continue to be top public health priorities. This emphasis stems from recognition that experimental drug use that often coincides with adolescent rebelliousness can easily transform into addiction if left unabated. Numerous trials have now shown that family-based drug prevention programs are an effective first line of defense preventing underage drinking and youth drug use [1-3]. The evidence base also reinforces the utility of combining parenting skills training with program content that emphasizes youth drug prevention [4]. An important consideration in this approach is that family dynamics, and especially parent-child social interactions, play a crucial role in a child's developmental outcome [2,5-7]. Most—if not all—of these programs have their conceptual roots in ecological [8], social interactional [9], and transactional [10] models of human development. All these approaches underscore the close interpersonal alliance formed between parent and child and its incipient role in developmental outcomes. Regardless of conceptual underpinnings, most family-based programs strongly reinforce that parents shape their child's behavior from a very early age.

The Strengthening Families Program (SFP) is one example of a parenting skills training program that combines youth drug prevention in a group-based, facilitator-led format for children aged 3 to 17 years. The program was first tested with 14 sessions designed to improve parenting skills among drug-addicted adults in treatment settings [11,12]. Since then, the program has been retooled for universal settings [13], applied to young children [14], delivered in rural settings [15], tested in urban schools [16], culturally adapted [17,18], and examined using rigorous randomized controlled trials in international settings [19,20]. There is now also evidence the program can promote parent-child reunification in child welfare settings [21] and produce other societal and economic benefits [22].

In keeping with well-established research traditions, the program's core components draw from family systems theory [23], relationship enhancement approaches [24], family-based therapeutic approaches [25], behavioral parent training [9], and social learning theory [26]. Collectively, these theoretical linchpins support the goal of teaching parents how to bond with their child, use more effective ways to communicate with their child, set reasonable boundaries and limits (controls and restrictions), reward their child for good behavior, and monitor their child's activities [27]. The child component addresses social and personal skills that will help them refuse negative peer pressure to use alcohol and drugs and improve their personal self-management skills. Children also practice ways to increase social-emotional regulation and impulse control and acquire better problem-solving and effective communication skills.

Similar to many other family-based programs, parents and children receive separate instruction in the first hour and then join together to practice and receive feedback on newly learned skills during a second hour. This portion includes supportive role play and games that the parent and child play together, which encourages parents to implement nonjudgmental dialog and increase parent-child attachment. Behavioral reinforcement techniques, such as situational role play and behavioral rehearsal with positive feedback given to families by highly trained facilitators, are a hallmark feature of SFP. The program uses curriculum guides, homework, and workbooks that can be practiced in the home to reinforce lesson material. Providing family meals before each weekly session, offering childcare for young siblings, and assisting with transportation are implemented to address attendance barriers that potentially inhibit participation and retention over time.

Internet-Delivered Parenting Programs

Notwithstanding evidence of their overall effectiveness, family-based programs remain resource-intensive with high costs and logistical demands. Any attempt to scale them for widespread dissemination has to address these implementation demands. In recent years, internet-based behavioral interventions have been making tremendous strides as alternatives to programs that customarily use face-to-face delivery, presenting a promising avenue for delivery [28,29]. Web-based programs ensure standardized delivery and implementation fidelity, are convenient for end users, and are self-paced, cost-efficient, and reduce social or personal demands. Some parenting skills programs have taken on this challenge, and a few internet-based programs have produced promising findings underscoring their efficacy [30-35]. Two programs that have successfully migrated to the Web include Familias Unidas [30,31] and Triple P Online [32,33]. Familias Unidas targets parents with goals of preventing substance use, sexual risk behaviors, and sexually transmitted infections in high-risk youth from immigrant Hispanic and Latino families; Triple P is a population-based intervention that targets families with young children with early-stage conduct disorder or behavioral or emotional problems. Both programs use formative evaluation strategies in the process of developing an eHealth prototype. These efforts included focus groups to identify the program features desired by consumers and discussed with families novel ways to structure program content when delivered on a Web platform.

Challenges With Web-Based Delivery

Repackaging group-based, facilitator-led programs for Web delivery faces several formidable challenges. To begin with, a program developer must consider how to preserve the active ingredients of a program as it is transitioned to a Web-based platform. This entails finding ways to transform behavioral practice and positive reinforcement techniques into digital content that preserves the integrity of the intervention strategies.

Then developers must consider finding the proper dosage and session length, particularly because internet-based programs are traditionally shorter than in-person, group-based formats. Developers must also find ways to assess exposure, including engagement, and use this information to establish a metric for fidelity of implementation and adherence—two important components related to program efficacy [36-38].

Engagement can be defined as the adoption of the program by consumers, and then continued use can be measured through session exposure. In both cases, recruitment to participate and retention throughout delivery in family-based programs have faced several hurdles [39-41]. Indeed, reviews of recruitment and retention in family-based programs highlight the problem of attrition, with a host of reasons cited by participants for their noncompliance [41,42]. Reasons for failing to attend or complete the program include that parents may not find the program accessible or they perceive the intervention demands as burdensome (ie, lengthy sessions) or intimidating. In the traditional group setting, additional logistical barriers to participation include transportation to and from the facility, costs associated with childcare if not provided by the program, and time away from family responsibilities [43]. Scheduling demands can also interfere with participation. In some cases, parents have expressed reluctance to attend because of stigma, loss of privacy, or the appropriateness of the intervention for their essential needs [30,41]. Regardless of their stated reasons, dropping out of the intervention lowers dosage and adversely affects program outcomes, reduces power, and limits external validity.

Involving Users in Program Design

Transparency in the process of meeting these challenges would provide a template to guide future research examining the design and implementation of Web-based parenting skills training and youth drug prevention programs [44]. Participatory action research strategies in the form of consumer preference studies represent one way to address the concerns that arise from the deployment of Web-based programs [45,46]. Consumer preference studies are a mainstay of community-based research, as they seek greater input from consumers in all phases of a program from its early conceptual stages through development, implementation, and even dissemination. There is now growing evidence that consumer preferences can be quite fruitful when designing programs for youth, especially those focusing on mental health [47]. The strategy of gleaning information from consumers was also the tactic used by the developers of Familias Unidas [30] and Triple P [32-33] when they solicited consumer's input before the developers designed a Web-based delivery system for their family programs. By allowing participants to contribute and provide input in the early program design stages as part of "collective making," consumers feel valued as co-creators [48]. The goal here is to have the target consumer population take a more vested interest in program development, feel empowered, and take ownership, leading the developer to construct a product that is more reflective of the target consumer's preferences and needs [49]. Motivation to attend sessions should be higher among consumers when they feel the session content and delivery format are well designed and congruous with their needs.

This study uses a mixed methods formative evaluation design to explore consumer preferences in the prototype build stage for SFP Online. In choosing to develop an online version, several pressing questions needed to be addressed, including whether consumers felt the group-based format could be transformed in a way to maintain the active ingredients with no loss of fidelity in a digital environment. We also wanted to know what levels of interactivity and what type of program content are needed to offset factors that contribute to noncompliance. We applied both formative and summative strategies to obtain this information before developing and testing a prototype. We also posed specific questions to consumers regarding the feasibility, perceived utility, technology acceptance, desired instructional format (ie, screen layout and navigational features), and other utilization factors that could influence engagement. This type of in-depth feedback is essential during the development process because it can be a foundation for the design of intervention modules and help structure their delivery. Although we solicited and received input from both parents and youth, we only focus on the adult consumer responses in this study.

We also extended prior work by including agency staff (stakeholders) and subject matter experts because they often have tangible insight that can shed light on and give shape to the product's functionality. Agency staff who routinely work with families offering services may have their finger on the pulse of what families want and need and may also recognize some of the barriers to attendance that affect program outcomes. Experts can weigh in on how to maintain close alignment with the developmental and family-based theories that drive most parenting skills training interventions. As detailed subsequently, a second angle that e-learning experts can weigh in on includes integrating various instructional design principles into a serious educational game to increase the program's overall attractiveness. In addition to the formative strategies, we used a consumer preference survey administered to focus group participants, which is explained in detail.

Methods

Table 1 shows the study design and gives a brief description of each arm corresponding to the formative and summative evaluation research plan. This study was part of a Phase I NIH-funded Small Business Innovative Research grant (NIDA; R43DA046238-01) and the study was approved by the Heartland IRB located in Belleville, IL (HIRB Project #180213-187) prior to engaging in any scientific work.

Table 1. Research plan for the different study arms.

Study arm	Sample size	Focus	Comments
Focus groups	85 adults in 11 separate groups	Necessary steps for translation of group-based program to a Web platform	Questions probed active ingredients, interactivity, cellular phone and computer efficacy, proposed content, games, and ways to connect parent and youth lessons
Consumer preference survey	85	Barriers and facilitators, computer technology familiarity, desired content and materials, interactivity, and desire for social media connections	45 questions assessing past experience with the Strengthening Families Program assessing whether prior program exposure guides consumer preferences for an interactive, multimedia Web platform
Key stakeholder interviews	20 agency staff	Suitability of Web platform for clientele, assessing required changes in content and delivery methods, whether program would be engaging (interactivity) and fit the agency goals	Semistructured interview with questions provided in advance, then teleconference call to expand on answers
Expert interviews	10 (5 prevention scientists, 5 e-learning experts)	Two themes: prevention science to bolster core active ingredients of program and e-learning emphasis to include recent technology and its influence on learning and behavior change	Semistructured interview with questions provided in advance, and then phone or video interview used to expand on answers

Focus Groups

We recruited 85 adults from nine geographically dispersed agencies in the US covering the western portion (NM, CA, WA, UT, and NV), the middle portion (KS, IN, and OH), the eastern seaboard (NC and NY), and the south (TX). These agencies were representative of the target agencies where SFP Online could be distributed, and they were willing to participate in the formative evaluation. The agencies conducted the focus groups between May and early June of 2018. Each agency had participated in prior SFP implementation training using the traditional group format or had implemented the DVD curriculum used in conjunction with group classes. An email was sent to all the agencies outlining the scope of the project and the requirements for participation. After the initial contact, we sent a memorandum of agreement to each agency outlining the participant recruitment requirements (emails, phone calls, and posted flyers) and what family participation would entail, including a description of incentives that would be provided to both the individuals and the agency.

We gauged the number of groups (approximately 4 to 10 members in each group) based on recent studies that confirmed thematic sufficiency with nonprobability sampling strategies can be achieved with three to six focus groups [50] and also accounted for potential dropout after recruitment. The inclusion criteria stipulated that parents (including guardians and caregivers) and youth had attended an SFP class or watched the SFP DVD. This requirement was intended to solicit input from families who had sufficient knowledge of the programs' core content and could comment on the suitability of an online version, make suggestions for future content, and discuss the utility and feasibility of a Web-based version.

Each focus group lasted between 60 and 90 minutes and involved a moderator; refreshments were provided at the midpoint. Agency staff assisted in obtaining written informed consent, conducting the groups, and collecting survey materials. Moderators were given a seven-page instructional guide outlining how to conduct the group, preserve confidentiality, and solicit balanced input from the group members. Moderator probes queried the parents' computer technology and mobile phone usage, their video game experience, desired program content, instructional materials, navigational features, interactive components, and whether parents wanted asynchronous social media connections (ie, blogs, chat, discussion board). Members of the research team were present during one of the focus groups and listened to others via GoToMeeting Web conferencing.

Key Informant Interviews

We also conducted 20 key informant semistructured interviews by teleconference with agency staff primarily drawn from the same agencies that participated in the focus group. One exception was a new agency in MA that participated in the interviews. One agency staff from each site participated, except for NY, UT, and KS, where more than one staff member was interviewed. Agency staff had been previously trained to deliver SFP as part of activities distinct from this project; they also provide, as part of their regular duties, additional psychoeducational and support services to the families. Their input was intended to address whether they felt the intended Web-based program could retain the core instructional strategies, how to structure the instructional modalities so they would be engaging, and what technology design features would be appropriate for the target families. Before the phone call, each stakeholder was provided a 20-item questionnaire with specific probes intended to elucidate their involvement with SFP, their organizational role, the demographic profile of the agency clientele (racial composition, income, risk level), their perceived feasibility and utility of implementing SFP Online, and the essential program components and skills training they felt should be included.

Expert Panel

We recruited 10 subject matter experts to weigh in on prototype development. Five were prevention scientists who we asked to address the procedures for translating core active ingredients into online programming and ways to ensure engagement. An additional five e-learning experts shared ideas on different ways to capitalize on current trends in multimedia interactive programming and whether this should influence the design of SFP Online. A brief survey was emailed to the experts before the teleconference interview, and their written responses helped guide the interview. The first part included 10 brief questions intended to elucidate their experience designing interactive programs. This was followed by 19 interview probes addressing the feasibility of SFP Online, implementation strategies, use of live family coaches, social media, interactive activities to promote family problem-solving skills, dashboard and navigation features that might be attractive, methods to stimulate engagement and increase session exposure, and perceived barriers or problems when implementing eHealth programs (eg, recruitment, retention, session length, interactivity, and structuring programmatic content).

Consumer Preference Survey

Parents in the focus groups also filled out a 45-item consumer preference survey. Following completion, the anonymous surveys were collected by agency staff; placed in sealed, preaddressed envelopes; and mailed to the research team. Table 2 shows the survey items (predictors) that were used in the summative analyses to examine consumer preferences. Additional outcome measures are explained subsequently. Fixed-choice or dichotomous yes or no response formats were used for some questions; the balance used five-point Likert response formats.

Table 2. Predictor measures from the consumer preference survey.

Survey question	Response format[a]
How many Strengthening Families Program (SFP) DVD lessons did you watch?	0, 1-3, 4-6, 7-9, 10-11
How comfortable are you using a computer?	Very, somewhat, not very
Would you prefer doing the SFP course online, or would you rather attend a class?	Online, attend class, undecided
Would you prefer to track your skills practice on your computer instead of a paper handout?	Yes, maybe, probably not, no
Are you interested in using a game-like online version of SFP?	Yes, maybe, not really, not interested at all
Would you practice the skills at home without a live family coach to remind you?	Yes, no, maybe, maybe with points
Do you have access to a computer that connects to the internet?	Yes/no
Would you use SFP Online even if you also took a class?	Yes, maybe, probably not, not
Do you have a mobile or smartphone where you could view SFP Online?	Yes/no
If you have a smartphone, would you view SFP Online on your phone?	Yes/no
If you could access SFP Online, how likely are you to record your home practice assignments?	Likely, somewhat likely, somewhat unlikely, unlikely
How many lessons should be included in SFP Online?	1-5, 6-10, 11-15
How long should each individual lesson be?	6-8, 9-12, 13-20, 21-30, >30 minutes
Would including fun games in SFP Online help your family learn new skills?	Yes, maybe, probably not, no
Would you like family members to earn reward points after completing SFP assignments?	Yes, maybe, probably not, no
Which of the following would you prefer in a family game?	Points to reward progress, experience points to move up a level in game, both
Would you like SFP Online games to include a family competition for points when you practice the skills?	Yes/no
Do you think you would practice the skills at home without a live family coach to remind you?	Yes, no, maybe, maybe, if we earned points for doing it

[a] Some variables recoded to avoid sparse cells. Age, gender, and race/ethnicity also asked. Race/ethnicity categories included African American or black, Asian, American Indian, Pacific Islander or Native Hawaiian, Alaska Native, white, Hispanic or Latino, and more than one race.

Qualitative Coding Schemes

Digital audiotapes from the focus groups and interviews were professionally transcribed verbatim, and the content of this material was then examined using thematic analysis as a research tool. This approach breaks down the open-ended answers into smaller units so the researcher can actively identify consistent and meaningful patterns in the text [51]. We used a data-driven inductive approach similar to grounded theory [52,53] to quantify themes based on explicit content analysis. For example, thematic content from the focus groups might reference the word “games” used in the context of programmatic features that will attract youth and contribute to engagement. Likewise, stating the need for interactive components modeled after a coach would be considered a theme (ie, “the program needs a coach”). Frequency of how often a word or phrase appeared in different contexts provided a quantified metric of thematic sufficiency. We repeated this coding procedure across all groups with two members of the investigative team descriptively tabulating key terms and phrases verbatim and then synthesizing across groups until thematic saturation was obtained. The coders read the transcripts and compiled their thematic results independently, and then convened for discussion of emerging themes. The results would then be summarized and tabulated into a product specification plan, a written document presented to the product design team before the alpha prototype build.

Consumer Preference Survey: Outcome Measures

The survey included additional items assessing issues that might arise in using SFP Online, which were modeled as outcomes. Seven items assessed technology barriers that might present concerns when using SFP Online, including data plans, visual content, screen size, and navigation, with responses coded on a five-point Likert scale ranging from not at all concerned to extremely concerned. An additional 20 items assessed what subject material parents felt was important to include in the online version (eg, brain development, mindfulness, communication skills, giving compliments, and monitoring) with responses coded on a five-point Likert scale ranging from very important to not at all important. An additional 16 items assessed the importance of different program components for learning parenting skills (eg, reporting progress, short videos, family goals, virtual coach, and tracking progress) with responses coded on a five-point Likert scale ranging from very important to not at all important. A 10-item scale assessed reasons for using SFP Online (eg, self-paced, review lesson materials, monitoring family progress, and tracking family performance) with responses coded on a five-point Likert scale ranging from very important to not at all important. A single item asked parents if they were interested in a game-like version of SFP, with responses recoded as yes or no.

Both factor analysis and estimates of internal consistency indicated low reliability and lack of conceptual purity in the outcome scales. This is likely because parents treated each subset of items as a checklist, and this prevented us from obtaining homogenous unidimensional scales. As a result, we created unit-weighted indexes to represent the different technological, logistic, programmatic, and motivational reasons for using SFP Online. We dichotomously recoded the Likert response scales to 1 or 0 by collapsing “not at all concerned” and “slightly concerned” to 1 and coded the remaining response categories to 0. Likewise, we collapsed “very important” and “important” to 1 and all other responses to 0. For all analyses, the significance level was set at $P<.05$ with two-sided tests.

Results

Adult Focus Group Characteristics

Eighty-five parents attended the 10 focus groups (77% female, 65/85). Their age breakdown included 11 parents (13%) between 18 and 31 years, 43 parents (51%) between 32 and 45 years, 19 parents (22%) between 46 and 52 years, 10 parents (12%) between 53 and 66 years, and 2 grandparents (2%) 67 years and older. In all, 56 (66%) were white, 8 (9%) were Native American or American Indian, 5 (6%) were Hispanic or Latino, 4 (5%) were African American, 9 (9%) were mixed race, 2 (2%) were Asian, and 1 (1%) indicated “other.” The family service agencies served a fairly homogeneous clientele, who were relatively poor, socially marginalized, experiencing family distress, and characteristically low in education. Therefore, we did not ask about their socioeconomic status using income or education.

Of the 85 parents, 79 (93%) said they had computer access with internet connection, and 69 (82%) said they were comfortable using computers, 11 (13%) said they were somewhat comfortable, and 5 (6%) said they were not comfortable. As anticipated, all the parents had some exposure to the SFP traditional group-based classes. Twenty-nine (34%) took a class with their family. Twenty (24%) viewed the SFP DVD video clips in a class setting, 21 (25%) viewed the DVD at home, and 11 (13%) viewed the SFP DVD with a coach. Thirty-three (39%) parents attended between zero and three lessons, 11 (13%) attended between four and six lessons, and 40 (48%) attended more than six lessons. Twenty-two (26%) parents said they preferred an online class, 35 (41%) said they wanted to attend an in-person class, and the remaining 23 (27%) were undecided between the two. Seventy-four (87%) parents said they would view SFP Online if it was available on a smartphone. When asked about including games in the new online version, 4 parents (4%) said they were not at all interested in games, 10 (12%) said they were slightly interested, 41 (48%) said they might be interested, and 30 (35%) said they would be very interested.

Qualitative Analyses

A total of 241 keywords were culled from the thematic content analysis of the focus groups and interviews. Two coders conducted the content analysis and reduced the pool of open-ended answers to 86 distinct keywords. An example of a word that parents provided that was highly prevalent was “activities,” and the corresponding commentary included “parent-and-child activities to do together.” Another example was “challenges,” and the commentary included “ability to customize challenges” and “challenges that spark our interest.” We then tallied the keywords by frequency of their use with the most popular terms, which included “games” (n=55), “unique” (n=40), “fun” (n=15), “challenges” (n=14), “points” (n=11), and “rewards” (n=9). The remaining terms had frequencies in the single digits. Using an inductive procedure, we then summarized these terms into three major categories: program content, logistics, and interactivity/engagement. Program content encompassed teaching points and lesson plans (eg, discussing risk and protective factors for drug use, practicing family management skills, and teaching families how to create strong bonds), and referenced specific SFP core activities (eg, communication boulders, mindfulness, and automatic negative thoughts).

Logistics encompassed program delivery features, such as what will make SFP Online engaging (technology and navigation control) and encourage future continued use (eg, natural-sounding voices, customization and feedback, accessibility on multiple devices, animation, and realistic role play). Interactivity and engagement encompassed multimedia features of the program (eg, avatars, chat or threaded discussion boards, expert coaching, pop-up notifications, and weekly expert blogs). Textboxes 1 to 3 provide several examples corresponding to each category created from the three different formative evaluation strategies.

Textbox 1. Thematic content analysis: SFP (Strengthening Families Program) Online summary categories for focus groups.

Content

1. Explain importance of sincere compliment (brain triggers); show Emotional Bank Account
2. Explain five steps of a reinforcing compliment
3. Keep a 4:1 ratio of compliments to corrections
4. Click to create compliments you can give family members
5. Short, realistic video clips
6. Download using SFP handouts

Logistics

1. Clear, simple navigation tools; effective search bar
2. Bright colors; animation; lots of graphics; no clutter
3. Customizable dashboard
4. App for tracking compliments
5. Interface linking parent and child responses
6. Points for completed lessons

Interactivity/Engagement

1. Virtual, customizable family coach who can ask questions, give advice, invite family to practice
2. Customizable Cheering Squad
3. Rewards for learning and practicing
4. Allow optional competition
5. Send chat messages to a family member (parent discussion board)
6. Use Scoreboard to track and see if achieving personal SFP goals
7. Earn “practice coins” to “buy” accessories for avatar and coach

Textbox 2. Thematic content analysis: SFP (Strengthening Families Program) Online summary categories for key informant interviews.

Content

1. Allow families to choose lessons; explain prerequisites
2. Ability to download and use SFP handouts
3. “Yes, but...” page to resolve concerns with a virtual family coach
4. Page with additional info: “To learn more, click here...”

Logistics

1. Simple, short, text language
2. Audio for nonreaders
3. Interface between parents and children to send messages
4. Frequently asked parenting questions with answers

Interactivity/Engagement

1. Text message reminders
2. Reward family completion of a lesson visually on dashboard (eg, assemble a family photo puzzle into a frame by joint practicing of skills)
3. Enable parent discussion board (eg, assemble a family photo puzzle into a frame by joint practicing of skills)

Textbox 3. Thematic content analysis: SFP (Strengthening Families Program) Online summary categories for expert interviews.

Content

1. Teach core essential SFP Skills that target key risk and protective factors
2. Include on-going review
3. Earn points for skill practice
4. Reward extra points and give congratulations on completion
5. Provide family and individuals Certificate of Completion

Logistics

1. Ability to practice and get quick feedback.
2. Low-stakes failure with immediate correcting feature
3. On sign up, receive a Family Page plus subpage for each child
4. Site autoreports results completed and skills practiced
5. Mobile phone accessible

Interactivity/Engagement

1. Customizable Family page with name, banner, motto, Dashboard
2. Customize learning modules to fit the age of children
3. Allow parents and youth to choose lesson topics; explain needed lesson prerequisites for each skill
4. Reward family goal setting

Consumer Preference Survey: Gender, Race, and Age Differences

There were no gender differences in the categorical measures assessing parents’ SFP course exposure, technology readiness, or course content preferences. Younger parents were more likely to want to view SPF Online using their mobile phone (χ^2_1=4.0, P=.04, ϕ=−0.22), were more likely to want to track their skills practice on a computer (χ^2_2=8.4, P=.01, ϕ=0.31), and were more likely to want a coach (χ^2_2=7.6, P=.02, ϕ=0.29). White parents (compared with all other race/ethnic groups) viewed more SFP lessons (χ^2_2=7.1, P=.03, ϕ=0.29) and preferred doing the classes online (χ^2_2=6.7, P=.03, ϕ=0.28). There were no significant race/ethnicity, age, or gender differences for desiring a gamified version of SFP Online. There were also no significant demographic differences in the mean scores for the four outcome measures (ie, concerns about online use, desired components of online program, level of interactivity, and reasons to use an online program).

Regression Models

Multimedia Appendix 1 shows the complete results of the regression models with the three blocks of predictors and four outcome measures. Three distinct blocks were used to model predictors of the four outcome measures. Each block was configured to include between four and five predictors to create balance. The first block contained the number of lessons the parent viewed using the DVD, comfort using computers, whether they preferred taking SFP online or attending a class, whether they wanted to track their skills practice on a computer or a handout, and whether they would practice the newly learned skills at home without a coach. Comfort using computers was significantly associated with the index summing concerns accessing SFP Online (F_2=5.34, P=.006). Parents who reported being very comfortable using computers expressed fewer concerns (mean 1.85, SE 0.42) compared to parents reporting they were somewhat comfortable (mean 3.03, SE 0.64) or not very comfortable (mean 3.68, SE 0.90), albeit these differences were not significantly different from each other by the Tukey-Kramer post hoc multiple comparison test.

The second block included whether parents had access to a computer with internet service, whether they would use SFP Online even if they had a class, if they had a mobile or smartphone to view SFP Online, would they view SFP Online if they had a smartphone, and whether they would record their home practice assignments on a computer or smartphone. Parents with access to a mobile phone reported significantly fewer concerns (F_1=10.94, P=.001; those owning a phone: mean 2.23, SE 0.63; those not owning one: mean 5.34, SE 1.07). Those wanting to view SFP on their mobile phone reported fewer concerns (F_1=4.76, P=.03, mean 2.97, SE 0.72) for parents who could view SFP on a mobile phone compared with those who could not view SFP on a mobile phone (mean 4.59, SE 0.82). The third block included a measure asking parents how many lessons should be included in SFP Online, the proposed length of the individual lessons, whether fun games would help their family learn new skills, and whether family members should earn reward points when engaging with the online program. None of the individual predictors were significantly associated with technological concerns using SFP Online.

The next model examined the associations between the same three blocks and an index assessing important programmatic components to include in SFP Online. For the first block, ease of using computers was the only significant measure in the model (F_2=4.07, *P*=.02). The post hoc multiple comparison showed that parents expressing some comfort using computers wanted significantly fewer lesson components (mean 16.67, SE 0.68) compared with parents feeling very comfortable (mean 18.53, SE 0.44). Parents wanting to view SFP Online using a mobile phone wanted more online components (F_1=10.46, *P*=.002; mean 20.19, SE 0.75) compared with parents who did not want to access the online program using a mobile phone (mean 18.24, SE 0.86). No predictors in the third block were individually significant.

The third index captured desired interactive components to include in the new online program. In the first block, the number of SFP lessons completed on the DVD was significant (F_2=5.87, *P*=.004). Parents who completed an intermediate number of lessons (4 to 6) wanted more interactive components (mean 12.31, SE 1.32) compared with parents who watched the most lessons (>6 lessons or more: mean 10.98, SE 0.98) and parents who watched the fewest lessons (0-3 lessons: mean 8.59, SE 1.04). The second block included one significant measure; whether parents desired to take a class even if they took SFP Online was significant (F_3=4.58, *P*=.005). Parents desiring to use SFP Online even with a class wanted more interactive components (mean 12.07, SE 1.31) compared with parents stating "no" (mean 6.18, SE 2.74, *P*=.09), although these post hoc comparisons were not significant.

The final block included three significant predictors assessing the number of proposed lessons for the new online program (F_2=4.14, *P*=.02), wanting SFP Online to include fun games (F_2=5.01, *P*=.009), and wanting rewards built into the program (F_3=2.82, *P*=.04). Parents viewing between 1 and 5 lessons wanted fewer interactive components (mean 10.18, SE 1.63) compared with parents viewing between 6 and 10 lessons (mean 12.95, SE 1.27). Parents stating "yes" to wanting fun games also reported wanting more interactivity (mean 12.82, SE 1.13) compared with those stating "maybe" (mean 10.53, SE 1.09).

The final model included the same three blocks individually predicting reasons (motivations) to use SFP Online. In the first block, only tracking SFP skills practice on the computer (versus handouts) was significant (F_2=4.38, *P*=.02). Parents not wanting to track their skills practice reported significantly fewer reasons to use SFP Online (mean 6.48, SE 0.81) compared with those saying "yes" (mean 8.17, SE 0.53) or "maybe" (mean 8.30, SE 0.47). In the second block, only wanting to take SFP Online in conjunction with a class was a significant predictor (F_3=5.11, *P*=.002). Parents stating they would probably not do the online class in addition to a group class reported significantly fewer reasons to use SFP Online (mean 4.56, SE 1.17) compared with those stating "maybe" (mean 8.49, SE 0.67) and "yes" (mean 8.77, SE 0.69). In the third block, wanting rewards while using the new online program was significant (F_3=5.17, *P*=.003). Parents who said they probably did not want rewards through the online program reported significantly fewer reasons to use SFP Online (mean 5.05, SE 0.99) compared with parents who did want rewards (mean 8.05, SE 0.47).

We then repeated the block entry procedure using logistic regression with a binary measure asking parents whether they were interested in using a game-like version of SFP (response categories "maybe" and "yes, very interested" combined to 1; "not really" and "not interested at all" combined to 0 as the reference category). To avoid convergence problems (maximum likelihood estimates that were not trustworthy), we examined each predictor one by one and collapsed response categories for the measures, reducing the number of cells in the estimation process. We then combined all the significant predictor measures culled from the individual models into a single model, controlling for race and age. Based on the individual models, parents who preferred taking SFP online (attend class was designated the reference or comparison class) were almost nine times as likely to want a game-like experience than those who were undecided (likelihood ratio [LR] χ^2_2=8.8, *P*=.01; OR 8.88, 95% CI 1.06-74.48) to want a game-like experience. Those parents who were undecided were five times as likely (OR 5.10, 95% CI 1.02-25.29) to want a game-like experience compared with the reference class. Unadjusted, the C-statistic (a variant of Somers' D) was 0.719, and the rescaled R^2 (coefficient of determination) accounted for 16.6% of the variance.

Parents reporting they were very comfortable using computers were eight times as likely to be interested in a game-like online version of SFP (LR χ^2_2=8.8, *P*=.01; rescaled R^2=16.6%, C-statistic=0.654; OR 7.90, 95% CI 1.18-52.97) compared with those reporting they were not at all comfortable using computers. Parents with mobile phones on which they could view SFP Online were 11 times as likely to want a game-like online version than parents lacking mobile phone access (LR χ^2_1=3.96, *P*=.046; OR 11.66, 95% CI 0.98-138.92; rescaled R^2=8%, C-statistic=0.564). Parents who had a smartphone were six times as likely to want fun games than parents who did not have smartphones (LR χ^2_1=6.1, *P*=.01; OR 6.02, 95% CI 1.52-23.84; rescaled R^2=12%, C-statistic=0.636). Parents who stated they would be likely or somewhat likely to record home practice assignments on their computer were 78% more likely to want a game-like version compared with parents who reported they would be somewhat unlikely and unlikely (LR χ^2_1=6.3, *P*=.01; OR 0.22, 95% CI 0.07-0.72; rescaled R^2=12%, C-statistic=0.673).

A model with all the individually significant predictors combined fit well (LR χ^2_6=23.2, *P*<.001; rescaled R^2=40%, C-statistic=0.846, Akaike information criterion=66.84, Hosmer-Lemeshow goodness-of-fit test, χ^2_5=1.0, *P*=.96). When all six predictors were modeled at once, only the measure of preferring the online version to a class was significant (LR χ^2_2=6.4, *P*=.04; OR 16.27, 95% CI 1.49-177.02); adjusting for race and age did not improve the model fit.

Discussion

Principal Findings

Thematic content analysis indicated three key areas of concern raised by all study participants when asked about designing an online version of the SFP: (1) desired content of the program, (2) logistics affecting program delivery, and (3) interactive features that can stimulate engagement. Interestingly, these themes reflected commonality in how consumers, agency stakeholders, and experts view the essential components of an online parenting skills training program that also involves youth drug prevention. Importantly, experts emphasized features consistent with cognitive behavioral principles, including skills practice, positive feedback, low-stakes failure exercises, and using reward structures (ie, proficiency “bars”) that help families move toward various programmatic benchmarks. They also outlined several implementation challenges, including program length and finding ways to use multimedia and interactive features to reinforce the focus and core active ingredients of the program.

Experts also felt the online program content should be realistic, simple, and use video examples to teach problem-focused coping skills, reflect the body of literature on risk and protective factors, and provide parents with ready solutions to their problems. The parents, in particular, stated that all this should be crafted using animation that can exemplify role-playing skills and reinforce lesson content. Experts also felt that SFP Online should stimulate “authentic” collaboration between parent and child and offer coaching, discussion boards, and virtual interactions—all with the goal of increasing engagement to maintain high exposure over time.

Agency staff agreed in principle with the experts and felt that the online version should maintain a skills-oriented approach and be configured around real family experiences. Given their familiarity delivering the group format of SFP, they emphasized preserving many of the current fun activities that work well in the group setting. From a logistical standpoint, they listed numerous features that would enhance SFP Online, including shorter sessions, graphics emphasized over text, dashboards for the entire family, text message reminders, home practice materials, printable worksheets, and a means for agency personnel to track family progress. They also mentioned that parents may still want to meet in person with a facilitator, even if they had online access, and that boosters might be essential to reinforce program content over time.

Parents felt very enthusiastic about helping to design the program content and delivery format for the online version. They wanted shorter sessions and state-of-the-art navigation features, with pop-up notifications, rewind capabilities, drop-down screens and scrollable information, graphic display with virtual animation, and realistic role plays. The parents who had the most prior exposure to SFP provided the most feedback on program components that they believed could work online. They felt the program should include bulletin boards, quizzes, pause-and-practice features, proficiency rewards, weekly blogs from experts, and virtual role plays with instruction. They also felt the online version should appeal to a culturally diverse population.

The consumer preference survey augmented the inductive procedures by adding more fine-grained information that helped to clarify what factors are associated with technological barriers, course content, interactivity, and motivations to use the online version. Interestingly, technology-savvy parents expressed fewer concerns about using an online program and showed a clear preference for interactivity and designing SFP Online like a serious educational game. The parents reluctant to track their performance online and those not wanting performance rewards were less inclined to want a Web-based platform to learn parenting skills. Mobile phone access was ubiquitous in this sample, underscoring the narrowing of the digital divide in terms of access to mobile technology and the potential for eventually creating a mobile app to complement SFP Online [54]. In all models tested, there was no evidence that any differences in the outcome could be attributed to race, gender, or age.

Even though the qualitative portion of the study reinforced tremendous overlap between the different groups, there were also unique contributions expressed by each group. For instance, agency stakeholders desired to find ways of connecting parent and child in the Web platform as this could be a strength of SFP Online. In addition, agency stakeholders were less inclined to address animation as crucial to the program, and more inclined to address interactive components that reinforce behavior change (ie, downloads, FAQs, discussion boards) and the use of technology to improve compliance (eg, text reminders). In this case, the focus of stakeholders appeared vested more in creating teachable moments and less focused on using technology to increase consumer involvement.

Strengths and Limitations

Several limitations should be mentioned. The data are cross-sectional, and we do not know, for example, whether any of the computer and mobile phone use questions actually predict subsequent use of SFP Online. Moreover, the purposive sampling intentionally recruited parents with prior exposure to SFP, including having used the DVD or participating in the group format. This can introduce bias because many of the parents who attended classes but desired an online version wanted the option to attend a class based on their prior exposure. Having the opportunity to receive immediate in-person feedback on their performance may be an attraction that is hard to duplicate on the Web. In addition, the small sample size altered the model testing sequence to avoid overfitting models. As a result, we could not test incremental variance attributed to each block of predictors. Related to this, some of the statistical relations may not have achieved significance merely because the study was underpowered. This problem underscores the trade-off between obtaining qualitative data from a small but sufficient number of focus groups and wanting a large enough and appropriately powered sample to produce robust parameter estimates. Notwithstanding, there is a wealth of information gathered using a mixed methods approach, and we certainly benefited from employing these different strategies.

One strength of this study is that we solicited collaborative input from prospective consumers as part of an active partnership preceding the development of the online program. This made participants feel valued and that they could contribute to program development in a meaningful way.

Conclusion

Curtailing the rising tide of addiction in America requires taking evidence-based family prevention programs to scale, which can best be achieved by providing high-quality, low-cost, online delivery. Creating an effective, online family-based prevention tool that parents and youth will enjoy and use requires knowledge of what parents and youth want to learn, how they want to acquire this knowledge, and what factors will keep them engaged so they can develop the appropriate skills and complete the course.

Acknowledgments

The authors would like to thank all the family service agencies that participated in the focus groups and assisted with the consumer preference surveys. We also want to thank the families who participated in this project, making it a collective effort that will ensure the final product is both engaging and addresses their parenting and child's developmental needs. This paper is dedicated to the memory of Steven P Schinke, a giant both in physical presence and in the field of prevention science. He was an inspiration and a true gentleman.

Conflicts of Interest

JLB and KLK have a principal financial interest in SFP 7-17. They receive financial remuneration from training group leaders, sales of the curriculum and evaluation contracts. Neither party has a financial interest in the SFP DVD, which is wholly owned by the Strengthening Families Foundation, a 501(c)3 not-for-profit corporation. All profits from sales of the DVD belong to the Foundation and no proceeds from sales are distributed in the form of wages or salary compensation. Further, neither author receives royalties from the SFP DVD. LMS is employed by LARS Research Institute, a not-for-profit company that offers professional consulting services in program evaluation, program development, and behavioral science technology transfer. Q-QH is a doctoral candidate in the Division of Public Health, University of Utah, and has no financial interest in the SFP product line.

Multimedia Appendix 1
Results of consumer preference survey regression models.
[PDF File (Adobe PDF File), 258 KB - formative_v3i4e14906_app1.pdf]

References

1. Kumpfer K, Hansen W. Family-based prevention programs. In: Scheier L, Hansen W, editors. Parenting and Teen Drug Use: The Most Recent Findings from Research, Prevention, and Treatment. New York, NY: Oxford University Press; 2014:166-192.
2. Lochman J, van den Steenhoven A. Family-based approaches to substance abuse prevention. J Prim Prev 2002;23(1):49-114. [doi: 10.1023/A:1016591216363]
3. Forgatch MS, Patterson GR, Gewirtz AH. Looking forward: the promise of widespread implementation of parent training programs. Perspect Psychol Sci 2013 Nov;8(6):682-694 [FREE Full text] [doi: 10.1177/1745691613503478] [Medline: 24443650]
4. Allen ML, Garcia-Huidobro D, Porta C, Curran D, Patel R, Miller J, et al. Effective parenting interventions to reduce youth substance use: a systematic review. Pediatrics 2016 Aug;138(2):195 [FREE Full text] [doi: 10.1542/peds.2015-4425] [Medline: 27443357]
5. Baumrind D. Rearing competent children. In: Damon W, editor. Child Development Today and Tomorrow. San Francisco, CA: Jossey-Bass; 1989:349-378.
6. Maccoby EE, Martin JA. Socialization in the context of the family: parent-child interactions. In: Mussen PH, Hetherington EM, editors. Handbook of Child Psychology: Volume 4. Socialization, Personality, Social Development. 4th edition. New York: John Wiley and Sons; 1983:1-101.
7. Bell RQ. Parent, child, and reciprocal influences. Am Psychol 1979;34(10):821-826 [FREE Full text] [doi: 10.1037//0003-066x.34.10.821]
8. Bronfenbrenner U. Ecology of the family as a context for human development: research perspectives. Dev Psychol 1986;22(6):723-742. [doi: 10.1037//0012-1649.22.6.723]
9. Patterson GR. Coercive Family Process. Eugene, OR: Castalia Pub Co; 1982.
10. Sameroff AJ. The Transactional Model of Development: How Children and Contexts Shape Each Other. Washington, DC: American Psychological Association; 2009.
11. DeMarsh JP, Kumpfer KL. Family-oriented interventions for the prevention of chemical dependency in children and adolescence. J Child Contemp Soc 1985;18(1-2):117-151 [FREE Full text] [doi: 10.1300/j274v18n01_07]
12. Kumpfer K, Demarsh J. Family environmental and genetic influences on children's future chemical dependency. J Children in Contemp Soc 1986 Nov 21;18(1-2):49-91 [FREE Full text] [doi: 10.1300/J274v18n01_04]

13. Kumpfer KL, Magalhães C, Whiteside H, Xie J. Strengthening families for middle/late childhood. In: Van Ryzin M, Kumpfer KL, Fosco G, Greenberg M, editors. Family-Based Prevention Programs for Children and Adolescents: Theory, Research, and Large-Scale Dissemination. New York, NY: Psychology Press; 2016:68-86.
14. Kumpfer KL, Whiteside HO, Greene JA, Allen KC. Effectiveness outcomes of four age versions of the Strengthening Families Program in statewide field sites. Group Dynamics: Theory, Research, and Practice 2010;14(3):211-229. [doi: 10.1037/a0020602]
15. Kumpfer KL, Alvarado R, Tait C, Turner C. Effectiveness of school-based family and children's skills training for substance abuse prevention among 6-8-year-old rural children. Psychol Addict Behav 2002 Dec;16(4, Suppl):S65-S71. [doi: 10.1037/0893-164x.16.4s.s65]
16. Gottfredson DC, Kumpfer KL, Polizzi-Fox D, Wilson D, Puryear V, Beatty P, et al. The Strengthening Washington DC Families project: a randomized effectiveness trial of family-based prevention. Prev Sci 2006 Mar;7(1):57-74. [doi: 10.1007/s11121-005-0017-y] [Medline: 16555144]
17. Kumpfer KL, Magalhães C, Xie J. Cultural adaptation and implementation of family evidence-based interventions with diverse populations. Prev Sci 2017 Aug;18(6):649-659. [doi: 10.1007/s11121-016-0719-3] [Medline: 27757773]
18. Kumpfer KL, Pinyuchon M, de Melo AT, Whiteside HO. Cultural adaptation process for international dissemination of the Strengthening Families Program. Eval Health Prof 2008 Jun;31(2):226-239. [doi: 10.1177/0163278708315926]
19. Bröning S, Baldus C, Thomsen M, Sack P, Arnaud N, Thomasius R. Children with elevated psychosocial risk load benefit most from a family-based preventive intervention: exploratory differential analyses from the German. Prev Sci 2017 Nov;18(8):932-942. [doi: 10.1007/s11121-017-0797-x] [Medline: 28643133]
20. Skärstrand E, Sundell K, Andréasson S. Evaluation of a Swedish version of the Strengthening Families Programme. Eur J Public Health 2014 Aug;24(4):578-584. [doi: 10.1093/eurpub/ckt146] [Medline: 24078373]
21. Brook J, McDonald TP, Yan Y. An analysis of the impact of the Strengthening Families Program on family reunification in child welfare. Child Youth Serv Rev 2012 Apr;34(4):691-695. [doi: 10.1016/j.childyouth.2011.12.018]
22. Johnson-Motoyama M, Brook J, Yan Y, McDonald TP. Cost analysis of the strengthening families program in reducing time to family reunification among substance-affected families. Child Youth Serv Rev 2013 Feb;35(2):244-252. [doi: 10.1016/j.childyouth.2012.11.008] [Medline: 30872122]
23. Forehand RL, McMahon RJ. Helping The Noncompliant Child: A Clinician's Guide To Parent Training. New York, NY: The Guilford Press; 1981.
24. Guerney BG, Coufal J, Vogelsong E. Relationship enhancement versus a traditional approach to therapeutic/preventative/enrichment parent-adolescent programs. J Consult Clin Psychol 1981 Dec;49(6):927-939. [doi: 10.1037//0022-006x.49.6.927] [Medline: 7309960]
25. Bowen M. Alcoholism as viewed through family systems theory and family psychotherapy. Ann NY Acad Sci 1991 Apr;1:94-102. [doi: 10.1111/j.1749-6632.1974.tb40288.x]
26. Bandura A. Social Learning Theory. Englewood Cliffs, NJ: Prentice-Hall; 1977.
27. Van Ryzin KL, Kumpfer KL, Fosco G, Greenberg M. Family-Based Prevention Programs for Children and Adolescents: Theory, Research, and Large-Scale Dissemination. New York, NY: Psychology Press; 2016.
28. Bennett GG, Glasgow RE. The delivery of public health interventions via the internet: actualizing their potential. Annu Rev Public Health 2009 Apr;30(1):273-292. [doi: 10.1146/annurev.publhealth.031308.100235]
29. Lustria ML, Cortese J, Noar SM, Glueckauf RL. Computer-tailored health interventions delivered over the web: review and analysis of key components. Patient Educ Couns 2009 Feb;74(2):156-173. [doi: 10.1016/j.pec.2008.08.023]
30. Estrada Y, Molleda L, Murray A, Drumhiller K, Tapia M, Sardinas K, et al. eHealth Familias Unidas: pilot study of an internet adaptation of an evidence-based family intervention to reduce drug use and sexual risk behaviors among Hispanic adolescents. Int J Environ Res Public Health 2017 Mar 04;14(3):264 [FREE Full text] [doi: 10.3390/ijerph14030264] [Medline: 28273862]
31. Prado G, Estrada Y, Rojas LM, Bahamon M, Pantin H, Nagarsheth M, et al. Rationale and design for eHealth Familias Unidas Primary Care: a drug use, sexual risk behavior, and STI preventive intervention for hispanic youth in pediatric primary care clinics. Contemp Clin Trials 2019 Jan;76:64-71 [FREE Full text] [doi: 10.1016/j.cct.2018.11.005] [Medline: 30453076]
32. Metzler CW, Sanders MR, Rusby JC, Crowley RN. Using consumer preference information to increase the reach and impact of media-based parenting interventions in a public health approach to parenting support. Behav Ther 2012 Jun;43(2):257-270 [FREE Full text] [doi: 10.1016/j.beth.2011.05.004] [Medline: 22440064]
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 M, Jorm A, Bazley R, Kelly C, Ryan S, Lubman D. Web-based parenting program to prevent adolescent alcohol misuse: rationale and development. Australas Psychiatry 2011 Aug;19(4):339-344. [doi: 10.3109/10398562.2011.603334] [Medline: 21861634]
35. Breitenstein SM, Gross D, Christophersen R. Digital delivery methods of parenting training interventions: a systematic review. Worldviews Evid Based Nurs 2014 Jun;11(3):168-176. [doi: 10.1111/wvn.12040] [Medline: 24842341]

36. Breitenstein SM, Gross D, Garvey CA, Hill C, Fogg L, Resnick B. Implementation fidelity in community-based interventions. Res Nurs Health 2010;33(2):164-173. [doi: 10.1002/nur.20373]
37. Byrnes HF, Miller BA, Aalborg AE, Plasencia AV, Keagy CD. Implementation fidelity in adolescent family-based prevention programs: relationship to family engagement. Health Educ Res 2010 Aug;25(4):531-541 [FREE Full text] [doi: 10.1093/her/cyq006] [Medline: 20142414]
38. Giles S, Jackson-Newsom J, Pankratz MM, Hansen WB, Ringwalt CL, Dusenbury L. Measuring quality of delivery in a substance use prevention program. J Prim Prev 2008 Nov;29(6):489-501. [doi: 10.1007/s10935-008-0155-7] [Medline: 19030993]
39. Heinrichs N, Bertram H, Kuschel A, Hahlweg K. Parent recruitment and retention in a universal prevention program for child behavior and emotional problems: barriers to research and program participation. Prev Sci 2005 Dec;6(4):275-286. [doi: 10.1007/s11121-005-0006-1] [Medline: 16075192]
40. Smokowski P, Corona R, Bacallao M, Fortson BL, Marshall KJ, Yaros A. Addressing barriers to recruitment and retention in the implementation of parenting programs: lessons learned for effective program delivery in rural and urban areas. J Child Fam Stud 2018 Jun 16;27(9):2925-2942. [doi: 10.1007/s10826-018-1139-8]
41. Prinz RJ, Smith EP, Dumas JE, Laughlin JE, White DW, Barrón R. Recruitment and retention of participants in prevention trials involving family-based interventions. Am J Prev Med 2001 Jan;20(1 Suppl):31-37. [Medline: 11146258]
42. Mendez JL, Carpenter JL, LaForett DR, Cohen JS. Parental engagement and barriers to participation in a community-based preventive intervention. Am J Community Psychol 2009 Sep;44(1-2):1-14. [doi: 10.1007/s10464-009-9252-x] [Medline: 19533328]
43. Axford N, Lehtonen M, Kaoukji D, Tobin K, Berry V. Engaging parents in parenting programs: lessons from research and practice. Child Youth Serv Rev 2012 Oct;34(10):2061-2071 [FREE Full text] [doi: 10.1016/j.childyouth.2012.06.011]
44. Danaher BG, Seeley JR. Methodological issues in research on web-based behavioral interventions. Ann Behav Med 2009 Aug;38(1):28-39 [FREE Full text] [doi: 10.1007/s12160-009-9129-0] [Medline: 19806416]
45. Boote J, Telford R, Cooper C. Consumer involvement in health research: a review and research agenda. Health Policy 2002 Aug;61(2):213-236. [doi: 10.1016/s0168-8510(01)00214-7]
46. Whyte WF, Greenwood DJ, Lazes P. Participatory action research. Am Behav Sci 2016 Jul 27;32(5):513-551 [FREE Full text] [doi: 10.1177/0002764289032005003]
47. 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 Human Factors 2015 Jul 09;2(2):e12. [doi: 10.2196/humanfactors.4361]
48. Sanders EB, Stappers PJ. Co-creation and the new landscapes of design. CoDesign 2008 Mar;4(1):5-18. [doi: 10.1080/15710880701875068]
49. Sanders MR, Kirby JN. Surviving or thriving: quality assurance mechanisms to promote innovation in the development of evidence-based parenting interventions. Prev Sci 2014 Mar 8;16(3):421-431. [doi: 10.1007/s11121-014-0475-1]
50. Guest G, Namey E, McKenna K. How many focus groups are enough? Building an evidence base for nonprobability sample sizes. Field Methods 2017;29:3-22 [FREE Full text] [doi: 10.1177/1525822X16639015]
51. Thomas DR. A general inductive approach for analyzing qualitative evaluation data. Am J Eval 2006 Jun 01;27(2):237-246. [doi: 10.1177/1098214005283748]
52. Glaser B, Strauss AL. The Discovery Of Grounded Theory: Strategies For Qualitative Research. Chicago, IL: Aldine; 1967.
53. Strauss J, Corbin J. Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory. Newbury Park, CA: Sage; 1990.
54. Jackson LA, Zhao Y, Kolenic A, Fitzgerald HE, Harold R, Von Eye A. Race, gender, and information technology use: the new digital divide. Cyberpsychol Behav 2008 Aug;11(4):437-442. [doi: 10.1089/cpb.2007.0157]

Abbreviations

NIDA: National Institute of Drug Abuse
NIH: National Institutes of Health
SFP: Strengthening Families Program

Edited by G Eysenbach; submitted 01.06.19; peer-reviewed by A Nguyen, K Mummery, J Brož; comments to author 16.07.19; revised version received 09.09.19; accepted 24.09.19; published 05.11.19.

Please cite as:
Scheier LM, Kumpfer KL, Brown JL, Hu Q
Formative Evaluation to Build an Online Parenting Skills and Youth Drug Prevention Program: Mixed Methods Study
JMIR Form Res 2019;3(4):e14906
URL: http://formative.jmir.org/2019/4/e14906/
doi:10.2196/14906
PMID:31687934

Original Paper

Feasibility and Acceptability of an Interactive Mental Well-Being Intervention for People With Intellectual Disabilities: Pilot Mixed Methods Study

Leen Vereenooghe[1*], BA, MSc, PhD; Kristian Westermann[1*], BSc, MSc

Faculty of Psychology and Sports Science, Bielefeld University, Bielefeld, Germany
*all authors contributed equally

Corresponding Author:
Leen Vereenooghe, BA, MSc, PhD
Faculty of Psychology and Sports Science
Bielefeld University
PO Box 10 01 31
Bielefeld, D-33501
Germany
Phone: 49 521106 ext 67521
Fax: 49 52110667521
Email: leen.vereenooghe@uni-bielefeld.de

Abstract

Background: The availability of both digital and traditional mental well-being interventions is rising, but these interventions typically do not consider people with intellectual disabilities as potential users.

Objective: The study aimed to explore the acceptability and feasibility of a new digital intervention, developed with and for people with intellectual disabilities, to improve their subjective well-being.

Methods: Using a single-group pre-post design, participants with intellectual disabilities and their caregivers completed the 4-week intervention. Mixed methods questionnaires assessed the acceptability of the intervention, in addition to self-report and proxy-report measures of subjective well-being and behavioral problems.

Results: A total of 12 men with mild to moderate intellectual disabilities enrolled in and completed the study alongside 8 caregivers. Participant acceptability of the intervention was high, and feedback covered multiple aspects of the intervention, including (1) program concept and design, (2) program content, and (3) intervention usage. Self-rated mood barometers indicated mood improvements for 5 participants, deteriorations for 2 participants, and no observed changes for the remaining participants. Statistical analyses yielded no difference from pretest (median=79; range 39-86) to posttest (median=79; range 21-96) for subjective well-being in people with intellectual disabilities (W=10.5; *P*=.17) and for behavioral problems (W=14; *P*=.05).

Conclusions: People with intellectual disabilities and their caregivers are receptive to using digital well-being interventions, and this research shows such interventions to be feasible in routine practice. Given the acceptability of the intervention, its potential efficacy can now be evaluated in people with intellectual disabilities and symptoms of reduced mental well-being.

(JMIR Form Res 2019;3(4):e15190) doi:10.2196/15190

KEYWORDS

intellectual disabilities; feasibility studies; tablet computer; mental health

Introduction

Background

People with intellectual disabilities have an increased vulnerability for developing mental health problems, in part owing to the frequency of negative life events and experiences that are common in this population, for example, experiencing stigma, isolation, dependency on caregivers, and experiences of abuse [1-3]. Although they constitute a heterogeneous population, people with intellectual disabilities generally present with a wide range of difficulties in communication, cognitive, and adaptive skills with age of onset before 18 years. These difficulties can then act as an additional burden on their ability to cope with such negative life events, thereby increasing their risk of poor mental health.

There is comparably little research and clinical attention for this population, in spite of their vulnerability and the resulting high prevalence of mental disorders. In terms of treatment or intervention options, pharmacological therapies are available, but also frequently prescribed in the absence of diagnosed mental disorders [4]. Psychological therapies offer a viable alternative, with cognitive behavioral therapies (CBT), in particular, appearing very promising in their treatment of anxiety, aggression, and mood disorders [5,6]. Unfortunately, therapeutic disdain for working with this population as well as diagnostic overshadowing and misdiagnoses may contribute to the slow uptake of therapists to offer psychological therapies to this population [7-9]. Meanwhile, the assumption that people with intellectual disabilities lack the cognitive skills to undertake therapy has been challenged, in particular with regard to CBT [10,11]. Although people with mild intellectual disabilities and better verbal and communicative abilities are more likely to understand the CBT framework and concepts of cognitive mediation [12], certain CBT skills can be trained using computerized or video-based training paradigms in people with mild and moderate intellectual disabilities [13-15]. This paves the way for a third potential intervention strategy in the form of computerized or Web-based mental health interventions as an adjunct or preparatory step for traditional talking therapies and pharmacotherapies.

New technologies present a range of applications that people with intellectual disabilities can benefit from, from mobile apps to improve problem-solving skills to virtual and augmented reality applications for social, vocational and self-determination skills [16-18]. The use of computers in psychological therapy for people with intellectual disabilities is still uncommon, despite such interventions being widely available for a wide range of mental health problems to people without intellectual disabilities. To date, only one full computerized CBT intervention has been evaluated in the treatment of anxiety and depression in adults with intellectual disabilities [19]. Although the cost-effectiveness of this intervention has not been evaluated, its acceptability among both patients with intellectual disabilities and their treating therapists was high. This is in line with research showing that both groups are interested in introducing technology as an additional means to improve the mental health and the therapy process involving people with intellectual disabilities [20].

A key challenge in delivering internet interventions to people with intellectual disabilities is the accessibility of such interventions. Barriers include economic factors related to the cost of computers and difficulties associated with potential cognitive and physical limitations [21]. Turning to internet use, access rates in people with intellectual disabilities are much lower than that in people without intellectual disabilities [22]. Here, caregivers play an important role in providing people with intellectual disabilities access to technology and the internet, as their support is often required to obtain the devices and set up user accounts. They may also be inclined to protect people with intellectual disabilities from the risks associated with internet use, although this can result in people with intellectual disabilities missing the benefits of internet use in relation to self-efficacy, empowerment, and social networking [23].

Objectives

Taking together the vulnerability of people with intellectual disabilities to experience mental health problems and the potential of digital interventions as an addition to traditional treatment interventions, this study presents an initiative to provide an accessible digital intervention to promote the mental well-being of people with mild to moderate intellectual disabilities. The driving objective is that successful mental health promotion might present an alternative and preventative strategy to improve the mental well-being of people with intellectual disabilities, in addition to or before the treatment of mental health problems through established pharmacological or psychotherapeutic interventions. To this extent, we set out to develop an intervention that fulfilled the following criteria:

1. Including people with intellectual disabilities in the development and design phases of the intervention to ensure its accessibility and its acceptance by people with intellectual disabilities.
2. Incorporating cognitive-behavioral components from evidence-based interventions.
3. Adopting a resource-oriented approach which involves the caregivers and support workers to assist with the implementation of the intervention in daily life.
4. Offering content that is personally relevant or can be customized.

In this paper, we briefly described the development of the intervention and presented a small pilot study to explore its potential feasibility and practicality as well as gave a first indication of its potential efficacy.

Methods

Phase I: Intervention Development

The intervention was intended to provide tips to improve the users' psychological well-being. Program content was derived from existing empirically tested programs and included both psychoeducational, behavioral activation and cognitive restructuring components. An advisory group of people with and without intellectual disabilities was established to provide feedback during the development process.

Content Development

The intervention comprised 8 modules which covered the following topics: (1) participation, (2) being active, (3) friendships, (4) relaxation, (5) self-acceptance, (6) communication, (7) self-actualization, and (8) cognitive restructuring (Textbox 1 provides an overview of the module aims).

Textbox 1. Outline and modules of the Pudelwohl intervention.

Module 1: Participation.

- Aim: Understanding that participation may lead to improved social contact and social support as well as being enjoyable.

Module 2: Being active.

- Aim: Improving mental well-being through increasing physical activity.

Module 3: Friendships.

- Aim: Highlighting the importance of a supportive network and tips for maintaining friendships.

Module 4: Relaxation.

- Aim: To acknowledge the need for relaxation and how this can be achieved.

Module 5: Self-acceptance.

- Aim: Addressing and accepting your strengths and weaknesses, including your disabilities.

Module 6: Communication.

- Aim: Being able to communicate your needs using a positive communication style.

Module 7: Self-actualization.

- Aim: Behavioral activation and finding of activities that support your personal development.

Module 8: Cognitive restructuring.

- Aim: Identifying and changing helpful and unhelpful thoughts to help change how we may feel about a situation.

Modules were spread over 4 sessions to be held at 1-week intervals. The first session started with an interactive introduction to learning to use the tablet, followed by module 1. Sessions 2 and 3 comprised 3 modules each. The fourth and final session presented module 8, followed by a reflection on all past modules and the instruction to choose which 3 tips they considered to be most important to them. These top 3 were then transferred onto a paper certificate, which participants received at the end of the study and acted as a self-help reminder. Furthermore, sessions 1 to 3 ended with a brief summary and homework assignment, whereas sessions 2 to 4 started with a reflection on the last session and a brief evaluation of the chosen homework assignment.

Technical Features

Adobe Captivate (2017, Version 10.0.0.192) was used to develop the intervention. The intervention was designed for use with 11.6-inch Odys Primo Win 12 2-in-1 tablet computers, which would be lent to the participants in the study. The images used within the intervention were either self-generated or license-free materials.

The intervention was introduced and led by an animal avatar: a poodle named *Wohl*, which references a German wordplay involving well-being. This poodle avatar explains how to navigate through the slides using the on-screen arrows and what users can expect from the program. This includes explaining symbols used throughout the program that introduce quizzes and discussions. Some quizzes prompt users to evaluate statements as helpful or unhelpful or good or bad, whereas in other quizzes users are given multiple-choice questions using drag-and-drop selection methods. Discussion slides present a question or statement that the user is encouraged to respond to or discuss together with their support worker. The main poodle avatar also provides an overview of his 8 animal friends, represented by 8 different animal avatars, that are associated with each module.

Overall, we aimed to minimize visual distractions by refraining from presenting a table of contents or navigation menu. To enhance user engagement, slides were animated where possible and were provided both on-screen text and a voice-over.

Advisory Group

An advisory group was established in collaboration with a local information technology workshop for people with and without disabilities or mental disorders. Flyers were distributed at the workshop to inform their users about our study. Workshop staff confirmed the time and date of the 4 planned consultations with users who expressed an interest. The consultations were held at 2-week intervals, with attendance varying between 4 and 7 participants, of which at least 3 presented with intellectual and developmental disabilities. We did not employ specific inclusion or exclusion criteria for people to attend the consultations. As these attendees were not participants in a research study, we did not routinely collect information regarding their sociodemographic characteristics, but all advisory group members were adults, with more men than women attending the consultations. During each session, the researchers presented drafts of the program modules and requested feedback regarding module content, design, accessibility, interactivity, and delivery. Changes, as a result of this feedback process, were presented at consecutive meetings for further evaluation and included changing font sizes, simplifying language, increasing and simplifying program interactivity, changes to program navigation options and page layout, and the use of supportive instead of corrective feedback in exercises. Feedback that was not incorporated for the pilot study because of limited resources included the provision of male and female voice-overs, more animations, and mutable background music.

A more detailed description of the co-development process for the intervention is given by Vereenooghe and Westermann [24].

Phase II: Pilot Study

Recruitment

People with intellectual disabilities and their support workers from 2 residential facilities in North Rhine-Westphalia, Germany, took part in the study. Both facilities provide services for children and young adults with intellectual and developmental disabilities. Support staff were informed about the study objectives and procedures and were asked to identify potential participants with intellectual disabilities among the residents whom they supported.

We aimed for minimal exclusion criteria to reflect the heterogeneity found within the population of people with intellectual disabilities and only excluded participants who (1) had a diagnosis of autism spectrum disorders without intellectual disabilities and (2) people with severe or profound intellectual disabilities who might not be able to operate the tablets and engage with the program content. Details regarding participants' level of intellectual disabilities was obtained by the participating support workers from the service users' records.

The sole inclusion criterion for support workers was the requirement of a minimum of 2 contact hours per week with the participant with intellectual disabilities to ensure both participants had sufficient time together to implement the intervention.

Design

A single-group pre-post design was used, with no control condition. Outcomes of mental well-being were assessed at pretest and posttest, in addition to mood barometers completed before each session. No follow-up data were obtained for this pilot study, for which recruitment took place from March 2018 to June 2018.

Measures

Primary Outcome

Program evaluation questionnaires were completed at posttest to assess the acceptability of the intervention. Questions were derived from the Client Change Interview Schedule [25] as implemented by Earley et al [26] and used both forced-choice and open-ended response formats. People with intellectual disabilities used a picture-based Likert scale to communicate their feedback regarding the program, similar to that used in the other outcome measures. The questionnaires were administered by the second author KW during a meeting organized with the participant with an intellectual disability after their support worker had indicated that they had completed all sessions. Posttest assessment took place within 2 weeks following the last session. Questions for participants with intellectual disabilities included "what in particular do you still remember about the intervention?" and "which aspects did you like?" Support workers' questionnaires asked whether they had noticed any change in the well-being of the person with an intellectual disability. Their version also asked more explicitly to identify specific content they considered to be helpful or not and what negative and positive aspects they have kept in mind.

Secondary Outcomes

The Personal Well-Being Index—Intellectual Disability (PWI-ID) [27] was translated to German based on the German version of the PWI [28] and administered to participants with an intellectual disability by their support worker. Administration took place within 1 week before the first session and within 2 weeks following the last session. The 8-item measure asks how people feel about their material, physical, mental, and general well-being. Responses are given on a Likert scale using 2, 3, or 5 facial expressions ranging from sad to happy and are then converted to numerical percent scores, with higher scores indicating better well-being. The number of response options (2, 3, or 5) is established for each participant before the assessment to optimize the validity of the responses.

Participants also completed mood barometers before each session, using the same picture-based Likert scale as used with the PWI-ID. To this extent, they received assistance from their support worker.

The Aberrant Behavior Checklist—Community (ABC-C) [29,30] was used as an indicator of informant-reported mental well-being, as it includes indicators of mental disorders in people with intellectual disabilities. Support workers rated the severity of the 58 behavioral problems over the last 4 weeks, using a 4-point Likert scale, with higher scores corresponding with more serious behavior problems. Support workers completed this measure independently following the instructions of the second author KW. The ABC-C was completed within 1 week before the first session and within 2 weeks following the last session.

Procedures

Participants were instructed to complete 1 session per week, with a maximum of 2 sessions per week. Tablets containing the program were lent to the participating organizations so people had instant access to the devices. The intervention was implemented in the regular working time of the support workers. Support workers had a facilitating role, and the intervention was intended to be used by the person with an intellectual disability as its primary user. Support workers could provide both practical and content support, for example, regarding the use of the tablet computers (eg, volume control or recharging the devices), clarifying the content (eg, repeating or reading the on-screen information), and assisting with program demands (eg, prompting the user to select an on-screen response or engaging the user in a discussion).

Ethical Concerns

Ethical approval for this study was obtained from the Bielefeld University Ethics Committee. All participants received information about the study in easy-to-read language and provided written informed consent. Participants without the capacity to provide informed consent were not enrolled in the study.

Data Analysis

Responses from the feedback questionnaires regarding the acceptability of the intervention were analyzed using quantitative content analysis of qualitative data [31]. First, all questionnaires were read to get an initial overview of the participant responses. Next, we generated codes to summarize each distinct piece of information, thereby being guided by the raw data rather than using pre-existing codes derived from the theory. For the qualitative analysis, we grouped these themes in overarching themes. For a further quantitative content analysis, the codes were then counted to identify themes that appeared more prevalent and to explore patterns in the data that could guide future development. Within this step, however, we made the decision to interpret the thematic codes as being positively or negatively valenced based on the questions that elicited these responses (eg, positively valenced questions included “which aspects did you like?” and “which content did you find helpful?” whereas negatively valenced questions, included “what did you notice negatively?” and “which content did you find unhelpful?”).

Differences in pretest and posttest scores were analyzed in SPSS (2017, Version 25) using nonparametric 1-tailed Wilcoxon signed rank tests.

Results

Participant Characteristics

Following eligibility checks and informed consent procedures, 12 pairs of participants completed baseline assessments (Figure 1). A total of 3 support workers completed the study with multiple participants with intellectual disabilities. Participants with intellectual disabilities were all male, whereas support workers included 5 women and 3 men who had worked with the participant with intellectual disabilities for at least 6 months (Table 1).

Intervention Feasibility, Acceptability, and Satisfaction

None of the enrolled participants, either people with intellectual disabilities or their support workers, dropped out during the study or withdrew their consent. The overall recruitment success (completed/informed) rates were 26% (12/46) for support workers and 80% (12/15) for people with intellectual disabilities (Figure 1). Owing to time constraints, 1 participant pair completed 2 sessions per week and completed the intervention in less than 4 weeks.

Content analysis of the program evaluation questionnaires was classified into 3 main areas: (1) program concept and design, (2) program content, and (3) intervention usage. Analysis of the positive and negative feedback, as shown in the quantitative content analysis in Figure 2, provided support for the conceptualization of the Pudelwohl intervention and its specific modules. It also indicated desired improvements for the implementation of Pudelwohl.

Figure 1. Study flowchart.

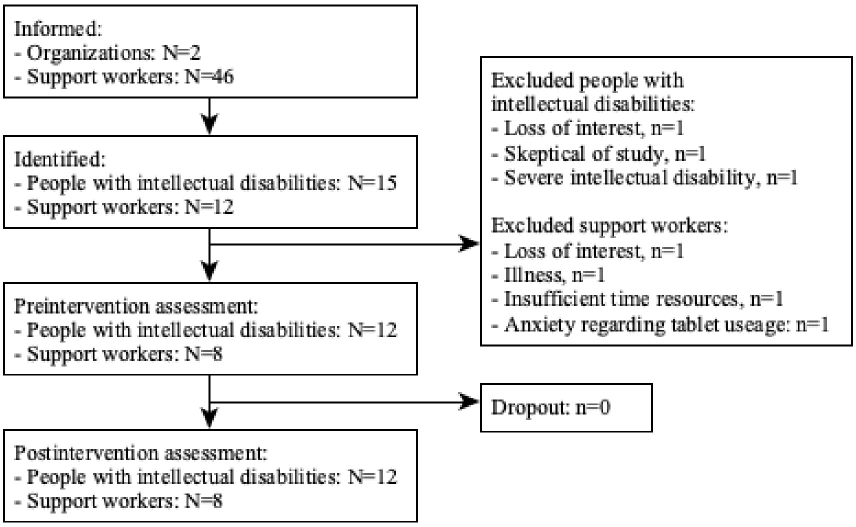
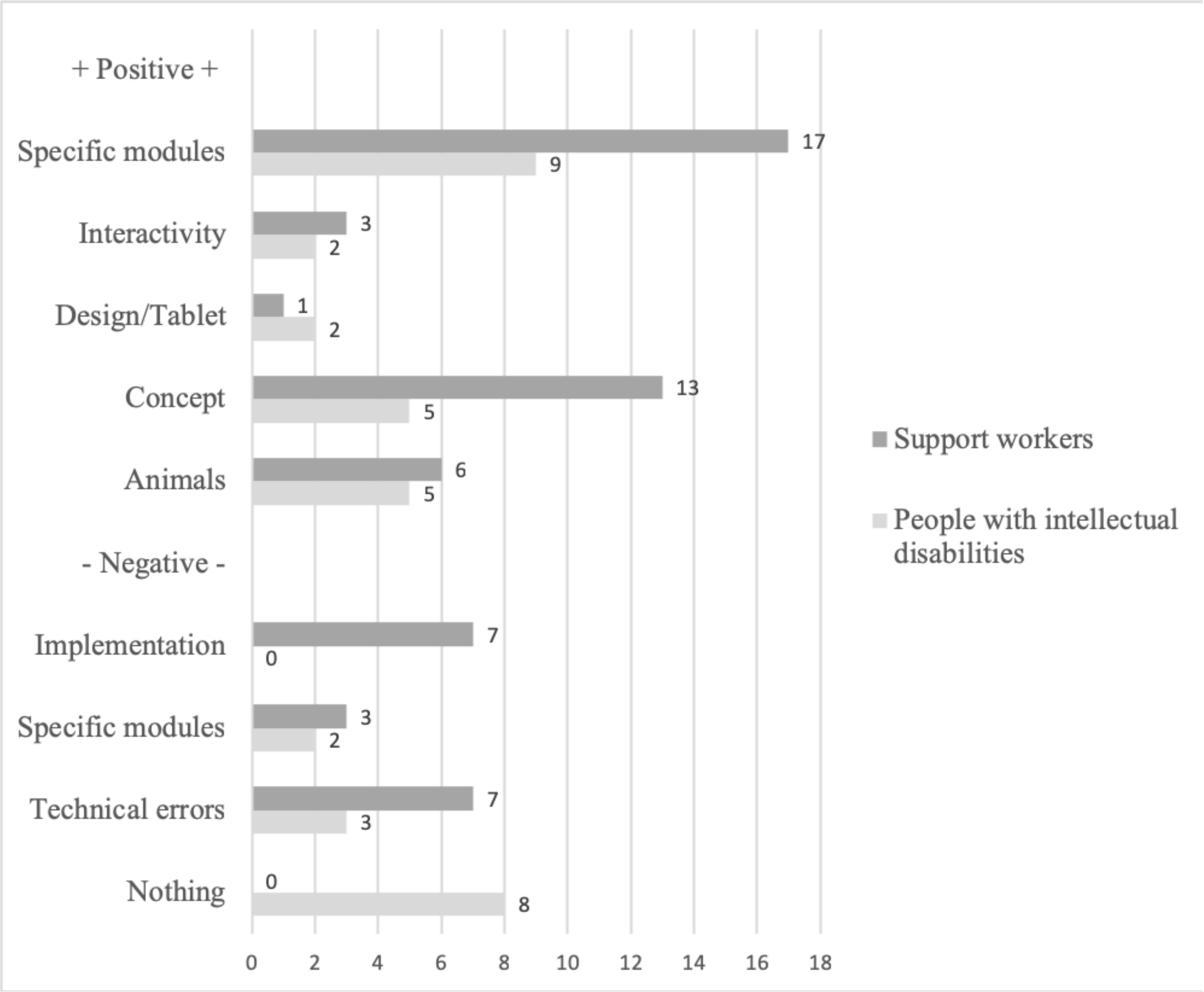


Table 1. Sociodemographic characteristics of participants with intellectual disabilities (N=12) and support workers (N=8).

Sociodemographic characteristics, variable	Value
People with intellectual disabilities	
Age (years), median (range)	19 (17-24)
Sex, male, n (%)	12 (100)
Mild intellectual disability, n (%)	9 (75)
Moderate intellectual disability, n (%)	3 (25)
Support workers	
Age (years), median (range)	42 (28-58)
Sex, male, n (%)	3 (38)
Years working with participant, median (range)	3 (0.5-6)

Figure 2. Acceptability of Pudelwohl as indicated by code frequency of the quantitative content analysis of the Program Evaluation Questionnaires.



Program Concept and Design

The use of animal avatars was liked by the majority of people with intellectual disabilities and support workers. The interactive design, with quizzes, and use of both images and videos were also positives of the program. Support workers valued the program as a user-appropriate concept based on its audiovisual approach, the use of repetition, encouragement and supportive feedback, and session length. Given the heterogeneity of people with intellectual disabilities in terms of cognitive and communicative abilities, some support workers also rated the intervention as either too childish or difficult with regard to language. By contrast, none of the participants with intellectual disabilities mentioned the program to be childish and instead were explicitly positive about the design of the intervention.

Program Content

The friendships, self-acceptance, and self-actualization modules were particularly well received by people with intellectual disabilities, as were specific tips to improve your well-being. Support workers approved the wide range of topics and the information regarding local service providers included in specific modules and also highlighted the need to provide more concrete and less abstract tips in the self-actualization module. They considered the content of the modules concerning maintaining friendships, being active, and relaxation as most helpful to the people they were supporting. Both user groups considered the communication module too difficult, which is likely because of the metaphor used to explain the different roles and aspects of good communication (ie, the talker, the listener, a conversation topic, switching talking-listening roles, and changing topics).

Intervention Usage

Technical errors in the use of the program were a main hindrance for both people with intellectual disabilities and their support workers. This applied to navigation arrows that were considered too small and buttons or interactions which were inactive. Session length was appropriate, according to support workers, and implementation should be limited to 1 session per week.

Overall satisfaction rates for the 4-week intervention were high. Participants with intellectual disabilities and support workers gave the program approval rates of 92% and 86%, respectively.

Evaluation Outcomes

The picture-based mood barometers for participants with intellectual disabilities administered before each session were converted to a 5-point Likert scale ranging from 0 to 4, with higher scores indicative of a better mood. Descriptive analysis showed improved moods for 5 participants and deteriorations for 2 participants (Figure 3). A total of 5 participants gave the maximum mood rating at each measurement point and are, therefore, not included in Figure 3.

Table 2 presents changes in self-reported personal well-being and proxy-report behavioral problems from pretest to posttest. Data are reported separately for the 2 organizations because of the observed differences in their baseline sample characteristics. Statistical analyses were conducted, as planned, with the full sample. PWI-ID data for 3 participants were excluded due to acquiescent responding. Wilcoxon signed rank tests yielded no statistical difference from pretest (median=79; range 39-86) to posttest (median=79; range 21-96) for subjective well-being in people with intellectual disabilities (W=10.5; *P*=.17). There was also a nonsignificant reduction in behavioral problems (W=14; *P*=.05) as reported by support workers at pretest (median=19; range 0-94) and posttest (median=14; range 1-73).

Finally, data from the program evaluation questionnaires showed that half of the support workers did not observe any changes in the behavior or well-being of the participants with intellectual disabilities. Other support workers did report an improved relationship with the participant with intellectual disabilities, as well as increased interests in daily activities associated with the program content (including planning activities, self-acceptance, and independent use of computers).

Figure 3. Mood ratings of participants with intellectual disabilities with fluctuating mood during the intervention (N=7).

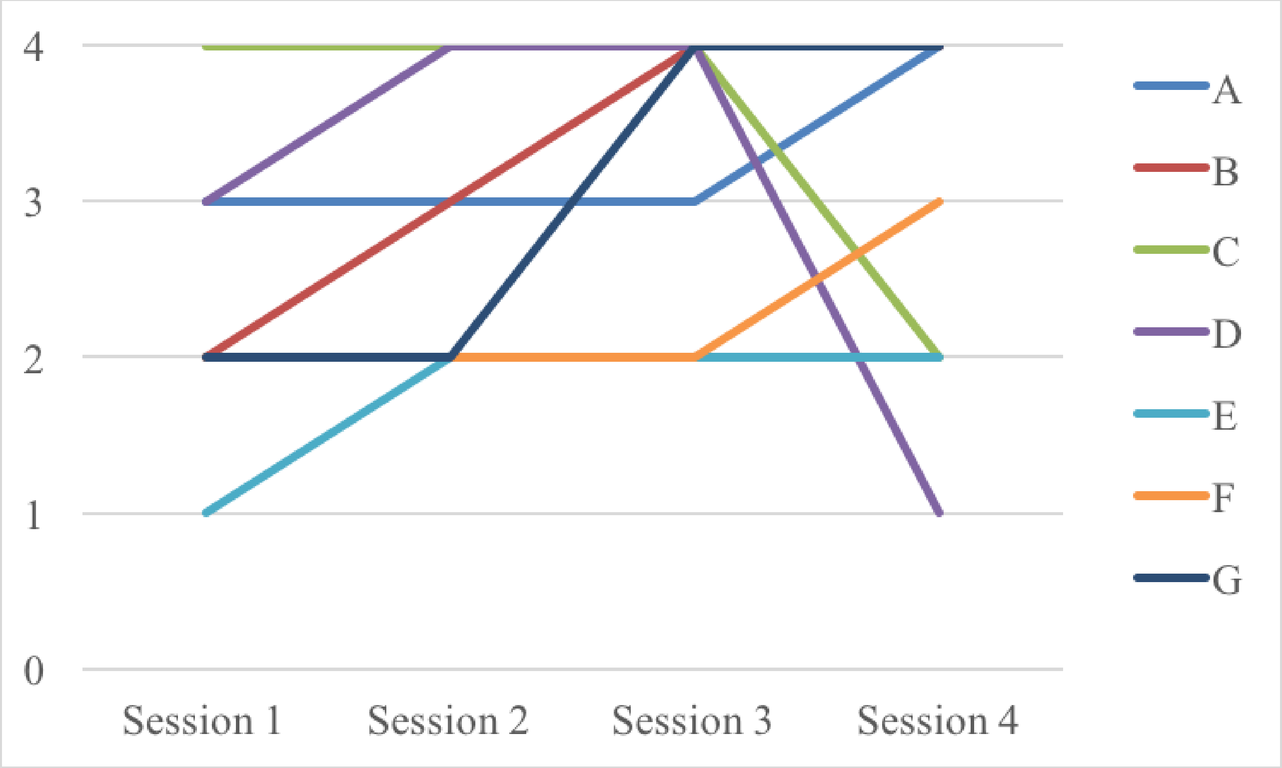


Table 2. Median and range for subjective well-being (Personal Well-Being Index—Intellectual Disability, N=9) and behavioral problems (Aberrant Behavior Checklist—Community, N=12) at baseline and post intervention.

Outcome measure	Baseline, median (range)	Post intervention, median (range)
Personal Well-Being Index—Intellectual Disability		
Organization A (n=6)	79 (39-82)	79 (21-96)
Organization B (n=3)	54 (54-86)	64 (57-79)
Aberrant Behavior Checklist—Community		
Organization A (n=8)	29 (10-94)	21 (11-73)
Organization B (n=4)	4 (0-10)	3 (1-3)

Discussion

Principal Findings

This pilot study shows that it is feasible to develop a digital psychoeducational intervention—with multiple brief sessions of interactive content—for use in routine practice. We also observed initial improvements in subjective well-being and reduced behavioral problems in some participants. Participants' interest in the study and the absence of any dropouts further indicate that an evaluation of the intervention's efficacy in a larger sample would be feasible.

The participatory approach in developing the intervention together with people with and without intellectual disabilities or mental health problems was both necessary and valuable in the changes that were made to the program. User-driven changes sometimes contradicted universal design recommendations that were otherwise expected to improve the accessibility of digital materials [21]. Their involvement in the intervention design is likely to have positively impacted upon the eventual uptake and acceptability of the intervention in this pilot study. The discrepancy between people with intellectual disabilities liking the use of animal avatars and other images and some support workers considering the intervention to be childish further highlights the need to involve all end users in the design of an intervention and not be satisfied with the input from only support workers or people with intellectual disabilities.

Study Limitations and Strengths

Key limitations of this study are the lack of randomization procedures and follow-up measures as well as the small sample size. These were deliberate decisions as the main objective was not to statistically evaluate the intervention's efficacy but to investigate, using a naturalistic study design, whether it would be acceptable and feasible to implement the proposed intervention in the weekly routine of people with intellectual disabilities, and whether the recruitment of people with intellectual disabilities for an interventions study would be feasible.

The absence of any statistical effects of improved mood and behavior is also not surprising as we did not employ a threshold score on either of the outcome measures to select a sample where improvements in the mental well-being and behavioral problems were most desirable.

We also noted that the 4-week duration of the study is likely to have been too short for participants to implement meaningful changes and to record their potential impact. Meanwhile, the quick succession of very brief modules on distinct topics did not allow for a comprehensive or thorough work-through for each module. Hence, participants may not have had sufficient time to explore and master the attitudes, knowledge, and skills related to each module. Although from a clinical perspective that would have been more useful, this study focused primarily on exploring whether such content could be delivered using a digital format. The brief modules in combination with the qualitative feedback, therefore, enabled us to identify which modules in particular could be considered more useful for expansion in the intervention's further development.

In spite of our efforts to recruit a representative sample, the final sample of people with intellectual disabilities was all male. It is not clear whether this reflected a bias by the support workers in identifying potentially interested participants, or an actual gender difference in interest to participate. Previous studies indicate, however, that technology use among people with intellectual disabilities is generally higher for adolescents and young adults aged below 30 years of age and for men [32,33].

The technical errors, as mentioned by the participants, were an additional negative aspect of the intervention. In spite of this, the overall approval ratings were nevertheless high.

Finally, although no participants dropped out during the study, it is unclear to what extent they adhered to the intervention contents and completed the homework tasks as we did not collect any data on actual use of the intervention. Given that usage characteristics may not be the ideal means of assessing adherence for electronic health technologies [34], future studies could explore to what extent participants experimented, tried, or internalized specific contents.

Future Research and Clinical Implications

Overview

The Pudelwohl intervention, in its content and design, most closely approaches a combination of the 2 intervention arms of the Beat It trial, behavioral activation and guided self-help, both of which led to patient improvement [35]. Direct comparison of both studies demonstrates that the pilot Pudelwohl intervention resulted in higher completion rates (95% vs 81%), but also yielded lower effect sizes as our study design was not set up for evaluating intervention efficacy. A larger trial of the present intervention in a similar sample of people with at least subclinical symptoms of reduced well-being is required to more adequately compare both the interventions. A shared finding, however, is that participants with intellectual disabilities in both studies reported a desire for more sessions or an overall longer intervention [36].

Focusing on the digital components of our intervention, we found similar levels of acceptability to that of a website aimed at informing people with intellectual disabilities about good mental health [37].

In terms of customizability and tailoring the intervention to individual participant needs, our intervention offers a compromise between a more rigid design of a website and the flexibility that can be offered in a face-to-face therapy, for example, by presenting options for local activities, giving participants the option to choose those tasks that are more relevant to them, and by involving support workers who could ensure program content was discussed in a way that was meaningful to the participant. However, it is not yet clear whether the acceptability and efficacy of individually tailored digital interventions is higher than that of standard packages [38].

Adapting Digital Interventions for People With Intellectual Disabilities

Contributing to the intervention's overall acceptability is our purposeful attempt to adapt all aspects of the intervention to the variable needs of people with intellectual disabilities. This includes incorporating all but one of the suggested adaptations to psychological therapies for people with intellectual disabilities, as put forward by Hurley et al [39]. Use of a digital platform should not exempt the intervention from addressing the users' intellectual disability; however, this disability in itself may present the person with considerable challenges and negative events in their daily lives that could contribute to poor mental well-being. In our intervention, we addressed this in Module 5 *Self-acceptance*, where participants were prompted to both say what they are good at and to reflect on questions such as *my weaknesses are not failures*, *you can laugh about some mistakes*, *I should feel bad about my weaknesses*, and *my weaknesses are a part of me*.

Defining the User Group: Should We Include Support Workers?

Involving support workers is generally viewed favorably when planning or delivering psychological interventions to people with intellectual disabilities [35,39,40]. When using digital interventions, their involvement can help overcome technical difficulties experienced by people with intellectual disabilities, as we found in this study. By contrast, their involvement could also present an extra barrier when support workers do not feel sufficiently confident themselves in using technologies or wish to protect the person from the potential risks associated with the use of Web-based technologies [23]. This approach could therefore lead to people with intellectual disabilities who might otherwise express an interest in digital interventions or who would have the capacity and skills to take part without the need for a support worker to be excluded from research or interventions built around the involvement of these gatekeepers. In this study, the challenges arising from such a design are apparent from the contrast between the low recruitment rate of support workers and the high interest of people with intellectual disabilities.

When a decision is made to include support workers, either as active participants or in a supporting role, attention should also be directed to any potential effects of the intervention on the support worker and their relationship with the person with an intellectual disability. This could include changes in attitudes, knowledge, or skills, as well as the quality of the relationship, which was reported to have improved by some of the support workers in this study.

Conclusions

Overall, and in spite of the methodological shortcomings of this pilot study, this study clearly shows that internet interventions to improve the mental well-being of people with intellectual disabilities are worth further exploring as an additional intervention or prevention strategy, alongside more traditional psychopharmacological and psychotherapeutic approaches. Involving people with intellectual disabilities in the design of the intervention and ensuring the involvement of support workers and personal relevance of the intervention contents were major contributors toward the acceptance and feasibility of this intervention. The higher interest and usage rates of digital technologies in young people with intellectual disabilities identify them as the likely initial target group for such interventions.

Acknowledgments

The authors would like to thank the staff and visitors of PIKSL Labor Bielefeld for their contribution and insights during the development of the intervention. They acknowledge support for the Article Processing Charge by the Deutsche Forschungsgemeinschaft and the Open Access Publication Fund of Bielefeld University.

Conflicts of Interest

None declared.

References

1. Biswas AB, Furniss FG. Life events and mental illness in people with learning disabilities. Psychiatry 2009 Nov;8(11):445-447. [doi: 10.1016/j.mppsy.2009.07.005]

2. Cooper SA, Smiley E, Morrison J, Williamson A, Allan L. Mental ill-health in adults with intellectual disabilities: prevalence and associated factors. Br J Psychiatry 2007 Jan;190:27-35. [doi: 10.1192/bjp.bp.106.022483] [Medline: 17197653]
3. Smiley E, Cooper SA, Finlayson J, Jackson A, Allan L, Mantry D, et al. Incidence and predictors of mental ill-health in adults with intellectual disabilities: prospective study. Br J Psychiatry 2007 Oct;191:313-319. [doi: 10.1192/bjp.bp.106.031104] [Medline: 17906241]
4. Sheehan R, Hassiotis A, Walters K, Osborn D, Strydom A, Horsfall L. Mental illness, challenging behaviour, and psychotropic drug prescribing in people with intellectual disability: UK population based cohort study. Br Med J 2015 Sep;351:h4326 [FREE Full text] [doi: 10.1136/bmj.h4326] [Medline: 26330451]
5. Vereenooghe L, Langdon PE. Psychological therapies for people with intellectual disabilities: a systematic review and meta-analysis. Res Dev Disabil 2013 Nov;34(11):4085-4102. [doi: 10.1016/j.ridd.2013.08.030] [Medline: 24051363]
6. Koslowski N, Klein K, Arnold K, Kösters M, Schützwohl M, Salize HJ, et al. Effectiveness of interventions for adults with mild to moderate intellectual disabilities and mental health problems: systematic review and meta-analysis. Br J Psychiatry 2016 Dec;209(6):469-474. [doi: 10.1192/bjp.bp.114.162313] [Medline: 27198481]
7. Bender M. The unoffered chair: The history of therapeutic disdain towards people with a learning difficulty. Clin Psychol Forum 1993;54:7-12.
8. Hayes S, McGuire B, O'Neill M, Oliver C, Morrison T. Low mood and challenging behaviour in people with severe and profound intellectual disabilities. J Intellect Disabil Res 2011 Feb;55(2):182-189. [doi: 10.1111/j.1365-2788.2010.01355.x] [Medline: 21129068]
9. Melville CA, Johnson PC, Smiley E, Simpson N, Purves D, McConnachie A, et al. Problem behaviours and symptom dimensions of psychiatric disorders in adults with intellectual disabilities: An exploratory and confirmatory factor analysis. Res Dev Disabil 2016;55:1-13. [doi: 10.1016/j.ridd.2016.03.007] [Medline: 27018744]
10. Willner P. Readiness for cognitive therapy in people with intellectual disabilities. J Appl Res Int Dis 2006;19(1):5-16. [doi: 10.1111/j.1468-3148.2005.00280.x]
11. Taylor JL, Lindsay WR, Willner P. CBT for people with intellectual disabilities: emerging evidence, cognitive ability and IQ effects. Behav Cogn Psychother 2008;36(6):723-733. [doi: 10.1017/s1352465808004906]
12. Cooney P, Tunney C, O'Reilly G. A systematic review of the evidence regarding cognitive therapy skills that assist cognitive behavioural therapy in adults who have an intellectual disability. J Appl Res Intellect Disabil 2017;31(1):23-42. [doi: 10.1111/jar.12365] [Medline: 28544303]
13. Vereenooghe L, Reynolds S, Gega L, Langdon PE. Can a computerised training paradigm assist people with intellectual disabilities to learn cognitive mediation skills? A randomised experiment. Behav Res Ther 2015 Aug;71:10-19. [doi: 10.1016/j.brat.2015.05.007] [Medline: 26004217]
14. Vereenooghe L, Gega L, Reynolds S, Langdon PE. Using computers to teach people with intellectual disabilities to perform some of the tasks used within cognitive behavioural therapy: A randomised experiment. Behav Res Ther 2016 Jan;76:13-23. [doi: 10.1016/j.brat.2015.11.002] [Medline: 26600556]
15. Tsimopoulou I, Kroese BS, Unwin G, Azmi S, Jones C. A case series to examine whether people with learning disabilities can learn prerequisite skills for cognitive behavioural therapy. Cogn Behav Ther 2018;11(e1). [doi: 10.1017/s1754470x1700023x]
16. Dattilo J, Cory L, Williams R. Effects of computerized leisure education on knowledge of social skills of youth with intellectual disabilities. Ther Recreation J 2003;37(2):142-155 [FREE Full text]
17. Standen PJ, Brown DJ, Cromby JJ. The effective use of virtual environments in the education and rehabilitation of students with intellectual disabilities. Br J Educ Technol 2001;32(3):289-299. [doi: 10.1111/1467-8535.00199]
18. Standen PJ, Brown DJ. Virtual reality in the rehabilitation of people with intellectual disabilities: review. Cyberpsychol Behav 2005 Jun;8(3):272-82; discussion 283. [doi: 10.1089/cpb.2005.8.272] [Medline: 15971976]
19. Cooney P, Jackman C, Coyle D, O'Reilly G. Computerised cognitive-behavioural therapy for adults with intellectual disability: randomised controlled trial. Br J Psychiatry 2017;211(2):95-102. [doi: 10.1192/bjp.bp.117.198630] [Medline: 28596245]
20. Vereenooghe L, Gega L, Langdon PE. Intellectual disability and computers in therapy: views of service users and clinical psychologists. Cyberpsychology J Psychosoc Res Cybersp 2017;11(1). [doi: 10.5817/CP2017-1-11]
21. Sheehan R, Hassiotis A. Digital mental health and intellectual disabilities: state of the evidence and future directions. Evid Based Ment Health 2017 Nov;20(4):107-111. [doi: 10.1136/eb-2017-102759] [Medline: 28947677]
22. Chadwick D, Wesson C, Fullwood C. Internet access by people with intellectual disabilities: inequalities and opportunities. Futur Internet 2013;5(3):376-397. [doi: 10.3390/fi5030376]
23. Chadwick DD, Quinn S, Fullwood C. Perceptions of the risks and benefits of internet access and use by people with intellectual disabilities. Br J Learn Disabil 2017;45(1):21-31. [doi: 10.1111/bld.12170]
24. Vereenooghe L, Westermann K. Co-development of an interactive digital intervention to promote the well-being of people with intellectual disabilities. Int J Dev Disabil 2019;65(3):128-134. [doi: 10.1080/20473869.2019.1599606]
25. Elliott R, Slatick E, Urman M. Qualitative change process research on psychotherapy: alternative strategies. In: Frommer J, Rennie DL, editors. Qualitative Psychotherapy Research: Methods and Methodology. Lengerich, Germany: Pabst Science Publishers; 2001:69-111.

26. 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;9:82-87 [FREE Full text] [doi: 10.1016/j.invent.2017.07.003] [Medline: 30135841]
27. Cummins R, Lau A. Personal Wellbeing Index - Intellectual Disability. Third Edition. Melbourne, Australia: Australian Centre on Quality of Life; 2005.
28. Renn D, Pfaffenberger N, Platter M, Mitmansgruber H, Cummins RA, Höfer S. International well-being index: The Austrian version. Soc Indic Res 2009;90(2):243-256. [doi: 10.1007/s11205-008-9255-2]
29. Aman M, Singh N. The Aberrant Behavior Checklist - Community: Supllementary Manual. East Aurora: Slosson Educational Publications; 1994.
30. Weber G, Fritsch A, Rögner M. Checkliste für abweichendes Verhalten: ABC- Community, deutschsprachige Version. Vienna: Universität Wien; 2003.
31. Morgan DL. Qualitative content analysis: a guide to paths not taken. Qual Health Res 1993 Feb;3(1):112-121. [doi: 10.1177/104973239300300107] [Medline: 8457790]
32. Tanis ES, Palmer S, Wehmeyer M, Davies DK, Stock SE, Lobb K, et al. Self-report computer-based survey of technology use by people with intellectual and developmental disabilities. Intellect Dev Disabil 2012 Feb;50(1):53-68 [FREE Full text] [doi: 10.1352/1934-9556-50.1.53] [Medline: 22316226]
33. Berger A, Caspers T, Croll J, Hofmann J, Kubicek H, Peter U. Web 2.0/Barrierefrei. Eine Studie zur Nutzung von Web2.0 Anwendungen durch Menschen mit Behinderungen. Germany: Aktion Mensch e.V; 2010.
34. 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;19(12):e402 [FREE Full text] [doi: 10.2196/jmir.8578] [Medline: 29212630]
35. Jahoda A, Hastings R, Hatton C, Cooper S, Dagnan D, Zhang R, et al. Comparison of behavioural activation with guided self-help for treatment of depression in adults with intellectual disabilities: a randomised controlled trial. Lancet Psychiatry 2017 Dec;4(12):909-919 [FREE Full text] [doi: 10.1016/S2215-0366(17)30426-1] [Medline: 29153873]
36. Knight R, Jahoda A, Scott K, Sanger K, Knowles D, Dagnan D, et al. 'Getting into it': people with intellectual disabilities' experiences and views of behavioural activation and guided self-help for depression. J Appl Res Intellect Disabil 2019 Jul;32(4):819-830. [doi: 10.1111/jar.12571] [Medline: 30762261]
37. Watfern C, Heck C, Rule C, Baldwin P, Boydell KM. Feasibility and acceptability of a mental health website for adults with an intellectual disability: qualitative evaluation. JMIR Ment Health 2019 Mar 28;6(3):e12958 [FREE Full text] [doi: 10.2196/12958] [Medline: 30920378]
38. 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]
39. Hurley A, Tomasulo D, Pfadt A. Individual and group psychotherapy approaches for persons with mental retardation and developmental disabilities. J Dev Phys Disabil 1998;10(4):365-386. [doi: 10.1023/A:1021806605662]
40. Willner P. The effectiveness of psychotherapeutic interventions for people with learning disabilities: a critical overview. J Intellect Disabil Res 2005;49(Pt 1):73-85. [doi: 10.1111/j.1365-2788.2005.00633.x] [Medline: 15634314]

Abbreviations

ABC-C: Aberrant Behavior Checklist—Community
CBT: cognitive behavioral therapies
PWI-ID: Personal Well-Being Index—Intellectual Disability

Edited by G Eysenbach; submitted 26.06.19; peer-reviewed by C Watfern, R Sheehan; comments to author 04.07.19; revised version received 06.08.19; accepted 17.08.19; published 14.11.19.

Please cite as:
Vereenooghe L, Westermann K
Feasibility and Acceptability of an Interactive Mental Well-Being Intervention for People With Intellectual Disabilities: Pilot Mixed Methods Study
JMIR Form Res 2019;3(4):e15190
URL: http://formative.jmir.org/2019/4/e15190/
doi:10.2196/15190
PMID:31724954

Original Paper

The Role of Age, Education, and Digital Health Literacy in the Usability of Internet-Based Cognitive Behavioral Therapy for Chronic Pain: Mixed Methods Study

Rosalie van der Vaart[1], PhD; Dorine van Driel[2], MSc; Kristel Pronk[1], MSc; Suzan Paulussen[1], MSc; Selma te Boekhorst[3], MSc; Judith G M Rosmalen[4,5], PhD; Andrea W M Evers[1], PhD

[1]Leiden University, Faculty of Social and Behavioural Sciences, Health, Medical and Neuropsychology Unit, Leiden, Netherlands
[2]SeniorBeter, Practice of Old Age Psychiatry, Gendt, Netherlands
[3]Dimence, Almelo, Netherlands
[4]University of Groningen, University Medical Center Groningen, Department of Psychiatry, Groningen, Netherlands
[5]University of Groningen, University Medical Center Groningen, Department of Internal Medicine, Groningen, Netherlands

Corresponding Author:
Rosalie van der Vaart, PhD
Leiden University
Faculty of Social and Behavioural Sciences
Health, Medical and Neuropsychology Unit
Wassenaarseweg 52
Leiden
Netherlands
Phone: 31 071 5276633
Email: r.vandervaart@fsw.leidenuniv.nl

Abstract

Background: Internet-based cognitive behavior therapy (iCBT) can be effective in mental and somatic health care. Research on the feasibility of internet interventions in clinical practice is, however, still scarce. Studies with a focus on the patient regarding usability of interventions and digital health literacy skills are especially lacking.

Objective: The goal of this study was to assess the usability of an iCBT for chronic pain, Master Your Pain, and the relationship between its usability outcomes and the factors age, educational level, and digital health literacy skills. The aims were to determine what changes were needed in the program for sufficient usability and which individual characteristics were related to the usability of the program.

Methods: Patients were recruited from two mental health care practices. A mixed methods approach was used in this study. A qualitative observational study comprising performance tasks in the iCBT program was used to test usability. A quantitative questionnaire was used to measure possible related constructs. Usability was operationalized as the number of tasks that could be completed and the type and number of problems that occurred while doing so. Performance tasks were set up to measure 6 digital skills: (1) operating the computer and internet browser, (2) navigation and orientation, (3) using search strategies, (4) evaluating relevance of content, (5) adding personal content, and (6) protecting and respecting privacy. Participants were asked to think aloud while performing the tasks, and screen activities and webcam recordings were captured. The qualitative observational data was coded using inductive analysis by two independent researchers. Correlational analyses were performed to test how usability relates to sociodemographics and digital health literacy.

Results: A total of 32 patients participated, with a mean age of 49.9 years and 84% (27/32) being female. All performance tasks except one (fill in a diary registration) could be completed independently by more than 50% of the participants. On operational, navigation, and search levels, participants struggled most with logging in, logging out, and finding specific parts of the intervention. Half of the sample experienced problems evaluating the relevance and adding content to the program to some extent. Usability correlated moderately negatively with age and moderately positively with digital health literacy skills but not with educational level.

Conclusions: The results provide insight into what is essential for proper usability regarding the design of an iCBT program considering variations in age, educational level, and digital health literacy. Furthermore, the results provide insight into what type of support is needed by patients to properly use the intervention. Tailoring support among the needs of certain age groups or skill levels could be beneficial and could range from no extra support (only online feedback, as intended) to practical support (an additional usability introduction session) to blended care (combined face-to-face sessions throughout the therapy).

(JMIR Form Res 2019;3(4):e12883) doi:10.2196/12883

KEYWORDS

internet-based cognitive behavior therapy; chronic pain; usability; digital health literacy skills; eHealth literacy

Introduction

Internet interventions are effective for treating patients with a broad range of mental health problems and supporting patients in their coping and self-management of chronic somatic conditions [1,2]. In general, these types of interventions are therapist-guided, Web-based programs. Patients log in using a personal account and enter a personal dashboard, from which they can follow modules consisting of psychoeducational texts, assignments, relaxation exercises, and diary registrations. Through the program they can also communicate with their care provider, commonly via secured email. In clinical trials, where participants are generally rather homogeneous, these programs have been found to be effective in reaching behavior change. However, less studied is the extent to which these interventions are feasible to use in clinical practice for a heterogeneous patient population [3]. This regards their fit in regular care processes and usability among individual patients and care providers. Over the past few years it has become clear that implementing internet interventions into regular health care is challenging [4-6], but the focus on the needs and skills of patients in this matter has been minimal.

Up until now, little has been known about for which (type of) patients internet interventions would be suitable [7]. A well-established precondition for effective and efficient use of internet interventions is that the programs are easy and intuitive to use for patients [8]. Furthermore, failure to meet the needs and skills of intended users can play a key role in disappointing use [9,10]. The need for assistance while using an internet intervention may imply that the usability of the application must be improved, which could be achieved by redesigning the program to simplify the user interface or offering guidance within the program. Nevertheless, in addition to the usability of an internet intervention being relevant to reach optimal effect, the skill level of the individual using the intervention must be adequate. These usage skills are called eHealth literacy, or digital health literacy, defined as the skills needed to “seek, find, understand, and appraise health information from electronic sources and apply the knowledge gained to addressing or solving a health problem” [11]. An in-depth examination of digital health literacy disclosed six types of health-related internet skills needed when using online health information and interactive interventions: (1) operating the computer and internet browser, (2) navigating and orienting, (3) using search strategies, (4) evaluating relevance and reliability of Web content, (5) adding personal content to the Web, and (6) protecting and respecting privacy [12].

Research on the use of health-related digital applications has shown that sociodemographics, age, and educational level are associated with the skills required to successfully use these applications [13]. Van Deursen et al [13] used performance-based assignments to measure usage skills of people in the general population, which showed that older age and lower educational level were predictive of lower operational and formal skills, the practical skills needed to use and navigate on a computer and on the internet. Additionally, educational level was predictive of informational and strategic skills, indicating that individuals with lower educational levels have more difficulties finding and critically appraising online information [13]. This suggests that elderly and less educated individuals might be more vulnerable groups when it comes to using internet interventions. These users might need more guidance or assistance in using an internet intervention, or they might benefit more from blended care, in which they also have regular face-to-face contact with their therapist [7].

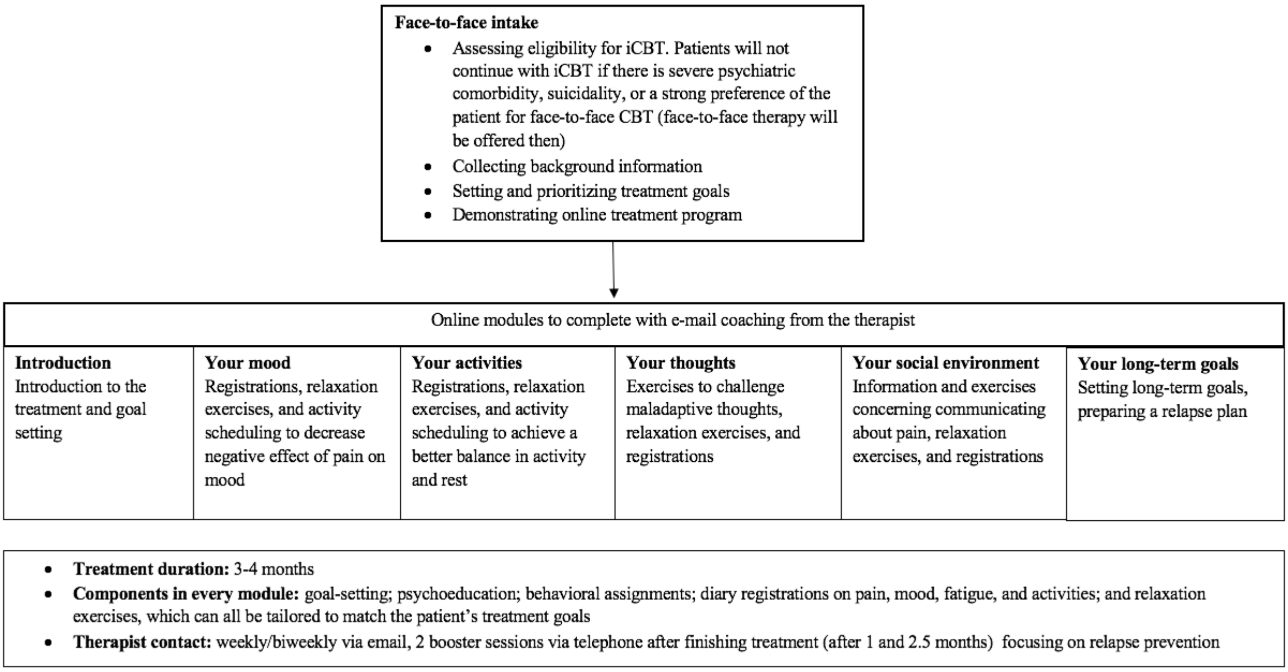
In this study, the usability of an internet-based cognitive behavioral therapy (iCBT) called Master Your Pain was assessed in patients with chronic pain to best facilitate patient needs and skills in using the program. First, we measured to what extent potential users were able to perform relevant assignments related to natural use of the intervention. This was operationalized by assessing the type and number of encountered problems. Second, since the population of chronic pain patients tends to be older and experience with the internet might be less, the associations between age, educational level, digital health literacy skills, and the performance and the number of occurred problems in the iCBT program were explored. This provided preliminary insight into the feasibility of using the intervention in clinical practice and the possible need for tailoring (ie, providing additional assistance or guidance to people working with the program).

Methods

Design

A mixed methods design was used in which a qualitative observational study was performed to test the usability of a Dutch iCBT called Master Your Pain for people with chronic pain (see Figure 1 for the content of the program and Figure 2 for a screenshot of the interface) [14], and a quantitative questionnaire was used to collect data on usability-related variables. Master Your Pain is derived from an evidence-based iCBT platform for chronic somatic conditions [15,16]. From the available treatment modules in the original platform, six modules (Your goals, Your mood, Your activities, Your thoughts, Your social environment, and Your long-term goals) were created to comprise an iCBT focused on coping with chronic pain using psychoeducation, assignments, diary registrations, and relaxation exercises. Communication with the therapist takes place via secure email messages.

Figure 1. Content of the internet-based cognitive behavioral therapy (iCBT) intervention Master Your Pain.



Face-to-face intake

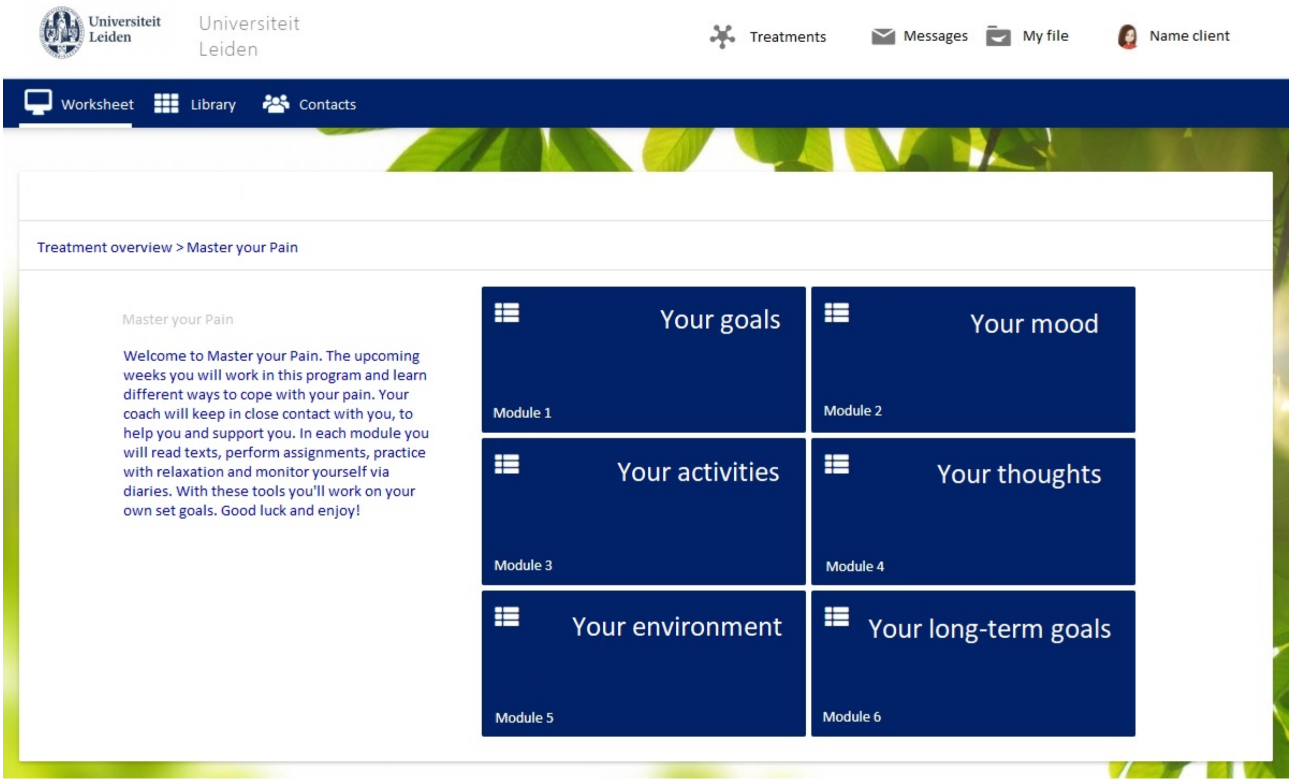
- Assessing eligibility for iCBT. Patients will not continue with iCBT if there is severe psychiatric comorbidity, suicidality, or a strong preference of the patient for face-to-face CBT (face-to-face therapy will be offered then)
- Collecting background information
- Setting and prioritizing treatment goals
- Demonstrating online treatment program

Online modules to complete with e-mail coaching from the therapist

Introduction	Your mood	Your activities	Your thoughts	Your social environment	Your long-term goals
Introduction to the treatment and goal setting	Registrations, relaxation exercises, and activity scheduling to decrease negative effect of pain on mood	Registrations, relaxation exercises, and activity scheduling to achieve a better balance in activity and rest	Exercises to challenge maladaptive thoughts, relaxation exercises, and registrations	Information and exercises concerning communicating about pain, relaxation exercises, and registrations	Setting long-term goals, preparing a relapse plan

- **Treatment duration:** 3-4 months
- **Components in every module:** goal-setting; psychoeducation; behavioral assignments; diary registrations on pain, mood, fatigue, and activities; and relaxation exercises, which can all be tailored to match the patient’s treatment goals
- **Therapist contact:** weekly/biweekly via email, 2 booster sessions via telephone after finishing treatment (after 1 and 2.5 months) focusing on relapse prevention

Figure 2. Master Your Pain homepage.



Welcome to Master your Pain. The upcoming weeks you will work in this program and learn different ways to cope with your pain. Your coach will keep in close contact with you, to help you and support you. In each module you will read texts, perform assignments, practice with relaxation and monitor yourself via diaries. With these tools you'll work on your own set goals. Good luck and enjoy!

Participants

To establish a representative sample, participants were recruited in two different mental health care institutions from two geographically distinct regions in the Netherlands. One institution specializes in treating patients with fibromyalgia and the other in treatment of medically unexplained symptoms and patients of older age with psychiatric problems. The inclusion criteria were being aged 18 years or above, having some internet experience (assessed by asking about home access to the internet and experience with common online activities such as emailing and online banking), having experience with a Word operating system, and mastery of reading and writing in the Dutch language. Health care providers from the institutions informed all their patients within a certain time frame (March and April 2016 and from June to December 2017, respectively) about the study by offering an information letter regarding the goal of the study, its procedure, location, estimated time investment, compensation, and participants’ rights during the study. If patients were interested, care providers would check the inclusion criteria. Upon eligibilty determination, the researcher was contacted by email to invite the patient for the study. Before the start of the study, participants were asked to sign an informed consent form. This informed consent form contained information about the permission to record video and audio material, anonymity of the participants, confidential use of data, and the possibility to stop at any time without having to state any reasons. After three consecutive assessments without new usability problems occurring, data saturation was reached. A total of 34 participants took part in the study. However, one participant appeared to have experience with the iCBT and one was incorrectly included, since this person never used the internet (and was not able to). Both were excluded from the analyses.

Procedure and Materials

The procedure took 60 to 90 minutes per participant. Participants were first asked to complete a questionnaire. The questionnaire started with items regarding demographic characteristics (age, sex, and educational level) and internet use (Through which medium do you have access to the internet? How often do you use the internet? Have you used the internet before to look up health-related information? Have you ever used an online self-help tool?). Digital health literacy skills were then measured using the Digital Health Literacy Instrument (DHLI) [17]. This 21-item questionnaire assesses 7 digital health literacy skill categories: operational skills, navigation skills, searching for information, evaluating information, applying information, adding content to the Web, and ensuring privacy. Questions are answered on a 4-point response scale, ranging from very easy to very difficult or from never to often. A mean score (ranging between 1 and 4) can be calculated if at least 18 items are completed. When calculating the mean scores, item scores are reversed, so that a higher score represents a higher level of digital health literacy. This instrument has been validated in a general sample of the Dutch population, with a mean of 3.11 (SD 0.45), Cronbach alpha=.87. Content and construct validity showed to be adequate [17].

After finishing the questionnaire, participants were asked to take place behind a PC (using Windows 10). They were asked to perform 8 tasks in the Master Your Pain program using a mock account created for the study. The tasks were set up as representative as possible, reflecting actual use during treatment and including all kinds of different functionalities of the intervention. The performance tasks were based on the health-related internet skills found by Van der Vaart et al [12]. All tasks required more than one skill, but they were designed so that one skill was predominantly required for each task: logging in (operating the computer and internet browser), navigating forward and back to the homepage (navigating and orienting), opening a specific assignment (using search strategies), performing an assignment (evaluating relevance), sending a message to the therapist (adding personal content to the Web), and logging out from the website (protecting and respecting privacy). One skill, evaluating reliability, was not explicitly tested since this skill is very relevant using the World Wide Web but is less relevant on a treatment platform provided by a university. The exact tasks and associated skills can be found in Multimedia Appendix 1. To prevent learning effects, performance tasks 3 to 7 were presented in random order. Tasks 1, 2, and 8 appeared in the same order for everyone, since they represented logging in (task 1), getting familiar with the program (task 2), and logging out (task 8).

Prior to the performance tasks, the researcher explained the instructions and basic information about the iCBT to the participants. It was emphasized that the focus of the study was on usability testing rather than on measuring performance in order to reduce any performance anxiety. During the tasks, participants were asked to vocalize their thoughts by thinking aloud [18,19]. This is a verbal report method that provides insight to understanding a person's decision-making and problem-solving process. If participants stopped talking during the tasks, the researcher provided the following instruction: "Please, keep talking." [18]. When participants had trouble finishing an assignment, a protocol was in place to allow the researcher to provide hints on how to continue. In the case a participant claimed he was not able to complete a certain task and was going to give up or when the researcher observed signs that the participant was not able to continue for whatever reason, the researcher asked, "Would you stop at this point if you were using the application on your own?" If the answer was yes, the researcher asked, "May I give you a hint in order to continue?" If the answer was yes, the researcher gave a hint. FlashBack Express screen recorder (Blueberry Software) was used to record screen activity including webcam video and audio.

Data Analysis

Analyses were performed using the SPSS Statistics 23 (IBM Corp). Descriptive analyses were used to describe the sample's demographic characteristics and actual internet use. Mean scores and Cronbach alpha were calculated for the DHLI. The video files with screen activity and webcam recordings were used to derive two task variables. First, the level of completion per task was rated as not completed (score 0), completed with help (score 1 when a hint was provided), or completed independently (score 2), and then summed with a possible range from 0 to 16. Second, all encountered problems per task were registered using qualitative inductive analysis [20] and classified into the 6 health-related internet skills [12]. This way, an overview was created of the types of problems that occurred regarding each type of internet skill. Two researchers at each institution coded all tasks for every participant independently, using the same coding scheme. Differences in coding were discussed until consensus was reached. Subsequently, researchers pooled and labeled similar problems. Categories that consisted of a single occurred problem were excluded, because their relevance was deemed too small. Once the assessed performance task variables were agreed upon, they were entered in SPSS, and the total number of encountered problems was calculated for each participant. Normality was tested with a Shapiro-Wilk test for all variables, considering the small sample size. For the second aim of the study, Spearman ρ correlations were calculated between the two outcome measures task completion and number of problems encountered and the three variables that were assumed to be associated: age, educational level, and DHLI scores.

Results

Participants

The sample of 32 participants (Table 1) was 84% women (27/32), and the sample's age range was 22 to 75 years (mean 49.9 [SD 15.69] years). The majority had a high school or vocational school education level (13/32, 41%) and lived with a partner, kids, or parents (20/32, 63%). A laptop was the main device (29/32, 91%) used to access the internet at home; the other 3 participants had internet access at home via a PC. Internet was used (almost) every day by most participants (26/32, 81%). Of the participants, 97% (31/32) had searched for health-related information on the internet before. None of the participants ever used an online health intervention (data not shown). The average score on the DHLI was 2.90 (SD 0.48) on a scale from 1 to 4. Minimum and maximum scores ranged from 1.21 to 3.62, and Cronbach alpha=.86.

Table 1. Sociodemographics and internet use (N=32).

Characteristics	Value
Gender, n (%)	
Men	5 (16)
Women	27 (84)
Age (years), mean (SD)	49.9 (16)
Living situation, n (%)	
Alone	12 (38)
With partner	11 (34)
With partner and kids	6 (19)
With parents	3 (9)
Educational level, n (%)	
Primary school	12 (38)
High school or vocation	13 (41)
College degree or higher	7 (22)
Internet access[a], n (%)	
Laptop	29 (91)
Smartphone	17 (53)
PC	6 (19)
Frequency of internet use, n (%)	
(Almost) every day	26 (81)
Multiple days a week	3 (9)
Once a week or less	3 (9)
Digital health literacy[b], mean (SD)	2.90 (0.48)

[a] Respondents could mark more than one answer on this item.

[b] Measured by the Digital Health Literacy Instrument.

Completion of the Performance Tasks

All performance tasks except one could be completed independently by more than 50% of the participants (Table 2). Task 7 (fill in a diary registration) was most often not completed, and participants most often needed help with performance task 1 (log in to the platform). Tasks 1, 4, and 6 could not be completed by a participant during testing (different participants).

Table 2. Completion and performance per performance task (N=32).

#	Assignment	Completed independently, n (%)	Completed with help, n (%)	Not completed, n (%)
1	Log in to platform[a]	15 (47)	12 (38)	1 (3)
2	Navigate to assignment and back to homepage	22 (69)	5 (16)	5 (16)
3	Send message via mailbox in platform	22 (69)	7 (22)	3 (9)
4	Search assignment	27 (84)	4 (13)	1 (3)
5	Perform an iCBT[b] assignment	24 (75)	5 (16)	3 (9)
6	Read a text and recall the core message	29 (91)	2 (6)	1 (3)
7	Fill in a diary registration	14 (44)	7 (22)	11 (34)
8	Log out of platform	18 (56)	6 (19)	8 (25)

[a] n=29 (due to a program error during start-up, assignment 1 could not be completed as requested by 4 participants; the research leader helped them to get onto the platform).

[b] iCBT: internet-based cognitive behavioral therapy.

Usability Problems Encountered in the Performance Tasks

A summary of usability problems is provided in Table 3. Usability problems occurred on all skill levels, some being more general (eg, regarding operating the computer and internet browser) and others being more specific to the iCBT (eg, regarding navigating and orientation skills to find a specific button or assignment).

Concerning operating the computer and internet browser, problems using the keyboard were observed in each performance task in which the keyboard was needed (12/32), and problems with general browser button knowledge were also relatively common (15/32). These operational problems ranged from typing errors to not knowing how to use the backspace or using the address bar incorrectly (eg, participants entered their log-in data instead of the Web address of the treatment program in the address bar). In navigating and orientation skills, problems with locating or properly using buttons on the iCBT platform were most common (such as returning to the homepage in task 2, finding the email box in task 3, finding the diary button in task 7, or logging out in task 8; 24/32). Concerning using search strategies, two main problems were observed related to selecting the right module and the right assignment (task 2). Selecting an incorrect assignment (choosing an irrelevant assignment or choosing an assignment participants were not supposed to select) was the most prevalent problem (15/32). Finding the diary registration (task 7) appeared to be especially problematic (14/32). In evaluating relevance and reliability, problems occurred in understanding the meaning of the text. For example, many participants were not able to understand and recall the core message of a text (task 6, 11/32). In the skill adding personal content to the Web, problems occurred writing a text or message as part of an exercise or in communication with the therapist (task 3 and 5). This could be related to writing skills (eg, using punctuation; 16/32), formulating a message (eg, incorporating all relevant information; 10/32), or using the appropriate fields for each part of the information (eg, distinguishing recipient and subject fields in task 3; 11/32). Concerning skill protecting and respecting privacy, one specific problem was experienced by more than half of all participants: neglecting to click on the close button in the browser to protect the action when asked to log out of the iCBT platform in task 8 (18/32).

Table 3. Observed problems per performance task and number of participants encountering these problems (N=32). Participants could have encountered a problem multiple times, so the rows are not cumulative.

Assignment	1. Log in, n (%)	2. Navigate, n (%)	3. Write, n (%)	4. Search, n (%)	5. Perform, n (%)	6. Recall, n (%)	7. Use diary, n (%)	8. Log out, n (%)
Problems with operational skills								
Limited general browser button knowledge	15 (47)	—[a]	2 (6)	—	—	—	4 (13)	
Limited general keyboard knowledge	8 (25)	—	2 (6)	—	3 (9)	—	1 (3)	
Typing errors	7 (21)	—	—	—	1 (3)	—	—	
Inadequate slider use		—	—	—	—	—	7 (22)	
Difficulty reading from monitor	3 (9)	—	—	—	1 (3)	1 (3)	—	
Problems with navigation skills								
Locating/proper use of buttons on platform	—	9 (28)	16 (50)	—	—	—	4 (13)	13 (41)
Not recognizing homepage as such	—	5 (16)	—	—	—	—	—	—
Problems using search strategies								
Selecting incorrect assignment/not finding correct assignment	—	4 (13)	—	2 (6)	—	—	14 (44)	—
Selecting incorrect module	—	1 (3)	—	2 (6)	—	—	—	—
Problems evaluating relevance								
Not understanding core message/goal of assignment	—	—	—	—	8 (25)	11 (34)	—	—
Problems adding content								
Language error or incorrect spelling	—	—	16 (50)	—	—	—	—	—
Uncommon use of email etiquette	—	—	14 (44)	—	—	—	—	—
Incorrect use of text fields	—	—	4 (13)	—	7 (22)	—	—	—
Providing incomplete information	—	—	10 (31)	—	—	—	—	—
Problems ensuring privacy								
Unprotected log-out	—	—	—	—	—	—	—	18 (56)

[a] Not applicable.

Correlations Between Usability Variables and Age, Educational Level, and Digital Health Literacy

Table 4 shows the correlations between the number of completed performance tasks and number of encountered problems with age, educational level, and the sum scores on the DHLI, respectively. Age showed to be moderately related to both dependent variables with significant correlations of ρ=–.64 and ρ=.65, respectively, meaning that older age was associated with a lower number of completed tasks and a higher number of encountered problems. Scores on the DHLI were moderately correlated with the number of completed tasks (ρ=.46), meaning that higher scores on this instrument were related to a higher number of completed tasks. Scores on the DHLI did not correlate with the number of encountered problems, and educational level did not show any significant correlations either. Age, educational level, and scores on the DHLI did not correlate with each other (not shown).

Table 4. Spearman ρ correlations between usability variables and age, educational level, and digital health literacy (N=32).

Completion and encountered problems	Age ρ	*P* value	Educational level ρ	*P* value	DHLI[a] ρ	*P* value
Number of completed performance tasks	–.638	<.001	.292	.13	.456	.02
Number of encountered problems	.645	<.001	–.103	.58	–.114	.54

[a] DHLI: Digital Health Literacy Instrument.

Discussion

Principal Findings

In this study, we investigated the usability of an iCBT and its relationship with age, educational level, and digital health literacy of potential users. This provided insight into both what needs to be improved in our intervention and what type of people will need extra assistance in using it.

By observing the rate at which the assignments in the study could be completed and the type of problems that occurred, several conclusions can be drawn. First, some operational, navigational, and search tasks (eg, asking for the more practical digital skills) caused difficulties for many participants. Summarizing the major repeated problems, these occurred during logging in, logging out, finding the email box, and finding the diary registrations. A large proportion of the participants needed assistance to complete these tasks. Considering the high number of people who struggled with these tasks, it would be reasonable to conclude that these issues are inherent in the usability of the program. These matters should be addressed by changing button locations or making sure that these buttons stand out more.

A number of problems occurred that were more related to the practical digital skills of the participants and could not be attributed to the design of the platform. These nonplatform-related problems were rooted in either lack of general orientation competence (eg, locating keys on a keyboard or finding buttons on an internet browser) or navigation on a website (eg, understanding what a homepage is and how to navigate through its pages). For participants who encountered these type of problems, an introduction to the platform given by the therapist, other support staff, or instructional video could probably help them overcome the large majority of these struggles. Showing patients how to log in and log out, how the platform is built, and where to find all the parts of the intervention would help them in getting started. Descriptive analyses showed that many of the navigational problems were present during the first encounter with a new part of the website and were not repeated in the later assignments. This validates the conclusion that some level of support would suffice to overcome these problems (especially when combined with the suggested design alterations to the platform, which would increase the intuitiveness on the website).

Another substantial number of observed problems related to more complex digital health literacy skills such as being able to properly understand and use the content of the intervention and being able to express oneself in writing (both in terms of grammar and content). Problems in these areas are not easily overcome with basic support. This indicates that an online iCBT program would not be suitable to use as a stand-alone replacement to regular face-to-face therapy for people experiencing these struggles. The results show that a subgroup of people would not be able to grasp the key messages from the therapy and would therefore probably not be able to use it beneficially. Moreover, they would not be able to communicate this to their therapist or be able to ask for feedback or support to do so. For this group, the platform might be used as support in a regular treatment (eg, using diary registrations, relaxation exercises, or simple assignments and psychoeducational texts). A blended version of therapy with both online and face-to-face sessions could meet these needs. The face-to-face sessions should then incorporate the core elements of CBT and focus on getting the theory across by helping patients in understanding and practicing it. The online platform could offer support for the more practical parts of the treatment [7]. Additionally, to deliver more benefits of online therapy, it would be valuable to look further than the rather general approach to iCBT, which provides information and assignments in a text-based manner. The use of online tools greatly supports other means of communication, using video, voiceover, and even gaming elements. By offering more creative options, users with lower (health) literacy skills could benefit more from iCBT, and it would additionally make the interventions more attractive and pleasurable to use.

Which level of support is needed by whom remains an unanswered question. A subgroup of patients could follow the program independently, a subgroup would need practical assistance, and a subgroup would need a blended version in which the theoretical content is provided face-to-face. The data in this study support a cautious proposal for this divide. First, it is must be emphasized that for this study, only those people with internet experience who used the computer and internet on a regular basis were included. This has shown to be the first prerequisite to consider starting iCBT. Second, the associations found in this study provide a starting point for deciding in which manner the iCBT could be used independently. Although the sample size in this study was small and the analyses can therefore only be explorative, the associations can give helpful indications. Correlations show that age is a large factor to consider. Our findings show that elderly patients with internet experience are able to use the iCBT but experience significantly more struggles than younger people. This corresponds with previous studies by Van Deursen [13,21] that showed older people more often have lower operational and navigational skills in comparison with younger people. This discrepancy could be resolved by offering eligible elderly patients an introductory session to the platform to help them through the program for the first time. This could also serve the purpose of screening whether their digital skills are sufficient. Furthermore, correlations were shown between the number of completed tasks in our intervention and scores on the DHLI. This corresponds with previous research, which shows that digital health literacy is a predictor of critical internet use [22]. Nevertheless, measuring digital health literacy is a challenge due to the complexity of the concept. The DHLI, however, might be a usable screening instrument. This would foremost be relevant to check whether patients possess complex skills such as evaluating relevance of information and adding content to the Web. In this study, correlations were only shown between the DHLI and number of completed tasks, not with the number of problems that occurred. This might be explained by the usability problems attributed to the interface on the website. Many participants struggled with these problems, so the skill level of the participants is most likely not the only issue. It could be hypothesized that people with better digital health literacy (and thus a higher score on the DHLI) were able to overcome these problems better and still complete the task, while people with lower digital health literacy (a lower score on the DHLI) were not able to.

The lack of association between educational level and usability of the platform was quite remarkable. An explanation for this could be that the study sample consisted of an overrepresentation of older people. Among this subgroup, the distribution of educational level is often different, since more people had lower education levels in previous decades, not because of their capabilities but because of other practical reasons [23]. Nevertheless, in our sample, no relationship between age and educational level was found, which would indicate that educational level is simply not related to the usability of this iCBT, implying that online interventions can be feasible to use regardless of the educational level of a patient.

Limitations

This study has a number of limitations. First of all, it should be considered that this paper describes the usability of a specific intervention, Master Your Pain, and these results cannot be generalized to all iCBT interventions. This also accounts for the relationships found between age and digital health literacy. However, we do feel that the level of usability is generally characterized by a combination of the ease of use of an online tool and the skills of those who use it. Therefore, while our results are not simply generalizable, they do provide insight into what to take into account on a broader level regarding iCBT interventions and other online (health) tools. Principles of user experience design are key to take into account in the developmental phase of any digital tool. Second, the usability study was not conducted in a naturalistic setting but in a lab with one of the researchers directly present. This may have influenced the performance of the participants. Third, this study comprised a small sample. For a usability test, this sample was actually quite large, as smaller groups are more common [24]. Conducting this type of research is labor intensive and regularly qualitative in nature. However, to also conduct statistical analyses on the data, as was done in this study, a larger number of participants is needed to draw firm conclusions. Also, from technology-design theory, it is advised to perform a usability study in several rounds, with necessary design alterations in between rounds in order to evaluate if previous operational and navigational usability issues were overcome [24]. This could have helped in accurately determining which problems were related to the design of the platform and which problems were related to the digital health literacy skills of the participants. Finally, the sample that was used is not free of bias, since only people with some internet experience were included; it could also be suggested that only people with an interest in online therapy would consider joining this type of study. Nevertheless, the sample was rather representative for eligible users of a chronic pain iCBT, since these patients are generally somewhat older and females are overrepresented. Moreover, the recruitment was done in two very different geographical regions in the Netherlands. Also, people without any internet experience or interest would not be suitable for iCBT in general, so it seems appropriate not to include them in this study.

Conclusion

All in all, when developing a new internet intervention it is key to test its usability on all domains of the digital health literacy spectrum, including the intended target group. Our study strongly supports this notion and asks for specific attention for older individuals and people with a low level of health literacy. In addition to the practical design alterations, such a test will provide insight into the level of digital health literacy needed to benefit from the program and provide indications on what type of support is needed by whom. Regarding our own interventions, we continuously strive to improve upon them and we will solve the easy to adjust usability issues in the next release of Master Your Pain. The issues regarding skills of users will be addressed in our therapist training so that therapists are keen on offering new patients a user experience that fits their needs and skills.

Acknowledgments

We would like to thank all the patients who participated in this study. We would like to thank Ferry van Putten and Erica Souverein from the Fibrocentrum Leiden for their involvement in participant recruitment. We thank Jason Doppenberg for his language review.

Conflicts of Interest

RV and AE have been involved in the design of the Master Your Pain program.

Multimedia Appendix 1
Usability assignments.
[DOCX File , 15 KB - formative_v3i4e12883_app1.docx]

References

1. 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]

2. 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;16(3):e88 [FREE Full text] [doi: 10.2196/jmir.2777] [Medline: 24675372]
3. 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]
4. 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. [doi: 10.1016/j.invent.2018.02.001]
5. Drozd F, Vaskinn L, Bergsund HB, Haga SM, Slinning K, Bjørkli CA. The implementation of internet interventions for depression: a scoping review. J Med Internet Res 2016 Sep 08;18(9):e236 [FREE Full text] [doi: 10.2196/jmir.5670] [Medline: 27608548]
6. 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 09;17:27 [FREE Full text] [doi: 10.1186/s12875-016-0424-0] [Medline: 26961547]
7. Wentzel J, van der Vaart R, Bohlmeijer ET, van Gemert-Pijnon JE. Mixing online and face-to-face therapy: how to benefit from blended care in mental health care. JMIR Ment Health 2016;3(1):e9 [FREE Full text] [doi: 10.2196/mental.4534] [Medline: 26860537]
8. van Gemert-Pijnen JEWC, Nijland N, van Limburg M, Ossebaard HC, Kelders SM, Eysenbach G, et al. A holistic framework to improve the uptake and impact of eHealth technologies. J Med Internet Res 2011;13(4):e111 [FREE Full text] [doi: 10.2196/jmir.1672] [Medline: 22155738]
9. Wallin E, Norlund F, Olsson EMG, Burell G, Held C, Carlsson T. Treatment activity, user satisfaction, and experienced usability of internet-based cognitive behavioral therapy for adults with depression and anxiety after a myocardial infarction: mixed-methods study. J Med Internet Res 2018 Mar 16;20(3):e87 [FREE Full text] [doi: 10.2196/jmir.9690] [Medline: 29549067]
10. Maguire M. Methods to support human-centred design. Int J Hum Comput Stud 2001 Oct;55(4):587-634. [doi: 10.1006/ijhc.2001.0503]
11. Norman CD, Skinner HA. eHealth literacy: essential skills for consumer health in a networked world. J Med Internet Res 2006 Jun;8(2):e9 [FREE Full text] [doi: 10.2196/jmir.8.2.e9] [Medline: 16867972]
12. van der Vaart R, Drossaert CHC, de Heus M, Taal E, van de Laar MAFJ. Measuring actual eHealth literacy among patients with rheumatic diseases: a qualitative analysis of problems encountered using Health 1.0 and Health 2.0 applications. J Med Internet Res 2013 Feb;15(2):e27 [FREE Full text] [doi: 10.2196/jmir.2428] [Medline: 23399720]
13. van Deursen AJAM, van Dijk JAGM. Internet skills performance tests: are people ready for eHealth? J Med Internet Res 2011 Apr;13(2):e35 [FREE Full text] [doi: 10.2196/jmir.1581] [Medline: 21531690]
14. Terpstra JA, van der Vaart R, Spillekom-van Koulil S, van Dam A, Rosmalen JGM, Knoop H, et al. Becoming an eCoach: training therapists in online cognitive-behavioral therapy for chronic pain. Patient Educ Couns 2018 Dec;101(9):1702-1707. [doi: 10.1016/j.pec.2018.03.029] [Medline: 29661704]
15. Ferwerda M, van Beugen S, van Middendorp H, Spillekom-van Koulil S, Donders ART, Visser H, et al. A tailored-guided internet-based cognitive-behavioral intervention for patients with rheumatoid arthritis as an adjunct to standard rheumatological care. PAIN 2017;158(5):868-878. [doi: 10.1097/j.pain.0000000000000845]
16. van Beugen S, Ferwerda M, Spillekom-van Koulil S, Smit JV, Zeeuwen-Franssen ME, Kroft EB, et al. Tailored therapist-guided internet-based cognitive behavioral treatment for psoriasis: a randomized controlled trial. Psychother Psychosom 2016 Aug 11;85(5):297-307. [doi: 10.1159/000447267]
17. 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]
18. Ericsson KA, Simon HA. Verbal reports as data. Psychol Rev 1980;87(3):215-251. [doi: 10.1037//0033-295x.87.3.215]
19. Jaspers MWM. A comparison of usability methods for testing interactive health technologies: methodological aspects and empirical evidence. Int J Med Inform 2009 May;78(5):340-353. [doi: 10.1016/j.ijmedinf.2008.10.002] [Medline: 19046928]
20. Thomas DR. A general inductive approach for analyzing qualitative evaluation data. Am J Eval 2006 Jun 01;27(2):237-246. [doi: 10.1177/1098214005283748]
21. van Deursen AJ, van Dijk JA. Internet skill levels increase, but gaps widen: a longitudinal cross-sectional analysis (2010–2013) among the Dutch population. Inf Commun Soc 2014 Dec 24;18(7):782-797. [doi: 10.1080/1369118X.2014.994544] [Medline: 26281194]
22. Diviani N, van den Putte B, Giani S, van Weert JC. Low health literacy and evaluation of online health information: a systematic review of the literature. J Med Internet Res 2015;17(5):e112 [FREE Full text] [doi: 10.2196/jmir.4018] [Medline: 25953147]
23. Vanderploeg R, Schinka J. Estimation of premorbid cognitive abilities: issues and approaches. In: Ricker JH, editor. Differential Diagnosis in Adult Neuropsychological Assessment. New York: Springer; 2004:27-65.

24. Nielsen J. Usability Engineering. 2nd Edition. Burlington: Morgan Kaufmann Publishers; 1993.

Abbreviations

DHLI: Digital Health Literacy Instrument
iCBT: internet-based cognitive behavioral therapy

Edited by G Eysenbach; submitted 22.11.18; peer-reviewed by K Stasiak, I Bongers, C Duncan, M Stellefson; comments to author 31.03.19; revised version received 15.05.19; accepted 19.08.19; published 21.11.19.

Please cite as:
van der Vaart R, van Driel D, Pronk K, Paulussen S, te Boekhorst S, Rosmalen JGM, Evers AWM
The Role of Age, Education, and Digital Health Literacy in the Usability of Internet-Based Cognitive Behavioral Therapy for Chronic Pain: Mixed Methods Study
JMIR Form Res 2019;3(4):e12883
URL: http://formative.jmir.org/2019/4/e12883/
doi:10.2196/12883
PMID:31750839

Original Paper

A Patient Decision Aid App for Patients With Chronic Kidney Disease: Questionnaire Study

Signe Bülow Therkildsen[1], RN; Linda Houlind Hansen[2], RN; Laura Emilie Dinesen Jensen[3], RN; Jeanette Finderup[4,5], DPhil

[1]Department of Orthopedics, Oslo University Hospital, Oslo, Norway
[2]Department of Nephrology, Odense University Hospital, Odense, Denmark
[3]Centre for Eating Disorders, Aarhus University Hospital, Aarhus, Denmark
[4]Department of Renal Medicine, Aarhus University Hospital, Aarhus, Denmark
[5]Department of Clinical Medicine, Aarhus University, Aarhus, Denmark

Corresponding Author:
Jeanette Finderup, DPhil
Department of Clinical Medicine
Aarhus University
Palle Juul-Jensens Boulevard 99
Aarhus, 8200
Denmark
Phone: 45 78452525
Email: jeajee@rm.dk

Abstract

Background: The Dialysis Guide (DG) is a patient decision aid (PDA) available as an app and developed for mobile phones for patients with chronic kidney disease facing the decision about dialysis modality.

Objective: The aim of this study was to uncover the applicability of the DG as a PDA.

Methods: The respondents completed a questionnaire before and after using the DG. The respondents' decisional conflicts were examined using the Decisional Conflict Scale, and the usability of the app was examined using the System Usability Scale (SUS). The change in decisional conflict was determined with a paired *t* test.

Results: A total of 22 respondents participated and their mean age was 65.05 years; 20 out of 22 (90%) had attended a patient school for kidney disease, and 13 out of 22 (59%) had participated in a conversation about dialysis choice with a health professional. After using the DG, the respondents' decisional conflicts were reduced, though the reduction was not statistically significant (P=.49). The mean SUS score was 66.82 (SD 14.54), corresponding to low usability.

Conclusions: The DG did not significantly reduce decisional conflict, though the results indicate that it helped the respondents decide on dialysis modality. Attending a patient school and having a conversation about dialysis modality choice with a health professional is assumed to have had an impact on the decisional conflict before using the DG. The usability of the DG was not found to be sufficient, which might be caused by the respondents' average age. Thus, the applicability of the DG cannot be definitively determined.

(JMIR Form Res 2019;3(4):e13786) doi:10.2196/13786

KEYWORDS

mobile phone; app; patient decision aid; dialysis; decisional conflict; usability

Introduction

This paper sheds light on the applicability of a patient decision aid (PDA), the Dialysis Guide (DG) [1], which is available as an app for mobile phones and is made in accordance with the International Patient Decision Aid Standards (IPDAS) [2]. The DG is for patients with chronic kidney disease facing the choice of dialysis modality. The DG is a further development of a PDA in paper format, currently used at four hospitals in Denmark [3].

According to IPDAS, the purpose of a PDA is to improve the quality of decisions to enable patients to make informed, value-based decisions [4]. PDAs must make the decision explicit, contain information, and clarify advantages and disadvantages. The aim is to create agreement between the decision and individual values and preferences [5]. PDAs contribute to reduction of decisional conflicts and positively affect patients' basis for making a decision [6]. Some PDAs have been developed for patients with chronic kidney disease; however, studies of the effects of PDAs are few [7]. Thus, it is relevant to measure whether the DG reduces patients' decisional conflicts and helps them decide.

A systematic literature search was performed in the Cumulative Index of Nursing and Allied Health Literature (CINAHL) and PubMed databases to identify evaluated PDAs in app format; one PDA in app format for iPads was found [8]. Multiple studies regarding the assessment of online PDAs were also found. Grey literature searches in The Ottawa Hospital Research Institute's database [9] and the Facebook group *Shared@ Shared Decision-Making Network* supported this, and one more PDA in app format for mobile phones was found [10]. Only a few studies about PDAs in app format were found; none of them covered choice of dialysis modality.

Mobile apps have become easily accessible. In 2017, 84% of Danish families had a mobile phone [11]. Thus, it can be questioned why there are so few PDAs in app format. This is supported by the concept *telemedicine*, which, according to the World Health Organization, includes Web-based apps. The aim of telemedicine is to improve health results and provide clinical support through the use of information and communication technology across physical and geographical barriers [12]. The European Commission regards telemedicine as a solution to demographic changes with more elderly patients and, thus, more chronic diseases [13]. It is important to consider the advantages and disadvantages of apps [14]. The app format makes it easy to update PDAs and the material is always easily accessible. However, not all patients will necessarily benefit from PDAs in the app format. Moreover, patients with no access to technology, as well as the elderly with limited technological knowledge, may find apps less beneficial [15]. A PDA for dialysis choice was required by the Renal Association in Denmark; this was a request made by the patients.

The aim of this study was to examine whether the app, the DG, is applicable as a PDA for patients with chronic kidney disease to decide on dialysis modality. The following hypotheses were made:

> *Hypothesis 1: The DG reduces the patient's decisional conflicts.*
>
> *Hypothesis 2: The DG has a high level of usability.*

Methods

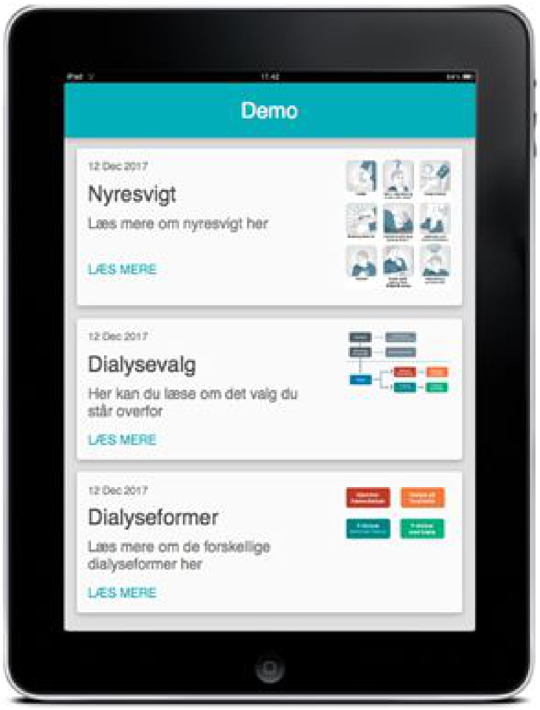
Study Design

Initially, we made a qualitative pilot study to adjust the DG. After this, a quantitative study design was used. This study design was developed as a hypothetico-deductive, cross-sectional study where data were collected through a pre- and postintervention questionnaire and used to examine the applicability of the DG. The method for assessing the DG was supported by the available literature [8,16-21].

The Dialysis Guide

The purpose of the DG is to support patients to clarify their values when choosing dialysis modality. The DG includes information and a test. The information focuses on kidney failure, dialysis choice, and dialysis modalities, and the app has a glossary (see Figure 1).

Figure 1. Screenshot of the Dialysis Guide index.



Demo

12 Dec 2017

Nyresvigt

Læs mere om nyresvigt her

LÆS MERE

12 Dec 2017

Dialysevalg

Her kan du læse om det valg du står overfor

LÆS MERE

12 Dec 2017

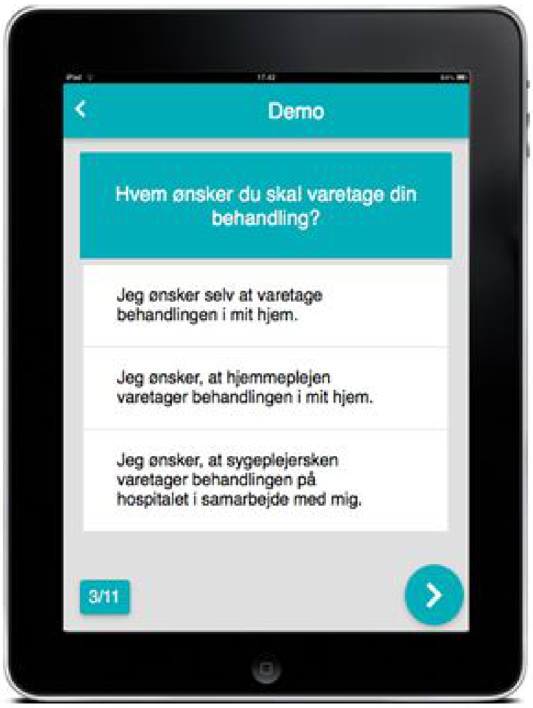
Dialyseformer

Læs mere om de forskellige dialyseformer her

LÆS MERE

The dialysis modalities to choose from include home hemodialysis, peritoneal dialysis, peritoneal dialysis with help, or hemodialysis at hospital. The test contains 11 questions to match the advantages and disadvantages of the different dialysis modalities to the patient's preferences for each dialysis modality (see Figure 2). The individual patient's test result is presented in the end.

Figure 2. Screenshot of the Dialysis Guide test.



Hvem ønsker du skal varetage din behandling?

Jeg ønsker selv at varetage behandlingen i mit hjem.

Jeg ønsker, at hjemmeplejen varetager behandlingen i mit hjem.

Jeg ønsker, at sygeplejersken varetager behandlingen på hospitalet i samarbejde med mig.

Technical Development

The DG is a responsive front-end Web app developed in HTML5. The front end communicates via http and JavaScript Object Notation (JSON) with a server-based back end, which is based on Drupal content management system (CMS), version 7.38 (Dries Buytaert). This is used to administer all the content published in the front-end app. The back end is executed via a Linux-based (The Linux Foundation) Apache Web server (The Apache Software Foundation) that runs a Hypertext Preprocessor (PHP), version 7.0.26 (The PHP Group), with two attached database servers that run the MySQL (Structured Query Language), version 5.7.20 (Oracle Corporation) [22].

Pilot Study

The pilot study was performed as two focus groups aimed at acquiring new knowledge [23]. These took place in September and October 2017 in two Danish regions; 6 and 11 patients participated, respectively. They were conducted by the second author in accordance with Kvale and Brinkmann [24]. Data analysis was conducted by all authors using systematic text condensation [25].

We found that the informants' ages and the severities of the disease were significant for a patient's approach to the DG. Some experienced that the DG was introduced too early in their disease trajectory. Multiple informants stated that apps were not for the elderly. Furthermore, several participants expressed a lack of time to try out the app. The DG information was found relevant and easy to understand. Constructive feedback in relation to the wording of some questions in the test was also given. In addition, a regional difference in the motivation for the use of the app was identified. The findings were used to adjust the DG as well as to construct the demographic questions for the evaluation of the app.

Selection of Respondents

People aged 18 years or older with chronic kidney disease and an estimated glomerular filtration rate (eGFR) between 10 and 20 mL/min/1.73 m^2 were included in the study. Patients unable to read and understand Danish, who did not use electronic devices, or with cognitive deficits were excluded. Likewise, patients who had previously received dialysis or who had already started dialysis were also excluded. Respondents were recruited by nurses from eight of 14 renal departments in Denmark between November 23, 2017, and June 30, 2018. The nurses informed the patients about the study and obtained patients' email addresses. The questionnaires were subsequently forwarded to the patients by email, including a link to the DG.

Decisional Conflict

To clarify whether the DG reduced patients' decisional conflicts, the validated 16-statement Decisional Conflict Scale (DCS) was used. The scale uses a 5-point Likert scale from 0 (*Strongly agree*) to 4 (*Strongly disagree*) [26,27]. The scale measures whether the PDA facilitates effective decision making as well as insecurities in relation to decision making [28]. SurveyXact (Rambøll) [29] was used for data collection. The preintervention questionnaire contained demographic questions (ie, age, gender, region, participation in a conversation regarding dialysis choice, and participation in a patient school) and measured the respondents' decisional conflicts using the DCS (DCS1). The respondents were recommended to use the DG for 20 minutes before they completed the postintervention questionnaire. Then, the respondents' decisional conflicts were measured again using the DCS (DSC2). The data analysis was performed using Excel (Microsoft Corporation) [30]. The means and SDs were calculated for the DCS. The results for DCS1 and DCS2 were compared using a paired *t* test. Data obtained from the DCS were rated using a total score between 0 and 100. In the DCS, a score of 0 is considered as no decisional conflict, a score of 100 is considered as an extremely high degree of decisional conflict, and a score lower than 25 is associated with making a decision. The subscores—*Uncertainty*, *Information*, *Values Clarity*, *Support*, and *Effective Decision*—were calculated [31].

Usability

To assess the usability of the DG, the System Usability Scale (SUS) was used. The SUS consists of 10 questions that also use a 5-point Likert scale from 1 (*Strongly disagree*) to 5 (*Strongly agree*) [32]. The scale was chosen as it was validated and developed to assess the usability of software [33].

Data were collected postintervention. Data obtained from the SUS were rated using a total score between 0 and 100, and means and SDs were calculated. In the SUS, a score of 0 is considered a low degree of usability and a score of 100 as a high degree of usability. A score above 68 is higher than the average SUS score and a score of 76 shows good usability [33].

The individual questions in the SUS were presented with the percentage of answers spread between the different choices on the Likert scale.

Ethical Considerations

The Central Denmark Region Committee on Health Research Ethics [34] and The Danish Data Protection Agency [35] have approved this study. All involved patients have given informed consent. Data were anonymized, safely stored, and shredded after use.

Results

Participant Flow and Characteristics

A total of 33 respondents were recruited; 28 (85%) of those completed the preintervention questionnaire and 22 (67%) also completed the postintervention questionnaire. Only the respondents who completed both questionnaires were included (22/33, 67%) (see Table 1). The mean age of the respondents was 65.06 years; 15 respondents were men (68%) and 7 were women (32%). Respondents from three out of five regions in Denmark were represented. A total of 20 out of 22 respondents (91%) had already attended a patient school, while 13 out of 22 (59%) had participated in a conversation about dialysis choice. Three electronic devices were used in the study: 12 out of 22 participants (55%) used a mobile phone, 5 out of 22 (23%) used an iPad or tablet, and 5 out of 22 (23%) used a computer.

Table 1. Demographic characteristics (N=22).

Characteristics	Value
Age (years), mean (range)	65.05 (44-86)
Gender, n (%)	
Female	7 (32)
Male	15 (68)
Region, n (%)	
North Denmark Region	0 (0)
Central Denmark Region	10 (45)
Region of Southern Denmark	0 (0)
Region Zealand	1 (5)
Capital Region of Denmark	11 (50)
Attended patient school, n (%)	
Yes	20 (91)
No	2 (9)
Attended a conversation about dialysis with a health practitioner, n (%)	
Yes	13 (59)
No	9 (41)
Electronic device used, n (%)	
Mobile phone	12 (55)
iPad or tablet	5 (23)
Computer	5 (23)

Decisional Conflict

A comparison of the results for DCS1 and DCS2 showed a difference in mean values and SDs (see Table 2). For DCS1, the total mean decisional conflict was 26.42 (SD 18.12) and for DCS2 the mean was 25.21 (SD 16.93). Thus, no significantly reduced decisional conflict was found (paired *t* test: difference=-1.21, *P*=.49). The remaining subscores were not significant either. The respondents' DCS1 scores for *Information* and *Support* were lower than 25, which was also the case for DCS2. The *Uncertainty* subscore was reduced the most with a fall of 3.41. Both the *Values Clarity* subscore and the *Support* subscore were higher after use of the DG.

The total mean for decisional conflicts was lower for respondents who had attended a conversation about dialysis choice before using the DG, while it was only reduced for those who had not attended one (see Table 3). The women's mean decisional conflict was lower compared to that of the men. The women's mean decisional conflict rose after having used the DG, whereas it was reduced among the men.

Table 2. Changes from preintervention Decisional Conflict Scale (DCS1) to postintervention Decisional Conflict Scale (DCS2) (N=22).

Score	DCS1, mean (SD)	DCS2, mean (SD)	Difference (paired *t* test)	*P* value
Total score	26.42 (18.12)	25.21 (16.93)	–1.21	.49
Uncertainty subscore	30.30 (24.87)	26.89 (19.57)	–3.41	.13
Informed subscore	22.73 (14.13)	21.97 (15.97)	–0.76	.67
Values Clarity subscore	25.38 (18.81)	28.03 (18.64)	2.65	.41
Support subscore	21.14 (17.45)	21.97 (18.10)	0.83	.83
Effective Decision subscore	29.26 (21.42)	26.70 (19.88)	–2.56	.29

Table 3. The meaning of demography for the Decisional Conflict Scale.

Demographic	DCS1[a] total score, mean (SD)	DCS2[b] total score, mean (SD)
Attended a conversation about dialysis with a health practitioner	20.67 (9.38)	21.75 (9.85)
Did not attend a conversation about dialysis with a health practitioner	34.72 (24.47)	31.60 (22.72)
Male	29.02 (15.91)	27.15 (13.71)
Female	20.76 (22.48)	20.98 (23.14)

[a] DCS1: preintervention Decisional Conflict Scale.

[b] DCS2: postintervention Decisional Conflict Scale.

Usability

The majority of the answers to the individual questions in the SUS ranged from 2 to 4 on the Likert scale. The mean for the overall SUS score was 66.82 (SD 14.54). The average SUS scores for the individual devices were 69.58 for mobile phones, 66.00 for iPads or tablets, and 61.00 for computers. See Table 4 for a summary of SUS response scores for each question.

Table 4. System Usability Scale (SUS) answers (N=22).

Question	Response, n (%)				
	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
I think that I would like to use the Dialysis Guide frequently	1 (5)	0 (0)	7 (32)	11 (50)	3 (14)
I found the Dialysis Guide unnecessarily complex	2 (9)	7 (32)	11 (50)	2 (9)	0 (0)
I thought the Dialysis Guide was easy to use	7 (32)	3 (14)	8 (36)	4 (18)	0 (0)
I think that I would need the support of a technical person to use the Dialysis Guide	1 (5)	0 (0)	5 (23)	14 (64)	2 (9)
I found the various functions in the Dialysis Guide were well integrated	2 (9)	10 (45)	8 (36)	1 (5)	1 (5)
I thought there was too much inconsistency in the Dialysis Guide	2 (9)	10 (45)	8 (36)	1 (5)	1 (5)
I would imagine that most people would learn to use the Dialysis Guide very quickly	0 (0)	2 (9)	4 (18)	13 (59)	3 (14)
I found the Dialysis Guide very cumbersome to use	5 (23)	10 (45)	7 (32)	0 (0)	0 (0)
I felt very confident using the Dialysis Guide	0 (0)	1 (5)	7 (32)	8 (36)	6 (27)
I needed to learn a lot of things before I could get going with the Dialysis Guide	4 (18)	5 (23)	8 (36)	5 (23)	0 (0)

Discussion

The reduction of the respondents' decisional conflicts as a result of using the DG was not significant. Thus, it is difficult to determine the extent to which the DG can help patients make a decision regarding the choice of dialysis modality. Therefore, it cannot be concluded that the DG meets the purpose of PDAs described by IPDAS, which is to improve the quality of decisions to facilitate informed and value-based decisions [3].

The usability of the DG is problematic because of the low number of respondents and their different answers in the SUS, with scores mainly around the middle of the Likert scale. Therefore, it cannot be confirmed that the DG has a high degree of usability. Likewise, it was difficult to shed light on the applicability of the app format for mobile phones, as all three types of devices were represented.

The results for the subscores in the DCS were not significant either, which might be due to the low number of respondents. However, the results indicated that the respondents received sufficient information, experienced an improved decision quality, and were less uncertain about their decision. In the evaluation of a similar PDA in paper format, the patients obtained an increased knowledge of the different dialysis modalities and became more prepared to make a decision [36]. This is in line with our study, as indications suggest that the DG improved respondents' levels of information as well as reduced their uncertainty.

The *Information* and *Support* subscores were below 25 in DCS1, which indicate that the patients had already received sufficient information and aid in making their decisions about dialysis modality before use of the DG [27]. Therefore, it is doubtful that these subscores could be reduced further through the use of the DG. However, a minor reduction of the *Information* subscore was seen. Already prior to using the DG, the respondents' decisional conflicts neared 25. This can be explained by the fact that 91% had participated in a patient school before using the DG. As the *Information* subscore in DCS1 was low, it is assumed that the respondents had received adequate information at the patient schools to make a decision. In other studies, it was also found that patients with chronic kidney disease who had already been educated on the subject had a lower degree of decisional conflict than patients who had not received any information [37,38]. One of the studies supports the notion that information at a patient school may have an impact on the applicability of a PDA. Whether or not patients had been exposed to a PDA—beyond being exposed to information—we found that the DG did not have an impact on patients' decisional conflicts. It was also found that decisional conflict only appeared among those respondents who had not yet had a conversation with a professional about dialysis choice [38]. This could indicate that the DG is mainly applicable for patients faced with the choice of dialysis modality, but who had not yet received information at a patient school or discussed dialysis choice with a professional.

Prior to use of the DG, the female respondents had a decisional conflict below 25. This indicates that they had already decided on a type of dialysis at this point [27]. On the other hand, the men had not decided on their choice before use of the DG, but their decisional conflict was reduced after its use. This indicates that the DG helped the men become more confident in making a decision. However, another study on significant factors for decisional conflict when choosing dialysis modality did not find that gender was a factor regarding decisional conflict [37].

The mean for the total SUS score was below the average SUS score, but the mean value when using a mobile phone was above this average score. Usability is considered highest when using a mobile phone, which is in contrast to a study evaluating a PDA for an iPad. Here, usability was higher on an iPad, as the size of the screen was bigger [8]. So far, there are only few

PDAs in app format for mobile phones. However, examples of patients finding PDAs for mobile phones usable and easy to navigate exist [10]. It is also presumed that a high degree of usability can be achieved for PDAs for mobile phones. However, the patients who used these had extensive knowledge of mobile phones, which might have impacted on the high usability [39,40]. This could also be the case in this study, as the respondents who chose to use a mobile phone would most likely choose it because they were familiar with this device.

The mean age for the respondents was 65.05 years, which is similar to the patient population's mean age in the Danish Nephrology Registry's Annual Report, 2016 [41]. It is presumed that the mean age may have an impact on the assessed usability of the DG, as the cognitive and physical abilities of older and younger participants differ. This may influence the older participants' use of Web-based telemedicine solutions [42]. In another study, researchers found that the usability of an online PDA was higher for patients under the age of 36 years [19].

The limitations of this study make it difficult to assess the applicability of the DG. The method of distribution entailed a weakness as it was not possible to make a dropout analysis. The number of respondents and the low response rate was also a weakness. The reason for the low number of respondents might be the lack of recruitment from some of the invited renal departments. Doubts about the validity of some of the email addresses might have impacted on the response rate. In other studies evaluating the use of PDAs in the same patient group, a response rate of 70% or lower was also seen [38,43]. One of these studies described that the nurses' recruitment of patients might be important to the outcome [36]. The nurses' approach to, and presentation of, the PDA may also have an impact on patients' views of the applicability [10]. Despite the inclusion and exclusion criteria, the nurses' subjective recruitment might have impacted on the results. The low number of respondents and low recruitment could mean that not all Danish regions were represented; therefore, it was not possible to assess whether the geographical location had an impact on the respondents' answers. Despite the recommended time for using the DG, not knowing how long the individual respondents used the app was a limitation.

On the basis of this study, it cannot be definitively concluded that the DG is applicable as a PDA for patients with chronic kidney disease deciding on dialysis modality. The DG did not reduce respondents' decisional conflicts significantly. Hence, Hypothesis 1 cannot be confirmed. However, a reduction of respondents' decisional conflicts was seen after use of the DG, but indications suggest that this is limited if the patient had received information at a patient school or attended a conversation about dialysis choice with a professional beforehand.

It can be concluded that the usability of the DG is not sufficiently clarified at present, which means that Hypothesis 2 is disproved. The low usability might be a result of the respondents' mean age. However, the usability was assessed as higher when using a mobile phone.

The limited knowledge about PDAs in app format and the number of respondents in this study mean that there is a need for further research to determine the applicability of the DG. It might be relevant to examine the difference between the applicability of PDAs in paper and app formats in a randomized controlled trial. If the DG is to be used in shared decision making, it is also relevant to examine the health professional's view of the applicability of the DG.

Conflicts of Interest

None declared.

References

1. The Dialysis Guide. URL: http://dialysetest.dk.linux69.unoeuro-server.com/www/ [accessed 2019-11-13]
2. International Patient Decision Aid Standards (IPDAS) Collaboration. URL: http://ipdas.ohri.ca/ [accessed 2017-11-13]
3. Finderup J, Jensen JKD, Lomborg K. Developing and pilot testing a shared decision-making intervention for dialysis choice. J Ren Care 2018 Sep;44(3):152-161. [doi: 10.1111/jorc.12241] [Medline: 29664179]
4. International Patient Decision Aid Standards (IPDAS) Collaboration. URL: http://ipdas.ohri.ca/what.html [accessed 2017-11-07]
5. Thompson R, Trevena L. Demystifying decision aids: A practical guide for clinicians. In: Elwyn G, Edwards A, Thompson R, editors. Shared Decision Making in Health Care: Achieving Evidence-Based Patient Choice. 3rd edition. Oxford, UK: Oxford University Press; 2016:51-56.
6. Stacey D, Légaré F, Lewis K, Barry MJ, Bennett CL, Eden KB, et al. Decision aids for people facing health treatment or screening decisions. Cochrane Database Syst Rev 2017 Apr 12;4:CD001431 [FREE Full text] [doi: 10.1002/14651858.CD001431.pub5] [Medline: 28402085]
7. Davis JL, Davison SN. Hard choices, better outcomes. Curr Opin Nephrol Hypertens 2017;26(3):205-213. [doi: 10.1097/mnh.0000000000000321]
8. Eden KB, Scariati P, Klein K, Watson L, Remiker M, Hribar M, et al. Mammography decision aid reduces decisional conflict for women in their forties considering screening. J Womens Health (Larchmt) 2015 Dec;24(12):1013-1020 [FREE Full text] [doi: 10.1089/jwh.2015.5256] [Medline: 26360918]
9. Ottawa Hospital Research Institute. Patient decision aids: Alphabetical list of decision aids by health topic URL: https://decisionaid.ohri.ca/AZlist.html [accessed 2017-09-20]

10. Korsbek L, Tønder ES. Momentum: A smartphone application to support shared decision making for people using mental health services. Psychiatr Rehabil J 2016 Jun;39(2):167-172. [doi: 10.1037/prj0000173] [Medline: 27030907]
11. Statistics Denmark. Home appliances URL: http://www.dst.dk/en/Statistik/emner/priser-og-forbrug/forbrug/elektronik-i-hjemmet [accessed 2019-11-13]
12. Telemedicine: Opportunities and Developments in Member States. Report on the Second Global Survey on eHealth. Geneva, Switzerland: World Health Organization; 2010. URL: https://www.who.int/goe/publications/goe_telemedicine_2010.pdf [accessed 2019-11-13]
13. Telemedicine for the Benefit of Patients, Healthcare Systems and Society. Brussels, Belgium: European Commission; 2009 Jun. URL: http://www.cupid-project.eu/sites/default/files/Staffworkingpaper-Telemedicineforthebenefitofpatientshealthcaresystemsandsociety.pdf [accessed 2017-11-09]
14. Mosa AS, Yoo I, Sheets L. A systematic review of healthcare applications for smartphones. BMC Med Inform Decis Mak 2012 Jul 10;12:67 [FREE Full text] [doi: 10.1186/1472-6947-12-67] [Medline: 22781312]
15. Abbasgholizadeh Rahimi S, Menear M, Robitaille H, Légaré F. Are mobile health applications useful for supporting shared decision making in diagnostic and treatment decisions? Glob Health Action 2017 Jun;10(sup3):1332259 [FREE Full text] [doi: 10.1080/16549716.2017.1332259] [Medline: 28838306]
16. Van Scoy LJ, Green MJ, Dimmock AE, Bascom R, Boehmer JP, Hensel JK, et al. High satisfaction and low decisional conflict with advance care planning among chronically ill patients with advanced chronic obstructive pulmonary disease or heart failure using an online decision aid: A pilot study. Chronic Illn 2016 Sep;12(3):227-235. [doi: 10.1177/1742395316633511] [Medline: 27055468]
17. Sivell S, Edwards A, Manstead AS, Reed MW, Caldon L, Collins K, BresDex Group. Increasing readiness to decide and strengthening behavioral intentions: Evaluating the impact of a Web-based patient decision aid for breast cancer treatment options (BresDex: www.bresdex.com). Patient Educ Couns 2012 Aug;88(2):209-217. [doi: 10.1016/j.pec.2012.03.012] [Medline: 22541508]
18. Hoffman AS, Llewellyn-Thomas HA, Tosteson AN, O'Connor AM, Volk RJ, Tomek IM, et al. Launching a virtual decision lab: Development and field-testing of a Web-based patient decision support research platform. BMC Med Inform Decis Mak 2014 Dec 12;14:112 [FREE Full text] [doi: 10.1186/s12911-014-0112-8] [Medline: 25495552]
19. Schackmann EA, Munoz DF, Mills MA, Plevritis SK, Kurian AW. Feasibility evaluation of an online tool to guide decisions for BRCA1/2 mutation carriers. Fam Cancer 2013 Mar;12(1):65-73 [FREE Full text] [doi: 10.1007/s10689-012-9577-8] [Medline: 23086584]
20. Wakefield CE, Watts KJ, Meiser B, Sansom-Daly U, Barratt A, Mann GJ, et al. Development and pilot testing of an online screening decision aid for men with a family history of prostate cancer. Patient Educ Couns 2011 Apr;83(1):64-72. [doi: 10.1016/j.pec.2010.05.025] [Medline: 20580521]
21. van Til JA, Drossaert CH, Renzenbrink GJ, Snoek GJ, Dijkstra E, Stiggelbout AM, et al. Feasibility of Web-based decision aids in neurological patients. J Telemed Telecare 2010;16(1):48-52. [doi: 10.1258/jtt.2009.001012] [Medline: 20086268]
22. MinMenu. URL: http://www.movesca.com/ [accessed 2019-11-23]
23. Halkier B. Planlægning af fokusgrupper. In: Kjølby A, Møllerhøj AB, editors. Fokusgrupper. 3rd edition. Frederiksberg C, Denmark: Samfundslitteratur; 2016:23.
24. Kvale S, Brinkmann S. Transcribing interviews. In: InterViews: Learning the Craft of Qualitative Research Interviewing. 3rd edition. Thousand Oaks, CA: SAGE Publications; 2015:203-214.
25. Malterud K. Systematic text condensation: A strategy for qualitative analysis. Scand J Public Health 2012 Dec;40(8):795-805. [doi: 10.1177/1403494812465030] [Medline: 23221918]
26. O'Connor AM. Validation of a decisional conflict scale. Med Decis Making 2016 Jul 02;15(1):25-30. [doi: 10.1177/0272989x9501500105]
27. O'Connor AM. Ottawa Hospital Research Institute. 1993. Decisional conflict scale URL: https://decisionaid.ohri.ca/docs/develop/Tools/DCS_Danish.pdf [accessed 2019-11-13]
28. Ottawa Hospital Research Institute. Patient decision aids: Decisional conflict scale URL: https://decisionaid.ohri.ca/eval_dcs.html [accessed 2019-11-13]
29. SurveyXact. URL: https://www.surveyxact.dk/ [accessed 2019-11-13]
30. Microsoft Office. Microsoft Excel URL: https://products.office.com/en-gb/excel [accessed 2019-11-13]
31. O'Connor AM. Ottawa Hospital Research Institute. 1993. User manual: Decisional conflict scale URL: https://decisionaid.ohri.ca/docs/develop/User_Manuals/UM_Decisional_Conflict.pdf [accessed 2019-11-13]
32. Brooke J. SUS: A Quick and Dirty Usability Scale. Reading, UK: Digital Equipment Corporation; 1986. URL: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.232.5526&rep=rep1&type=pdf [accessed 2019-11-13]
33. Sauro J. A Practical Guide to the System Usability Scale: Background, Benchmarks & Best Practices. Denver, CO: Measuring Usability; 2011.
34. National Videnskabetisk Komité. URL: http://www.nvk.dk/ [accessed 2019-11-13]
35. Datatilsynet. Om Datatilsynet URL: https://www.datatilsynet.dk/om-datatilsynet/ [accessed 2019-11-13]
36. Prichard A, Thomas N. The option grid: A shared decision-making tool for renal patients. J Ren Nurs 2013 Jan;5(1):6-11. [doi: 10.12968/jorn.2013.5.1.6]

37. Chen N, Lin Y, Liang S, Tung H, Tsay S, Wang T. Conflict when making decisions about dialysis modality. J Clin Nurs 2018 Jan;27(1-2):e138-e146. [doi: 10.1111/jocn.13890] [Medline: 28543737]
38. Winterbottom AE, Gavaruzzi T, Mooney A, Wilkie M, Davies SJ, Crane D, et al. Patient acceptability of the Yorkshire Dialysis Decision Aid (YoDDA) booklet: A prospective non-randomized comparison study across 6 predialysis services. Perit Dial Int 2016;36(4):374-381 [FREE Full text] [doi: 10.3747/pdi.2014.00274] [Medline: 26429419]
39. Wood J, Jenkins S, Putrino D, Mulrennan S, Morey S, Cecins N, et al. High usability of a smartphone application for reporting symptoms in adults with cystic fibrosis. J Telemed Telecare 2017 Aug 11;24(8):547-552. [doi: 10.1177/1357633x17723366]
40. Zia J, Schroeder J, Munson S, Fogarty J, Nguyen L, Barney P, et al. Feasibility and usability pilot study of a novel irritable bowel syndrome food and gastrointestinal symptom journal smartphone app. Clin Transl Gastroenterol 2016 Mar 03;7:e147 [FREE Full text] [doi: 10.1038/ctg.2016.9] [Medline: 26938478]
41. Danish Nephrology Registry (DNR): Nationwide Database for Patients With Chronic Renal Failure. Annual Report 2016. Copenhagen, Denmark: Danish Society of Nephrology; 2017. URL: http://nephrology.dk/wp-content/uploads/2017/08/%C3%85rsrapport-2016.pdf [accessed 2019-11-13]
42. Chun Y, Patterson P. A usability gap between older adults and younger adults on interface design of an Internet-based telemedicine system. Work 2012;41 Suppl 1:349-352. [doi: 10.3233/WOR-2012-0180-349] [Medline: 22316747]
43. Fortnum D, Grennan K, Smolonogov T. End-stage kidney disease patient evaluation of the Australian 'My Kidneys, My Choice' decision aid. Clin Kidney J 2015 Aug;8(4):469-475 [FREE Full text] [doi: 10.1093/ckj/sfv050] [Medline: 26251720]

Abbreviations

CINAHL: Cumulative Index of Nursing and Allied Health Literature
CMS: content management system
DCS: Decisional Conflict Scale
DCS1: preintervention Decisional Conflict Scale
DCS2: postintervention Decisional Conflict Scale
DG: Dialysis Guide
eGFR: estimated glomerular filtration rate
IPDAS: International Patient Decision Aid Standard
JSON: JavaScript Object Notation
PDA: patient decision aid
PHP: Hypertext Preprocessor
SQL: Structured Query Language
SUS: System Usability Scale

Edited by G Eysenbach; submitted 22.02.19; peer-reviewed by D Schatell, R Ciptaningtyas; comments to author 09.10.19; revised version received 22.10.19; accepted 31.10.19; published 21.11.19.

Please cite as:
Therkildsen SB, Hansen LH, Jensen LED, Finderup J
A Patient Decision Aid App for Patients With Chronic Kidney Disease: Questionnaire Study
JMIR Form Res 2019;3(4):e13786
URL: http://formative.jmir.org/2019/4/e13786/
doi:10.2196/13786
PMID:31750836

Original Paper

A Behavioral Activation Mobile Health App for Smokers With Depression: Development and Pilot Evaluation in a Single-Arm Trial

Jaimee L Heffner[1], PhD; Noreen L Watson[1], PhD; Edit Serfozo[1], MS; Kristin E Mull[1], MS; Laura MacPherson[2], PhD; Melissa Gasser[1,3], BA; Jonathan B Bricker[1,3], PhD

[1]Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, WA, United States

[2]Greenebaum Comprehensive Cancer Center, Baltimore, MD, United States

[3]University of Washington, Seattle, WA, United States

Corresponding Author:
Jaimee L Heffner, PhD
Public Health Sciences Division
Fred Hutchinson Cancer Research Center
1100 Fairview Ave N, M3-B232
PO Box 19024
Seattle, WA,
United States
Phone: 1 206 667 7314
Email: jheffner@fredhutch.org

Abstract

Background: The integration of Behavioral Activation Treatment for Depression (BAT-D) into smoking cessation interventions is a promising approach to address depression as a barrier to quitting. However, this approach has only been tested as a face-to-face intervention, which has low reach.

Objective: The aims of the study were to develop a BAT-D mobile health app with high potential reach and determine its feasibility, acceptability, and preliminary effects on theory-based behavioral processes of behavioral activation, reduced depressive symptoms, and smoking cessation.

Methods: Following a user-centered design process consisting of competitive analysis, focus groups, and prototype testing, we conducted a single-arm pilot trial of Actify!, a BAT-D app for depressed smokers. Participants used SmokefreeTXT along with Actify! to provide cessation content that had not yet been built into the app for this initial phase of pilot testing. Participants in the trial were current, daily smokers with mild to moderate depressive symptoms. We examined use outcomes for all enrolled participants and process and cessation outcomes at 6 weeks postenrollment for study completers (16/17, 94% retention).

Results: Regarding acceptability, average number of log-ins per participant was 16.6 (SD 13.7), and 63% (10/16) reported being satisfied overall with the app. Posttreatment interviews identified some usability challenges (eg, high perceived burden of planning and scheduling values-based activities). There was a significant decrease in depressive symptoms from baseline to follow-up (mean change in Patient Health Questionnaire–9 scores was –4.5, 95% CI –7.7 to –1.3; *P*=.01). Additionally, carbon monoxide (CO)-confirmed, 7-day point prevalence abstinence (PPA) at 6-week follow-up was 31% (5/16), and the 30-day PPA was 19% (3/16).

Conclusions: Results demonstrate promising engagement with Actify! and potential for impact on theory-based change processes and cessation outcomes. Preliminary quit rates compare favorably to previous trials of smoking cessation apps for the general population (ie, short-term, self-reported 30-day quit rates in the 8% to 18% range) and a previous trial of face-to-face BAT-D for depressed smokers (ie, CO-confirmed, 7-day PPA rate of 17% at end of treatment).

(JMIR Form Res 2019;3(4):e13728) doi:10.2196/13728

KEYWORDS

tobacco; nicotine; smoking cessation; depression; smartphone

Introduction

Cigarette smoking remains the leading preventable cause of death in the United States as well as the leading cause of cancer-related mortality, accounting for 32% of all cancer deaths [1,2]. Only 6% of smokers each year are able to successfully quit [3]. Novel interventions are needed to target the key factors that maintain smoking, including depressive symptoms, which are present in 40% to 55% of treatment-seeking smokers [4,5]. Depressive symptoms at the time of a quit attempt, including high negative affect (ie, aversive internal states such as anxiety or sadness) and low positive affect (ie, reduced experience of pleasure or enthusiasm), reduce the odds of smoking cessation by as much as 50% [4,6-8]. Failure to address depressive symptomatology as a barrier to quitting is a significant problem with the current standard treatment approach.

Behavioral Activation Treatment for Depression (BAT-D) is a promising approach for improving smoking cessation outcomes among smokers with depressive symptoms [9]. The primary goal of BAT-D is to increase rewarding and valued experiences in daily activities by planning and scheduling these activities to increase their frequency. These changes in daily activities are expected to increase positive affect and reduce negative affect and avoidance behaviors [10]. BAT-D provides an integrative, behavioral theory–based approach to address depressive symptoms in the context of a smoking cessation intervention, as reinforcement theory holds that both smoking and depressed behaviors are maintained by limited reward for alternative, healthy behaviors [11]. In support of the theory, previous work has demonstrated that increased involvement in rewarding, nonsmoking behaviors (eg, physical activity, playing games with family) is associated with improved depressive symptoms and successful quitting [12].

Adding BAT-D to standard cessation counseling (eg, Behavioral Activation Treatment for Smoking, or BATS) can improve quit rates. In a pilot randomized controlled trial (RCT), smokers with depressive symptoms were assigned to receive either group-delivered BATS (n=35) or standard cessation treatment (n=33), and both arms received transdermal nicotine patches [9]. Participants were assessed at baseline, throughout treatment, and at 4, 16, and 26-week follow-ups for depressive symptoms and carbon monoxide (CO)–verified 7-day PPA. At 4 weeks postquit date (ie, end of behavioral treatment), BATS participants were 2.1 times more likely to be abstinent (P=.02) than standard treatment participants (17.1% vs 9.1%). As expected, abstinence rates declined over time in both groups, yet there remained a significant effect of BATS treatment at the 26-week follow-up (14.3% vs 0%; P=.02). Participants receiving BATS also evidenced greater improvements in depressive symptoms across the 26-week follow-up period. Another recently published pilot study showed promising results of a BATS-based cessation treatment targeted at smokers recently hospitalized with acute coronary syndrome (Behavioral Activation Treatment for Cardiac Smokers [BAT-CS]), a population with a high prevalence of depression [13]. Participants received either the BAT-CS intervention (n=28) or an intervention using standard cessation counseling (standard of care, n=31). Results of this feasibility trial indicated that BAT-CS was highly acceptable to participants and also produced greater improvements in positive affect, negative affect, and stress than standard-of-care counseling at 24-week follow-up. Additionally, the effect size for BAT-CS on smoking cessation was potentially promising: odds ratio 1.27 (95% CI 0.41 to 3.93) for 7-day PPA at 24 weeks. The study was not powered to detect statistically significant effects, but cessation and mood outcomes were all in the hypothesized direction, favoring BAT-CS over standard of care.

At present, BAT-D–based smoking cessation interventions have solely been tested as a face-to-face treatment, which would reach only 4% to 6% of smokers in the United States based on current use of the traditional treatment modalities of individual and group counseling [3,14,15]. We developed a BAT-D mobile health (mHealth) app for smoking cessation to extend the reach of this promising intervention to the estimated 13 to 16 million smokers with smartphones who have depressive symptoms [8,15-17], the vast majority of whom will not seek face-to-face counseling. BAT-D is a promising treatment approach for a targeted mHealth smoking cessation intervention because it has shown promise helping smokers with depressive symptoms quit smoking in the face-to-face format [9,13], its components are well suited to translation into an app (eg, scheduling valued activities and tracking their completion), and it has a strong foundation in behavioral theory [11]. Indeed, other groups have begun developing and testing BAT-D apps for depression either to be used in conjunction with therapist support [18,19] or in self-guided format and designed for a primary care context [20,21]. With the most advanced evaluation of a BAT-D app to date being pilot RCTs not adequately powered to test efficacy [18,22,23] and the remainder of the published work limited to intervention descriptions or single-arm feasibility studies [19-21], there remains much to be learned about how to effectively translate BAT-D principles into mHealth interventions.

In this report, we describe the development process and preliminary evaluation of Actify!—a BAT-D smoking cessation mHealth app for depressed smokers. The aims of the study were to develop an mHealth app with great potential for high reach built upon BAT-D principles and determine its feasibility, acceptability, and preliminary effects on theory-based behavioral processes of behavioral activation, depressive symptomatology reduction, and the primary outcome of smoking cessation.

Methods

Development Overview

With the support of a pilot grant from the Fred Hutchinson Cancer Research Center (principal investigator JLH), we developed a working prototype of a smartphone app, Actify!, that delivers BAT-D as part of a smoking cessation intervention. The treatment frameworks guiding the app development were the BAT-D manual by Lejuez et al [10] for the treatment of depression [10] and the BATS integrative treatment for smoking cessation and depression by McPherson et al [9]. Key components of these treatments that were included in the app were (1) providing the treatment rationale for behavioral activation, (2) identifying values (ie, developing awareness of what matters most to each individual), (3) planning and scheduling values-based activities that are important or enjoyable as a means of increasing the frequency of these activities, and (4) monitoring progress at implementing values-based activities. Although they are included in the treatment manuals for face-to-face BAT-D, we decided that the initial version of Actify! would not include detailed, hour-by-hour monitoring of daily activities (including weeks where intervention activity is limited to self-monitoring) and formal social contracts for completing activities, as we learned from early user research that minimizing data entry burden (for daily monitoring) and privacy concerns (for social contracts) were critical design requirements. However, we did create content that prompts users, in a less structured manner, to consider how important or enjoyable their current daily activities are and reach out for social support as a means of overcoming barriers to completing activities.

The app was created following an iterative user-centered design process [24] including competitive analysis, focus groups, and usability testing of low- and high-fidelity prototypes followed by a single-arm pilot trial.

User-Centered Design Process

We first conducted a competitive analysis by reviewing a small number (5-6) of both paid and free apps and Web-based programs (eg, Mood Coach, Joyable) for depression as well as apps that include elements similar to the components of behavioral activation, such as scheduling and tracking activities or apps for setting goals and marking them complete (eg, Iwish, Habitica). The purpose of a competitive analysis is to review other programs currently on the market to generate (via the personal experience of the design team) and assess (via user comments in app stores) potential user sentiment toward them in relation to their design and functionality. Such information is useful as part of the ideation phase of development to understand desirable features for a new app as well as pain points (ie, problems that reduce the quality of the user's experience). Examples of desirable features identified through the competitive analysis to reduce pain points were making valued activities customizable and offering menus of options to keep text entry fields to a minimum and reduce user burden.

Two focus groups of 4 and 6 persons were conducted in November 2016 to assess potential users' previous experiences using depression apps (eg, likes and dislikes); develop user personas, fictional characters based on typical users, a common user-centered design strategy [25]; elicit ideas for valued activities in several life domains (eg, relationships, health, leisure) the app could suggest to the user; and understand users' associations between emotions and possible color palettes that might appear in the app. User personas were developed in order to guide design decisions (Multimedia Appendix 1), and lists of common values (eg, having close relationships, being a good citizen/community member) and valued activities (eg, helping others, visiting friends and family) were generated for inclusion in the app. Focus group data were then used to generate two distinct prototypes of a BAT-D app. We used prototyping software to allow user research participants to interact with these prototypes on a smartphone in order to provide feedback on design and functionality. Individual interviews were conducted to evaluate each prototype (7 for prototype A and 6 for prototype B).

Based on users' reactions to these low-fidelity prototypes and general satisfaction with the interface concept of prototype B (vs more mixed feedback about prototype A), the study team decided to continue developing that prototype for the first version of Actify! The user experience design team then conducted three iterative rounds of testing between January and March 2017 with higher fidelity prototypes. In each round, users were given 4 to 5 tasks to complete. Through these tasks, the study team tested users' ability to navigate the major app components of onboarding, values identification, activity scheduling, activity tracking, calendar, progress, and resources. These rounds of testing also included assessments of users' expectations and desires for app functionality. Refinements to the app at this phase focused on more fine-grained details such as text size and readability, size of visual elements (eg, buttons), and clarity of instructions included in the app onboarding screens. Development of the programmed version of the app was completed by the study team's programming partner, 2Morrow, in July 2017. For evaluation purposes, Actify! was developed as a native app for Apple's iOS operating system. We chose a native app (as opposed to a Web app) for two main reasons: the option of potentially including features restricted to native apps (eg, interactive push notifications) and better responsiveness of native apps compared with Web apps (ie, less wait time between user action and app response). Supporting our choice to develop for iOS, a recent study demonstrated that the majority (77%) of individuals who download smoking cessation apps download them on iOS devices as opposed to Android (23%) [26]. However, we plan to develop later versions of Actify! so that it is accessible on multiple platforms.

Description of Actify!

Guided by the BAT-D treatment model of Lejuez and colleagues [10] as well as the BATS smoking cessation protocol developed by author LM and found to show promise in a pilot RCT [9], the core functions of the current version of Actify! are as follows:

- Treatment rationale: introducing the user to the program by describing how behavioral activation works and providing instructions on how to use the program (Figure 1)
- Values identification: helping the user identify what matters most to them (eg, family, fun) within six core life areas—relationships, health/self-care, spirituality, leisure, career and education, and daily responsibilities (Figures 2 and 3)
- Activity scheduling: planning and scheduling of values-based activities that are important or enjoyable (eg, reading to my children) and organized by life area (Figures 4 and 5), including life area-specific activity suggestions (ie, a bank of 10 to 20 suggestions for each) and the ability to set an alert for each scheduled activity
- Progress monitoring: tracking completed activities and mood (Figure 6) and visualizing overall progress toward

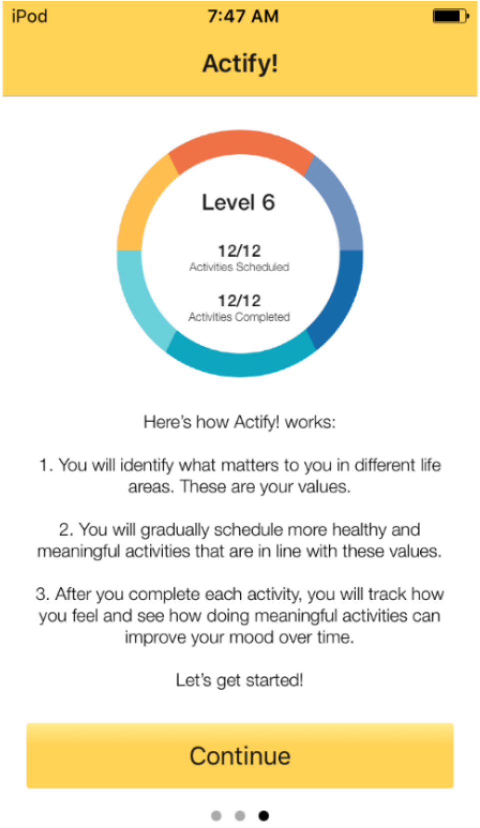
meeting activity goals in calendar form (Figure 7) and across time, broken down by life area (Figure 8)

The app also contains a Resource section with Frequently Asked Questions (eg, “What if I increase my activity and still don’t feel better?”) as well as a Depression Help section (eg, “What is depression?” and “What if I am thinking of harming myself?”).

To increase engagement, the app has a game-like structure [27] with a level system that, over time, encourages escalating activity goals and increasing breadth of life areas represented in the chosen activities. For the pilot trial, 6 levels were available. Level 1 encourages users to complete 2 activities in one life area. By Level 6, the user has been prompted to identify values in all 6 life areas and must complete a total of 12 activities to finish the level. Each level begins with a user story of a person who improved their mood and quit smoking using the program. User stories illustrating how values can be translated into actions were associated with cessation in our prior work [28]. We also used the content of user stories to address common barriers to implementing behavioral activation, such as setting goals that aren’t realistic, having difficulty reaching out to others for support, and feeling overwhelmed and having difficulty getting started. An example user story is that of Elisa: “A clean home makes me feel proud and like I have self-respect. But feeling depressed, I let my house get really messy. I felt overwhelmed by all of the cleaning I needed to do. I put washing dishes in my schedule the first night. I felt so much better afterwards and started scheduling more things to clean.” Actify! also includes motivational messages that appear each time the user opens the home screen (eg, “Having trouble getting started on activities? Try asking a friend to join you”) and notifications that appear in response to certain user action or inaction (eg, a suggestion to try new activities if the user enters three consecutive low mood ratings; a suggestion to start with small activities if three consecutive scheduled activities are not tracked as being completed; a notification to suggest scheduling an activity after three consecutive days of no activity scheduling or tracking).

Because the other components to support smoking cessation were not included in this early development phase due to budgetary constraints, participants used the SmokefreeTXT text messaging program from the US National Cancer Institute (NCI) alongside Actify! to provide content similar to what will be included in the later versions of the app. SmokefreeTXT was developed by NCI’s Smokefree.gov development team and includes 6 weeks of daily messages delivered 2 to 4 times per day, with more frequent messaging on and around the quit date. Messages are designed to prepare smokers for their quit day and support cessation maintenance. Some messages are interactive, prompting users to respond to a query (eg, “To deal with cravings: breathe in, hold for 5 seconds, breathe out, and repeat. What is your current craving level? Reply: HI, MED, LOW”). Programmed messages provide tailored feedback based on the user’s responses to these queries.

Figure 1. Onboarding.



Here’s how Actify! works:

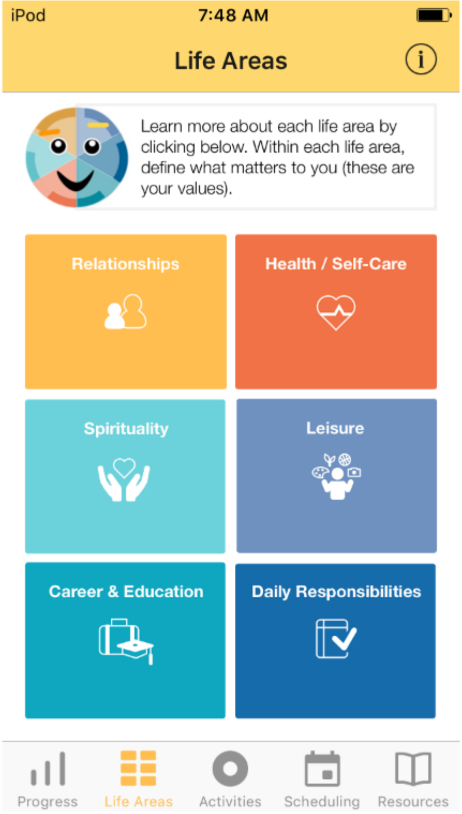
1. You will identify what matters to you in different life areas. These are your values.

2. You will gradually schedule more healthy and meaningful activities that are in line with these values.

3. After you complete each activity, you will track how you feel and see how doing meaningful activities can improve your mood over time.

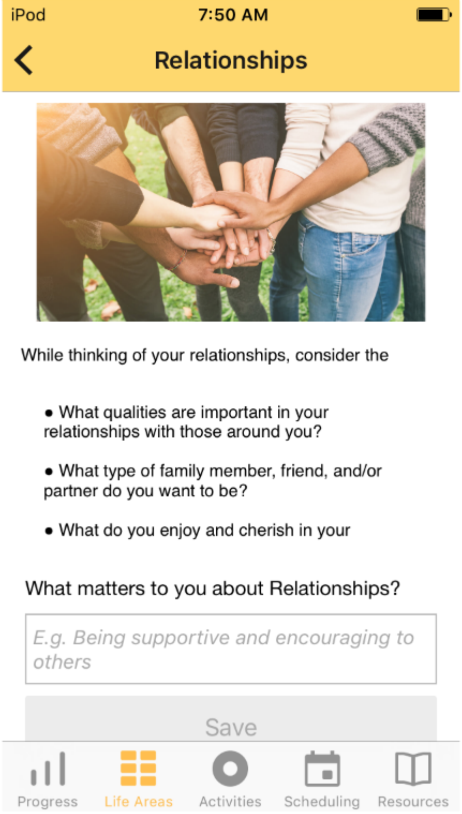
Let’s get started!

Figure 2. Life areas.



Learn more about each life area by clicking below. Within each life area, define what matters to you (these are your values).

Figure 3. Values identification within life areas.



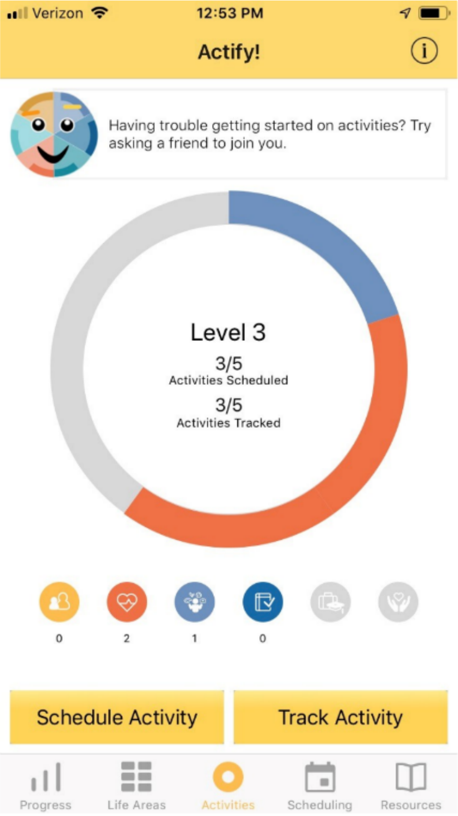
While thinking of your relationships, consider the

- What qualities are important in your relationships with those around you?
- What type of family member, friend, and/or partner do you want to be?
- What do you enjoy and cherish in your

What matters to you about Relationships?

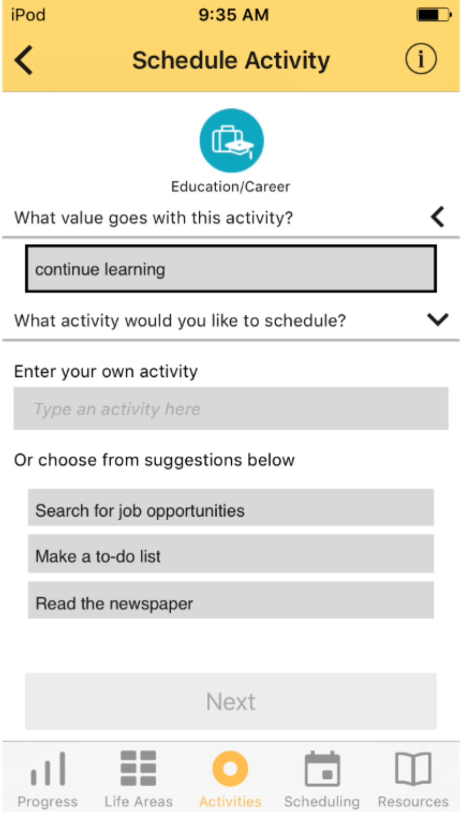
E.g. Being supportive and encouraging to others

Figure 4. Activity wheel.



Having trouble getting started on activities? Try asking a friend to join you.

Figure 5. Scheduling.



What value goes with this activity?

continue learning

What activity would you like to schedule?

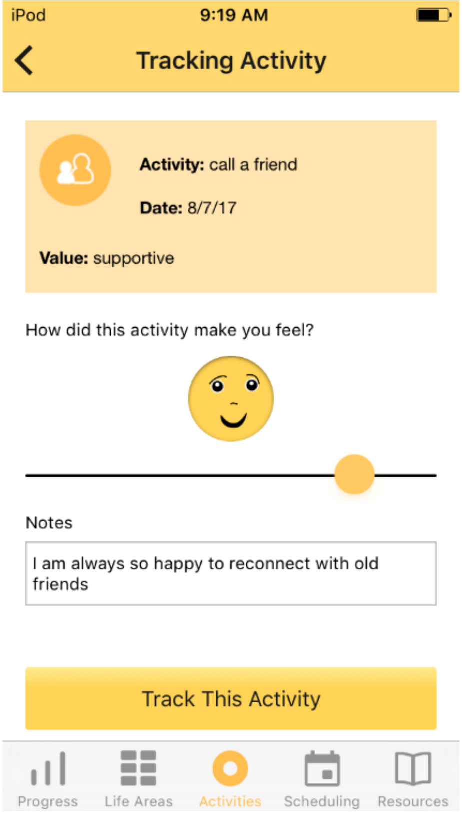
Enter your own activity

Type an activity here

Or choose from suggestions below

- Search for job opportunities
- Make a to-do list
- Read the newspaper

Figure 6. Tracking.



iPod 9:19 AM

Tracking Activity

Activity: call a friend

Date: 8/7/17

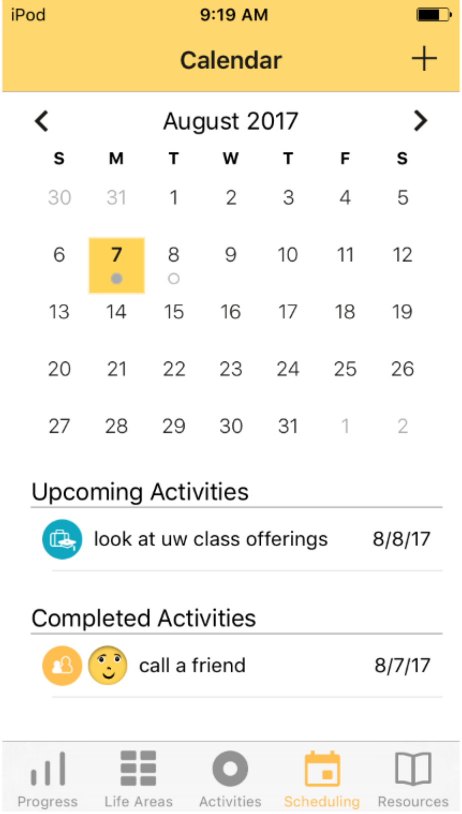
Value: supportive

How did this activity make you feel?

Notes

I am always so happy to reconnect with old friends

Figure 7. Calendar.



iPod 9:19 AM

Calendar

August 2017

S	M	T	W	T	F	S
30	31	1	2	3	4	5
6	7	8	9	10	11	12
13	14	15	16	17	18	19
20	21	22	23	24	25	26
27	28	29	30	31	1	2

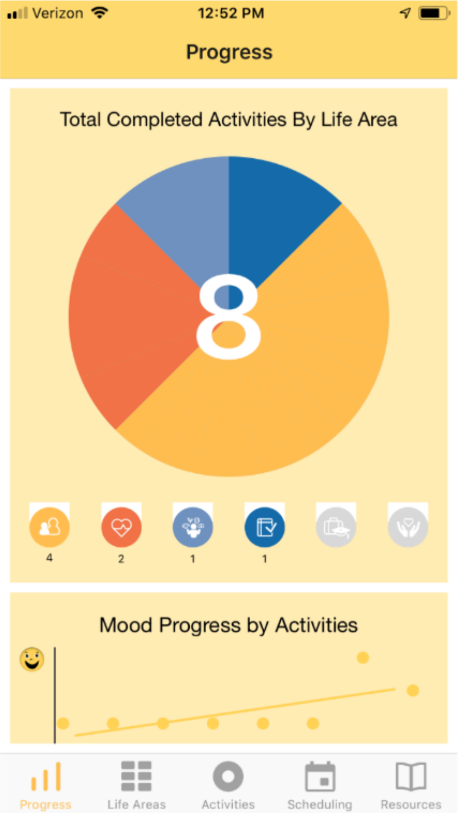
Upcoming Activities

look at uw class offerings 8/8/17

Completed Activities

call a friend 8/7/17

Figure 8. Progress.



Pilot Trial Procedures

Trial Overview

Between November 2017 and June 2018 we completed a single-arm pilot trial of the initial programmed version of Actify! with functionality limited to the BAT-D component. Primary aims were to obtain preliminary data on user engagement and satisfaction and evaluate effects on the target behaviors of increased engagement in valued activities, smoking cessation, and improved mood. To achieve these aims, we conducted a 6-week evaluation to assess engagement and short-term efficacy outcomes and obtain detailed user feedback about the intervention.

Recruitment and Informed Consent

Participants were recruited from the Seattle area primarily via no-cost Craigslist ads (53% of sample) and paid Facebook advertisements (41% of sample). Interested potential participants were screened for participation using a 2-stage process. First they completed a Web-based preliminary screening form that assessed only the nonsensitive eligibility criteria (ie, all selection criteria except questions about mental health). Those who were potentially eligible based on the initial screening were asked for contact information to follow up with a phone call from trained research staff. During this call, the purpose and procedures of the study were explained in further detail and remainder of the eligibility criteria (ie, the mental health–related criteria) were assessed. Those who remained interested and eligible were invited to participate in the study; those who were not eligible were given referral information for smoking cessation assistance. Eligible participants were invited to attend an initial face-to-face meeting, during which the full details of the study were explained and participants provided informed consent. Participants were instructed to download the Actify! app during the initial meeting and enroll in the SmokefreeTXT messaging program. They were encouraged to use the program as much as possible but told that the frequency and duration of use was ultimately their choice. The study was reviewed and approved by the Institutional Review Board of the Fred Hutchinson Cancer Research Center.

Participants

Inclusion criteria for the single-arm pilot trial were (1) age 18 or older, (2) current smoker, averaging at least 5 cigarettes per day for the last 30 days, (3) interested in quitting smoking in the next 30 days, (4) own an iPhone version 5 or higher (necessary to run the current version of Actify! without technical problems), (5) have experience downloading and using one or more apps on their iPhone, and (6) screen positive for mild to moderate current depressive symptoms (9-item Patient Health Questionnaire [PHQ-9] score 5-19 [29]). Exclusion criteria were (1) currently using a depression app or receiving other treatment for depression, including psychotherapy or medication as this would be a significant source of bias in a single-arm assessment of intervention effects, (2) mood symptoms too severe to be addressed with a self-guided app (ie, severe depression [PHQ-9>20] or current suicidal ideation [PHQ-9 item 9 score >1]), (3) receiving other treatment for smoking cessation, or (4) previous use of the SmokefreeTXT program.

Measures

Demographic items assessed at baseline included age, gender, education, employment, income, and marital status. The 6-item Fagerström Test for Nicotine Dependence (FTND) [30] assessed degree of physical dependence on nicotine. The Patient Health Questionnaire–9 (PHQ-9) is a widely used assessment of depressive symptoms that has demonstrated ability to detect symptom changes over time as a result of treatment [29]. The scale scores range from 1 to 27, with severity thresholds as follows: 5-9=mild, 10-14=moderate, 15-19=moderately severe, 20-27=severe. A score of 10 has 88% sensitivity and 88% specificity for a diagnosis of major depression [31]. The smoking timeline follow-back (TLFB) [32] is a calendar-based method of obtaining retrospective estimates of daily smoking. For establishing baseline smoking levels, the 30-day period prior to the baseline visit was used.

Treatment acceptability was assessed at the 6-week follow-up visit using measures of treatment use and satisfaction. The primary outcome of treatment use was defined as the number of times Actify! was opened on the phone, determined objectively via time-stamped, server-recorded page views. App page views separated by 10 minutes or more were counted as unique log-ins. Treatment satisfaction was assessed with 8 study-specific items on the 6-week outcome survey. Example items include “Overall, how satisfied are you with the Actify! app?” and “Overall, how useful was the app for helping you engage more in activities that are important to you?” Follow-up questions probed satisfaction with specific components of the program (eg, app organization, ease of use, and usefulness of text messages).

Efficacy for increasing behavioral activation was assessed using the 25-item Behavioral Activation for Depression Scale (BADS) [33], which is designed to detect changes in activation resulting from treatment. Scores for individual items range from 0 to 6, with total scores ranging from 0 to 150. Scores on the 7-item activation subscale range from 0 to 42. Efficacy for reducing depression symptoms was assessed via change in PHQ-9 scores. Scores for individual items range from 0 to 3, with total scores ranging from 0 to 27. Efficacy for smoking cessation was CO-confirmed 7-day PPA at 6 weeks postenrollment. The secondary efficacy end point was 30-day PPA at 6 weeks postenrollment. Smoking outcomes were assessed using the smoking TLFB [32]. Expired-air CO levels were taken at baseline and at the follow-up visit as a means of biochemically verifying smoking self-report. Consistent with the recommendations of the Society for Research on Nicotine and Tobacco Workgroup on Biochemical Verification [34], a CO cutoff of ≤8 ppm was used as a basis for confirming self-reported abstinence.

Adverse Events

Any adverse events volunteered by participants at the 6-week follow-up visit were recorded.

User Experience Interview

We collected qualitative data on participant experiences with the intervention via a semistructured interview (approximately 45 minutes) conducted at the 6-week follow-up visit. Objectives were to evaluate the (1) context of use, including time of day the app was used, what prompted the user to open the app, and how the user tended to interact with the app; (2) satisfaction with the user interface and app design, including the aesthetics, ease of use, and specific BAT-D components such as activity scheduling and tracking; and, (3) perceived effects of the app on behavior and mood. All interviews were completed by the first author, and detailed notes were recorded.

Follow-Up Procedures and Compensation

The 6-week follow-up was an in-person visit of approximately 1 hour that was scheduled at the time of the baseline visit. Participants received reminder calls and/or emails (based on individual preference) to increase adherence to their scheduled appointments. Compensation for completing the follow-up visit was higher than for the baseline visit ($70 vs $30) due to the greater length of the follow-up visit.

Statistical Analysis Plan

Consistent with the developmental nature of the study, statistical analyses were primarily descriptive. Changes in behavioral activation and depression symptoms from baseline to 6-week follow-up were evaluated using paired *t* tests, with *P*<.05 as the threshold for statistical significance.

Results

Recruitment, Retention, and Sample Description

We prescreened 250 individuals via Web-based survey to identify 177 potentially eligible participants. We were able to reach 74 of these individuals by phone to complete additional screening and enrolled 17 eligible participants in the single-arm trial. The most common reasons for ineligibility on the Web-based survey were not providing a name or contact information (73/250), not owning an iPhone 5 or higher (32/250), and smoking fewer than 5 cigarettes per day (20/250). On the phone screen, the most common reasons for ineligibility were inability to attend in-person visits (15/74, of which 14/74 were due to a Craigslist technical error that resulted in the ad being posted in multiple cities), currently receiving other treatment for depression (13/74), and not owning an iPhone 5 or higher (7/74).

Of the enrolled participants, 94% (16/17) completed the in-person 6-week follow-up visit and are included in the subsequent complete-case results. These participants were 50% (8/16) men, 38% (6/16) racial minority (two or more races [4/16], Asian [1/16], or other [1/16]), and 25% (4/16) Hispanic, with a mean age of 36.1 (SD 9.7) years. The majority identified as heterosexual (14/16, 88%), never married (10/16, 63%), employed (12/16, 75%), and had a bachelor’s degree or higher level of educational attainment (11/16, 69%). Average smoking at baseline was 6.5 (SD 3.8) cigarettes per day, and average FTND score was 2.6 (SD 2.1), indicative of low nicotine dependence. Severity of depressive symptoms at baseline was moderate, on average PHQ-9 10.2 (SD 5.0).

Engagement and Satisfaction

As shown in Table 1, the average number of app openings per participant was 16.6 (SD 13.7) over the 6-week study period, with a minimum of 6, a maximum of 55, and a median of 13.

Table 1. Actify! use metrics for enrolled participants (n=17).

Characteristic	App use			Retention[a], %	Rolling retention[b], %
	Mean (SD)	Median	Range		
Overall use					
Number of log-ins	16.6 (13.7)	13	6-55	—[c]	—
Time per session (minutes)	4.2 (2.3)	4.5	0.7-8.5	—	—
Unique number of days of use	10.9 (8.8)	8	4-37	—	—
Log-ins, week 1	5.7 (3.2)	5	1-12	100	100
Log-ins, week 2	2.4 (3.0)	1	0-9	71	94
Log-ins, week 3	1.2 (1.3)	1	0-4	59	88
Log-ins, week 4	1.6 (2.6)	0	0-9	47	65
Log-ins, week 5	1.0 (2.4)	0	0-10	41	53
Log-ins, week 6	0.8 (1.8)	0	0-7	35	35
App component use					
Progress screen viewing	13.5 (14.2)	9	1-57	—	—
Resources section viewing	2.5 (3.0)	1	0-8	—	—
Calendar viewing	13.9 (14.3)	8	1-54	—	—
Life areas added (maximum possible: 6)	4.8 (1.6)	5	0-6	—	—
Activity scheduling	18.2 (23.9)	10	0-93	—	—
Activity tracking	15.4 (23.3)	9	0-93	—	—

[a] Retention: proportion of users who logged in at least once that week.

[b] Rolling retention: proportion of users who logged at least once that week or in subsequent weeks.

[c] Not applicable.

The average numbers of times users accessed major components of the app are listed in Table 1 along with the medians and ranges for each metric. As shown in the table, all areas of the app were accessed at least once, on average, and the most-accessed features were activity scheduling (mean 18.2 [SD 23.9], median 10) and activity tracking (mean 15.4 [SD 23.3], median 9). Week-by-week retention (ie, proportion of users accessing the app at least once that week) and rolling retention (ie, proportion of users accessing the app at least once that week or in a subsequent week) are also provided in Table 1. As 30-day rolling retention is a commonly reported metric among app developers, the most comparable metrics for Actify! are a rolling retention of 65% (11/17) for week 4 and 53% (9/17) for week 5.

On the measure of treatment satisfaction, 63% (10/16) reported being satisfied overall, 63% (10/16) thought that the app was useful for increasing important activities, 63% (10/16) thought the app was organized, and 75% (12/16) found the accompanying text messages helpful.

Efficacy for Smoking Cessation

As shown in Table 2, the CO-confirmed 7-day PPA rate at 6-week follow-up was 31% (5/16), and the 30-day PPA at 6-week follow-up was 19% (3/16). In a post hoc sensitivity analysis using a conservative cutoff for CO levels indicative of abstinence (3 ppm or less [35]), the findings were unchanged.

Efficacy for Decreasing Depressive Symptoms and Increasing Behavioral Activation

Average change in depression on the PHQ-9 was a 5-point decrease from baseline (mean –4.5, 95% CI –7.7 to –1.3), which was a statistically significant improvement in symptoms (*P*=.01) as well as a clinically significant change from moderate symptoms to mild symptoms (Table 2). Average change in BADS overall scale scores and activation subscale scores (BAT-D’s theory-based change mechanism) were not statistically significant but were in the hypothesized direction: BADS total scores increased by 14 points (mean 14.4, 95% CI –2.6 to 31.3; *P*=.09) and activation subscale scores increased by 4 points (mean 3.6, 95% CI –1.7 to 8.9; *P*=.17; Table 2).

Table 2. Actify! cessation and process outcomes for the complete-case sample (n=16).

Variable	Value	*P* value
Biochemically confirmed cessation at end of treatment		
7-day PPA[a], complete (n=16), n (%)	5 (31)	—[b]
30-day PPA, complete case, n (%)	3 (19)	—
Mechanisms of change		
Change in BADS[c] activation subscale, mean (95% CI)	3.6 (–1.7 to 8.9)	.17
Change in BADS total, mean (95% CI)	14.4 (–2.6 to 31.3)	.09
Change in PHQ-9[d], mean (95% CI)	–4.5 (–7.7 to –1.3)	.009

[a] PPA: point prevalence abstinence.

[b] Not applicable.

[c] BADS: Behavioral Activation for Depression Scale.

[d] PHQ-9: 9-item Patient Health Questionnaire.

Adverse Events

No adverse events were reported during the 6-week evaluation period.

User Experience Interview Findings

Regarding context of use, participants tended to interact with the app at home, either in the morning or the evening. They liked the concept of the app (ie, thinking about what’s most important to them and scheduling activities in line with those values) but indicated that implementation could be improved. The vast majority of user-identified pain points were in setting up the hierarchy of life areas, values, and activities to be scheduled. This hierarchical process, which was integrated into the scheduling of each activity (Figure 5), was perceived as overly burdensome and sometimes confusing to use. Participants had considerable difficulty distinguishing values from activities and therefore became confused when prompted to identify a value and select an activity that goes with that value. Some participants indicated that they scheduled and tracked activities at the time the activity was completed rather than prescheduling and tracking later, which was the intent of the intervention. They also wanted more activity suggestions (beyond the 3 they were offered each time an activity was scheduled) and the option to schedule repeat activities and/or select from a group of saved activities. Because of the data entry burden, some participants reported that they stopped using the app and switched to low-tech methods of scheduling and tracking activities (eg, writing activities on a calendar and using different highlighter colors to represent different life areas).

None of the participants could recall any specific detail of the introduction, which provided the treatment rationale and instructions for using the program, and many indicated that they didn’t remember seeing any introduction. Feedback on other specific app features was mixed. For example, some users liked the level system, noting that it helped create a sense of playfulness and reward as well as a feeling of accomplishment, whereas others didn’t find the level system engaging or motivating.

Discussion

Principal Findings

Collectively, our results demonstrate promising engagement with the app and a potential signal for impact on theory-based change processes, depression symptoms, and cessation outcomes. Taken together with the results of two published pilot studies on BAT-D for smoking cessation [9,13], these findings support the feasibility of a BAT-D app to improve smoking cessation outcomes and mood among smokers with mild to moderate depression. They also support future planned work to complete development of Actify!, modify it to improve usability, and evaluate the complete intervention in a controlled pilot trial.

No prior studies have investigated the efficacy of smoking cessation apps for depressed smokers; however, our efficacy data compare favorably to that of previous reports of smoking cessation apps more broadly as well as face-to-face BAT-D for cessation. Although a fully powered effectiveness trial has yet to be completed, mHealth apps for smoking cessation have been evaluated in a number of pilot studies, demonstrating encouraging results for acceptability and quit rates [36-42]. The literature base to support apps for smoking cessation is still emerging and currently includes only two RCTs—both of which are pilot trials with wide confidence intervals for the point estimates of cessation. In a pilot RCT comparing SmartQuit to the NCI’s QuitGuide app (n=196), 30-day quit rates at 2 months were 13% (95% CI 6% to 22%) in SmartQuit versus 8% (95% CI 3% to 16%) in QuitGuide [36]. In a pilot trial comparing a text message intervention to the REQ-Mobile smartphone app (n=102) for young adult smokers, the REQ-Mobile app produced a 30-day PPA rate of 18% (95% CI 7% to 28%) at 12 weeks compared with the text message quit rate of 31% (95% CI 18% to 45%) [38]. Compared with previous pilot trial data showing short-term, 30-day quit rates of 8% to 18% in a general population of smokers, our 30-day quit rate of 19% is encouraging, especially given that the Actify! quit rates represent biochemically confirmed abstinence, whereas the two published RCTs did not include biochemical verification of abstinence to validate self-reported cessation (nor did the majority of the observational and single-arm pilot feasibility studies). Although biochemical confirmation is not universally necessary or feasible for population-level intervention trials [34], an assessment of biochemically confirmed quit rates for cessation apps provides support for the validity of the findings. We also note that biochemically confirmed abstinence rates (ie, 19% for 30-day PPA and the more comparable 31% for 7-day PPA) for Actify! at 6 weeks postenrollment compare favorably with short-term cessation outcomes for the face-to-face study of Behavioral Activation Treatment for Smoking (BATS) for depressed smokers (ie, 17% biochemically confirmed 7-day PPA at 4 weeks postquit date) [9]. Finally, given that previous work has reported a 6-week end-of-treatment quit rate of 7% for SmokefreeTXT [43], the text messaging component alone seems unlikely to fully account for the quit rate of the Actify! program.

Limitations and Future Directions

As an early pilot trial of an mHealth app, this study has a number of limitations. The sample size and study design (eg, single-arm pilot, n=16 in complete case analysis, no control group), albeit consistent with recommendations for early pilot intervention work [44], do not allow for strong causal conclusions about efficacy to be made, and estimates of cessation outcomes are inherently less reliable than would be the case in larger studies. We did not track nonstudy treatment use, so we cannot determine the extent to which use of other treatments may have influenced outcomes. We also used an existing text message intervention (SmokefreeTXT) to simulate content that will be added to Actify! in future iterations. The cessation content integrated into future versions of the app could result in differential use and impact. Actify! also included very minimal coverage of two elements of BAT-D that were deemed by the team to be potentially problematic in the mHealth app format—hour-by-hour activity monitoring and social contracts. This decision was based on our early Actify! user research as well as our broader experience with health app development, which has suggested that programs that are too burdensome or inadequately sensitive to the social stigma still attached to mental health conditions (ie, lengthy, time-intensive self-monitoring prior to initiating activity planning; use of social contracts to engage the user’s social network) increase the risk of attrition. However, in future iterations of the app, we will continue to explore methods to improve self-awareness and social support while minimizing risk of attrition. Finally, our sample, on average, was of high socioeconomic status, had a low level of nicotine dependence, and only included smokers with mild to moderate depression; thus, results may not generalize to smokers who are more severely depressed, more nicotine dependent, and of lower socioeconomic status.

Conclusion

While primarily designed to inform further refinements to Actify!, our findings from the user experience interviews contain important lessons learned that may be useful to others who are developing digital interventions based on BAT-D. For example, prior experience with in-person BAT-D interventions has demonstrated that the distinction between values and goals or activities is difficult for people to understand [10]. Based on our user experiences, this remains at least as challenging, if not more so, in self-guided mHealth interventions, and creative solutions are needed. This pilot trial also highlights the utility of implementing a user-centered design framework to develop mHealth interventions. Insights from posttreatment interviews indicate that the acceptability, and likely the outcomes, of Actify! were limited by usability issues not identified in our initial lab-based user testing phase. Future research could focus on allowing participants to use the app in their natural environment (eg, using diary study methodology to assess usability) to identify critical usability issues with activity planning and scheduling—core elements of the BAT-D treatment model—prior to a pilot trial. Nevertheless, the interview data gathered from pilot trial participants provide clear direction for the next iteration of the app. Key planned improvements include integrating therapeutic rationale and instructions for use into the main features of the app rather than in introductory onboarding screens, restructuring values identification to clarify the distinction between values and activities, and simplifying activity scheduling and tracking.

Acknowledgments

This work was funded by a grant to JLH from the Fred Hutchinson Cancer Research Center (Fred Hutch). The funding organization had no role in the study design and analysis, interpretation of the results, or decision to submit this work for publication. The authors would like to thank the University of Washington Human Centered Design and Engineering students who worked on app design and usability testing either as a class project (Evan Feenstra, Sang Woo Nam, Janet Ng, Anna Serra, Julie Stuart, Mia Suh, Brooks Tiffany, and Audrey Wu) under the direction of instructor Daniella Kim, PhD, or as a Fred Hutch user experience design and research intern (Audrey Wu and Anna Serra) under the direction of user experience researcher Kim Le. We would also like to acknowledge Seattle-based 2Morrow for their consultation and development work in support of this project; Josh Belzman, Fred Hutch social media manager, for his assistance with recruiting for the study via Facebook; and the participants who volunteered for the study.

Conflicts of Interest

JLH has received research support from Pfizer. JBB has served as a consultant for GlaxoSmithKline and serves on the advisory board of Chrono Therapeutics. Other authors have no financial conflicts to disclose.

Multimedia Appendix 1

Primary and secondary user personas.
[DOCX File, 284 KB - formative_v3i4e13728_app1.docx]

References

1. Jacobs EJ, Newton CC, Carter BD, Feskanich D, Freedman ND, Prentice RL, et al. What proportion of cancer deaths in the contemporary United States is attributable to cigarette smoking? Ann Epidemiol 2015 Mar;25(3):179-182 [FREE Full text] [doi: 10.1016/j.annepidem.2014.11.008] [Medline: 25487970]
2. Centers for Disease Control and Prevention (CDC). Smoking-attributable mortality, years of potential life lost, and productivity losses—United States, 2000-2004. MMWR Morb Mortal Wkly Rep 2008 Nov 14;57(45):1226-1228 [FREE Full text] [Medline: 19008791]
3. Centers for Disease Control and Prevention (CDC). Quitting smoking among adults—United States, 2001-2010. MMWR Morb Mortal Wkly Rep 2011 Nov 11;60(44):1513-1519 [FREE Full text] [Medline: 22071589]
4. Hebert KK, Cummins SE, Hernández S, Tedeschi GJ, Zhu S. Current major depression among smokers using a state quitline. Am J Prev Med 2011 Jan;40(1):47-53 [FREE Full text] [doi: 10.1016/j.amepre.2010.09.030] [Medline: 21146767]
5. Bricker JB, Mull KE, McClure JB, Watson NL, Heffner JL. Improving quit rates of web-delivered interventions for smoking cessation: full-scale randomized trial of WebQuit.org versus Smokefree.gov. Addiction 2018 May;113(5):914-923. [doi: 10.1111/add.14127] [Medline: 29235186]
6. Berlin I, Covey LS. Pre-cessation depressive mood predicts failure to quit smoking: the role of coping and personality traits. Addiction 2006 Dec;101(12):1814-1821. [doi: 10.1111/j.1360-0443.2006.01616.x] [Medline: 17156181]
7. Cook J, Spring B, McChargue D, Doran N. Effects of anhedonia on days to relapse among smokers with a history of depression: a brief report. Nicotine Tob Res 2010 Sep;12(9):978-982 [FREE Full text] [doi: 10.1093/ntr/ntq118] [Medline: 20709727]
8. Heffner JL, Mull KE, McClure JB, Bricker JB. Positive affect as a predictor of smoking cessation and relapse: does it offer unique predictive value among depressive symptom domains? Subst Use Misuse 2018 May 12;53(6):980-988 [FREE Full text] [doi: 10.1080/10826084.2017.1387569] [Medline: 29161212]
9. MacPherson L, Tull MT, Matusiewicz AK, Rodman S, Strong DR, Kahler CW, et al. Randomized controlled trial of behavioral activation smoking cessation treatment for smokers with elevated depressive symptoms. J Consult Clin Psychol 2010 Feb;78(1):55-61 [FREE Full text] [doi: 10.1037/a0017939] [Medline: 20099950]
10. Lejuez CW, Hopko DR, Acierno R, Daughters SB, Pagoto SL. Ten year revision of the brief behavioral activation treatment for depression: revised treatment manual. Behav Modif 2011 Mar;35(2):111-161. [doi: 10.1177/0145445510390929] [Medline: 21324944]
11. Hopko DR, Lejuez C, Ruggiero KJ, Eifert GH. Contemporary behavioral activation treatments for depression: procedures, principles, and progress. Clin Psychol Rev 2003 Oct;23(5):699-717. [Medline: 12971906]
12. Goelz PM, Audrain-McGovern JE, Hitsman B, Leone FT, Veluz-Wilkins A, Jepson C, et al. The association between changes in alternative reinforcers and short-term smoking cessation. Drug Alcohol Depend 2014 May 01;138:67-74 [FREE Full text] [doi: 10.1016/j.drugalcdep.2014.02.007] [Medline: 24598122]
13. Busch AM, Tooley EM, Dunsiger S, Chattillion EA, Srour JF, Pagoto SL, et al. Behavioral activation for smoking cessation and mood management following a cardiac event: results of a pilot randomized controlled trial. BMC Public Health 2017 Apr 17;17(1):323 [FREE Full text] [doi: 10.1186/s12889-017-4250-7] [Medline: 28415979]
14. Shiffman S, Brockwell SE, Pillitteri JL, Gitchell JG. Use of smoking-cessation treatments in the United States. Am J Prev Med 2008 Feb;34(2):102-111. [doi: 10.1016/j.amepre.2007.09.033] [Medline: 18201639]
15. The health consequences of smoking—50 years of progress: a report of the Surgeon General. Atlanta: Centers for Disease Control and Prevention; 2014. URL: https://www.ncbi.nlm.nih.gov/books/n/surgsmoke50/pdf/ [accessed 2019-10-02]
16. Mobile fact sheet. Washington: Pew Internet and American Life Project; 2019 Jun 12. URL: https://www.pewinternet.org/fact-sheet/mobile/ [accessed 2019-10-02]
17. Heffner JL, Mull KE. Smartphone ownership among US adult cigarette smokers: 2014 Health Information National Trends Survey (HINTS) Data. J Med Internet Res 2017 Aug 31;19(8):e305 [FREE Full text] [doi: 10.2196/jmir.7953] [Medline: 28860108]
18. Ly KH, Trüschel A, Jarl L, Magnusson S, Windahl T, Johansson R, et al. Behavioural activation versus mindfulness-based guided self-help treatment administered through a smartphone application: a randomised controlled trial. BMJ Open 2014 Jan 09;4(1):e003440 [FREE Full text] [doi: 10.1136/bmjopen-2013-003440] [Medline: 24413342]
19. Dahne J, Kustanowitz J, Lejuez CW. Development and preliminary feasibility study of a brief behavioral activation mobile application (behavioral apptivation) to be used in conjunction with ongoing therapy. Cogn Behav Pract 2018 Feb;25(1):44-56 [FREE Full text] [doi: 10.1016/j.cbpra.2017.05.004] [Medline: 29725198]
20. Dahne J, Lejuez CW, Kustanowitz J, Felton JW, Diaz VA, Player MS, et al. Moodivate: a self-help behavioral activation mobile app for utilization in primary care—development and clinical considerations. Int J Psychiatry Med 2017 Mar;52(2):160-175 [FREE Full text] [doi: 10.1177/0091217417720899] [Medline: 28792292]

21. Dahne J, Collado A, Lejuez CW, Risco C, Diaz VA, Kustanowitz J, et al. ¡Aptívate!: a Spanish-language behavioral activation mobile application for delivery via primary care. Psychol Serv 2019 May;16(2):271-275. [doi: 10.1037/ser0000304] [Medline: 30431309]
22. Dahne J, Lejuez C, Diaz V, Player M, Kustanowitz J, Felton J, et al. Pilot randomized trial of a self-help behavioral activation mobile app for utilization in primary care. Behav Ther 2019 Jul;50(4):817-827. [doi: 10.1016/j.beth.2018.12.003] [Medline: 31208690]
23. Dahne J, Collado A, Lejuez CW, Risco CM, Diaz VA, Coles L, et al. Pilot randomized controlled trial of a Spanish-language Behavioral Activation mobile app (¡Aptívate!) for the treatment of depressive symptoms among united states Latinx adults with limited English proficiency. J Affect Disord 2019 May 01;250:210-217. [doi: 10.1016/j.jad.2019.03.009] [Medline: 30870770]
24. 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]
25. LeRouge C, Ma J, Sneha S, Tolle K. User profiles and personas in the design and development of consumer health technologies. Int J Med Inform 2013 Nov;82(11):e251-e268. [doi: 10.1016/j.ijmedinf.2011.03.006] [Medline: 21481635]
26. Ubhi HK, Kotz D, Michie S, van Schayck OCP, West R. A comparison of the characteristics of iOS and Android users of a smoking cessation app. Transl Behav Med 2017 Jun;7(2):166-171 [FREE Full text] [doi: 10.1007/s13142-016-0455-z] [Medline: 28168609]
27. El-Hilly AA, Iqbal SS, Ahmed M, Sherwani Y, Muntasir M, Siddiqui S, et al. Game on? Smoking cessation through the gamification of mhealth: a longitudinal qualitative study. JMIR Serious Games 2016 Oct 24;4(2):e18 [FREE Full text] [doi: 10.2196/games.5678] [Medline: 27777216]
28. Heffner JL, Vilardaga R, Mercer LD, Kientz JA, Bricker JB. Feature-level analysis of a novel smartphone application for smoking cessation. Am J Drug Alcohol Abuse 2015 Jan;41(1):68-73 [FREE Full text] [doi: 10.3109/00952990.2014.977486] [Medline: 25397860]
29. Titov N, Dear BF, McMillan D, Anderson T, Zou J, Sunderland M. Psychometric comparison of the PHQ-9 and BDI-II for measuring response during treatment of depression. Cogn Behav Ther 2011 Jun;40(2):126-136. [doi: 10.1080/16506073.2010.550059] [Medline: 25155813]
30. 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. [doi: 10.1111/j.1360-0443.1991.tb01879.x] [Medline: 1932883]
31. Kroenke K, Spitzer RL, Williams JBW. 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]
32. Brown RA, Burgess ES, Sales SD, Whiteley JA, Evans DM, Miller IW. Reliability and validity of a smoking timeline follow-back interview. Psychol Addict Behav 1998;12(2):101-112. [doi: 10.1037/0893-164X.12.2.101]
33. Kanter JW, Mulick PS, Busch AM, Berlin KS, Martell CR. The Behavioral Activation for Depression Scale (BADS): psychometric properties and factor structure. J Psychopathol Behav Assess 2006 Oct 25;29(3):191-202. [doi: 10.1007/s10862-006-9038-5]
34. SRNT Subcommittee on Biochemical Verification. Biochemical verification of tobacco use and cessation. Nicotine Tob Res 2002 May;4(2):149-159. [doi: 10.1080/14622200210123581] [Medline: 12028847]
35. Cropsey KL, Trent LR, Clark CB, Stevens EN, Lahti AC, Hendricks PS. How low should you go? Determining the optimal cutoff for exhaled carbon monoxide to confirm smoking abstinence when using cotinine as reference. Nicotine Tob Res 2014 Oct;16(10):1348-1355 [FREE Full text] [doi: 10.1093/ntr/ntu085] [Medline: 24891552]
36. Bricker JB, Mull KE, Kientz JA, Vilardaga R, Mercer LD, Akioka KJ, et al. Randomized, controlled pilot trial of a smartphone app for smoking cessation using acceptance and commitment therapy. Drug Alcohol Depend 2014 Oct 01;143:87-94 [FREE Full text] [doi: 10.1016/j.drugalcdep.2014.07.006] [Medline: 25085225]
37. Bricker JB, Copeland W, Mull KE, Zeng EY, Watson NL, Akioka KJ, et al. Single-arm trial of the second version of an acceptance & commitment therapy smartphone application for smoking cessation. Drug Alcohol Depend 2017 Jan 01;170:37-42. [doi: 10.1016/j.drugalcdep.2016.10.029] [Medline: 27870987]
38. Buller DB, Borland R, Bettinghaus EP, Shane JH, Zimmerman DE. Randomized trial of a smartphone mobile application compared to text messaging to support smoking cessation. Telemed J E Health 2014 Mar;20(3):206-214 [FREE Full text] [doi: 10.1089/tmj.2013.0169] [Medline: 24350804]
39. Ubhi HK, Michie S, Kotz D, Wong WC, West R. A mobile app to aid smoking cessation: preliminary evaluation of SmokeFree28. J Med Internet Res 2015 Jan 16;17(1):e17 [FREE Full text] [doi: 10.2196/jmir.3479] [Medline: 25596170]
40. Iacoviello BM, Steinerman JR, Klein DB, Silver TL, Berger AG, Luo SX, et al. Clickotine—a personalized smartphone app for smoking cessation: initial evaluation. JMIR Mhealth Uhealth 2017 Apr 25;5(4):e56 [FREE Full text] [doi: 10.2196/mhealth.7226] [Medline: 28442453]
41. Gordon JS, Armin J, D Hingle M, Giacobbi P, Cunningham JK, Johnson T, et al. Development and evaluation of the See Me Smoke-Free multi-behavioral mHealth app for women smokers. Transl Behav Med 2017 Jun;7(2):172-184 [FREE Full text] [doi: 10.1007/s13142-017-0463-7] [Medline: 28155107]

42. Businelle MS, Ma P, Kendzor DE, Frank SG, Vidrine DJ, Wetter DW. An ecological momentary intervention for smoking cessation: evaluation of feasibility and effectiveness. J Med Internet Res 2016 Dec 12;18(12):e321 [FREE Full text] [doi: 10.2196/jmir.6058] [Medline: 27956375]
43. Augustson E, Cole-Lewis H, Sanders A, Schwarz M, Geng Y, Coa K, et al. Analysing user-reported data for enhancement of SmokefreeTXT: a national text message smoking cessation intervention. Tob Control 2017 Nov;26(6):683-689. [doi: 10.1136/tobaccocontrol-2016-052945] [Medline: 27852892]
44. Rounsaville B, Carroll K, Onken L. A stage model of behavioral therapies research: getting started and moving on from stage I. Clin Psychol Sci Pract 2001;8(2):133-142. [doi: 10.1093/clipsy.8.2.133]

Abbreviations

BADS: Behavioral Activation for Depression Scale
BAT-CS: Behavioral Activation Treatment for Cardiac Smokers
BAT-D: Behavioral Activation Treatment for Depression
BATS: Behavioral Activation Treatment for Smoking
CO: carbon monoxide
Fred Hutch: Fred Hutchinson Cancer Research Center
FTND: Fagerström Test for Nicotine Dependence
mHealth: mobile health
NCI: National Cancer Institute
PHQ-9: 9-item Patient Health Questionnaire
PPA: point prevalence abstinence
RCT: randomized controlled trial
TLFB: timeline follow-back

Edited by G Eysenbach; submitted 15.02.19; peer-reviewed by J Dahne, Y Li; comments to author 23.07.19; revised version received 21.08.19; accepted 25.09.19; published 27.11.19.

Please cite as:
Heffner JL, Watson NL, Serfozo E, Mull KE, MacPherson L, Gasser M, Bricker JB
A Behavioral Activation Mobile Health App for Smokers With Depression: Development and Pilot Evaluation in a Single-Arm Trial
JMIR Form Res 2019;3(4):e13728
URL: http://formative.jmir.org/2019/4/e13728/
doi:10.2196/13728
PMID:31774405

Original Paper

Acceptability and Use of Interactive Voice Response Mobile Phone Surveys for Noncommunicable Disease Behavioral Risk Factor Surveillance in Rural Uganda: Qualitative Study

Charles Ssemugabo[1], MPH; Elizeus Rutebemberwa[2], MPH, MBChB, PhD; Dan Kajungu[3], PhD; George W Pariyo[4], MBChB, PhD; Adnan A Hyder[5], MPH, MD, PhD; Dustin G Gibson[4], MS, PhD

[1] Department of Disease Control and Environmental Health, Makerere University School of Public Health, Makerere University College of Health Science, Kampala, Uganda

[2] Department of Health Policy, Planning and Management, Makerere University School of Public Health, Makerere University College of Health Science, Kampala, Uganda

[3] Iganga Mayuge Health and Demographic Surveillance Site, Makerere University Centre for Health and Population Research, Kampala, Uganda

[4] Department of International Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, United States

[5] Milken Institute School of Public Health, George Washington University, Washington, DC, United States

Corresponding Author:
Charles Ssemugabo, MPH
Department of Disease Control and Environmental Health
Makerere University School of Public Health
Makerere University College of Health Science
New Mulago Hill Road
Kampala
Uganda
Phone: 256 779625182
Email: cssemugabo@gmail.com

Abstract

Background: There is need for more timely data to inform interventions that address the growing noncommunicable disease (NCD) epidemic. With a global increase in mobile phone ownership, mobile phone surveys can bridge this gap.

Objective: This study aimed to explore the acceptability and use of interactive voice response (IVR) surveys for surveillance of NCD behavioral risk factors in rural Uganda.

Methods: This qualitative study employed user group testing (UGT) with community members. The study was conducted at the Iganga-Mayuge Health and Demographic Surveillance Site (IM-HDSS) in Eastern Uganda. We conducted four UGTs which consisted of different categories of HDSS members: females living in urban areas, males living in urban areas, females living in rural areas, and males living in rural areas. Participants were individually sent an IVR survey, then were brought in for a group discussion using a semistructured guide. Data were analyzed thematically using directed content analysis.

Results: Participants perceived that IVR surveys may be useful in promoting confidentiality, saving costs, and raising awareness on NCD behavioral risk factors. Due to the clarity and delivery of questions in the local language, the IVR survey was perceived as easy to use. Community members suggested scheduling surveys on specific days and sending reminders as ways to improve their use for surveillance. Social issues such as domestic violence and perceptions toward unknown calls, technological factors including poor network connections and inability to use phones, and personal issues such as lack of access to phones and use of multiple networks were identified as barriers to the acceptability and use of mobile phone surveys. However, incentives were reported to motivate people to complete the survey.

Conclusions: Community members reflected on contextual and sociological implications of using mobile phones for surveillance of NCD behavioral risk factors. The opportunities and challenges that affect acceptability and use of IVR surveys should be considered in designing and implementing surveillance programs for NCD risk factors.

(JMIR Form Res 2019;3(4):e15000) doi:10.2196/15000

KEYWORDS

mobile phones; interactive voice response; surveillance; behavioral risk factors; noncommunicable diseases; Uganda

Introduction

The use of mobile phones has a huge potential to promote health in low- and middle-income countries (LMIC) [1-3]. Mobile phones have been used in the surveillance of infectious diseases including malaria, tuberculosis, and influenza-like illness; child malnutrition and maternal health cases; and routine surveillance of various diseases and symptoms [4,5]. In Uganda, mobile phones have been adopted in health programs and used for improving patient care and clinic attendance among HIV/AIDS patients, communication of laboratory results among HIV/AIDS patients, and surveillance [6-9]. However, there is limited literature on use of mobile phones in prevention and control of noncommunicable diseases (NCDs) such as cancer, diabetes, and cardiovascular disease.

NCDs rank fourth among the 10 most common causes of death in Uganda [10]. The increasing burden of NCDs in Uganda and globally is attributed partly to behavioral risk factors including tobacco use, alcohol use, dietary intake, and physical activity [11]. In order to strengthen the health systems response to NCD prevention and control in Uganda, there is a need for frequent data on NCD behavioral risk factors. Surveillance of NCD behavioral risk factors in Uganda has been carried out using the World Health Organization Stepwise Approach to Surveillance (STEPS) NCD risk factor surveillance [12]. However, nationally representative household surveys are resource intensive and therefore are typically only conducted approximately every 5 years [13]. Of note, Uganda last conducted a STEPS survey in 2014 [14]. In order to have more up-to-date data, novel methodologies that can generate data at more frequent intervals need to be explored. Given the high levels of mobile phone ownership and access [15], NCD risk factor surveillance could be explored.

In high income countries, mobile phone surveys have been used to conduct surveillance for NCDs [16-18], but their use in LMICs is limited [19]. Interactive voice response (IVR) is one type of mobile phone survey. In IVR, participants use their mobile phone keypad to answer prerecorded questions. The responses are submitted to a Web database or server to ensure timely data synchronization and analysis. In order to ensure adoption and acceptance of mobile phone surveys for NCD behavioral risk factors in Uganda, there is need to consider the contextual issues in the setting where surveillance will be implemented.

The technology acceptance model (TAM) can help guide the development and adaptation of mobile phone surveys by offering a framework to predict the acceptance and use of information systems and technology by an individual user [20]. It demonstrates that behavioral intention to accept and use the technology is dependent on the perceived usefulness and perceived ease of use; however, these are influenced by external factors within the setting under which the technology is used. Therefore, this paper explores perceptions on the acceptability and use of IVR surveys for surveillance of NCD behavioral risk factors in rural Uganda.

Methods

Study Design

This was a qualitative study that employed user group testing (UGT) to inform the development of an IVR survey that can be used to collect NCD behavioral risk factors. Participants were sent an IVR survey to their personal mobile phone and then convened for a group discussion. The UGT approach allowed participants to provide feedback after interacting with the technology [21,22]. Study participants were purposively selected from four categories of people in the study area: females living in urban areas, males living in urban areas, females living in rural areas, and males living in rural areas. One UGT was conducted for each category. This study was designed and conducted in accordance with the Consolidated Criteria for Reporting Qualitative studies (COREQ) [23].

Study Area and Participants

This study was conducted within the Iganga-Mayuge Health and Demographic Surveillance Site (IM-HDSS). The IM-HDSS is located in two districts, Iganga and Mayuge, in eastern Uganda. The site is about 120 kilometers or 2.5 hours’ drive east of Kampala, the capital city of Uganda, along the Uganda-Kenya highway. A population of 90,000 people resides in 17,000 households with approximately 59% living in rural areas [24]. The site spans 65 villages from 7 subcounties, and it is served by 16 health centers and one hospital. The IM-HDSS conducts 2 regular surveillance rounds of vital statistics per year. Participants were eligible for the study if they were 18 years or older, owned a mobile phone, and had been HDSS residents for at least 6 months.

Development of the Interactive Voice Response Survey

The IVR survey consisted of questions selected from standardized household surveys such as World Health Organization STEPS, Tobacco Questions for Surveys, and the Behavioral Risk Factor Surveillance System [12,25,26]. Questions that highlighted the indicators in the global monitoring framework for NCDs and covered the four main behavioral risk factors for NCDs (physical activity, alcohol consumption, tobacco use, and dietary intake) were selected [27]. To ensure content validity, the questions were reviewed by mobile health (mHealth), public health, health systems, and HDSS experts.

The selected questions were audio-recorded into Lusoga and uploaded to the Viamo (IVR provider) platform. The IVR survey began with an introduction that included the purpose, duration, benefits, and sponsoring agency. After the survey introduction, participants were asked a series of demographic and NCD risk factor questions where participants use the key pad of their mobile phone to respond (eg, Do you currently smoke tobacco? If yes, press 1. If no, press 3.) Respondent’s data from the IVR survey are stored on a cloud server. The IVR survey was pretested with community members in Kampala prior to the UGTs to ensure that audio files were properly uploaded and that the programming (skip logic and response options) was correct.

Recruitment

An HDSS staff member with the help of a local community health worker (CHW) approached potential participants at their household and assessed their interest in study participation. The HDSS staff worked with a CHW from the rural area to identify rural UGT participants and an urban-based CHW for urban UGT groups. For those who were interested, the HDSS staff screened community members for study eligibility as previously described. Recruitment for study participation occurred one day prior to the UGT. Recruited participants were told the time, date, and location of the UGT and were asked to bring their mobile phone.

User Group Testing

The UGTs were carried out in the midmorning at a nearby primary school that was convenient for the study participants. Upon arrival, an HDSS staff member obtained informed consent and collected basic demographic information and mobile phone numbers from each participant. Participants' mobile phone numbers were then uploaded to the IVR survey platform. In a group setting, the participants were informed about the purpose of the UGT and how to take an IVR survey. Next, all participants were sent the IVR survey to their personal mobile phone by a co-investigator (CS). HDSS staff were available to assist the study participants if they had difficulty answering the survey. In cases where participants did not complete the IVR survey, the IVR survey was resent up to 3 times. This process lasted 30 minutes to 1 hour depending on how fast participants adapted to the exercise.

After participants interacted with the IVR survey, a group discussion was conducted. A semistructured guide developed by the study team was employed to explore perceptions on three thematic areas: (1) acceptability and willingness to using IVR surveys, (2) opportunities and challenges to using IVR surveys, (3) opportunities for improving use of IVR surveys for NCD behavioral risk factor surveillance (Multimedia Appendix 1). The guide was translated to the local language (Lusoga) and back-translated to English for accuracy. The final tool was approved for use by the HDSS site leader. Group discussions were moderated by a public health specialist (CS) and supported by an mHealth specialist (DGG) and two research assistants. The moderator and research assistants were familiar with the IM-HDSS site, spoke the local language (Lusoga), and had experience collecting qualitative data. UGTs were conducted until saturation was reached. All group discussions were audio recorded.

Data Management and Analysis

The audio recordings of the group discussions were labeled and stored on a secured server. They were transcribed verbatim and translated to English by two experienced research assistants fluent both in English and Lusoga. Two people read the transcripts and developed a codebook. This was done by rereading all the transcripts, assigning meaning units to each response and codes to each meaning unit, and taking note on the emerging subthemes. They combined the descriptive codes and discussed them to produce the code book. Data were entered in ATLAS.ti software version 7.0 (ATLAS.ti Scientific Software Development GmbH) for coding. Qualitative data were analyzed using directed content analysis. We used the TAM as the guiding framework to identify recurrent themes and subthemes by categorizing the codes [20,28]. Emerging themes are presented in the results and supported with quotes. Participant responses in IVR surveys were wirelessly transmitted to a cloud server on the Viamo platform. All IVR surveys taken during the UGT were immediately deleted after the group discussion.

Ethical Considerations

Ethical approval was obtained from the Makerere University School of Public Health Higher Degrees, Research, and Ethics Committee (protocol 526), Uganda National Council of Science and Technology (registration number SS 4477), and Johns Hopkins Bloomberg School of Public Health. Participation in the study was voluntary, and participants provided written informed consent. Other than the participant's telephone number, no name or identifying information was used during data collection. Phone numbers in the IVR database were deleted upon completion of the UGT. Data were only assessed by the study team and used only for the purposes of the study. Participants were provided 5000 Ugandan Shillings (US $1.36) for travel reimbursement.

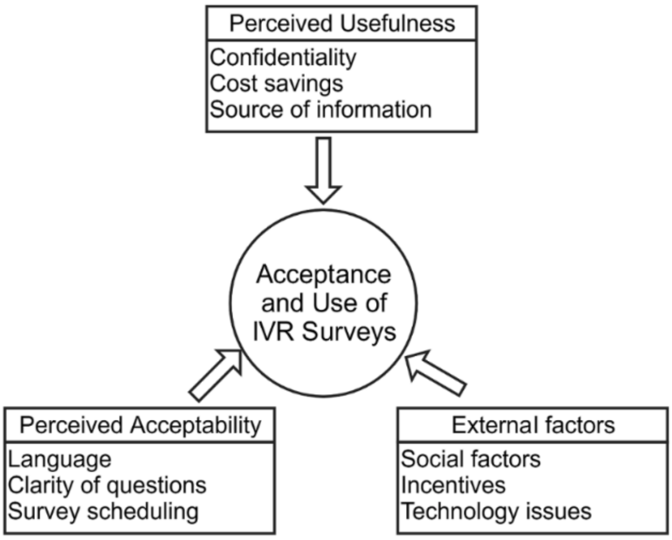
Results

Demographic Characteristics

In October 2017, four UGTs with a total of 43 adults selected purposively from the community were carried out. On average, each UGT was composed of 11 people (range 10 to 12 participants). Of the 43 UGT participants, 22 were females (51%), 29 had completed primary school (70%), and the median age was 41 years. The median time to complete the 32-question survey was 18 minutes and 31 seconds.

The findings are organized and presented based on TAM themes: perceived usefulness, perceived ease of use, and external factors (Figure 1). Under perceived usefulness, confidentiality, saving costs, and source of information on NCD risk factors emerged. Language used to administer questions, clarity of questions, and survey scheduling were the themes that emerged under perceived ease of use. External factors included social, technological, personal, and incentives.

Figure 1. Perceived usefulness and ease of use and external factors that inform use and acceptability of mobile phone surveys in rural Uganda. IVR: interactive voice response.



Perceived Usefulness

Confidentiality

UGT participants expressed how the IVR-administered questionnaire promotes confidentiality during the interview, commenting that one responds to questions while they are alone. As such, they are likely to give honest and accurate answers.

> *What I enjoyed most is the fact that the information will remain confidential, a person responds when he/she is alone.* [UGT, rural male]

Saving Costs

Most study participants thought that the IVR-administered questionnaires were saving costs, as one can respond to them at their current location as opposed to traditional face-to-face interviews where they would have to walk back home and meet the interviewer. IVR-administered questionnaires were also viewed as a mechanism of cutting fuel costs used by data enumerators, as the survey is delivered to everyone in a short time.

Source of Information on Noncommunicable Disease Risk Factors

Many participants viewed the IVR questionnaire as a source of information on NCD risk factors. They said that the questions highlight their lifestyle choices including diet and work and how they affect their health. The majority said that the surveys reminded them of traditional fruits and vegetables that they had not recently been eating.

> *The questions asked were concerned with the kind of food that we eat, like fruits such as pawpaw, jackfruit that we had abandoned.* [UGT, urban male]

Perceived Ease of Use

Language Used to Administer the Interactive Voice Response Survey

All participants perceived the use of local language in the delivery of the IVR-administered questionnaire as one of the motivators to complete the survey. They said that the questions are asked in the language they understand.

> *It gives you options of selecting the language you want to use to respond to questions, either Lusoga or English. So, the questions are asked in a language that you understand, as such you are able to understand everything very well.* [UGT, rural female]

Clarity of Questions

Almost all UGT participants said that the IVR survey questions were administered in a precise and clear voice, which increased their understanding of questions and hence the quality of data collected. They also mentioned that the questions were administered with clear instructions and sufficient time intervals between questions in a chronological manner.

> *I like the IVR technology because the voices were coming out clearly as though you are talking to a person physically. I could respond to the questions that were asked very well because I was understanding them.* [UGT, rural male]

UGT participants viewed introduction to the IVR questionnaire as very important in setting the stage for the respondent and drawing their attention to the subject being surveyed. In fact, they said that the introduction determines the interviewee's response.

However, some participants expressed challenges in quantifying or estimating the amounts of food, number of steps taken (for physical activity), and number of drinks taken among others. This was attributed to the format in which the questions were designed.

> *Questions regarding the types of heavy work we do and the time we take doing such work are hard to respond to because some of us don't have specific work that we do. We do work that comes along. In fact, I can even take a full week without getting heavy work. I may land on an old man and he tells me to go*

> *and help him dig in his garden, so answering that question was difficult.* [UGT, urban male]

Survey Scheduling

Participant preferences about scheduling of IVR questionnaire delivery varied with some preferring a specific day and others any day of the week. Participants who preferred a specific day of the week suggested weekends in the morning and others Sunday after lunch. Given the short duration of the survey (18 minutes and 31 seconds), some participants thought they would block off this time to complete the survey.

> *There should be specific day when the survey is delivered. I have a phone with a weak battery, so this will help me to prepare and charge my phone so that they can call me when I am ready.* [UGT, rural female]

Community members also thought receiving reminders would increase their chances of completing the IVR questionnaire. Reminders would help them to prepare for the call by charging their phones and ensuring that they are on at the time of the call. However, they had varied opinions on the format in which the reminders should be sent. Some preferred short message service (SMS) messages while others preferred voice messages, especially because many community members cannot read. Another issue was with the period the reminder should be sent prior to the call. Most participants preferred the reminder to be sent one day prior to the call.

External Factors

Social Factors

Participant preference for the voice of the IVR survey was dependent on gender. Females thought that if the survey was recorded in a male voice, their husbands might think they are having an extramarital relationship and thus cause violence in the home. However, some male participants preferred the female voice as a motivator for them to complete the survey.

> *If the call is recorded in a male voice and the survey comes to a female at night, it may hurt the husband. By the time you explain to him that it is a survey, domestic violence might have occurred already. In such cases, the call should be sent out during the day.* [UGT, urban female]

Societal constructions were also reported as potential hindrances to acceptance of the IVR survey. Study participants reported that people fear answering unknown calls because they are associated with the Illuminati or perceived evil schemes that can result in death.

> *Some people fear to receive unknown calls unless you inform them before the survey is delivered that the DSS will call you.* [UGT, rural female]

Incentives

Most participants reported that providing a small incentive would motivate people to complete the IVR interview. The incentives could be provided in either airtime or mobile money and would be given only if the participant completed the survey.

Technological Issues

The majority of the UGT participants said that people don’t know how to use mobile phones. In fact, they reported that most people only know how to receive and end calls. Given the survey has questions that might require them to use multiple keypads on their phones, they might not be in position to complete the survey without any training on how the system works. In addition to technology literacy, there may be problems with the respondents’ mobile phone, such as dysfunctional buttons which would prevent them from answering the question.

> *I was frustrated by the phone. I would understand the instruction but some of the buttons on my phone were not working.* [UGT, urban female]

Most participants also mentioned network connection as a challenge to attempting and completing the IVR survey.

Personal Factors

Lack of access to phones was also reported by most participants as a challenge to using the IVR questionnaire for surveillance of NCD risk factors. They said that some community members, especially in the rural areas, do not have phones while others might have left them charging or in the house at the time the survey is sent. Others don’t have electricity and their phones might not be charged.

> *Some homes don’t have phones. I don’t know how they will be receiving the IVR survey.* [UGT, rural female]

It was also reported that some people are on multiple networks yet they have a single network phone. So, they keep switching from one to another. At the time of the call, the known number might be off.

Discussion

Principal Findings

Our study provides insights into the acceptability and use of mobile phones for NCD behavioral risk factor surveillance in Uganda. Overall, participants were accepting of the IVR survey and perceived it as a useful tool to monitor a population’s health while also highlighting external factors as potential challenges to its implementation.

The UGTs and group discussions indicated a high level of acceptability given that the survey was narrated in the local language and the questions were clearly asked with appropriate examples as necessary. To improve the likelihood of someone taking an IVR survey, study participants suggested scheduling survey calls on specific days and sending reminders prior to the calls. In Australia, evidence from a randomized controlled trial found that sending text message reminders prior to a mobile phone survey significantly improved the response and completion rates [29]. Although reminders for mobile survey participation have not been assessed in LMICs, there is a large body of evidence supporting their use for other health outcomes [30].

A consensus emerged that IVR surveys are useful in ensuring confidentiality, potentially saving costs, and raising awareness on NCD behavioral risk factors. Whereas delivery of the survey in local language and clarity of questions were perceived to ease the use of the IVR survey, scheduling survey calls on specific days and sending reminders prior to the calls were proposed to ease its use. Community members reported that incentives improve use and acceptability of mobile phone surveys, but social, technological, and personal factors were perceived as barriers to their acceptability and use for NCD behavioral risk factor surveillance.

Privacy and safety of research participants is essential for any research undertaken. Mobile phone surveys using IVR may provide more opportunities to ensure privacy and safety of research participants by removing the human interviewer [31]. In our study, participants perceived that IVR surveys promote confidentiality. This is expected given that face-to-face interviews involve two parties as compared to the IVR questionnaire. Our findings are similar to studies carried out in Uganda and Australia [7,32]. However, studies in Burkina Faso, Uganda, and the United States highlight confidentially as a concern of mHealth research [6,33,34]. Our findings imply that there may be greater acceptance of this new technology as it ensures confidentiality.

Some of the IVR features that facilitated acceptance of the survey were how precise and clearly the questions were asked. Community members mentioned that the clarity of the voice used in the survey's narration and that instructions were given in a chronological manner prompted them to accept that the survey can be adopted and scaled for use in the HDSS. Given most study participants had not undertaken a survey using the IVR platform before, it is not surprising that clear voices and instructions easily familiarized community members with the technology. Community members also thought that the survey was introduced in a manner that drew the respondent's attention and eased its completion. Our findings corroborate the observations from another study by members of our group that piloted an IVR survey among university students [35].

Despite the willingness to use the IVR-administered questionnaire, community members felt that some questions that required estimations of their daily activities were difficult to answer. This could be due to lack of proper measurements for the different activities or their failure to make the right estimates or remember details of the activities they were involved with. These findings are also in line with those established during the IVR questionnaire pilot [35]. These findings have implications on the quality of data collected on NCD behavioral risk factors, requiring careful design and selection of questions [36].

The time of day, frequency, and timing of mHealth interventions have been found to influence their acceptance and efficacy elsewhere [30]. In our study, community members shared varied opinions on when the IVR survey should be sent out. While some preferred a specific day of the week (especially weekends), others were happy to receive the survey on any day of the week. This has implications on use and completion of the survey given people have different schedules. It is important that in scheduling the delivery of the IVR questionnaire, community norms and calendar are considered to ensure ease of use of the technology. Giving respondents the ability to schedule when the survey is sent may mitigate this challenge, especially since people are busy at work or their mobile phone may not be charged. It was also evident from our study that reminders may increase community acceptance and use of the IVR-delivered surveys. Most participants preferred voice to SMS reminders because of concerns about illiteracy.

Community members thought that IVR-administered surveys may result in domestic violence between family members if the IVR recording was in a voice opposite from the respondent's gender. Given that the majority of domestic violence is incurred by women, using a female voice may lessen the likelihood of this event. Participants also thought that societal constructions toward receiving unknown calls would hinder acceptance of the interventions. Communities hold beliefs that some unknown calls are sent by evil spirits that can claim the life of the receiver during the calls. Television, radio, or newspaper campaigns may sensitize communities to inform them about potentially receiving a survey. Moreover, including information on who is sponsoring the survey and offering a phone number for participants to call and verify the survey source may help improve participation. Our findings are similar to others who have explored acceptance of e-interventions such as students' use of e-learning platforms [37] and the general population's use of e-banking [38]. The use of financial incentives has been found to increase the behavioral intentions of students, farmers, and the general population to adopt a new technology [39-41]. It was established from our study that financial incentives may increase its acceptance, thereby motivating people to complete the survey. However, this needs to be tested and evaluated as it has implications on sustainability of the surveillance platform.

Use of an IVR-administered questionnaire largely depends on individuals' ability to use mobile phones and possession of functional phones [42]. Findings from our study hint at potential technology literacy challenges as we found that many community members can only receive calls and switch off their phones. In addition, where individuals can use phones, they are hindered by lack of network connectivity especially in the remote areas. Community members also reported that some of their colleagues do not own mobile phones and others have multiple networks and as such keep switching from one to another. All these technology challenges would prevent individuals from participating in the IVR survey and have been described in other mHealth studies conducted in Uganda. [6]. Like other mHealth interventions, plans to scale-up the use of an IVR-administered questionnaire needs to consider this limitation and may require alternative sampling strategies and survey modalities to reach all potential participants.

Limitations

The limitations to this study include that perceptions on acceptability and use were explored immediately after subjecting participants to one IVR survey call. This means that their views may have been mainly influenced by the experiences of the interactions during the session. In addition, such a short period of interaction with the innovation might have limited the type of experiences and perceptions that community members shared about acceptability and use of mobile phone surveys during the

UGT. In fact, some participants could not respond to or complete the interview due to dysfunctional phones and network challenges. This might have affected the shaping of their opinions toward acceptability and use of mobile phone surveys for NCD risk factor surveillance. Lastly, participants were required to own a mobile phone to participate. Given that phone ownership is not universal, our findings may be biased.

Conclusion

In conclusion, community members reflected on contextual and sociological implications of using mobile phones for surveillance of NCD behavioral risk factors. Using the TAM, we identified factors under three key themes—perceived usefulness, perceived ease of use, and external factors (technology and personal)—that should be considered in designing and implementing mobile phone surveillance programs for NCD risk factors. Findings from the UGT will be used to refine and adapt the IVR survey and its delivery prior to deployment.

Acknowledgments

The authors would like to thank the IM-HDSS team for their support in planning and organizing data collection as well as the community members for participating in the UGTs. The authors would also like to thank Viamo (formerly VOTOmobile) for their technical assistance with the IVR platform. Research reported in this publication was supported by the Fogarty International Center of the US National Institutes of Health under award number R21TW010415.

Conflicts of Interest

None declared.

Multimedia Appendix 1
User group discussion guide.
[DOCX File , 16 KB - formative_v3i4e15000_app1.docx]

References

1. Betjeman TJ, Soghoian SE, Foran MP. mHealth in Sub-Saharan Africa. Int J Telemed Appl 2013;2013:482324 [FREE Full text] [doi: 10.1155/2013/482324] [Medline: 24369460]
2. Porter G, Hampshire K, Milner J, Munthali A, Robson E, de Lannoy A, et al. Mobile phones and education in Sub-Saharan Africa: from youth practice to public policy. J Int Dev 2015 Jun 11;28(1):22-39. [doi: 10.1002/jid.3116]
3. Hampshire K, Porter G, Owusu SA, Mariwah S, Abane A, Robson E, et al. Informal m-health: how are young people using mobile phones to bridge healthcare gaps in Sub-Saharan Africa? Soc Sci Med 2015 Oct;142:90-99 [FREE Full text] [doi: 10.1016/j.socscimed.2015.07.033]
4. Brinkel J, Krämer A, Krumkamp R, May J, Fobil J. Mobile phone-based mhealth approaches for public health surveillance in sub-Saharan Africa: a systematic review. IJERPH 2014 Nov 12;11(11):11559-11582. [doi: 10.3390/ijerph111111559] [Medline: 25396767]
5. Deribe K. Mobile phones used for public health surveillance. Forced Migr Rev 2011;38:33.
6. Chang LW, Kagaayi J, Arem H, Nakigozi G, Ssempijja V, Serwadda D, et al. Impact of a mHealth intervention for peer health workers on AIDS care in rural Uganda: a mixed methods evaluation of a cluster-randomized trial. AIDS Behav 2011 Nov;15(8):1776-1784 [FREE Full text] [doi: 10.1007/s10461-011-9995-x] [Medline: 21739286]
7. Kunutsor S, Walley J, Katabira E, Muchuro S, Balidawa H, Namagala E, et al. Using mobile phones to improve clinic attendance amongst an antiretroviral treatment cohort in rural Uganda: a cross-sectional and prospective study. AIDS Behav 2010 Dec;14(6):1347-1352. [doi: 10.1007/s10461-010-9780-2] [Medline: 20700644]
8. Siedner MJ, Haberer JE, Bwana MB, Ware NC, Bangsberg DR. High acceptability for cell phone text messages to improve communication of laboratory results with HIV-infected patients in rural Uganda: a cross-sectional survey study. BMC Med Inform Decis Mak 2012 Jun 21;12:56 [FREE Full text] [doi: 10.1186/1472-6947-12-56] [Medline: 22720901]
9. Asiimwe C, Gelvin D, Lee E, Ben AY, Quinto E, Katureebe C, et al. Use of an innovative, affordable, and open-source short message service-based tool to monitor malaria in remote areas of Uganda. Am J Trop Med Hyg 2011 Jul;85(1):26-33 [FREE Full text] [doi: 10.4269/ajtmh.2011.10-0528] [Medline: 21734120]
10. Centers for Disease Control and Prevention. 2018. Global Health—Uganda URL: https://www.cdc.gov/globalhealth/countries/uganda/default.htm [accessed 2019-10-22]
11. Schwartz JI, Guwatudde D, Nugent R, Kiiza CM. Looking at non-communicable diseases in Uganda through a local lens: an analysis using locally derived data. Global Health 2014 Nov 19;10(1). [doi: 10.1186/s12992-014-0077-5] [Medline: 25406738]
12. STEPS: A framework for surveillance—The WHO STEPwise Approach to Surveillance of Noncommunicable Diseases. Geneva: World Health Organization; 2003. URL: https://www.who.int/ncd_surveillance/en/steps_framework_dec03.pdf [accessed 2019-10-22]

13. Riley L, Guthold R, Cowan M, Savin S, Bhatti L, Armstrong T, et al. The World Health Organization STEPwise approach to noncommunicable disease risk-factor surveillance: methods, challenges, and opportunities. Am J Public Health 2016 Jan;106(1):74-78. [doi: 10.2105/AJPH.2015.302962]
14. World Health Organization. Non-communicable disease risk factor baseline survey: Uganda 2014 URL: http://www.who.int/ncds/surveillance/steps/Uganda_2014_STEPS_Report.pdf [accessed 2019-10-22]
15. World Bank. 2017. Mobile cellular subscriptions (per 100 people) URL: https://data.worldbank.org/indicator/IT.CEL.SETS.P2 [accessed 2019-10-22]
16. Echouffo-Tcheugui JB, Yaya S, Joshi R, Venkat Narayan KM, Kengne AP. Population surveillance of cardiovascular diseases in low-income to middle-income countries should leverage existing international collaborations. BMJ Glob Health 2018 Oct 01;3(5):e000866. [doi: 10.1136/bmjgh-2018-000866]
17. Kamis K, Janevic MR, Marinec N, Jantz R, Valverde H, Piette JD. A study of mobile phone use among patients with noncommunicable diseases in La Paz, Bolivia: implications for mHealth research and development. Global Health 2015 Jul;11:30 [FREE Full text] [doi: 10.1186/s12992-015-0115-y] [Medline: 26141528]
18. Hyder AA, Wosu AC, Gibson DG, Labrique AB, Ali J, Pariyo GW. Noncommunicable disease risk factors and mobile phones: a proposed research agenda. J Med Internet Res 2017 May 05;19(5):e133. [doi: 10.2196/jmir.7246]
19. Gibson DG, Pereira A, Farrenkopf BA, Labrique AB, Pariyo GW, Hyder AA. Mobile phone surveys for collecting population-level estimates in low- and middle-income countries: a literature review. J Med Internet Res 2017 May 05;19(5):e139. [doi: 10.2196/jmir.7428]
20. Lai P. The literature review of technology adoption models and theories for the novelty technology. J Info Sys Technol Manag 2017 Apr 30;14(1). [doi: 10.4301/s1807-17752017000100002]
21. Tausch AP, Menold N. Methodological aspects of focus groups in health research. Global Qual Nurs Res 2016 Mar 14;3. [doi: 10.1177/2333393616630466]
22. Pope C, Ziebland S, Mays N. Qualitative research in health care. Analysing qualitative data. BMJ 2000 Jan 08;320(7227):114-116 [FREE Full text] [doi: 10.1136/bmj.320.7227.114] [Medline: 10625273]
23. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357 [FREE Full text] [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
24. Makerere University Center for Health and Population Research. 2019. Demographic surveillance URL: http://www.muchap.org/what-we-do/demographic-surveillance [accessed 2019-01-09]
25. Global Adult Tobacco Survey Collaborative Group. Tobacco Questions for Surveys: A Subset of Key Questions from the Global Adult Tobacco Survey (GATS) 2nd ed. Atlanta: Centers for Disease Control and Prevention; 2011.
26. Behavioral Risk Factor Surveillance System. Atlanta: Centers for Disease Control and Prevention; 2016. URL: https://www.cdc.gov/brfss/ [accessed 2019-01-10]
27. World Health Organization. NCD global monitoring framework: ensuring progress on noncommunicable diseases in countries. Geneva: World Health Organization URL: http://www.who.int/nmh/global_monitoring_framework/en/ [accessed 2019-01-10]
28. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MISQ 1989 Sep;13(3):319-340. [doi: 10.2307/249008]
29. Dal Grande E, Chittleborough CR, Campostrini S, Dollard M, Taylor AW. Pre-survey text messages (SMS) improve participation rate in an australian mobile telephone survey: an experimental study. PLoS One 2016 Feb;11(2):e0150231 [FREE Full text] [doi: 10.1371/journal.pone.0150231] [Medline: 26919558]
30. Gibson DG, Tamrat T, Mehl G. The state of digital interventions for demand generation in low- and middle-income countries: considerations, emerging approaches, and research gaps. Glob Health Sci Pract 2018 Oct 10;6(Supplement 1):S49-S60. [doi: 10.9745/GHSP-D-18-00165]
31. Arora S, Yttri J, Nilse W. Privacy and security in mobile health (mHealth) research. Alcohol Res 2014;36(1):143-151 [FREE Full text] [Medline: 26259009]
32. Baffour B, Roselli T, Haynes M, Bon JJ, Western M, Clemens S. Including mobile-only telephone users in a statewide preventive health survey—differences in the prevalence of health risk factors and impact on trends. Prev Med Rep 2017 Sep;7:91-98. [doi: 10.1016/j.pmedr.2017.05.009]
33. Duclos V, Yé M, Moubassira K, Sanou H, Sawadogo NH, Bibeau G, et al. Situating mobile health: a qualitative study of mHealth expectations in the rural health district of Nouna, Burkina Faso. Health Res Policy Sys 2017 Jul 12;15(S1):47. [doi: 10.1186/s12961-017-0211-y]
34. Ali J, DiStefano MJ, Coates McCall I, Gibson DG, Al Kibria GM, Pariyo GW, et al. Ethics of mobile phone surveys to monitor non-communicable disease risk factors in low- and middle-income countries: A global stakeholder survey. Global Public Health 2019 Jan 10;14(8):1167-1181. [doi: 10.1080/17441692.2019.1566482]
35. Gibson DG, Farrenkopf BA, Pereira A, Labrique AB, Pariyo GW. The development of an interactive voice response survey for noncommunicable disease risk factor estimation: technical assessment and cognitive testing. J Med Internet Res 2017 May 05;19(5):e112. [doi: 10.2196/jmir.7340]

36. Pariyo GW, Greenleaf AR, Gibson DG, Ali J, Selig H, Labrique AB, et al. Does mobile phone survey method matter? Reliability of computer-assisted telephone interviews and interactive voice response non-communicable diseases risk factor surveys in low and middle income countries. PLoS ONE 2019 Apr 10;14(4):e0214450. [doi: 10.1371/journal.pone.0214450]
37. Abbas T. Social factors affecting students' acceptance of e-learning environments in developing and developed countries. J Hospit Tourism Technol 2016 May 03;7(2):200-212. [doi: 10.1108/jhtt-11-2015-0042]
38. Korwatanasakul W. Factors influencing technology adoption of people with visual impairment: case study of financial transactions through an automatic teller machine (ATM). Kasetsart J Soc Sci 2018 Aug;7(5):e5 [FREE Full text] [doi: 10.1016/j.kjss.2018.07.018]
39. Roumani Y, Nwankpa JK, Roumani YF. The impact of incentives on the intention to try a new technology. Technol Analysis Strat Manag 2014 Sep 04;27(2):126-141. [doi: 10.1080/09537325.2014.952625]
40. Jones LR, Cherry CR, Vu TA, Nguyen QN. The effect of incentives and technology on the adoption of electric motorcycles: a stated choice experiment in Vietnam. Transport Res A Policy Prac 2013 Nov;57:1-11 [FREE Full text] [doi: 10.1016/j.tra.2013.09.003]
41. Hoffman V. Market incentives for technology adoptionxperimental evidence from Kenyan maize farmers. 2018 Presented at: International Association of Agricultural Economists; July 28-August 2, 2018; Vancouver.
42. Ma Q, Chan AH, Chen K. Personal and other factors affecting acceptance of smartphone technology by older Chinese adults. Appl Ergonom 2016 May;54:62-71 [FREE Full text] [doi: 10.1016/j.apergo.2015.11.015]

Abbreviations

CHW: community health worker
COREQ: Consolidated Criteria for Reporting Qualitative Studies
IM-HDSS: Iganga-Mayuge Health and Demographic Surveillance Site
IVR: interactive voice response
LMIC: low- and middle-income countries
mHealth: mobile health
NCD: noncommunicable disease
SMS: short message service
STEPS: Stepwise Approach to Surveillance
TAM: technology acceptance model
UGT: user group testing

Edited by G Eysenbach; submitted 11.06.19; peer-reviewed by N Kaonga, A McNaghten; comments to author 09.07.19; revised version received 14.08.19; accepted 29.08.19; published 03.12.19.

Please cite as:
Ssemugabo C, Rutebemberwa E, Kajungu D, Pariyo GW, Hyder AA, Gibson DG
Acceptability and Use of Interactive Voice Response Mobile Phone Surveys for Noncommunicable Disease Behavioral Risk Factor Surveillance in Rural Uganda: Qualitative Study
JMIR Form Res 2019;3(4):e15000
URL: http://formative.jmir.org/2019/4/e15000/
doi:10.2196/15000
PMID:31793889

Original Paper

Beliefs Related to Participation in a Large Web-Based Prospective Survey on Diet and Health Among Individuals With a Low Socioeconomic Status: Qualitative Study

Mélina Côté[1]; Annie Lapointe[1], PhD; Catherine Laramée[1], MSc; Simone Lemieux[1,2], PhD; Sophie Desroches[1,2], PhD; Ariane Belanger-Gravel[1,3], PhD; Benoît Lamarche[1,2], PhD

[1]Institute of Nutrition and Functional Foods, Université Laval, Québec, QC, Canada
[2]School of Nutrition, Université Laval, Québec, QC, Canada
[3]Department of Communication, Université Laval, Québec, QC, Canada

Corresponding Author:
Benoît Lamarche, PhD
Centre Nutrition, Santé, et Société
Institute of Nutrition and Functional Foods
Université Laval
2440 Hochelaga Boulevard
Québec, QC, G1V 0A6
Canada
Phone: 1 418 656 2131 ext 404355
Email: benoit.lamarche@fsaa.ulaval.ca

Abstract

Background: NutriQuébec is a Web-based prospective study on the relationship between diet and health as well as the impact of food-related health policies in the adult population of Québec, Canada. Recruitment and retention of individuals with a low socioeconomic status (SES) in such a study are known to be challenging, yet critical for achieving representativeness of the entire population.

Objective: This study aimed to identify the behavioral, normative, and control beliefs of individuals with a low SES regarding participation in the NutriQuébec project and to identify their preferences regarding recruitment methods.

Methods: A total of four focus groups were conducted in community centers located in low-income areas of Québec City, Canada. On the basis of the theory of planned behavior, participants' beliefs associated with attitude, subjective norm, and perceived behavioral control regarding hypothetical participation in the NutriQuébec project were identified. Focus groups were recorded, transcribed, and coded by two analysts.

Results: Participants (16 men and 12 women) were aged between 28 and 72 years, and a majority of the participants had an annual household income of Can $19,999 or less. The main perceived advantages of participating in the NutriQuébec project were contributing to improved collective health and supporting research. The only disadvantage identified was the risk of having to fill out too many questionnaires. Participants could not, in general, identify persons from their entourage who would approve or disapprove their participation in the study. The main facilitators identified were obtaining a brief health assessment and the ability to complete questionnaires in a way that is not Web-based. The main barrier was the lack of internet access. The preferred means of recruitment were through social media, television, and community centers.

Conclusions: These results provide insightful information regarding the best methods and messages to use in order to recruit and retain individuals with a low SES in a population-based prospective study on lifestyle and health on the internet.

(JMIR Form Res 2019;3(4):e13854) doi:10.2196/13854

KEYWORDS

focus groups; qualitative research; social class; research subject; retention

Introduction

Background

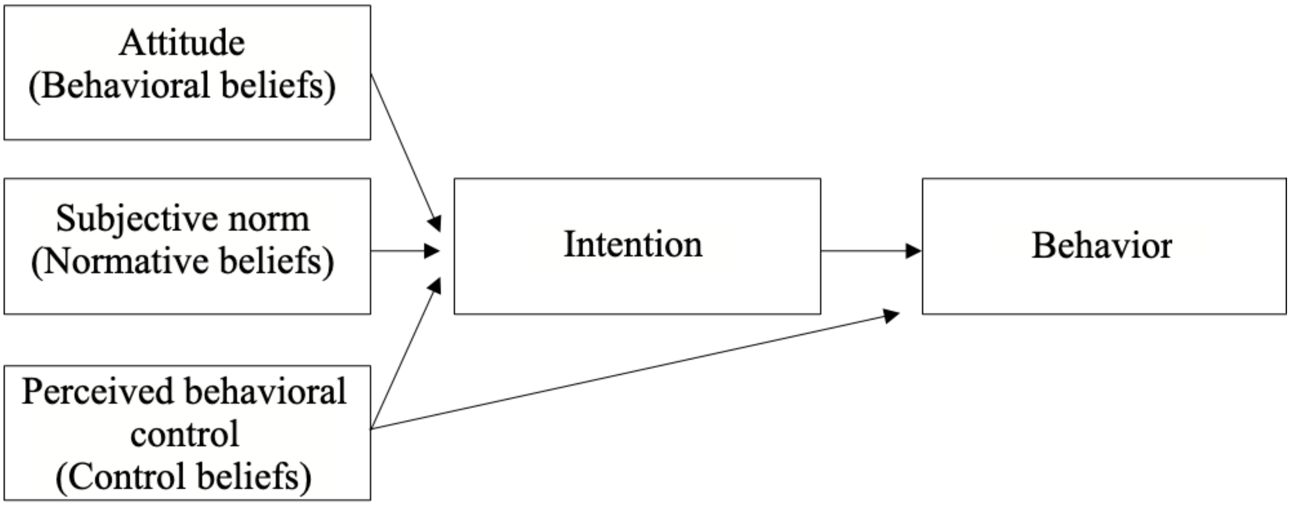
Evaluating and monitoring lifestyle habits, including dietary habits, at the population level is imperative in implementing impactful policies, regulations, and educational programs to improve public health. This requires collection of data from population-based samples that are as representative of the general population as possible. Vulnerable populations, including individuals with a low socioeconomic status (SES), are at the greatest risk of developing chronic diseases as well as adopting poor lifestyle habits such as inadequate dietary habits, sedentary lifestyles, smoking, and alcoholism [1-3]. Such populations are generally underrepresented in health studies [2,4]. Therefore, identifying beliefs such as facilitators and barriers among low SES populations is essential for increasing their participation in health studies and, hence, increasing the generalizability of results from population-based research.

Recruiting individuals with a low SES may represent an even greater challenge in Web-based studies, even if access to the internet and the use of social media have broadened significantly over the recent years [5]. This is of significance because the use of Web-based surveys has become an attractive way of conducting large prospective epidemiological studies at lower costs [6,7]. Perceived barriers to participation among hard-to-reach populations have been reported, including transportation and time constraints [8-10], economic constraints [9], fear of exploitation, and lack of knowledge [9-11]. Yet, the perceived barriers reported in these studies were not documented in the context of a Web-based prospective study and were not conducted according to a validated theoretical framework.

Theoretical Framework

According to the theory of planned behavior (Figure 1) [12], adoption of a given behavior is determined by one's intention (or motivation) to perform the behavior. In turn, intention is determined by the attitude toward this behavior; the subjective norm related to this behavior; and perceived behavioral control (PBC), which reflects how one perceives oneself as being capable of performing the behavior. Each of these primary constructs are determined by one's beliefs. Attitude reflects behavioral beliefs, which are the perceived advantages and disadvantages of performing the behavior. Subjective norms reflect normative beliefs, which are the perceived social pressures of performing or not performing the behavior. PBC reflects control beliefs, which are the perceived facilitators and barriers to one's ability to perform the behavior. Developing knowledge on beliefs underlying attitudes, subjective norms, and PBC among individuals with a low SES regarding a specific behavior, in this case, the intention to participate in the Web-based NutriQuébec project, will contribute to a better understanding of factors related to the intention to perform the behavior in this population.

Figure 1. The Ajzen theory of planned behavior.



NutriQuébec Project

The NutriQuébec project is a Web-based prospective study aimed at documenting the impact of nutrition-related public health policies and action plans on the diet of the Québec adult population in Canada. More broadly, the NutriQuébec project aims to examine the associations between diet and health outcomes. To this effect, male and female adults (aged ≥18 years) are invited to complete a series of yearly core questionnaires through the Web, assessing dietary habits, physical activity, other lifestyle habits, sociodemographic characteristics, and general health. Time required to complete all questionnaires on a yearly basis is approximately 2 hours (unpublished data). Participants may be invited to complete additional questionnaires on other nutrition-related issues between each yearly core measurements. A brief personalized assessment of dietary habits is returned yearly to each participant as a token of appreciation for their involvement.

To the best of our knowledge, perceptions regarding participation in a Web-based prospective study on nutrition and health have not yet been documented. Thus, the aim of this study was to identify the salient beliefs of individuals with a low SES toward the participation in a prospective Web-based study by using the theory of planned behavior as a framework [12]. We also wanted to identify the methods that were more likely to be successful to recruit and retain individuals with a low SES in such a Web-based study.

Methods

Participants and Recruitment

This study used a qualitative approach with focus groups and the theory of planned behavior framework to achieve its objectives. The research team recruited participants through personal contact by going to nonprofit organizations (food banks and community organizations) known to serve populations living in sectors of Québec City (Canada) with a low SES. These sectors are known to have a high rate of unemployment and low level of education. We included adults who were (1) living in a low-income sector, (2) living in the Québec City metropolitan area, and (3) able to speak French. Men and women were invited to participate in a 2-hour focus group to gather opinions about the NutriQuébec project. To accommodate participants, the focus groups were conducted in the community centers visited by the participants. Inclusion criteria were validated on site, and participants were given a date for when the focus group would occur. Participants were called back the day before the focus group to confirm their presence. Participants who completed the focus group received a Can $50 honorarium redeemed in cash and were reimbursed for transportation costs. This study was approved by the Ethics Committee of Laval University (2018-042 A-1/18-05-2018).

Data Collection

Focus groups were preferred over individual interviews or surveys because they encourage participants to interact with each other, leading to a wider range of salient beliefs, particularly among populations with a low SES who, in general, have lower education levels and lower literacy compared to the rest of the population. A total of four focus groups of six to eight participants each were conducted in three preselected community centers between May 29 and June 12, 2018. Focus groups were audiotaped and videotaped to facilitate verbatim transcription and to document nonverbal behavior when relevant. Before starting the focus group, all participants provided written and informed consent. The participants also completed a questionnaire inquiring about their sociodemographic situation as well as use of the internet. Only a brief description of the soon-to-be-implemented NutriQuébec project was provided to participants to limit biases toward views and perceptions on the potential benefits of participation. A moderator, an assistant moderator, and an observer sat at the table with the participants. The moderator conducted the discussion using a semistructured interview guide (Multimedia Appendix 1), which was developed based on the constructs of the theory of planned behavior [12]. The interview questionnaire was used to identify the beliefs that determine the participants' attitudes (advantages and disadvantages), subjective norms (approval and disapproval of significant others), and PBC (facilitators and barriers) related to a hypothetical participation in the NutriQuébec project. Participants were also questioned about their concerns regarding the security, use, and sharing of data as well as their preferences regarding recruitment methods. The moderator guide was pretested during the first focus group and needed no modification thereafter. Hence, the first focus group was included in the analysis. The moderator asked the questions in the same order for each of the focus groups to increase comparability of responses and encouraged all participants to express their opinions. The moderator also asked insightful questions and provided prompts when necessary to help move the discussion forward. Responses from the participants in the first focus groups were not used as prompts in subsequent focus groups. Once the participants left, the moderator, assistant moderator, and observer briefly met to comment on the focus group and modal beliefs.

Data Analysis

The SAS software (SAS Institute, Cary, NC) was used for descriptive analyses. A deductive content analysis [13] of the focus groups was conducted. The verbatim transcriptions were used to code the quotes and categorize the beliefs expressed by the participants according to the constructs of the theory of planned behavior, namely, behavioral beliefs, normative beliefs, and control beliefs or according to their relevance to additional established themes, namely, security and sharing of data and preferred recruitment methods. If a belief was named by several participants in the same group, it was considered as a single belief. Moreover, two analysts (MC and AL) coded the verbatim transcript of each focus group independently using NVivo10 qualitative research software (QSR International Pty Ltd, Melbourne, Australia) before comparing and discussing the identified themes to reach consensus. A third person was available to consult in case of a disagreement. Data saturation, based on informational redundancy, was attained after the fourth focus group [14]. Hence, no additional focus groups were planned.

Results

Participants' Characteristics

A total of 28 participants (16 men and 12 women) aged between 28 and 72 years participated in the study. The majority of participants had an annual household income of Can $19,999 or less, reflecting a low SES. Moreover, three-quarters (21/27, 78%) of the participants had access to the internet at home or elsewhere within their community. Computer skills were self-rated as moderate (6/27, 22%), advanced (10/27, 37%), or expert (3/27, 11%; Table 1).

Table 1. Sociodemographic characteristics of the participants from the four focus groups (N=28).

Characteristics	Values
Age (years), mean (SD)	50 (12)
Gender, n (%)	
Male	16 (57)
Female	12 (43)
Education, n (%)	
Primary school	5 (18)
High school	11 (39)
College	8 (29)
University	3 (11)
I prefer not to answer	1 (4)
Household income (Can $), n (%)	
0-9999	13 (46)
10,000- 19,999	11 (39)
20,000- 29,999	3 (11)
I prefer not to answer	1 (4)
Computer skills, n (%)[a]	
Not at all competent	5 (19)
Hardly competent	3 (11)
Moderately and rather competent	6 (22)
Very competent	10 (37)
Expert	3 (11)
Employment situation, n (%)	
Full-time worker	0 (0)
Part-time worker	4 (14)
Unemployed	6 (21)
Stay-at-home parent	3 (11)
Retired	7 (25)
Unable to work	4 (14)
Other	3 (11)
I prefer not to answer	1 (4)
Internet access, n (%)[a]	
Yes	21 (78)
No	6 (22)

[a] N=27.

Results of the Analysis of the Focus Groups

All reported beliefs are listed along with quotes to illustrate the key messages and beliefs. The quotes were originally in French and have been translated.

Behavioral Beliefs: Perceived Advantages and Disadvantages of Participating in the NutriQuébec Project

Frequently cited advantages of participating in the NutriQuébec project pertained to supporting research, contributing to improving collective health, receiving a brief health assessment, and improving one's lifestyle habits:

> *Improve living conditions of people also, not just ourselves. The population in general.* [Group 1, participant #3]

Other advantages that were mentioned include improving the family's lifestyle habits, making time for oneself, and acquiring and sharing knowledge.

No clear disadvantages of participating in the NutriQuébec project emerged in most focus groups. In fact, some participants indicated that although they could not identify personal advantages, they did not believe that participating in the project would do any harm or present any disadvantages. The only disadvantage discussed was the risk of having to fill out too many questionnaires.

Normative Beliefs: Approval and Disapproval of Significant Others Toward Participation in the NutriQuébec Project

Children, family members, and friends were identified as individuals who would approve participants' involvement in the NutriQuébec project, and the spouse was identified as an individual who would disapprove. However, half of the focus groups did not identify any individuals who would approve or disapprove their participation in the project. In fact, participants mentioned that they were not influenced by their acquaintances and that their decision to participate in the project would only depend on themselves.

Control Beliefs: Facilitators and Barriers to Participating in the NutriQuébec Project

The perceived facilitators were having the possibility to complete the questionnaires using alternative methods to internet and receiving a nonmonetary incentive. The alternative means to internet that were mentioned were having group sessions in predetermined places to complete questionnaires, having the possibility to complete paper questionnaires, receiving questionnaires by regular mail, and filling out questionnaires over the phone. The nonmonetary incentives included a point system, gift cards, a random drawing or contest, and a brief health assessment. Furthermore, participants also acknowledged that they would take part in the project even if no incentive was offered, because the project was important and interesting to them or because they simply did not consider the incentive to be necessary.

Other facilitators discussed were receiving a financial incentive, using simple questions, receiving general information related to food and health, having anonymous questionnaires, having access to a picture of the research team, having a reasonable time frame for completing questionnaires, having tutorials, and providing internet access:

> *Well, to clearly write "anonymous." You know, so that...yes, in bold, in capitals, you know so people say "ok, I am nobody when I answer, so I can be myself..."* [Group 2, participant #9]

The barrier that was most frequently mentioned in the focus groups was the lack of internet access. Other barriers discussed were the absence of a nonmonetary incentive, the lack of time (long-term commitment and questionnaires too long to complete), having limited access to internet, feeling that the study is useless, using complex vocabulary, or the 1-year follow-up leading to forgetting the study.

Security and Sharing of Data

Participants were asked to share their thoughts regarding security aspects related to data management. They mentioned that they would not be concerned about data security because of the signed consent form, if the data collected remained anonymous, and if less personal data such as dietary habits were to be collected as compared with more personal data such as presence of a disease:

> *No, anyways, it's always about food, there is nothing too personal about that. It doesn't matter that much.* [Group 3, participant #16]

Nevertheless, participants also stated that they felt more concerned about data security if health insurance numbers were collected as part of the NutriQuébec project, if the information was provided over the internet, or because they were afraid of data hacking.

Participants expressed no particular concern regarding the sharing of data among researchers as long as the data provided were not modified, because it would allow researchers to have a representative sample of the population of Québec and it would limit the number of research projects (and further solicitation). On the other hand, participants mentioned that they would not be comfortable with data sharing between researchers because it would increase the risk of data leakage.

Preferred Recruitment Methods

Participants were also asked to identify the recruitment methods that would most likely be successful in encouraging them to participate in the NutriQuébec project. Channeling recruitment efforts through community groups, social media, television, and brochures were identified as the most promising approaches by the participants. Social media included Facebook, YouTube, and Instagram. Television included news reports as well as other specific channels or television shows. Other recruitment methods mentioned were via posters, newsletters, schools, and hospitals.

Discussion

Summary

Our objective in undertaking this qualitative study in a low SES population was to identify the salient beliefs underlying attitudes, subjective norms, and PBC toward participation in a prospective Web-based study on nutrition and health. We also wanted to identify the most promising methods to maximize recruitment of individuals with a low SES in such a project. Results of this study are of relevance not only to the NutriQuébec project but also potentially to any other Web-based surveys that aim to include populations with a low SES.

Behavioral Beliefs: Perceived Advantages and Disadvantages to Participating in the NutriQuébec Project

Perceived advantages should be part of the key messages used to promote participation in a Web-based project. In our study, contributing to improving collective health, supporting research efforts, receiving a brief health assessment, acquiring and sharing knowledge, improving one's and the family's lifestyle habits, and making time for oneself have been identified by these focus groups as key advantages to participating in the NutriQuébec project. However, addressing these advantages is challenging in prospective studies on diet and health such as NutriQuébec because of the risk of interference with the normal trajectory of a cohort. Thus, the means chosen to promote participation in the Web-based project must not encourage nor promote any drastic changes in lifestyle. Yet, data have shown that providing information related to healthy lifestyle habits or a brief health assessment has only trivial long-term impact on a person's behavior [15].

Normative Beliefs: Approval and Disapproval of Significant Others Toward Participation in the NutriQuébec Project

The majority of focus groups were unable to identify persons who would approve or disapprove of their participation in the NutriQuébec project. In other words, participants felt that their decision to participate in the project would not be influenced by others. Therefore, acknowledging the subjective norms related to participation among individuals with a low SES in a prospective health study may not be necessary when choosing recruitment and retention methods.

Control Beliefs: Facilitators and Barriers to Participating in the NutriQuébec Project

The lack of internet access was identified as an important barrier to taking part in the NutriQuébec project. This is not surprising, considering the fact that groups with a low SES show the lowest proportions of internet access in their household [16]. This is consistent with data from Whitaker et al, who have also identified internet access as a barrier to taking part in Web-based studies [17]. In the province of Québec, 90% of households have declared having access to internet, but there is still a gradient related to SES within the population [16]. Indeed, it was estimated that only 59% of low-income families living in the province of Québec had access to internet compared with 98% of moderate-income families [16]. All focus groups suggested the need to allow alternative means to the internet to facilitate participation in the NutriQuébec project, thus highlighting a significant challenge in the context of conducting Web-based epidemiological studies. Nevertheless, the majority of participants (78%) easily had access to the internet at home or elsewhere in their community through local libraries, for example. Although internet access may be perceived as a barrier to participation of individuals with a low SES in Web-based studies, its impact in the coming years will most likely diminish as the internet becomes more broadly available for all segments of the population.

Receiving incentives was mentioned in the focus groups as a facilitator for taking part in the Web-based NutriQuébec project. These data are consistent with results from a previous study indicating that offering an incentive, such as a financial compensation or a gift card, is an effective method for encouraging low-income persons to participate in a study [18]. Hernando et al [19] also reported that the majority of participants would accept a monetary compensation, and this form of incentive was of greater interest among unemployed individuals. Even the use of a modest US $10 monetary incentive was shown to increase the number of returned surveys in a longitudinal study [20]. However, this is not a consistent finding, as Edwards et al [21] have shown that there is no evidence demonstrating that monetary incentives encourage participation in Web-based surveys. It is noteworthy that the monetary incentive was provided via an electronic fund transfer (PayPal), which may have been less accessible for some participants compared with a cash incentive. Providing nonmonetary incentives such as Amazon gift cards, lottery participation, and early grade feedback increased the odds of participants completing Web-based surveys by two-fold [21]. These data are consistent with data from our study in which focus groups expressed a stronger interest in receiving a nonmonetary incentive than a monetary incentive for participating in a lifestyle and health Web-based survey. Receiving a brief health assessment was also mentioned as a facilitator by the focus groups and, thus, is a good example of a promising nonmonetary incentive to facilitate participation in future Web-based studies.

The lack of time was identified as a barrier to possible participation in the NutriQuébec project. This was expressed in the context of the long-term commitment and the time required to complete all questionnaires, which is estimated to be 2 hours. Therefore, the idea of being considerably committed to the study may discourage participation and increase the likelihood of dropping out of the study. This is consistent with data from Hernando et al, who identified time commitment as a key barrier to participation in a longitudinal health study among vulnerable populations, more specifically among migrants [19]. This suggests that retention methods will need to be applied to favor long-term participation. A recent meta-analysis and systematic review on retention strategies in longitudinal cohort studies identified barrier reduction strategies as the most efficient predictor of improved retention [22]. Similarly, Méjean et al reported that reduced participant burden (convenience, simplicity, and not feeling judged) in a Web-based longitudinal study could help reducing attrition [23]. Strategically choosing effective retention strategies rather than increasing the number of strategies was also mentioned as a way to help increase retention rates [22].

All other facilitators and barriers have already been considered in the context of the NutriQuébec project and are applied when possible. For example, the NutriQuébec website uses simple sentences and avoids any complex vocabulary, it offers a thorough explanation of the participant's implication, and it introduces and describes the research team.

No participant spontaneously identified data security and data sharing among researchers as potential barriers to participating in a lifestyle and health Web-based survey. Concerns regarding security and data sharing were discussed only after the moderator of the focus groups specifically asked the participants their thoughts on these issues. Still, focus groups revealed no particular concern that suggests that only minimal information regarding data security and management needs to be provided to potential participants when undertaking a Web-based study on lifestyle and health.

Preferred Recruitment Methods

Social media such as Facebook, Instagram, YouTube, and other social media websites were identified in all focus groups as privileged recruitment channels among populations with a low SES. This is consistent with data from studies in other populations where the use of social media contributed to higher enrollment rates compared with other recruitment methods and was successful in recruiting hard-to-reach populations [24]. Indeed, 44% of low-income individuals in Québec connect to social media websites [5]. Using Facebook specifically has also been shown to maximize participation when recruiting hard-to-reach individuals [17,25,26]. In fact, Facebook is not only the most used social media among Canadians but also the most used among low-income families [5]. Studies have shown that using Facebook for recruitment is more cost-effective than traditional recruitment methods [26,27]. Nevertheless, data on the relative efficiency of social media to recruit hard-to-reach populations compared with more traditional recruitment methods remain limited and need further investigation [24].

Recruiting through community centers was identified as another method to reach out to the participants. This is consistent with data from non–Web-based studies that channeled recruitment through community centers to reach out to a vulnerable population. More specifically, advertising a study in community centers was more successful than advertising one in clinical facilities to recruit participants from an ethnic minority [28].

Television was mentioned as an efficient way to promote the study in the focus groups and, hence, recruitment and participation in the NutriQuébec project. This is consistent with a previous study in which the use of advertisement and promotion of a study on television reached the largest number of subjects and was an adequate recruitment method for unemployed individuals, low-income individuals, and individuals with only primary or secondary education compared with other channels [29].

Limitations

The 28 participants of the focus groups included both men and women and were heterogeneous in terms of computer skills. However, only one participant was aged under 30 years. We do not know the extent to which similar facilitating factors and barriers associated with participating in a Web-based project such as NutriQuébec will be identified by younger individuals with a low SES population. The data from this focus group study also need to be replicated in surveys with larger samples of individuals who will be representative of low SES populations. Finally, although there was evidence of saturation in terms of information yielded from the focus groups, there is a possibility that other facilitators and barriers toward participation in a Web-based study on diet and health would have been discussed if we had included more groups and individuals.

Conclusions

To our knowledge, this is the first study to document the beliefs and preferred recruitment methods of individuals with a low SES related to their intent to participate in a Web-based prospective study on lifestyle and health. This work is important because populations with a low SES, who are at greater risk of developing chronic diseases, are known to be underrepresented in health studies. Using the theory of planned behavior as a framework, we identified the lack of time, the absence of a nonmonetary incentive, and the lack of internet access as the main barriers. Receiving a brief health assessment (as a nonmonetary incentive) and having alternative means to internet for the completion of questionnaires were mentioned as the main facilitators. Preferred recruitment methods included social media, community centers, and television, and these methods should be considered when recruiting groups with a low SES. Although results from this study are specific to the NutriQuébec project, they are potentially relevant to other Web-based prospective studies aiming to recruit participants with a low SES. However, participants were recruited in the metropolitan area of Québec City and all spoke French. Additional studies in populations with a low SES from other sociocultural backgrounds may yield different results.

Acknowledgments

Financial support for this study was provided by the Réseau de recherche en santé cardiométabolique, diabète et obésité. The authors express their gratitude to the study participants, without whom the study would not have been possible.

Authors' Contributions

MC wrote a first draft of this paper. SD, SL, AB-G, and BL designed the study and obtained the funding. AL and CL coordinated the study. AL contributed to the transcriptions and content analysis of the focus groups with MC. All authors have reviewed and accepted the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview guide.
[DOCX File, 23 KB - formative_v3i4e13854_app1.docx]

References

1. Katikireddi SV, Whitley E, Lewsey J, Gray L, Leyland AH. Socioeconomic status as an effect modifier of alcohol consumption and harm: analysis of linked cohort data. Lancet Public Health 2017 Jun;2(6):e267-e276 [FREE Full text] [doi: 10.1016/S2468-2667(17)30078-6] [Medline: 28626829]
2. Department of Health and Social Services Publications. Plan d'action interministériel 2017-2020 : politique gouvernementale de prévention en santé : un projet d'envergure pour améliorer la santé et la qualité de vie de la population 2018 URL: http://publications.msss.gouv.qc.ca/msss/fichiers/2017/17-297-02W.pdf [accessed 2019-10-09] [WebCite Cache ID 76HXZ8Duq]
3. Martinez SA, Beebe LA, Thompson DM, Wagener TL, Terrell DR, Campbell JE. A structural equation modeling approach to understanding pathways that connect socioeconomic status and smoking. PLoS One 2018;13(2):e0192451 [FREE Full text] [doi: 10.1371/journal.pone.0192451] [Medline: 29408939]
4. Rolnitsky A, Kirtsman M, Goldberg HR, Dunn M, Bell CM. The representation of vulnerable populations in quality improvement studies. Int J Qual Health Care 2018 May 1;30(4):244-249. [doi: 10.1093/intqhc/mzy016] [Medline: 29447405]
5. Bourget C, Badraoui G. Cefrio. Médias sociaux et économie de partage en ligne au Québec 2016 URL: https://cefrio.qc.ca/media/1228/netendances_2016-medias-sociaux-et-economie-de-partage-en-ligne-au-quebec.pdf [accessed 2019-10-09] [WebCite Cache ID 76HU9Uz7I]
6. Toledano MB, Smith RB, Brook JP, Douglass M, Elliott P. How to Establish and Follow up a Large Prospective Cohort Study in the 21st Century--Lessons from UK COSMOS. PLoS One 2015;10(7):e0131521 [FREE Full text] [doi: 10.1371/journal.pone.0131521] [Medline: 26147611]
7. Smith B, Smith TC, Gray GC, Ryan MA, Millennium Cohort Study Team. When epidemiology meets the internet: web-based surveys in the Millennium Cohort Study. Am J Epidemiol 2007 Dec 1;166(11):1345-1354. [doi: 10.1093/aje/kwm212] [Medline: 17728269]
8. Nagler RH, Ramanadhan S, Minsky S, Viswanath K. Recruitment and retention for community-based eHealth interventions with populations of low socioeconomic position: strategies and challenges. J Commun 2013 Feb 1;63(1):201-220 [FREE Full text] [doi: 10.1111/jcom.12008] [Medline: 23439871]
9. Ejiogu N, Norbeck JH, Mason MA, Cromwell BC, Zonderman AB, Evans MK. Recruitment and retention strategies for minority or poor clinical research participants: lessons from the Healthy Aging in Neighborhoods of Diversity across the Life Span study. Gerontologist 2011 Jun;51(Suppl 1):S33-S45 [FREE Full text] [doi: 10.1093/geront/gnr027] [Medline: 21565817]
10. Ellard-Gray A, Jeffrey NK, Choubak M, Crann SE. Finding the hidden participant. Int J Qual Methods 2015;14(5):160940691562142. [doi: 10.1177/1609406915621420]
11. Bonevski B, Randell M, Paul C, Chapman K, Twyman L, Bryant J, et al. Reaching the hard-to-reach: a systematic review of strategies for improving health and medical research with socially disadvantaged groups. BMC Med Res Methodol 2014 Mar 25;14:42 [FREE Full text] [doi: 10.1186/1471-2288-14-42] [Medline: 24669751]
12. Ajzen I. The theory of planned behavior. Organ Behav Hum Decis Process 1991 Dec;50(2):179-211. [doi: 10.1016/0749-5978(91)90020-T]
13. Elo S, Kyngäs H. The qualitative content analysis process. J Adv Nurs 2008 Apr;62(1):107-115. [doi: 10.1111/j.1365-2648.2007.04569.x] [Medline: 18352969]
14. Saunders B, Sim J, Kingstone T, Baker S, Waterfield J, Bartlam B, et al. Saturation in qualitative research: exploring its conceptualization and operationalization. Qual Quant 2018;52(4):1893-1907 [FREE Full text] [doi: 10.1007/s11135-017-0574-8] [Medline: 29937585]
15. Kelly MP, Barker M. Why is changing health-related behaviour so difficult? Public Health 2016 Jul;136:109-116 [FREE Full text] [doi: 10.1016/j.puhe.2016.03.030] [Medline: 27184821]
16. Bourget C. Cefrio. 2017. Portrait numérique des foyers québécois URL: https://cefrio.qc.ca/fr/enquetes-et-donnees/netendances2017-portrait-numerique-foyers-quebecois/ [accessed 2019-10-09] [WebCite Cache ID 76HUcPI6k]
17. 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 Aug 28;19(8):e290 [FREE Full text] [doi: 10.2196/jmir.7071] [Medline: 28851679]
18. Nicholson LM, Schwirian PM, Groner JA. Recruitment and retention strategies in clinical studies with low-income and minority populations: Progress from 2004-2014. Contemp Clin Trials 2015 Nov;45(Pt A):34-40. [doi: 10.1016/j.cct.2015.07.008] [Medline: 26188163]
19. Hernando C, Sabidó M, Casabona J. Facilitators and barriers of participation in a longitudinal research on migrant families in Badalona (Spain): A qualitative approach. Health Soc Care Community 2018 Jan;26(1):e64-e74. [doi: 10.1111/hsc.12478] [Medline: 28737003]
20. Yu S, Alper HE, Nguyen A, Brackbill RM, Turner L, Walker DJ, et al. The effectiveness of a monetary incentive offer on survey response rates and response completeness in a longitudinal study. BMC Med Res Methodol 2017 Apr 26;17(1):77 [FREE Full text] [doi: 10.1186/s12874-017-0353-1] [Medline: 28446131]
21. 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]

22. Teague S, Youssef GJ, Macdonald JA, Sciberras E, Shatte A, Fuller-Tyszkiewicz M, SEED Lifecourse Sciences Theme. Retention strategies in longitudinal cohort studies: a systematic review and meta-analysis. BMC Med Res Methodol 2018 Nov 26;18(1):151 [FREE Full text] [doi: 10.1186/s12874-018-0586-7] [Medline: 30477443]
23. Méjean C, Edelenyi F, Touvier M, Kesse-Guyot E, Julia C, Andreeva VA, et al. Motives for participating in a web-based nutrition cohort according to sociodemographic, lifestyle, and health characteristics: the NutriNet-Santé cohort study. J Med Internet Res 2014 Aug 7;16(8):e189 [FREE Full text] [doi: 10.2196/jmir.3161] [Medline: 25135800]
24. Topolovec-Vranic J, Natarajan K. The use of social media in recruitment for medical research studies: a scoping review. J Med Internet Res 2016 Nov 7;18(11):e286 [FREE Full text] [doi: 10.2196/jmir.5698] [Medline: 27821383]
25. Kayrouz R, Dear BF, Karin E, Titov N. Facebook as an effective recruitment strategy for mental health research of hard to reach populations. Internet Interv 2016 May;4:1-10 [FREE Full text] [doi: 10.1016/j.invent.2016.01.001] [Medline: 30135786]
26. 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]
27. Prescott TL, Phillips G II, DuBois LZ, Bull SS, Mustanski B, Ybarra ML. Reaching Adolescent Gay, Bisexual, and Queer Men Online: Development and Refinement of a National Recruitment Strategy. J Med Internet Res 2016 Aug 04;18(8):e200 [FREE Full text] [doi: 10.2196/jmir.5602] [Medline: 27492781]
28. Ebenibo S, Edeoga C, Ammons A, Egbuonu N, Dagogo-Jack S, Pathobiology of Prediabetes in a Biracial Cohort (POP-ABC) Research Group. Recruitment strategies and yields for the Pathobiology of Prediabetes in a Biracial Cohort: a prospective natural history study of incident dysglycemia. BMC Med Res Methodol 2013 May 10;13:64 [FREE Full text] [doi: 10.1186/1471-2288-13-64] [Medline: 23663750]
29. Kesse-Guyot E, Andreeva V, Castetbon K, Vernay M, Touvier M, Méjean C, et al. Participant profiles according to recruitment source in a large Web-based prospective study: experience from the Nutrinet-Santé study. J Med Internet Res 2013 Sep 13;15(9):e205 [FREE Full text] [doi: 10.2196/jmir.2488] [Medline: 24036068]

Abbreviations

PBC: perceived behavioral control
SES: socioeconomic status

Edited by G Eysenbach; submitted 28.02.19; peer-reviewed by A Nguyen, A Cyr; comments to author 22.06.19; revised version received 15.08.19; accepted 01.10.19; published 10.12.19.

Please cite as:
Côté M, Lapointe A, Laramée C, Lemieux S, Desroches S, Belanger-Gravel A, Lamarche B
Beliefs Related to Participation in a Large Web-Based Prospective Survey on Diet and Health Among Individuals With a Low Socioeconomic Status: Qualitative Study
JMIR Form Res 2019;3(4):e13854
URL: http://formative.jmir.org/2019/4/e13854/
doi:10.2196/13854
PMID:31821149

Original Paper

Web-Based Module for the Collection of Electronic Patient-Reported Outcomes in People Living With HIV in Nouvelle Aquitaine, France: Usability Evaluation

Diana Barger[1], BA, MSc; Olivier Leleux[1], MSc; Valérie Conte[2], MSc; Vincent Sapparrart[2], MSc; Marie Gapillout[2], BSc; Isabelle Crespel[3], RN; Marie Erramouspe[4], BA; Sandrine Delveaux[3], RN; Linda Wittkop[1,5], MD, PhD; François Dabis[1,3], MD, PhD; Fabrice Bonnet[1,3,6], MD, PhD

[1]University of Bordeaux, ISPED, Inserm, Bordeaux Population Health Research Center, Team MORPH3EUS, UMR 1219, Bordeaux, France
[2]Centre de Recherche et Développement en Informatique Médicale, University of Bordeaux, Bordeaux, France
[3]CHU de Bordeaux, COREVIH Nouvelle Aquitaine, Bordeaux, France
[4]AIDES Nouvelle Aquitaine, Bordeaux, France
[5]CHU de Bordeaux, Service d'information médicale, Pôle de sante publique, Bordeaux, France
[6]CHU de Bordeaux, Service de médecine interne et maladie infectieuses, Bordeaux, France

Corresponding Author:
Diana Barger, BA, MSc
University of Bordeaux
ISPED, Inserm, Bordeaux Population Health Research Center
Team MORPH3EUS, UMR 1219
146 rue Leo Saignat CS61292
Bordeaux, F-33000
France
Phone: 33 0557579291
Email: diana.barger@u-bordeaux.fr

Abstract

Background: Patient-reported outcomes (PROs) can be of great value for both research and chronic disease management. We developed a new module of the ANRS CO3 Aquitaine cohort study's Web-based data capture and visualization solution (APPEGE 2.0) for the collection of electronic PROs among people living with HIV cared for in Nouvelle Aquitaine, France.

Objective: This study aimed to evaluate the usability of 2 successively developed prototypes of ARPEGE 2.0's electronic PROs module before launching a pilot study, owing to the novelty of the proposed data collection method for our setting and specific characteristics of the target population.

Methods: A total of 2 sequential rounds of empirical, task-based usability evaluations were conducted, involving 8 research staff and then 7 people living with HIV. Evaluators provided written feedback during round 1 and oral feedback during round 2. Evaluators who completed the full set of tasks responded to the System Usability Scale (SUS). We assessed changes in SUS scores between rounds and concluded usability testing when SUS scores reached a ceiling effect, defining good usability a priori as a usability score of 70.

Results: Insights were generated regarding the visibility of system status and the match between the system and the real world that improved the module's usability. Research staff evaluators reported mean SUS scores of 65 (SD 18.87) and patient evaluators reported mean SUS scores of 85 (SD 5.4; *P*=.032).

Conclusions: Software modifications, informed by successive rounds of usability testing, resulted in sufficient gains in usability to undertake piloting. Insights generated during evaluations prompted us to find the appropriate balance between optimal security and ease of use.

Trial Registration: ClinicalTrials.gov NCT03296202; https://clinicaltrials.gov/ct2/show/NCT03296202

International Registered Report Identifier (IRRID): RR2-10.2196/10.2196/resprot.9439

(JMIR Form Res 2019;3(4):e15013) doi:10.2196/15013

KEYWORDS

patient reported outcome measures; patient generated health data; quality of life

Introduction

HIV, once fatal, is now a manageable chronic illness [1]. In Western Europe, the majority of individuals who received a diagnosis of HIV are in care and on potent antiretroviral therapy, which prevents serious diseases both related and unrelated to AIDS [2]. The improved prognosis and the increased life expectancy of people living with HIV (PLWH) makes preserving health and ensuring good quality of life the cornerstone of their care [3-5]. One strategy to help providers respond to PLWH's evolving needs and improve the quality and efficiency of their overall care is collecting and using patient-reported outcomes (PROs) [6].

PROs or "any report of the status of the patient's health condition that comes directly from the patient, without interpretation of the patient's response by a clinician or anyone else" [7] have been used extensively in clinical research [6]. PROs can be used at the population level for research and to improve the quality of care or at the individual level to support clinical decision making [8]. Their use may allow for more accurate symptom detection, better patient-provider communication, and improved outcomes [9]. Logistical, technical, and ideological barriers have nevertheless limited their use in routine care [10]. The adoption of electronic medical records coupled with the adaptation of paper questionnaires to computerized and internet-based formats may help overcome these barriers [10,11].

With evidence from the United States suggesting that the collection of PROs by using touchscreen-based information technology was both feasible and of value for both research and clinical HIV care [12-14], a prototype of an electronic PRO module linked to the ANRS CO3 Aquitaine cohort's data capture and visualization system (ARPEGE 2.0) was developed in 2017 [15]. As the overall usefulness of interactive health care applications or their usability is likely to affect their acceptability and adoption, usability evaluations of 2 successively developed prototypes of the ARPEGE 2.0 solution were conducted in preparation for a pilot study [15].

Methods

This formative research study took place in Bordeaux, France, at the Inserm UMR 1219-Bordeaux Population Health Research Centre and the St André Bordeaux University Hospital. It was designed as part of the ANRS CO3 Aquitaine study, an open, prospective hospital-based cohort of PLWH in care in 13 clinics in southwestern France. A local institutional review board approved the study's protocol (Comité de Protection de Personnes Sud-Ouest et Outre-Mer III) on September 18, 2017.

Description of the Electronic Patient-Reported Outcome Module Powered by ARPEGE 2.0

ARPEGE 1.0 is a proprietary, secure, electronic case report form developed in Microsoft ASP.NET (WebForm). Data are stored within a Microsoft SQL Server 2014–based data management system. The ANRS CO3 Aquitaine cohort relies on ARPEGE 1.0 for data capture. Clinical data, extracted from both medical records and laboratory data, derived from the hospital's laboratory information management systems, have been collected systematically since 1987 and electronically via ARPEGE 1.0 since 2013 with the support of Clinical Research Associates. ARPEGE 2.0 is a generic Web-based data capture and visualization system also developed in Microsoft ASP.NET (WebForm). ARPEGE 2.0 has enabled the creation of the module for the collection of electronic PROs in routine care for observational research and, ultimately, clinical care.

The content of ARPEGE 2.0's initial electronic PRO module is based on current treatment guidelines for people being treated for HIV and associated comorbidities [16]. Prototyping was carried out over 2017 with the support and regular feedback from a working group comprising research staff, local stakeholders, and end users (clinicians and patient representatives). The questionnaires were evaluated individually according to their psychometric properties, administration method, and length. The following areas are covered by the electronic PRO module: socioeconomic status and individual social and material deprivation [17], multidimensional quality of life (WHOQOL-HIV BREF) [18], treatment burden (Treatment Burden Questionnaire) [19], physical activity (the Short Version of the International Physical Activity Questionnaire), alcohol use and screening for at-risk drinking behavior (Alcohol Use Disorders Identification Test Consumption, Fast Alcohol Consumption Evaluation) [20], tobacco and nicotine use and screening for tobacco dependency (Fagerström), cannabis (Cannabis Abuse Screening Test) and drug use, and, finally, depression (Patient Health Questionnaire) [21].

Conditional branching was used where appropriate. The module also allows patients to report any other treatment-related issues in a free text field. Where applicable, the International Society for Pharmacoeconomics and Outcomes Research ePRO Task Force's recommendations on adapting paper-based instruments were followed, ensuring that data produced are equivalent or superior to those generated from paper-based administration methods [22].

Recruitment

Nielsen's recommendations that favor conducting several iterative studies, each with a small number of participants, were adopted [23]. In round 1 (May 2018), evaluators were employees of the Inserm UMR 1219 Bordeaux Population Health Research Center or affiliated with the project, referred to herein as *research staff*. In round 2 (June 2018), a convenience sample of PLWH being cared for at the St André Bordeaux University Hospital was identified by clinical staff either before or during their routine visit.

Procedure

The evaluation procedure differed between round 1 and round 2. However, for both rounds, oral consent was obtained. It was then explained that each study participant (evaluator) would be provided with a unique identifier, which would allow him/her to create a personal account and access the questionnaires. Evaluators were shown the study-specific brochure where the number would be written on a detachable coupon (Multimedia Appendix 1). They were asked to complete 5 tasks: (1) navigate between pages on the publicly available website and locate key information, (2) create a user account, (3) confirm their account, (4) initiate the electronic PRO assessment, and (5) complete the electronic PRO assessment. Whether or not each task was completed with ease, assistance or not was monitored, and a score of 2 to 0 was attributed (2=the task was completed with ease and 0=it was not completed). The highest possible score was therefore 10, and the lowest score was 0. Neither round 1 nor round 2 evaluators were compensated.

In round 1, research staff were provided with instructions detailing the background of the study and how it would be implemented in a clinical setting. Evaluators were given a link to a staging version of the electronic PRO module. They were asked to complete the previously described tasks. They then responded to an Web-based questionnaire that included the System Usability Scale (SUS), a widely used robust tool for measuring usability. It consists of 10 items with 5 response options, from strongly agree to strongly disagree [24,25]. Evaluators provided written feedback in an open text field and by email.

In round 2, patients participated in one-on-one testing sessions, lasting between 1 and 2 hours, with a researcher in a dedicated, private space at the hospital (June 2018). The researcher based each session on a standardized qualitative interview guide. A personal computer (Mac Book Air) with access to the staging site was provided to complete the study tasks. Patient evaluators were also allowed to complete the questionnaire on their personal smartphones, matching how the electronic PRO module might be accessed in routine care. Evaluators were instructed to use the *think aloud* method, in which users are asked to verbalize all thoughts as they interact with the system while carrying out tasks. Subsequently, those who completed all tasks responded orally to the SUS and provided open-ended feedback [24]. All sessions were audio recorded, and field notes were taken.

Analysis

Task completion and SUS scores were calculated for each evaluator, and means and standard deviations were calculated for each round. We performed a *t* test assuming unequal variance to determine if each round of testing produced significant difference in the mean SUS scores. A priori, we defined *success* in usability when the SUS score reached a ceiling effect, with a minimum score of 70—generally accepted as a cut-off for *good* usability [26].

Qualitative analysis included review of written feedback, audio recording–enhanced field notes, and responses to open-ended questions. We performed thematic content analysis on written feedback and audio recording–enhanced field notes, abstracting and compiling emerging themes from each round of testing. These are reported according to Nielsen’s usability heuristic categories [27].

Results

Overview

Table 1 presents evaluators’ characteristics and mean task completion scores for rounds 1 and 2. The majority of round 1 evaluators were women (7/8). They reported using computers either regularly (5/8) or often (3/8). In all, 5 out of 7 round 2 evaluators were men. A total of 3 reported using a computer regularly, 3 often, and 1 never. Overall, mean task completion scores were 7.8 (out of 10) in round 1 and 7.1 in round 2. In round 1, 7 evaluators completed all tasks compared with 4 out of 7 in round 2. Task completion was hampered owing to 2 evaluators being locked out of their accounts and 1 evaluator being unable to complete tasks owing to poor eyesight. This evaluator was attributed 0 on all tasks.

The usability insights uncovered during the 2 rounds of usability evaluations together with the solutions adopted are presented in Table 2.

Table 1. Evaluator characteristics and task scores.

Evaluator characteristics	Task 1—information found	Task 2—account created	Task 3—account confirmed	Task 4—PRO[a] assessment initiated	Task 5—PRO assessment completed	Total
Round 1 (N=8)	2.0	1.1	1.4	1.5	1.8	7.8
Male (n=1)	2.0	0.0	0.0	0.0	0.0	2.0
30-40	2.0	0.0	0.0	0.0	0.0	2.0
Female (n=7)	2.0	1.3	1.6	1.7	2.0	8.6
<30	2.0	1.5	1.5	1.5	2.0	8.5
30-40	2.0	1.3	1.3	1.7	2.0	8.3
41-50	2.0	1.0	2.0	2.0	2.0	9.0
>50	2.0	1.0	2.0	2.0	2.0	9.0
Round 2 (N=7)	1.7	1.1	1.4	1.7	1.1	7.1
Male (n=5)	1.6	1.2	1.6	1.6	1.6	7.6
<30	2.0	1.5	2.0	2.0	2.0	9.5
30-40	2.0	2.0	2.0	2.0	2.0	10.0
>50	1.0	0.5	1.0	1.0	1.0	4.5
Female (n=2)	2.0	1.0	1.0	2.0	0.0	6.0
>50	2.0	1.0	1.0	2.0	0.0	6.0

[a] PRO: patient-reported outcome.

Table 2. Usability insights per round and solution adopted according to Nielsen's usability heuristics.

Usability categories	Round 1—research staff	Round 2—patients	Solution
Visibility of system status	Login procedure was confusing owing to the complexity of password, requiring 2 symbols	Challenges adhering to password requirements for certain patients	Password requirements were spelled out for users in bold. A password visualization button was also added to the password field to allow users to ensure that passwords created matched before registering their account
	—[a]	Unclear whether the QuAliV number (required for creating the account) is case sensitive	Information incorporated into the presentation of the study to participants
	Validation of questionnaire unclear	—	Information buttons added to the home page of the electronic PRO[b] module instructing users on how the questionnaires functioned and reminding them to *submit* their completed questionnaires. The button was also relabeled to make its functionality clearer
Match between system and the real world	Date picker was in English and began in 2018, requiring users to click to go back in time	—	The date picker was replaced with a French version. It allowed users to type in their birth dates without using the calendar
	—	Nonmutually exclusive modalities or response missing	Minor modifications made to question modalities to ensure clarity
	—	Issues stemming from the translation of questionnaire from English to French	Further cognitive debriefing with native speakers to identify the best translation of the item in question
	—	Difficulties understanding the meaning of certain questions	Less formal language substituted where possible and examples given to facilitate the comprehension of certain questions
	—	Confirmation of account on one's smartphone (email) resulted in being locked out of one's account on another device	Automatic connection to the site after creating one's account deleted (temporarily) to avoid users locking themselves out of their account. Users must reenter their username and password
User control and freedom	Need for returning back to last page completed in the questionnaire	—	The user is now redirected back to the most recent page completed within each questionnaire. Scrolling from one page of a questionnaire to another automatically saves entered data
	Radio button could not be unclicked or erased	—	A refresh button was added to each item to allow users to erase their responses and therefore leave items unanswered
	Questionnaire opens in a pop-up window whose size cannot be modified	—	Double checked to ensure that text could be easily read in each window
	—	Unclear whether users had to provide first and last name	We added text indicating that typing one's first and last name was optional
Consistency and standards	Format of certain questions was noted as being inconsistent between questionnaires. Yes/No questions appeared in a table format as soon as they used the same response thesaurus	—	Minor improvements in formatting were made where possible. Further development required to accommodate this change in the longer run
	Typos in certain questions were identified	—	—
Error prevention	Aberrant response possible for certain free text fields	—	Stricter constraints added
	The password required was complex. Instructions on password requirements were missing from the account creation page	—	Instructions on password requirements added
	Need to clarify units in free text fields		Units added in gray in each text field

Usability categories	Round 1—research staff	Round 2—patients	Solution
	Need to indicate which questions were mandatory in the questionnaire. Need to indicate when multiple answers could be given	—	An asterisk was added to indicate which questions were mandatory. The user is sent back to mandatory questions before being allowed to progress in the questionnaire. These questions were marked in red to indicate that they were mandatory
Recognition rather than recall	Automatic logout obligations meant that users could not reconnect to their accounts for 20 min if they left the page, resulting in certain evaluators being locked out of their account	—	Error message added to the module explaining that users would be able to reaccess their accounts after 20 min
Flexibility and efficiency of use	Errors encountered with the progress bar depending on responses to questions	—	Questionnaires are programmed to open successively
	Errors on certain Web browsers	—	Further trouble shooting using full array of browsers and devices
Aesthetic and minimalist design	Methods for completing a visual analog scale unclear as definition of extreme values was missing, and a *not applicable* box was not included	—	An 11-point radio button scale was proposed as a temporary solution
	The IPAQ[c] questionnaire was difficult to read on the pop-up screen	—	Alternative formatting used to improve readability
Help users with errors	Need to flag missed items	—	Progression bar for each questionnaire goes from orange to green as soon as all nonconditional questions are answered. Users are directed to unanswered obligatory questions upon attempting to go on to the next page of the questionnaire
Help and documentation	Information missing from different links (contact and preferences)	—	—
	Print button of informed nonopposition did not function correctly	—	—

[a] Not applicable.

[b] PRO: patient-reported outcome.

[c] IPAQ: International Physical Activity Questionnaire.

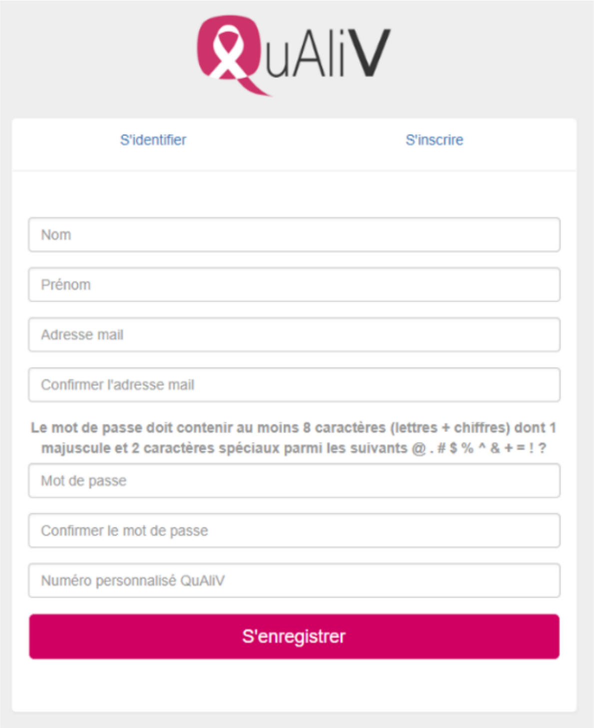
What Worked

The first task involved navigating the external website that patients would access from home, unassisted, to create their account. Users found the information provided on the external website quickly and found its structure clear. All users quickly understood how the attributed unique identifier would be used to create their personal account. Once users had created their account, efforts to guide him/her through electronic PROs by having each questionnaire open one after the other appeared to work well. The use of stoplight-style color coding and a progress bar allowed users to see if they had missed a question and helped them recognize, diagnose, and recover from errors seamlessly. The order of the PROs was received positively by users and therefore remained unchanged between prototype versions.

What Did Not Work

The account creation task was the most challenging for users. One of the issues identified was the complexity of the password requirements. The password had to be entered twice and contain at least 8 alphanumeric characters, including 2 special characters and a capital letter (Figure 1). Many evaluators, both research staff and patients, attempted this step more than once. We clarified the password requirements and ensured that error messages were informative regarding the system status, and we made it possible to visualize the password after round 1 (Figure 2). As errors still occurred, we added additional error prevention features. The password is validated as the user types as opposed to the user receiving an error message upon clicking *register* (Multimedia Appendix 2).

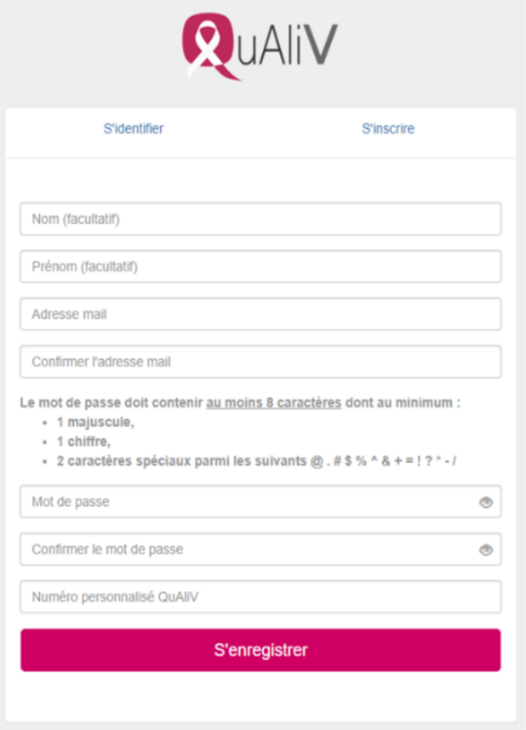
Figure 1. Initial log-in page (round 1).



Le mot de passe doit contenir au moins 8 caractères (lettres + chiffres) dont 1 majuscule et 2 caractères spéciaux parmi les suivants @ . # $ % ^ & + = ! ?

Fields (top to bottom) : Name, First Name, E-mail address, Confirm e-mail address, Password, Confirm password, QuAliV Number.
Password instructions (in bold): The password must contain at least 8 characters (letters and numbers) of which one must be uppercase and two special characters among the following @ . # $ ^ & + = ! ?

Figure 2. Revised log-in page (round 2).



Le mot de passe doit contenir <u>au moins 8 caractères</u> dont au minimum :

- **1 majuscule,**
- **1 chiffre,**
- **2 caractères spéciaux parmi les suivants @ . # $ % ^ & + = ! ? * - /**

Fields (top to bottom) : Name (optional), First Name (optional), E-mail address, Confirm e-mail address, Password, Confirm password, QuAliV Number.
Password instructions (in bold): The password must contain at least 8 characters : 1 uppercase, 1 number and 2 special characters among the following @ . # $ ^ & + = ! ? * /

A login problem, also detected during the second round of usability testing, was being *locked out* of one's account accidentally. This issue arose from a security measure included in the electronic PRO module's design. Users were logged out automatically after a period of 20 min of inactivity. If users accidently left the page without logging out of their accounts, they could no longer log back in owing to the Bordeaux University servers' restrictions. If the user attempted to return to their account, they received an error message indicating that they were already connected. This issue could not be resolved without completely relaxing the automatic logout timeframe (shortening it). We therefore modified the error message indicating that the user could access their account again in 20 min.

System Usability Scale Scores

In round 1, experts reported mean SUS scores of 65 (SD 18.87), and patients, in round 2, reported mean SUS scores of 85 (SD 5.4) (P=.032).

Discussion

Principal Findings

Iterative usability evaluations of 2 successively developed prototypes allowed us to see how easy our electronic PRO module was to use and identify when and where users encountered problems or experienced confusion. We were able to improve the module's usability markedly, specifically the visibility of system status and the match between the system and the real world, and take into account the specific needs of our patient population (their level of computer literacy and age) and the specificities of our clinical setting. Finally, we were pushed to find the appropriate balance between optimal security and ease of use.

Unlike PRO collection methods employed in clinics in the United States [12-14], where patients complete an electronic PRO assessment by using touchscreen information technology with the assistance of a research assistant/administrator at clinics, we aimed to design a Web-based *Bring Your Own Device* solution. We therefore assumed that the majority of users would have access to a smartphone or personal computer with a reliable internet connection. The proposed solution, developed in-house, had to work well enough to allow a group of users, with varying levels of computer familiarity, to use it with little to no assistance.

Strengths and Limitations

Some caveats should be considered in the interpretation of our results. We conducted the first round of usability testing in a sample of research staff who may not fully represent end users. This strategy, recognized as an easy way of catching obvious usability issues, resulted in high-quality, detail-oriented, and exhaustive feedback, allowing for a number of basic usability problems to be resolved before evaluations with patients. Most evaluators were comfortable using computers and the internet. They may not fully reflect the diversity of the cohort of PLWH in the region. More purposeful sampling of evaluators with lower computer literacy may have resulted in the detection of additional usability insights.

In round 2, we used the *think aloud* method. This method has been known to slow the thought process and increase mindfulness, which might prevent errors that might have normally occurred [28]. However, when evaluators are asked to perform simple tasks, the method has been shown to have no effect on user performance [29]. We opted for this method as the tasks were not considered complex.

Conclusions

Nevertheless, software modifications, informed by successive rounds of usability testing, resulted in sufficient gains in usability to undertake piloting.

Acknowledgments

Alain Volny-Anne (European AIDS Treatment Group) and Eugenie Destandau provided valuable feedback on the design and content and the QuAliV website and electronic PRO module. The authors would like to thank all of the evaluators, both research staff and patients, who contributed to this study. They thank the scientific committee, clinical sites and their investigators, and clinical research associates, further cited, who have been involved in the successful implementation of the subsequent phase of the study.

ANRS CO3 Aquitaine Cohort—Scientific Committee: F Bonnet (Principal Investigator), L Wittkop (Methodologist); C Cazanave, V Gaborieau, M Hessamfar, E Lazaro, G Le Moal, D Malvy, P Mercié, D Neau, MO Vareil, I Pellegrin, P Blanco, ME Lafon, P Bellecave, S Bouchet, D Breilh, D Lacoste, S Lawson-Ayayi, A Gimbert, S Desjardin, L Lacaze-Buzy, V Petrov-Sanchez, L Marchand, A Perrier, F Le Marec, and O Leleux.

Clinical Sites and Investigators: Hôpital Saint André, CHU de Bordeaux, Médecine Interne et Maladies Infectieuses, (F Bonnet, N Bernard, D Dondia, P Duffau, I Faure, M Hessamfar, D Lacoste, P Mercié, P Morlat, F Paccalin, MC Pertusa, MA Vandenhende, E Riebero, and C Rivoisy); Hôpital Pellegrin, CHU de Bordeaux, Maladies Infectieuses et Tropicales, (C Cazanave, FA Dauchy, A Desclaux, M Dupon, H Dutronc, D Neau, D Malvy, A Ochoa, T Pistone, MC Receveur, G Wirth); Hôpital Haut-Lévêque, CHU de Bordeaux, Médecine Interne et Maladies Infectieuses, (C Greib, E Lazaro, JL Pellegrin, JF Viallard); Hôpital d'Agen, Médecine Interne (Y Imbert, M Thierry-Mieg, P Rispal); Hôpital de Libourne, Médecine Interne (O Caubet, H Ferrand, S Tchamgoué); Hôpital de Bayonne, Maladies Infectieuses (S Farbos, MO Vareil, H Wille); Hôpital de Dax, Médecine Interne et Maladies Infectieuses, (K Andre, L Caunegre, Y Gerard, F Osorio-Perez); Hôpital Saint-Cyr/Villeneuve-sur-Lot, Maladies Infectieuses, (I Chossat) ; Hôpital de Mont de Marsan, Médecine Interne et Maladies Infectieuses, (G Iles, Y Gerard, M

Labasse-Depis, F Lacassin); Hôpital d'Arcachon, Médecine Interne, (A Barret, C Courtault) ; Hôpital de Périgueux, Médecine Interne et Maladies Infectieuses, (N Berthol, B Cougoul, P Lataste, J Marie, N Rouanes); Hôpital de Pau, Médecine Interne et Maladies Infectieuses, (G Dumondin, V Gaborieau); Hôpital d'Orthez, Médecine Interne, (Y Gerard).

Clinical Research Associates: S Delveaux, B Uwamaliya, K Zara, A Pougetoux, F Diarra, C Hanapier, MJ Blaizeau, M Decoin, E Lenaud, and S Lawson-Ayayi.

Project Team: A Perrier (Data Manager), F Le Marec (Statistician), and O Leleux (Project Leader).

The Aquitaine ANRS Cohort is sponsored by the Bordeaux University Hospital and funded by the ANRS (France REcherche Nord&Sud Sida-hiv Hépatites) and the Bordeaux University Hospital. The cohort is coordinated from within the Inserm UMR 1219—Bordeaux Population Health Research Centre. Seed funding was granted by the ANRS in 2017 via the CSS-5 call to develop the electronic PRO module. Diana Barger was awarded a 36-month *young researcher* grant from Sidaction to design and conduct a study on quality of life in PLWH within the ANRS CO3 Aquitaine cohort as part of her doctoral research.

Conflicts of Interest

DB has received a speaking fee from Gilead. FB declares to have received reimbursement for attending a symposium from ViiV Healthcare, Gilead, Bristol-Myers Squib, Merck and Janssen, speaking fee and consultancy fee from ViiV Healthcare, Gilead, Bristol-Myers Squib, Merck and Janssen, and funds for research from Gilead and ViiV Healthcare. The other authors do not have any conflict of interest to declare.

Multimedia Appendix 1
Patient brochure with unique identifier.
[PNG File , 478 KB - formative_v3i4e15013_app1.png]

Multimedia Appendix 2
Login page with password verification.
[PNG File , 234 KB - formative_v3i4e15013_app2.png]

References

1. Deeks SG, Lewin SR, Havlir DV. The end of AIDS: HIV infection as a chronic disease. Lancet 2013 Nov 2;382(9903):1525-1533 [FREE Full text] [doi: 10.1016/S0140-6736(13)61809-7] [Medline: 24152939]
2. INSIGHT START Study Group, Lundgren JD, Babiker AG, Gordin F, Emery S, Grund B, et al. Initiation of antiretroviral therapy in early asymptomatic HIV infection. N Engl J Med 2015 Aug 27;373(9):795-807 [FREE Full text] [doi: 10.1056/NEJMoa1506816] [Medline: 26192873]
3. Rasmussen LD, Obel N. How do we preserve health among adults living with HIV? Lancet HIV 2019 Feb;6(2):e69-e70. [doi: 10.1016/S2352-3018(18)30336-9] [Medline: 30683626]
4. Antiretroviral Therapy Cohort Collaboration. Survival of HIV-positive patients starting antiretroviral therapy between 1996 and 2013: a collaborative analysis of cohort studies. Lancet HIV 2017 Aug;4(8):e349-e356 [FREE Full text] [doi: 10.1016/S2352-3018(17)30066-8] [Medline: 28501495]
5. Lazarus JV, Safreed-Harmon K, Barton SE, Costagliola D, Dedes N, Valero J, et al. Beyond viral suppression of HIV - the new quality of life frontier. BMC Med 2016 Jun 22;14(1):94 [FREE Full text] [doi: 10.1186/s12916-016-0640-4] [Medline: 27334606]
6. Engler K, Lessard D, Lebouché B. A review of HIV-specific patient-reported outcome measures. Patient 2017 Apr;10(2):187-202. [doi: 10.1007/s40271-016-0195-7] [Medline: 27637488]
7. Food and Drug Administration. 2009 Oct. Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims URL: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/patient-reported-outcome-measures-use-medical-product-development-support-labeling-claims [accessed 2019-11-12]
8. Snyder CF, Aaronson NK, Choucair AK, Elliott TE, Greenhalgh J, Halyard MY, et al. Implementing patient-reported outcomes assessment in clinical practice: a review of the options and considerations. Qual Life Res 2012 Oct;21(8):1305-1314. [doi: 10.1007/s11136-011-0054-x] [Medline: 22048932]
9. Greenhalgh J, Long AF, Flynn R. The use of patient reported outcome measures in routine clinical practice: lack of impact or lack of theory? Soc Sci Med 2005 Feb;60(4):833-843. [doi: 10.1016/j.socscimed.2004.06.022] [Medline: 15571900]
10. Jones JB, Snyder CF, Wu AW. Issues in the design of internet-based systems for collecting patient-reported outcomes. Qual Life Res 2007 Oct;16(8):1407-1417. [doi: 10.1007/s11136-007-9235-z] [Medline: 17668293]
11. Gensheimer SG, Wu AW, Snyder CF, PRO-EHR Users' Guide Steering Group, PRO-EHR Users' Guide Working Group. Oh, the places we'll go: patient-reported outcomes and electronic health records. Patient 2018 Dec;11(6):591-598. [doi: 10.1007/s40271-018-0321-9] [Medline: 29968179]

12. Crane HM, Lober W, Webster E, Harrington RD, Crane PK, Davis TE, et al. Routine collection of patient-reported outcomes in an HIV clinic setting: the first 100 patients. Curr HIV Res 2007 Jan;5(1):109-118. [doi: 10.2174/157016207779316369] [Medline: 17266562]
13. Kozak MS, Mugavero MJ, Ye J, Aban I, Lawrence ST, Nevin CR, et al. Patient reported outcomes in routine care: advancing data capture for HIV cohort research. Clin Infect Dis 2012 Jan 1;54(1):141-147 [FREE Full text] [doi: 10.1093/cid/cir727] [Medline: 22042879]
14. Schumacher JE, McCullumsmith C, Mugavero MJ, Ingle-Pang PE, Raper JL, Willig JH, et al. Routine depression screening in an HIV clinic cohort identifies patients with complex psychiatric co-morbidities who show significant response to treatment. AIDS Behav 2013 Oct;17(8):2781-2791 [FREE Full text] [doi: 10.1007/s10461-012-0342-7] [Medline: 23086427]
15. Barger D, Leleux O, Conte V, Sapparrart V, Gapillout M, Crespel I, et al. Integrating electronic patient-reported outcome measures into routine HIV care and the ANRS CO3 Aquitaine Cohort's Data Capture and Visualization System (QuALiV): protocol for a formative research study. JMIR Res Protoc 2018 Jun 7;7(6):e147 [FREE Full text] [doi: 10.2196/resprot.9439] [Medline: 29880467]
16. Ministry of Solidarity and Health. 2013. Medical management of people living with HIV, recommendations of the expert group, 2013 report URL: http://solidarites-sante.gouv.fr/IMG/pdf/Rapport_Morlat_2013_Mise_en_ligne.pdf [accessed 2017-11-16] [WebCite Cache ID 6v1OhYDdI]
17. Labbe E, Blanquet M, Gerbaud L, Poirier G, Sass C, Vendittelli F, et al. A new reliable index to measure individual deprivation: the EPICES score. Eur J Public Health 2015 Aug;25(4):604-609. [doi: 10.1093/eurpub/cku231] [Medline: 25624273]
18. O'Connell KA, Skevington SM. An international quality of life instrument to assess wellbeing in adults who are HIV-positive: a short form of the WHOQOL-HIV (31 items). AIDS Behav 2012 Feb;16(2):452-460. [doi: 10.1007/s10461-010-9863-0] [Medline: 21181253]
19. Tran V, Montori VM, Eton DT, Baruch D, Falissard B, Ravaud P. Development and description of measurement properties of an instrument to assess treatment burden among patients with multiple chronic conditions. BMC Med 2012 Jul 4;10:68 [FREE Full text] [doi: 10.1186/1741-7015-10-68] [Medline: 22762722]
20. Dawson DA, Grant BF, Stinson FS, Zhou Y. Effectiveness of the derived Alcohol Use Disorders Identification Test (AUDIT-C) in screening for alcohol use disorders and risk drinking in the US general population. Alcohol Clin Exp Res 2005 May;29(5):844-854. [doi: 10.1097/01.alc.0000164374.32229.a2] [Medline: 15897730]
21. 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]
22. Coons SJ, Gwaltney CJ, Hays RD, Lundy JJ, Sloan JA, Revicki DA, ISPOR ePRO Task Force. Recommendations on evidence needed to support measurement equivalence between electronic and paper-based patient-reported outcome (PRO) measures: ISPOR ePRO Good Research Practices Task Force report. Value Health 2009 Jun;12(4):419-429 [FREE Full text] [doi: 10.1111/j.1524-4733.2008.00470.x] [Medline: 19900250]
23. Nielsen J, Landauer TK. A Mathematical Model of the Finding of Usability Problems. In: Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems. 1993 Presented at: CHI'93; April 24-29, 1993; Amsterdam, The Netherlands p. 206-213.
24. Brooke J. Hell - Jens Oliver Meiert. 1996. SUS - A Quick and Dirty Usability Scale URL: https://hell.meiert.org/core/pdf/sus.pdf [accessed 2019-11-12]
25. Bangor A, Kortum PT, Miller JT. An empirical evaluation of the system usability scale. Int J Human-Comput Interact 2008;24(6):574-594. [doi: 10.1080/10447310802205776]
26. Sauro J. User Experience Magazine. 2011 Aug. SUStisfied? Little-Known System Usability Scale Facts URL: https://uxpamagazine.org/sustified/ [accessed 2019-06-15]
27. Nielsen J. Nielsen Norman Group: UX Training, Consulting, & Research. 2012 Jan 3. Usability 101: Introduction to Usability URL: https://www.nngroup.com/articles/usability-101-introduction-to-usability/ [accessed 2019-04-12]
28. Dumas JS, Redish JC. A Practical Guide To Usability Testing. Bristol: Intellect Ltd; 1994.
29. Ericsson AK, Herbert SA. Protocol Analysis: Verbal Reports As Data. Boston: The Mit Press; 1993.

Abbreviations

PLWH: people living with HIV
PRO: patient-reported outcome
SUS: System Usability Scale

Edited by G Eysenbach; submitted 12.06.19; peer-reviewed by K Engler, C Johnson, C Chen; comments to author 19.07.19; revised version received 07.08.19; accepted 07.09.19; published 18.12.19.

Please cite as:
Barger D, Leleux O, Conte V, Sapparrart V, Gapillout M, Crespel I, Erramouspe M, Delveaux S, Wittkop L, Dabis F, Bonnet F
Web-Based Module for the Collection of Electronic Patient-Reported Outcomes in People Living With HIV in Nouvelle Aquitaine, France: Usability Evaluation
JMIR Form Res 2019;3(4):e15013
URL: http://formative.jmir.org/2019/4/e15013/
doi:10.2196/15013
PMID:31850847

Original Paper

Understanding End-User Perspectives of Mobile Pulmonary Rehabilitation (mPR): Cross-Sectional Survey and Interviews

Rosie Dobson[1], PhD; Pauline Herbst[2], PhD; Sarah Candy[3], PGDipHealthScience; Tamzin Brott[2], MHSc, MBA; Jeffrey Garrett[3], MBChB, FRAC; Gayl Humphrey[1], MSocSc; Julie Reeve[4], PhD; Merryn Tawhai[5], PhD; Denise Taylor[4], PhD; Jim Warren[6], PhD; Robyn Whittaker[1,2], PhD

[1]National Institute for Health Innovation, University of Auckland, Auckland, New Zealand
[2]Waitemata District Health Board, Auckland, New Zealand
[3]Counties Manukau District Health Board, Auckland, New Zealand
[4]School of Clinical Sciences, Auckland University of Technology, Auckland, New Zealand
[5]Auckland Bioengineering Institute, University of Auckland, Auckland, New Zealand
[6]School of Computer Science, University of Auckland, Auckland, New Zealand

Corresponding Author:
Rosie Dobson, PhD
National Institute for Health Innovation
University of Auckland
Private Bag 92019
Auckland Mail Centre
Auckland, 1072
New Zealand
Phone: 64 9 3737599
Fax: 64 9 3731710
Email: r.dobson@auckland.ac.nz

Abstract

Background: Pulmonary rehabilitation (PR) is an effective intervention for the management of people with chronic respiratory diseases, but the uptake of and adherence to PR programs is low. There is potential for mobile health (mHealth) to provide an alternative modality for the delivery of PR, overcoming many of the barriers contributing to poor attendance to current services.

Objective: The objective of this study was to understand the needs, preferences, and priorities of end users for the development of an adaptive mobile PR (mPR) support program.

Methods: A mixed methods (qualitative and quantitative) approach was used to assess the needs, preferences, and priorities of the end users (ie, patients with chronic respiratory disorders) and key stakeholders (ie, clinicians working with patients with chronic respiratory disorders and running PR). The formative studies included the following: (1) a survey to understand the preferences and priorities of patients for PR and how mobile technology could be used to provide PR support, (2) ethnographic semistructured interviews with patients with chronic respiratory disorders to gain perspectives on their understanding of their health and potential features that could be included in an mPR program, and (3) key informant interviews with health care providers to understand the needs, preferences, and priorities for the development of an mPR support program.

Results: Across all formative studies (patient survey, n=30; patient interviews, n=8; and key stakeholder interviews, n=8), the participants were positive about the idea of an mPR program but raised concerns related to digital literacy and confidence in using technology, access to technology, and loss of social support currently gained from traditional programs. Key stakeholders highlighted the need for patient safety to be maintained and ensuring appropriate programs for different groups within the population. Finding a balance between ensuring safety and maximizing access was seen to be essential in the success of an mPR program.

Conclusions: These formative studies found high interest in mHealth-based PR intervention and detailed the potential for an mPR program to overcome current barriers to accessing traditional PR programs. Key considerations and features were identified, including the importance of technology access and digital literacy being considered in utilizing technology with this population.

(JMIR Form Res 2019;3(4):e15466) doi:10.2196/15466

KEYWORDS

mHealth; rehabilitation; COPD

Introduction

Chronic obstructive pulmonary disorder (COPD), an umbrella term for a range of debilitating respiratory diseases [1], is the fourth leading cause of mortality worldwide [2]. In New Zealand (NZ), COPD affects approximately 14% of adults aged >40 years [3]. Maori, the NZ indigenous population, as well as ethnic minority groups and those from socioeconomically deprived groups are disproportionately affected by COPD with higher prevalence rates and hospitalizations and are more likely to die from the condition [2,3].

One of the most effective interventions for COPD is pulmonary rehabilitation (PR), an evidence-based, interdisciplinary intervention that is a key component in the management of people with respiratory diseases [4]. PR is a formalized structured program comprising, but not limited to, exercise training, education, and behavior change, and it is designed to improve a patient's physical and psychological health and encourage engagement with health-enhancing behaviors [5]. PR is an individually tailored intervention based on thorough patient assessment, which is typically delivered in group programs in hospital or community settings. PR has been clearly demonstrated to improve breathlessness and health-related quality of life and reduce hospital admissions for exacerbations of COPD [4,6]. Clinical guidelines strongly recommend the uptake of PR by all patients with COPD, particularly following hospital admissions [7]. Despite this, and PR programs being available across nearly all regions of NZ, the uptake of, and adherence to, PR programs in NZ is poor [8]. It was estimated that in 2012, <1% of all patients with COPD were participating in PR in NZ [8]. Poor attendance and adherence to PR programs is common internationally. Previous literature has suggested that this is because of transportation, lack of perceived benefit, depression, and the interruption to the patient's daily routines [9-11]. In addition, many patients experience significant barriers to accessing PR services, especially those living in rural areas and where transportation to a central service may be unaffordable or unavailable. Home-based PR programs have been shown to overcome some of these access barriers for people with chronic respiratory diseases [12-14].

There is potential for mobile health (mHealth) to provide an alternative modality for the delivery of PR, overcoming many of the barriers contributing to poor attendance to current services. By utilizing mobile technology, PR can be made available to people within their everyday lives at times and places most suitable to the patient, removing the barriers associated with transport, timing, and location. There is a wealth of evidence for the use of mobile technology to deliver health interventions to people with chronic conditions including the delivery of rehabilitation interventions, self-management support programs, behavior change interventions, and supportive care [15-18]. Not only does mHealth allow for individually tailored interventions to be easily delivered in a cost-effective way but also it has potential for interventions to be adapted over time as individual needs and characteristics change [19].

When developing new mHealth tools, engagement with end users in the design is essential. By incorporating the perspectives of end users, it ensures the intervention will meet the population need and enables it to be tailored to specific cultural needs, contexts, and levels of technology access [20]. When end users are not considered in the design, it can contribute to poor uptake and use of tools [21]. Furthermore, it is important that formative research, including adequate description of the population context, is reported in the development of new mHealth interventions [22]. Formative research provides the basis for designing tools to meet user needs within system constraints and the local context.

This study aimed to understand the needs, preferences, and priorities of end users for the development of a mobile PR (mPR) support program. Specifically, it aimed to understand (1) the preferences and priorities of patients for PR and how mobile technology could be used to provide PR support and (2) the needs, preferences, and priorities of health care providers including physiotherapists, respiratory physicians, primary care nurses, and general practitioners (GPs).

Methods

Study Design

This cross-sectional study utilized a mixed method (qualitative and quantitative) approach incorporating surveys and interviews to assess the needs, preferences, and priorities of end users and key stakeholders. The study was split into 3 parts—(1) patient survey, (2) patient interviews, and (3) key stakeholder interviews—and was conducted between July and December 2018.

Ethical Approval

Ethics approval was obtained from the NZ Health and Disability Ethics Committee (18/NTA/105). Research approval from each District Health Board (DHB) from which patients were recruited was also obtained. Written informed consent was obtained from survey and interview participants before their participation in the studies.

Part 1: Patient Survey

Inclusion Criteria

Inclusion criteria were adults with chronic respiratory disease who would be eligible for PR, able to read and understand English, and provide informed consent.

Recruitment and Procedures

Potential participants were identified by clinicians at 2 DHBs in the Auckland region, NZ, through respiratory outpatient clinics and inpatient services. Eligible participants were given a letter about the study, the information sheet, and consent form. By selecting patients from this defined subgroup, we were more likely to define patients who had either been invited to attend or had attended a rehabilitation program. Those interested in participating provided informed consent before completing the survey via 1 of 3 methods: (1) on the web via the study website, (2) over the phone with a member of the research team, or (3) on paper. Surveys completed by phone or on paper were entered into the Web-based version of the survey by the researcher.

Survey Design

The survey comprised 4 parts and contained both closed- and open-ended questions to allow participants to elaborate on their answers: (1) your health including diagnosis, attendance at PR, and barriers to attendance and completion of PR; (2) technology access, including current use and access to technology and devices; (3) technology and pulmonary support, including perceptions of mPR, preferences for technology-based support, and perceived barriers and benefits to technology-based pulmonary support; and (4) demographics, specifically age group, gender, and ethnicity.

The survey was designed in paper format and then uploaded into REDCap software (v8.5.0). It was pretested by the research team and a selection of patients before finalization.

Part 2: Patient Interviews

Inclusion Criteria

Survey participants who had consented to be contacted for future research following completion of the survey were eligible for inclusion in an interview.

Procedures

Patients who completed part 1 of this study and identified that they were happy to be contacted about further research were contacted to invite them to take part in an interview. Participants provided informed consent to participate as well as consent to access medical records related to their eligibility for PR. Interviews were conducted in person, or over the phone, by a trained interviewer (PH) at a time and location convenient to the participant. If needed, interviews were split over multiple sessions to reduce burden for participants. Notes were taken by the researcher during the interviews, and the interviews were recorded to supplement the notes. At the end of the interview process, the participants were offered a NZ $20.00 (approximately US $13.00 or €11.50) voucher for their time.

Interview Design

Ethnographic, semistructured interviews were undertaken. The interviews were designed to explore in-depth beliefs and perceptions of chronic respiratory disorders and their treatments, perspectives on potential features of an mPR intervention, and patient understanding of their health.

Part 3: Key Stakeholder Interviews

Inclusion Criteria

Inclusion criteria were clinicians managing patients with COPD and other chronic pulmonary disorders (eg, bronchiectasis and interstitial lung disorders), able to read and understand English, and able to provide informed consent.

Procedures

Potential participants were identified through respiratory and PR services by coinvestigators RW, PH, SC, and JR and invited to take part via email. Interviews were conducted in person or via phone at the participant's preference by a trained interviewer (RD) and were between 30 min and 1 hour in duration. Notes were taken by the interviewer, and the interviews were recorded to supplement the notes.

Interview Design

Interviews were semistructured and designed to cover the current use of technology in care of people with chronic respiratory conditions, experiences with PR, and perceptions of mPR including enablers and barriers.

Statistical Analysis

Quantitative data from the survey were analyzed and summarized using descriptive quantitative analyses including means, standard deviations, and proportions. Qualitative comments were analyzed using a simple, general inductive thematic approach to identify common themes and meanings from the data. Only completed surveys were included in the analysis. Prioritized ethnicity was used as recommended by the NZ Ministry of Health for the reporting of ethnicity data; only 1 of the ethnic categories nominated by the participant was used according to a predetermined hierarchy (Maori, Pacific Islander, Asian, European, and other ethnic groups, in order of prioritization).

Results

Part 1: Patient Survey

There were 34 entries to the survey site from which 30 people consented to take part and completed the survey. Slightly more than half of the sample were male (17/30, 57%), and the majority were aged >65 years (22/30, 73%; Table 1). Over one-third of the participants identified as Maori or Pacific Islander (11/30, 37%). Approximately half the sample reported that they had been diagnosed with COPD (16/30, 53%). A total of 7 (7/30, 23%) participants reported that they were unsure, or did not know, their diagnosis.

Table 1. Demographic and clinical characteristics (N=30).

Characteristics	Participants, n (%)
Gender	
Male	17 (57)
Female	12 (40)
Did not answer	1 (3)
Ethnicity	
New Zealand European	17 (57)
Maori	7 (23)
Pacific	4 (13)
Asian	0 (0)
Other	1 (3)
Did not answer	1 (3)
Age (years)	
<45	0 (0)
45-54	3 (10)
55-64	4 (13)
65-74	10 (33)
75-84	12 (40)
>85	0 (0)
Did not answer	1 (3)
Patient-reported diagnosis	
Chronic obstructive pulmonary disorder	16 (53)
Emphysema	3 (10)
Bronchitis	1 (3)
Pulmonary fibrosis	1 (3)
Bronchiectasis	1 (3)
Asbestosis	1 (3)
Do not know or unsure	7 (23)

Pulmonary Rehabilitation

Participants were asked if they had attended a PR program. Almost half (14/30, 47%) reported that they had completed PR or were intending to complete it in the future. Nearly a quarter of participants (7/30, 23%) had started a PR program but had not completed it. The reasons for not completing the program included issues related to location and transport (n=2), timing of the sessions (n=2), being hospitalized (n=2), and other commitments (n=1). A total of 3 participants (3/30, 10%) were offered PR but did not attend because of transport (n=2) or other commitments (eg, having to care for grandchildren; n=1). The remaining 6 (6/30, 20%) participants reported that they had not been offered PR or did not remember being offered PR.

Access to Technology

Participants were asked about their access to digital devices (eg, mobile phones, tablets, computers, and sensors) for personal use, with all but 4 (4/30, 13%) participants reporting having access to a mobile phone. Of those who had a mobile phone, the majority had a smartphone (20/26, 77%) and the remainder (6/26, 23%) a mobile phone without internet capability. Of those that had a smartphone, only 60% (12/20) reported having access to the internet on the smartphone all the time. A total of 5 (5/20, 25%) had access to the internet on the phone sometimes, 2 (2/20, 10%) not at all, and 1 (1/20, 5%) did not answer.

One-third of the sample (10/30, 33%) reported having access to a tablet for personal use, 14 (14/30, 47%) a computer (ie, laptop or desktop), 2 (2/30, 7%) a Fitbit or other fitness tracking device, and only 1 participant reported access to no devices for personal use. A total of 23 (23/30, 77%) participants reported access to internet at home, and an additional 3 (3/30, 10%) reported they sometimes had access to internet at home. There were 4 (4/30, 13%) that had no access to the internet at home.

In relation to the use of technology-based devices and tools to manage health, there was only 1 participant (1/30, 3%) that reported using an app, 1 (1/30, 3%) who reported using a smart watch, 3 (3/30, 10%) used a fitness tracking device, and 4 (4/30, 13%) a peak flow meter. A total of 20 (20/30, 67%) participants reported no use of these tools to manage their health.

Technology and Respiratory Health Support

Participants were asked about their perceptions of a technology/mobile phone program that would allow people to receive PR support at home. The majority (23/30, 77%) reported that they liked this idea, and 7 (7/30, 23%) did not like the idea. The proportion of those that liked the idea was highest in Maori participants (6/7, 86%) compared with NZ European (13/17, 77%) or Pacific (2/4, 50%). Those that liked the idea (n=23) were asked what they liked about the idea. The most common theme was around the access at home meaning no need to travel (n=10), and others reported the cost reduction related to parking and travel (n=3), being able to access the program wherever and whenever (n=1), less embarrassment within the home (n=1), and that family would be able to help and be involved (n=1):

> *Privacy—don’t have to be embarrassed.* [45-54-year-old male, ID number 27]
>
> *Would make it easier as would not have to travel in bad weather.* [75-84-year-old male, ID number 8]
>
> *The convenience. At the moment, travel outside the home is quite onerous for me.* [65-74-year-old male, ID number 9]
>
> *Family would be able to help Dad to do this at home.* [75-84-year-old male, ID number 25]
>
> *Would have saved me some trips... Transport is hard for people, don't have a car.* [45-54-year-old male, ID number 27]

There were 5 participants whose responses were moderately supportive of it being a good idea and 3 who did not answer the question. In addition, 3 participants identified concerns around the proposed idea potentially resulting in less social contact and less access to health care professionals:

> *Sounds good although I do like meeting other people.* [65-74-year-old female, ID number 33]
>
> *But it is good to meet other people and learn from others.* [65-74-year-old female, ID number 22]

Those that did not like the idea (n=7) were asked the reason, with the most common theme being that they found technology too difficult/hard (n=3). Other themes included not having internet (n=1), wanting the company/social aspect of an in-person group program (n=1), or preferring to go to the hospital (n= 1; there was 1 participant who did not provide a reason):

> *Too hard to use computer.* [75-84-year-old male, ID number 18]
>
> *No internet at home.* [55-64-year-old male, ID number 19]
>
> *I would feel isolated, when you are older you need company.* [75-84-year-old female, ID number 21]

Participants were asked about the features they thought would be helpful to include in the proposed program. Of those that liked the proposed idea and who responded to the question (n=22), 19 (19/22, 86%) selected tips and suggestions for managing their breathing, 17 (17/22, 77%) information and education about their condition, 12 (12/22, 54%) access to their health information, 11 (11/22, 50%) tracking of their health data, and 9 (9/22, 41%) motivational and support messages. A total of 23 (23/30, 77%) reported that they would consider wearing a sensor (eg, a Fitbit or activity tracker) as part of an mPR program, and 22 (22/30, 73%) reported that they would want the program to be linked into their existing health record and information so that the program could be tailored to their individual health condition and treatment.

Participants were asked to identify the barriers to using the proposed mPR program (Table 2). The most common barrier was concern about not being comfortable or confident using technology.

When asked about the perceived benefits of using technology to support people with health conditions in the community, the majority (25/29, 86%) identified the convenience of having access at any time or anywhere as a benefit. There were 15 (15/29, 52%) participants who identified the benefit of being able to involve/include family/whanau, and 4 (4/29, 14%) identified feeling more comfortable not being in the group or clinical environment.

Table 2. Barriers to using a mobile pulmonary rehabilitation program (N=30).

Barriers	Participants, n (%)
No barriers	6 (20)
No access to technology	5 (17)
Do not feel comfortable using technology	11 (37)
Security concerns regarding my health information	3 (10)
Not interested in this type of support	3 (10)
Could increase worry and anxiety	1 (3)
Would not be useful if not translated	1 (3)

Part 2: Patient Interviews

A total of 8 patient interviews were conducted. Participants included 2 women (2/8, 25%) and 6 men (6/8, 75%), with 2 of the 8 (2/8, 25%) individuals identifying as Maori and the remaining (6/8, 75%) as NZ European. The majority were aged 75 to 84 years (4/8, 50%), 2 (2/8, 25%) aged 65 to 74 years, and the remaining (2/8, 25%) aged 45 to 54 years. Participants were registered across 2 DHBs and interviewed at home (2/8, 25%), in the hospital (5/8, 63%), or via phone (1/8, 13%). Analysis of the interview data is summarized in 4 major themes.

Condition and Health

Participants identified themselves as having COPD, emphysema, and asbestosis, but only 6 participants were able to clearly describe their diagnosis and treatment. Diagnosis was reported as being a shock to receive; however, this shock did not necessarily change behavior:

> *Well I didn't know much about it, so it didn't make much of a difference to me. I continued smoking, didn't I?* [75-84-year-old male, ID number 4]
>
> *I knew from the day he said that, that things were not so good for me.* [75-84-year-old male, ID number 6]

Participants described COPD affecting all facets of their lives negatively. Despite this, all expressed some form of resilience and stoicism during the interviews:

> *It affects everything: social, mental, physical.* [75-84-year-old female, ID number 5]
>
> *It can't be cured you just have to live with it and get on with life and do what you've got to do.* [75-84-year-old male, ID number 2]
>
> *I'm pretty humble with it. I've lived a pretty good life.* [45-54-year-old male, ID number 3]

Only 2 participants showed some degree of understanding of their test results, and the remaining participants did not understand the results of their tests; they knew what their diagnosis was but found the presentation of the test results difficult to comprehend. All participants struggled to remember which tests had been conducted and at which point in their illness trajectory:

> *Honestly, I just go when I am told to go and have it done, and that's it.* [75-84-year-old female, ID number 5]

Of the 8 participants, 7 were ex-smokers and acknowledged this as a contributing factor to their current respiratory condition; however, 5 offered additional environmental factors as contributors including asbestos and work environments.

Pulmonary Rehabilitation

All patients interviewed had some experience of hospital-based PR programs. A total of 2 had completed programs within the past 3 years, 2 had started but stopped indefinitely because of hospitalization, 3 were in the process of completing an 8-week program, and 1 had tried a program over a decade previously. All but 1 reported enjoying the programs:

> *I wish I had done it earlier.* [75-84-year-old male, ID number 2]
>
> *I love it. I wish they had it 12 months of the year.* [65-74-year-old female, ID number 7]
>
> *Yeah, I enjoy the exercise. It's just a matter of getting your mind in the right place. I get up and I do it and I feel a lot better afterwards.* [65-74-year-old male, ID number 1]

Reasons for this included socializing during the training, being able to compare progress and daily lived experiences with others with a similar condition, compassionate and dedicated staff who were known to the participants, and both a measurable and perceived improvement in physical ability. However, a participant reported anxiety and an increased sense of vulnerability after he was transferred to community support, following completion of the hospital-based PR program, arising from inconsistency in times and scheduling, different staff members at each visit, and a decreased sense of support.

Technology-Based Pulmonary Rehabilitation

All but 1 of the participants (7/8, 88%) described that they preferred to be offered a PR program in person at the moment of diagnosis by a medical professional they trusted and had an existing relationship with, rather than when in hospital with an exacerbation:

> *When you're really sick and they come at you, like the physio comes, and this one comes, and that one comes, and you just feel like being left alone, I felt, just leave me alone sort of thing... maybe after being first diagnosed.* [75-84-year-old female, ID number 5]

The main barrier to mPR was a perceived incompetence with technology and a fear that this would be difficult to work with; however, 6 participants appeared to be confident texting during interviews, and 4 talked about using Facebook to communicate with family. A participant enjoyed using health apps to measure aspects of his health.

Feature suggestions for mPR included knowing which part of the body exercises were targeting, with the most ambitious proposals suggesting:

> *I think you should give them every feature that you can and maybe give them the option to choose... incorporating videos of people doing it and how they started and then how they finished... a motivational start, and then a talk and showing "look at me, now I can run two kilometers.* [45-54-year-old male, ID number 3]

Imaging

During the interview, participants were shown a model of the lung [23,24] and an interactive website [25]. The subsequent discussion revealed the underlying power of visual imagery to reveal the *truth* of diagnosis, with patients reporting that they struggled to accept their diagnosis until they saw an image and that seemed more *real* than paper, which they saw as a tool for the medical professional:

> *I didn't get how bad my lungs were until I saw it.* [45-54-year-old male, ID number 3]
>
> *She said "have you ever seen your lung?" and showed me this x-ray and I thought, "whoa."* [75-84-year-old male, ID number 2]

A participant said she had never been shown an x-ray or image of her lungs and insisted that this would have caused her to stop smoking earlier.

Part 3: Key Stakeholder Interviews

A total of 8 key stakeholder interviews were completed. Key stakeholders included 2 doctors (ie, GP and respiratory specialist), 2 nurses, 3 physiotherapists, and 1 health psychologist. Participants worked across either predominantly urban (5/8, 63%) or predominantly rural (3/8, 37%) populations. A total of 5 (5/8, 63%) participants were directly involved in the delivery of PR, whereas the remaining 3 (3/8, 37%) were involved in referring patients to PR services.

When asked about their perceptions of patients' understanding of COPD, all acknowledged that it was generally poor. A total of 3 participants acknowledged that understanding varied, and those that had attended PR or sought health information through the internet had a better understanding. Common tools for explaining COPD to patients included handouts and pamphlets (n=3), Web resources (ie, websites and YouTube videos; n=2), drawings and models (n=2), and patient scans (n=1).

Mobile Pulmonary Rehabilitation

All participants thought that a technology- or mobile-based PR program was a good idea, particularly for overcoming the barriers their patients currently experienced to attending PR. However, the majority of the participants also raised concerns that an mPR program would lack the social element and not be suitable for some groups such as older patients or those with limited confidence and access to technology.

Participants were asked how they felt an mPR program would fit into current models of care. Responses could be grouped into 3 main categories: (1) an alternative service to increase the access to PR for those that were not able to attend current services, (2) a maintenance program following traditional in-person PR programs, and (3) a combination/mixed model of care where patients used an mPR program to complement the in-person program. All participants reported that an mPR program should include the following components: (1) education, (2) exercise information, (3) motivation and support, (4) the ability to view personal health information/data, and (4) monitoring of health behaviors.

Other components identified by the participants for inclusion in an mPR program included health psychology content, medication reminders, personalized action plans, step-by-step videos, social features, sharing of information between patient and health care professional, and general self-care information. Participants also identified that it would be essential for a program to be both individually and culturally tailored.

When asked about the different technologies the potential program should utilize, many different technologies were identified including sensors, smart inhalers, apps, and text messages. But consistent across all participants was that the technologies used needed to consider differing access to technology, devices, and data, as well as confidence with using technology:

> *Important to use mobile phone as a lot of people won't have anything else.* [ID number 1]
>
> *An app would be a good way of doing it... Sensors are useful for helping patients to see progress to goals.* [ID number 2]
>
> *Text messages are easy for patients to get, older patients to get, minority populations do not tend to have a lot of money on their phones.* [ID number 3]
>
> *A lot of people don't have internet at all. We text or phone them, they aren't able to text, call as they have no money on their phones. Some people don't even have mobile coverage...* [ID number 6]
>
> *We already use SMS, people are used to getting text messages.* [ID number 7]

There were conflicting views from participants around how an mPR program should be accessed. A participant felt that referrals should be through the same avenues as current services to ensure safety was prioritized, and only people who were appropriate were accessing the program. Others felt that in addition to clinician referrals, patients should be able to access it directly to reduce barriers to access. There was only 1 participant who felt that clinicians should not be involved in referring to the program at all because of concerns that the clinician would need to then provide technical support:

> *Referred [by a clinician] is best rather than self-referral, ensures that the info will be right for that patient. Same criteria as current groups.* [ID number 1]
>
> *Shouldn't put up a barrier of the clinician referring them. Especially if they can't afford to see their GP.* [ID number 2]
>
> *Anyone should be able to have access to it but there needs to be some sort of way of knowing people are safe with it, that they are safe to exercise.* [ID number 4]
>
> *Patients should be able to access it themselves. Health care professionals could promote it but don't want to become IT support.* [ID number 5]

Finally, participants were asked about the potential barriers/downsides to an mPR program. All but 1 participant (7/8, 88%) reported concerns relating to the digital divide, including access to technology and the confidence to use it. A full list of barriers identified can be seen in Table 3.

Table 3. Barriers to using a mobile pulmonary rehabilitation program (N=8).

Barriers	Participants, n (%)
Digital divide	7 (88)
The lack of the social/group environment	6 (75)
Lack of relationship between clinician and patient	4 (50)
Safety for patients not being supervised	2 (25)
Patient compliance to the program	2 (25)
Health care professional digital literacy	1 (13)
Successful marketing and promotion of the program	1 (13)
Patient access to exercise equipment	1 (13)

Discussion

Principal Findings

This study aimed to identify the needs, preferences, and priorities of end users in the development of an adaptive mPR support program. A survey of patients, together with interviews of patients and key stakeholders, found a common interest in an mPR program. The potential for mPR overcoming barriers to accessing traditional PR programs was highlighted. These findings are consistent with previous research reporting high acceptability of digital health tools in patients with chronic respiratory conditions [26-28].

This formative study has identified important aspects of our target audience and their diversity of needs. Some have low access to digital technology as well as low digital literacy and confidence, although most would like the option of a technology-delivered program. Needs include support during PR programs (between group sessions), support after a PR program to maintain behavior change, and a complete program for those unable to attend a traditional inpatient PR program. The study also identified a difference in clinicians' perceptions of patients' understanding of their condition (poor) versus patients' actual comprehension and management of their condition.

From these findings, our key considerations in designing a user-centered mPR support program will be in replicating the benefits of social support provided by the in-person group sessions, ensuring options are available for differing levels of digital literacy and confidence with technology while not providing a *second class* program for those with lower technology access is essential. Providing a program that not only considers individual technology access and literacy but also considers personal preferences and characteristics is also needed to ensure that it will be positively received by users. As indigenous populations, ethnic minority groups, and those from socioeconomically deprived groups suffer worse outcomes from COPD, ensuring an mPR program that strives for equity is essential. A strength of the survey sample was that it comprised over one-third Maori and Pacific participants. Key stakeholders highlighted the importance of considering culture and of culturally tailoring the mHealth tool.

Key considerations in terms of satisfying referring clinicians are ensuring safety can be maintained and providing appropriate programs for different groups within the population. Finding a balance between ensuring safety and maximizing access is vital to ensure that an mPR program overcomes the barriers and increases access to PR support. Although the majority of the participants in the survey and interviews had attended PR or intended to, this study has identified a range of barriers to traditional in-person PR services that align with the previous studies [9-11]. These barriers related to transport and timing can be overcome with an mHealth alternative.

It is important that these findings are interpreted in light of the study's limitations, including the small number of participants surveyed and interviewed, the potential sampling bias in those who agreed to participate, and the patients who participated having a higher proportion accessing PR services than the general population categories. It is likely that those patients who did participate had more interest in mHealth; therefore, engagement with this type of tool may be lower in the wider population. Pretesting of an mPR program with a wider patient sample will be essential to understand the acceptability of this type of intervention.

Although there is a proliferation of technology in health care and increasingly innovative technologies being embraced for patient self-management, our study, consistent with previous studies, has shown that a digital divide exists, contributed to by differing access to data and tools as well as confidence and digital literacy to effectively use them [29-31]. Although our study demonstrated interest in using technology for PR support, the current use of many tools such as sensors and apps was low among the participants. If an mPR program is to utilize more than basic modalities, such as text messaging, then there is likely a need for these tools to be provided to patients with training and ongoing support when issues arise. This equates to additional resources and associated costs that must be considered not only in setting up an mPR program but in sustaining it.

Conclusions

We have developed a prototype mHealth-based PR program based on the results of this mixed methods research. This includes different components suggested in the paper that will be pretested with people with chronic respiratory disorders. Feedback will be provided, and funding will be sought for the development of a full mPR program. The findings support the need for involving patients in the initial design and development of an mHealth intervention and the need to conduct a feasibility pilot study, once the intervention is developed, to try and better understand the degree of support required by health professionals and the degree of technical support required.

Acknowledgments

The authors would like to thank the study participants. They would also like to acknowledge Koray Atalag, Neela Bhana, Kelly Bohot, and Ian Warren for contributions to study design and procedures. The study was funded by the research grant from the Medical Technologies Center of Research Excellence (MedTech CoRE).

Authors' Contributions

TB, SC, RD, JG, GH, PH, JR, MT, DT, JW, and RW contributed to the study design and procedures. Data collection was performed by SC, RD, and PH. RD and PH analyzed and interpreted the data. RD and RW prepared the draft for the paper. TB, SC, JG, PH, and DT provided critical feedback on the paper.

Conflicts of Interest

None declared.

References

1. Pauwels RA, Buist AS, Calverley PM, Jenkins CR, Hurd SS, GOLD Scientific Committee. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease. NHLBI/WHO Global Initiative for Chronic Obstructive Lung Disease (GOLD) Workshop summary. Am J Respir Crit Care Med 2001 Apr;163(5):1256-1276. [doi: 10.1164/ajrccm.163.5.2101039] [Medline: 11316667]
2. Ministry of Health NZ. 2016. Privately funded hospital discharges – 1 July 2013 to 30 June 2014 URL: https://www.health.govt.nz/publication/privately-funded-hospital-discharges-1-july-2013-30-june-2014 [accessed 2019-10-15]
3. Telfar BL, Zhang J. Asthma and Respiratory Foundation NZ. 2016. The impact of respiratory disease in New Zealand 2016 update URL: https://www.asthmafoundation.org.nz/research/the-impact-of-respiratory-disease-in-new-zealand-2016-update [accessed 2019-03-01]
4. McCarthy B, Casey D, Devane D, Murphy K, Murphy E, Lacasse Y. Pulmonary rehabilitation for chronic obstructive pulmonary disease. Cochrane Database Syst Rev 2015 Feb 23(2):CD003793. [doi: 10.1002/14651858.CD003793.pub3] [Medline: 25705944]
5. Spruit MA, Singh SJ, Garvey C, ZuWallack R, Nici L, Rochester C, ATS/ERS Task Force on Pulmonary Rehabilitation. An official American Thoracic Society/European Respiratory Society statement: key concepts and advances in pulmonary rehabilitation. Am J Respir Crit Care Med 2013 Oct 15;188(8):e13-e64. [doi: 10.1164/rccm.201309-1634ST] [Medline: 24127811]
6. Laviolette L, Bourbeau J, Bernard S, Lacasse Y, Pepin V, Breton M, et al. Assessing the impact of pulmonary rehabilitation on functional status in COPD. Thorax 2008 Feb;63(2):115-121. [doi: 10.1136/thx.2006.076844] [Medline: 17901158]
7. Alison JA, McKeough ZJ, Johnston K, McNamara RJ, Spencer LM, Jenkins SC, Lung Foundation Australiathe Thoracic Society of Australia and New Zealand. Australian and New Zealand pulmonary rehabilitation guidelines. Respirology 2017 May;22(4):800-819. [doi: 10.1111/resp.13025] [Medline: 28339144]
8. Levack WM, Weatherall M, Reeve JC, Mans C, Mauro A. Uptake of pulmonary rehabilitation in New Zealand by people with chronic obstructive pulmonary disease in 2009. N Z Med J 2012 Jan 20;125(1348):23-33. [Medline: 22282274]
9. Guo S, Bruce A. Improving understanding of and adherence to pulmonary rehabilitation in patients with COPD: a qualitative inquiry of patient and health professional perspectives. PLoS One 2014;9(10):e110835 [FREE Full text] [doi: 10.1371/journal.pone.0110835] [Medline: 25357128]
10. Harrison SL, Robertson N, Apps L, Steiner M, Morgan MD, Singh SJ. 'We are not worthy'--understanding why patients decline pulmonary rehabilitation following an acute exacerbation of COPD. Disabil Rehabil 2015;37(9):750-756. [doi: 10.3109/09638288.2014.939770] [Medline: 25009949]
11. Hayton C, Clark A, Olive S, Browne P, Galey P, Knights E, et al. Barriers to pulmonary rehabilitation: characteristics that predict patient attendance and adherence. Respir Med 2013 Mar;107(3):401-407 [FREE Full text] [doi: 10.1016/j.rmed.2012.11.016] [Medline: 23261311]
12. Maltais F, Bourbeau J, Shapiro S, Lacasse Y, Perrault H, Baltzan M, Chronic Obstructive Pulmonary Disease Axis of Respiratory Health Network, Fonds de recherche en santé du Québec. Effects of home-based pulmonary rehabilitation in patients with chronic obstructive pulmonary disease: a randomized trial. Ann Intern Med 2008 Dec 16;149(12):869-878. [doi: 10.7326/0003-4819-149-12-200812160-00006] [Medline: 19075206]
13. Liu X, Tan J, Wang T, Zhang Q, Zhang M, Yao L, et al. Effectiveness of home-based pulmonary rehabilitation for patients with chronic obstructive pulmonary disease: a meta-analysis of randomized controlled trials. Rehabil Nurs 2014;39(1):36-59. [doi: 10.1002/rnj.112] [Medline: 23780865]

14. Holland AE, Mahal A, Hill CJ, Lee AL, Burge AT, Cox NS, et al. Home-based rehabilitation for COPD using minimal resources: a randomised, controlled equivalence trial. Thorax 2017 Jan;72(1):57-65 [FREE Full text] [doi: 10.1136/thoraxjnl-2016-208514] [Medline: 27672116]
15. Whittaker R, McRobbie H, Bullen C, Rodgers A, Gu Y. Mobile phone-based interventions for smoking cessation. Cochrane Database Syst Rev 2016 Apr 10;4:CD006611 [FREE Full text] [doi: 10.1002/14651858.CD006611.pub4] [Medline: 27060875]
16. Dobson R, Whittaker R, Jiang Y, Maddison R, Shepherd M, McNamara C, et al. Effectiveness of text message based, diabetes self management support programme (SMS4BG): two arm, parallel randomised controlled trial. Br Med J 2018 May 17;361:k1959 [FREE Full text] [doi: 10.1136/bmj.k1959] [Medline: 29773539]
17. Pfaeffli Dale L, Whittaker R, Jiang Y, Stewart R, Rolleston A, Maddison R. Text message and internet support for coronary heart disease self-management: results from the Text4Heart randomized controlled trial. J Med Internet Res 2015 Oct 21;17(10):e237 [FREE Full text] [doi: 10.2196/jmir.4944] [Medline: 26490012]
18. Free C, Phillips G, Galli L, Watson L, Felix L, Edwards P, et al. The effectiveness of mobile-health technology-based health behaviour change or disease management interventions for health care consumers: a systematic review. PLoS Med 2013;10(1):e1001362 [FREE Full text] [doi: 10.1371/journal.pmed.1001362] [Medline: 23349621]
19. L'Engle K, Plourde KF, Zan T. Evidence-based adaptation and scale-up of a mobile phone health information service. mHealth 2017 Mar 23;3:11-11. [doi: 10.21037/mhealth.2017.02.06]
20. Velardo C, Shah SA, Gibson O, Clifford G, Heneghan C, Rutter H, EDGE COPD Team. Digital health system for personalised COPD long-term management. BMC Med Inform Decis Mak 2017 Feb 20;17(1):19 [FREE Full text] [doi: 10.1186/s12911-017-0414-8] [Medline: 28219430]
21. Goodyear-Smith F, Jackson C, Greenhalgh T. Co-design and implementation research: challenges and solutions for ethics committees. BMC Med Ethics 2015 Nov 16;16:78 [FREE Full text] [doi: 10.1186/s12910-015-0072-2] [Medline: 26573410]
22. Agarwal S, LeFevre AE, Lee J, L'Engle K, Mehl G, Sinha C, WHO mHealth Technical Evidence Review Group. Guidelines for reporting of health interventions using mobile phones: mobile health (mHealth) evidence reporting and assessment (mERA) checklist. Br Med J 2016 Mar 17;352:i1174. [doi: 10.1136/bmj.i1174] [Medline: 26988021]
23. Tawhai MH, Hunter P, Tschirren J, Reinhardt J, McLennan G, Hoffman EA. CT-based geometry analysis and finite element models of the human and ovine bronchial tree. J Appl Physiol (1985) 2004 Dec;97(6):2310-2321 [FREE Full text] [doi: 10.1152/japplphysiol.00520.2004] [Medline: 15322064]
24. Burrowes KS, Hunter PJ, Tawhai MH. Anatomically based finite element models of the human pulmonary arterial and venous trees including supernumerary vessels. J Appl Physiol (1985) 2005 Aug;99(2):731-738 [FREE Full text] [doi: 10.1152/japplphysiol.01033.2004] [Medline: 15802366]
25. Tawhai M. MedTech CoRE. 2017. Lung Simulation: 'My Smoking Lung' URL: https://sites.bioeng.auckland.ac.nz/rotary/lungs/ [accessed 2018-10-01]
26. Orme MW, Weedon AE, Saukko PM, Esliger DW, Morgan MD, Steiner MC, et al. Findings of the chronic obstructive pulmonary disease-sitting and exacerbations trial (COPD-SEAT) in reducing sedentary time using wearable and mobile technologies with educational support: randomized controlled feasibility trial. JMIR Mhealth Uhealth 2018 Apr 11;6(4):e84 [FREE Full text] [doi: 10.2196/mhealth.9398] [Medline: 29643055]
27. Kwon H, Lee S, Jung EJ, Kim S, Lee J, Kim DK, et al. An mHealth management platform for patients with chronic obstructive pulmonary disease (efil breath): randomized controlled trial. JMIR Mhealth Uhealth 2018 Aug 24;6(8):e10502 [FREE Full text] [doi: 10.2196/10502] [Medline: 30143475]
28. Loeckx M, Rabinovich RA, Demeyer H, Louvaris Z, Tanner R, Rubio N, et al. Smartphone-based physical activity telecoaching in chronic obstructive pulmonary disease: mixed-methods study on patient experiences and lessons for implementation. JMIR Mhealth Uhealth 2018 Dec 21;6(12):e200 [FREE Full text] [doi: 10.2196/mhealth.9774] [Medline: 30578215]
29. McCloud RF, Okechukwu CA, Sorensen G, Viswanath K. Beyond access: barriers to internet health information seeking among the urban poor. J Am Med Inform Assoc 2016 Nov;23(6):1053-1059 [FREE Full text] [doi: 10.1093/jamia/ocv204] [Medline: 27206459]
30. Mossberger K, Tolbert CJ, Bowen D, Jimenez B. Unraveling different barriers to internet use. Urban Affairs Rev 2012 Aug 13;48(6):771-810. [doi: 10.1177/1078087412453713]
31. Sarkar U, Karter AJ, Liu JY, Adler NE, Nguyen R, López A, et al. Social disparities in internet patient portal use in diabetes: evidence that the digital divide extends beyond access. J Am Med Inform Assoc 2011 May 1;18(3):318-321 [FREE Full text] [doi: 10.1136/jamia.2010.006015] [Medline: 21262921]

Abbreviations

COPD: chronic obstructive pulmonary disease
DHB: District Health Board
GP: general practitioner
mHealth: mobile health

mPR: mobile pulmonary rehabilitation
NZ: New Zealand
PR: pulmonary rehabilitation

Edited by G Eysenbach; submitted 14.07.19; peer-reviewed by D Ryan, L Spencer; comments to author 01.10.19; revised version received 01.10.19; accepted 02.10.19; published 20.12.19.

Please cite as:
Dobson R, Herbst P, Candy S, Brott T, Garrett J, Humphrey G, Reeve J, Tawhai M, Taylor D, Warren J, Whittaker R
Understanding End-User Perspectives of Mobile Pulmonary Rehabilitation (mPR): Cross-Sectional Survey and Interviews
JMIR Form Res 2019;3(4):e15466
URL: http://formative.jmir.org/2019/4/e15466/
doi:10.2196/15466
PMID:31859681

Review

A Comprehensive Framework to Evaluate Websites: Literature Review and Development of GoodWeb

Rosalie Allison[1], BSc, MSc; Catherine Hayes[1], BSc; Cliodna A M McNulty[1], MBBS, FRCP; Vicki Young[1], BSc, PhD

Public Health England, Gloucester, United Kingdom

Corresponding Author:
Rosalie Allison, BSc, MSc
Public Health England
Primary Care and Interventions Unit
Gloucester, GL1 1DQ
United Kingdom
Phone: 44 0208 495 3258
Email: rosie.allison@phe.gov.uk

Abstract

Background: Attention is turning toward increasing the quality of websites and quality evaluation to attract new users and retain existing users.

Objective: This scoping study aimed to review and define existing worldwide methodologies and techniques to evaluate websites and provide a framework of appropriate website attributes that could be applied to any future website evaluations.

Methods: We systematically searched electronic databases and gray literature for studies of website evaluation. The results were exported to EndNote software, duplicates were removed, and eligible studies were identified. The results have been presented in narrative form.

Results: A total of 69 studies met the inclusion criteria. The extracted data included type of website, aim or purpose of the study, study populations (users and experts), sample size, setting (controlled environment and remotely assessed), website attributes evaluated, process of methodology, and process of analysis. Methods of evaluation varied and included questionnaires, observed website browsing, interviews or focus groups, and Web usage analysis. Evaluations using both users and experts and controlled and remote settings are represented. Website attributes that were examined included usability or ease of use, content, design criteria, functionality, appearance, interactivity, satisfaction, and loyalty. Website evaluation methods should be tailored to the needs of specific websites and individual aims of evaluations. GoodWeb, a website evaluation guide, has been presented with a case scenario.

Conclusions: This scoping study supports the open debate of defining the quality of websites, and there are numerous approaches and models to evaluate it. However, as this study provides a framework of the existing literature of website evaluation, it presents a guide of options for evaluating websites, including which attributes to analyze and options for appropriate methods.

(JMIR Form Res 2019;3(4):e14372) doi:10.2196/14372

KEYWORDS

user experience; usability; human-computer interaction; software testing; quality testing; scoping study

Introduction

Background

Since its conception in the early 1990s, there has been an explosion in the use of the internet, with websites taking a central role in diverse fields such as finance, education, medicine, industry, and business. Organizations are increasingly attempting to exploit the benefits of the World Wide Web and its features as an interface for internet-enabled businesses, information provision, and promotional activities [1,2]. As the environment becomes more competitive and websites become more sophisticated, attention is turning toward increasing the quality of the website itself and quality evaluation to attract new and retain existing users [3,4]. What determines website quality has not been conclusively established, and there are many different definitions and meanings of the term quality, mainly in relation to the website's purpose [5]. Traditionally, website evaluations have focused on usability, defined as "the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use [6].” The design of websites and users’ needs go beyond pure usability, as increased engagement and pleasure experienced during interactions with websites can be more important predictors of website preference than usability [7-10]. Therefore, in the last decade, website evaluations have shifted their focus to users’ experience, employing various assessment techniques [11], with no universally accepted method or procedure for website evaluation.

Objectives

This scoping study aimed to review and define existing worldwide methodologies and techniques to evaluate websites and provide a simple framework of appropriate website attributes, which could be applied to future website evaluations.

A scoping study is similar to a systematic review as it collects and reviews content in a field of interest. However, scoping studies cover a broader question and do not rigorously evaluate the quality of the studies included [12]. Scoping studies are commonly used in the fields of public services such as health and education, as they are more rapid to perform and less costly in terms of staff costs [13]. Scoping studies can be precursors to a systematic review or stand-alone studies to examine the range of research around a particular topic.

The following research question is based on the need to gain knowledge and insight from worldwide website evaluation to inform the future study design of website evaluations: what website evaluation methodologies can be robustly used to assess users’ experience?

To show how the framework of attributes and methods can be applied to evaluating a website, e-Bug, an international educational health website, will be used as a case scenario [14].

Methods

This scoping study followed a 5-stage framework and methodology, as outlined by Arksey and O’Malley [12], involving the following: (1) identifying the research question, as above; (2) identifying relevant studies; (3) study selection; (4) charting the data; and (5) collating, summarizing, and reporting the results.

Identifying Relevant Studies

Following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines [15], studies for consideration in the review were located by searching the following electronic databases: Excerpta Medica dataBASE, PsycINFO, Cochrane, Cumulative Index to Nursing and Allied Health Literature, Scopus, ACM digital library, and IEEE Xplore SPORTDiscus. The keywords used referred to the following:

- Population: websites
- Intervention: evaluation methodologies
- Outcome: user’s experience.

Table 1 shows the specific search criteria for each database. These keywords were also used to search gray literature for unpublished or working documents to minimize publication bias.

Table 1. Full search strategy used to search each electronic database.

Database	Search criteria	Note	Hits (n)
EMBASE[a]	(web* OR internet OR online) AND (user test* OR heuristic	Search on the field “Title”	689
PsychINFO	evaluation OR usability OR evaluation method* OR measur* OR eye-track* OR eye track*	Search on the field “Title”	816
Cochrane	OR metric* OR rat* OR rank* OR question* OR survey OR stud* OR thinking aloud OR think aloud OR observ* OR complet* OR evaluat*	Search on the fields “Title, keywords, abstract”	1004
CINAHL[b]	OR attribut* OR task*) AND (satisf* OR quality OR efficien* OR task efficiency OR effective* OR appear* OR content	Search on the field “Title”	263
Scopus	OR loyal* OR promot* OR adequa* OR eas* OR user* OR experien*);	Search on the field “Title”	3714
ACM[c] Digital Library	Publication date=between 2006 and 2016; Language published in=English	Search on the field “Title”	89
IEEE[d] Xplore	(web) AND (evaluat) AND (satisf* OR user* OR quality*); Publication date=between 2006 and 2016; Language published in=English	Search on the field “Title”	82

[a] EMBASE: Excerpta Medica database.

[b] CINAHL: Cumulative Index to Nursing and Allied Health Literature.

[c] ACM: Association for Computing Machinery.

[d] IEEE: Institute of Electrical and Electronics Engineers.

Study Selection

Once all sources had been systematically searched, the list of citations was exported to EndNote software to identify eligible studies. By scanning the title, and abstract if necessary, studies that did not fit the inclusion criteria were removed by 2 researchers (RA and CH). As abstracts are not always representative of the full study that follows or capture the full scope [16], if the title and abstract did not provide sufficient information, the full manuscript was examined to ascertain whether they met all the inclusion criteria, which included (1) studies focused on websites, (2) studies of evaluative methods (eg, use of questionnaire and task completion), (3) studies that reported outcomes that affect the user's experience (eg, quality, satisfaction, efficiency, effectiveness without necessarily focusing on methodology), (4) studies carried out between 2006 and 2016, (5) studies published in English, and (6) type of study (any study design that is appropriate).

Exclusion criteria included (1) studies that focus on evaluations using solely experts and are not transferrable to user evaluations; (2) studies that are in the form of electronic book or are not freely available on the Web or through OpenAthens, the University of Bath library, or the University of the West of England library; (3)studies that evaluate banking, electronic commerce (e-commerce), or online libraries' websites and do not have transferrable measures to a range of other websites; (4) studies that report exclusively on minority or special needs groups (eg, blind or deaf users); and (5) studies that do not meet all the inclusion criteria.

Charting the Data

The next stage involved *charting* key items of information obtained from studies being reviewed. *Charting* [17] describes a technique for synthesizing and interpreting qualitative data by sifting, charting, and sorting material according to key issues and themes. This is similar to a systematic review in which the process is called data extraction. The data extracted included general information about the study and specific information relating to, for instance, the study population or target, the type of intervention, outcome measures employed, and the study design.

The information of interest included the following: type of website, aim or purpose of the study, study populations (users and experts), sample size, setting (laboratory, real life, and remotely assessed), website attributes evaluated, process of methodology, and process of analysis.

NVivo version 10.0 software was used for this stage by 2 researchers (RA and CH) to chart the data.

Collating, Summarizing, and Reporting the Results

Although the scoping study does not seek to assess the quality of evidence, it does present an overview of all material reviewed with a narrative account of findings.

Ethics Approval and Consent to Participate

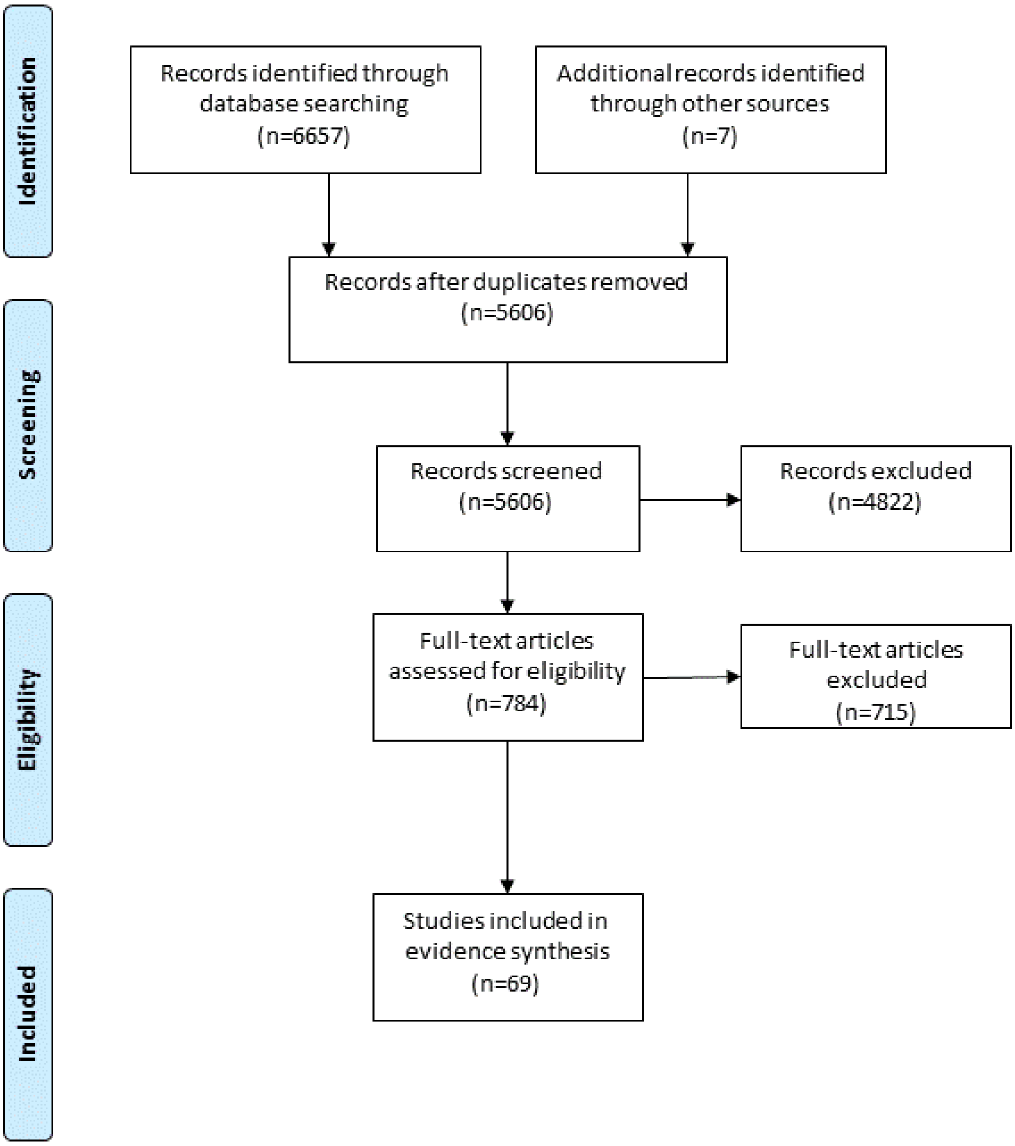
As no primary research was carried out, no ethical approval was required to undertake this scoping study. No specific reference was made to any of the participants in the individual studies, nor does this study infringe on their rights in any way.

Results

Study Selection

The electronic database searches produced 6657 papers; a further 7 papers were identified through other sources. After removing duplicates (n=1058), 5606 publications remained. After titles and abstracts were examined, 784 full-text papers were read and assessed further for eligibility. Of those, 69 articles were identified as suitable by meeting all the inclusion criteria (Figure 1).

Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses flowchart of search results.



Study Characteristics

Population

Studies referred to or used a mixture of users (72%) and experts (39%) to evaluate their websites; 54% used a controlled environment, and 26% evaluated websites remotely (Multimedia Appendix 1 [2-4,11,18-85]). Remote usability, in its most basic form, involves working with participants who are not in the same physical location as the researcher, employing techniques such as live screen sharing or questionnaires. Advantages to remote website evaluations include the ability to evaluate using a larger number of participants as travel time and costs are not a factor, and participants are able to partake at a time that is appropriate to them, increasing the likelihood of participation and the possibility of a greater diversity of participants [18]. However, the disadvantages of remote website evaluations, in comparison with a controlled setting, are that system performance, network traffic, and the participant's computer setup can all affect the results.

A variety of types of websites evaluated were included in this review including government (9%), online news (6%), education (1%), university (12%), and sports organizations (4%). The aspects of quality considered, and their relative importance varied according to the type of website and the goals to be achieved by the users. For example, criteria such as ease of paying or security are not very important to educational websites, whereas they are especially important for online shopping. In this sense, much attention must be paid when evaluating the quality of a website, establishing a specific context of use and purpose [19].

The context of the participants was also discussed, in relation to the generalizability of results. For example, when evaluations used potential or current users of their website, it was important that computer literacy was reflective of all users [20]. This could mean ensuring that participants with a range of computer abilities and experiences were used so that results were not biased to the most or least experienced users.

Intervention

A total of 43 evaluation methodologies were identified in the 69 studies in this review. Most of them were variations of similar methodologies, and a brief description of each is provided in Multimedia Appendix 2. Multimedia Appendix 3 shows the methods used or described in each study.

Questionnaire

Use of questionnaires was the most common methodology referred to (37/69, 54%), including questions to rank or rate attributes and open questions to allow text feedback and suggested improvements. Questionnaires were used in a combination of before or after usability testing to assess usability and overall user experience.

Observed Browsing the Website

Browsing the website using a form of task completion with the participant, such as cognitive walkthrough, was used in 33/69 studies (48%), whereby an expert evaluator used a detailed procedure to simulate task execution and browse all particular solution paths, examining each action while determining if expected user's goals and memory content would lead to choosing a correct option [30]. Screen capture was often used (n=6) to record participants' navigation through the website, and eye tracking was used (n=7) to assess where the eye focuses on each page or the motion of the eye as an individual views a Web page. The think-aloud protocol was used (n=10) to encourage users to express out loud what they were looking at, thinking, doing, and feeling, as they performed tasks. This allows observers to see and understand the cognitive processes associated with task completion. Recording the time to complete tasks (n=6) and mouse movement or clicks (n=8) were used to assess the efficiency of the websites.

Qualitative Data Collection

Several forms of qualitative data collection were used in 27/69 studies (39%). Observed browsing, interviews, and focus groups were used either before or after the use of the website. Pre-website-use, qualitative research was often used to collect details of which website attributes were important for participants or what weighting participants would give to each attribute. Postevaluation, qualitative techniques were used to collate feedback on the quality of the website and any suggestions for improvements.

Automated Usability Evaluation Software

In 9/69 studies (13%), automated usability evaluation focused on developing software, tools, and techniques to speed evaluation (rapid), tools that reach a wider audience for usability testing (remote), and tools that have built-in analyses features (automated). The latter can involve assessing server logs, website coding, and simulations of user experience to assess usability [42].

Card Sorting

A technique that is often linked with assessing navigability of a website, card sorting, is useful for discovering the logical structure of an unsorted list of statements or ideas by exploring how people group items and structures that maximize the probability of users finding items (5/69 studies, 7%). This can assist with determining effective website structure.

Web Usage Analysis

Of 69 studies, 3 studies used Web usage analysis or Web analytics to identify browsing patterns by analyzing the participants' navigational behavior. This could include tracking at the widget level, that is, combining knowledge of the mouse coordinates with elements such as buttons and links, with the layout of the HTML pages, enabling complete tracking of all user activity.

Outcomes (Attributes Used to Evaluate Websites)

Often, different terminology for website attributes was used to describe the same or similar concepts (Multimedia Appendix 4). The most used website attributes that were assessed can be broken down into 8 broad categories and further subcategories:

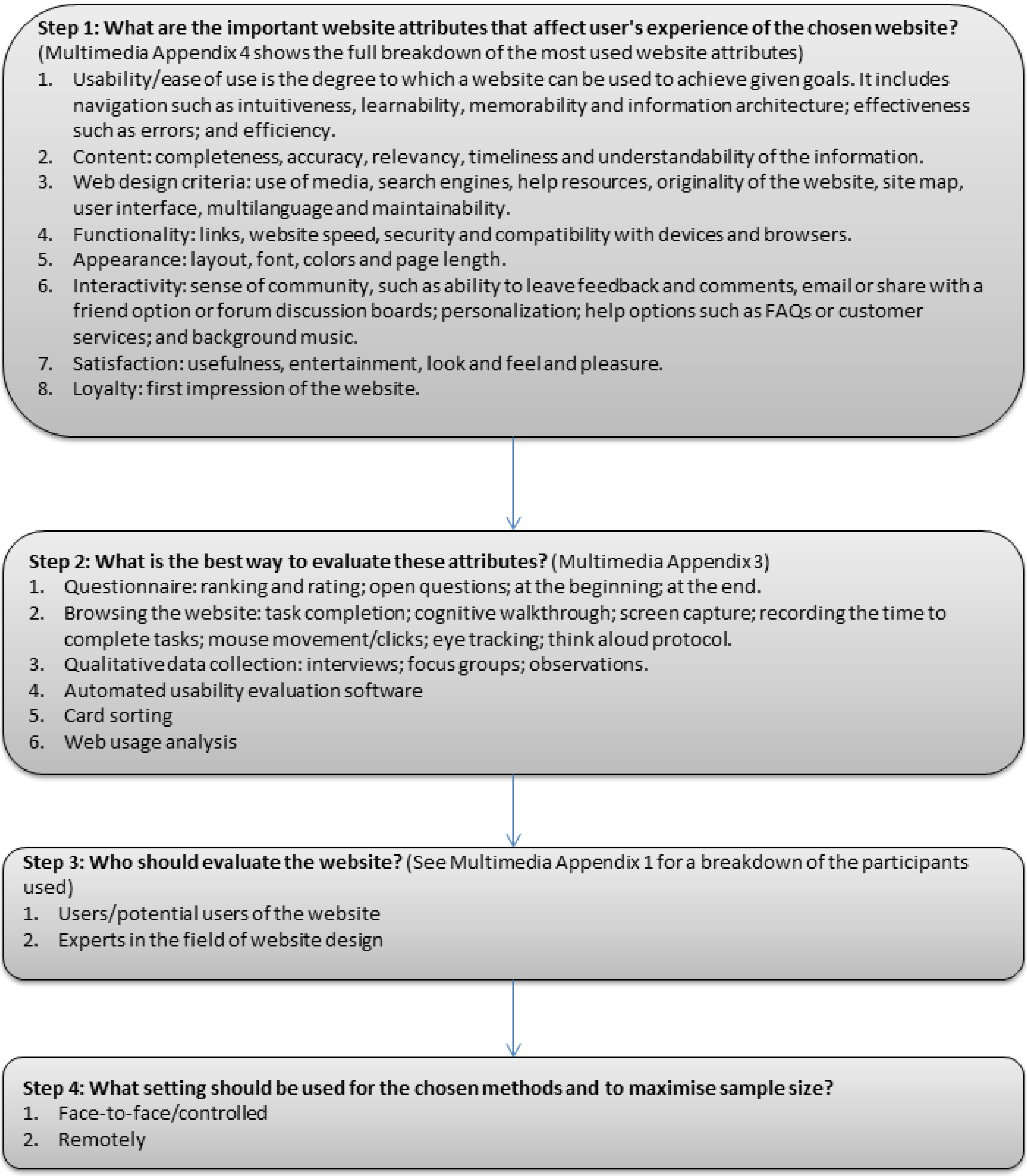
1. Usability or ease of use is the degree to which a website can be used to achieve given goals (n=58). It includes navigation such as intuitiveness, learnability, memorability, and information architecture; effectiveness such as errors; and efficiency.
2. Content (n=41) includes completeness, accuracy, relevancy, timeliness, and understandability of the information.
3. Web design criteria (n=29) include use of media, search engines, help resources, originality of the website, site map, user interface, multilanguage, and maintainability.
4. Functionality (n=31) includes links, website speed, security, and compatibility with devices and browsers.
5. Appearance (n=26) includes layout, font, colors, and page length.
6. Interactivity (n=25) includes sense of community, such as ability to leave feedback and comments and email or share with a friend option or forum discussion boards; personalization; help options such as frequently answered questions or customer services; and background music.
7. Satisfaction (n=26) includes usefulness, entertainment, look and feel, and pleasure.
8. Loyalty (n=8) includes first impression of the website.

Discussion

GoodWeb: Website Evaluation Guide

As there was such a range of methods used, a suggested guide of options for evaluating websites is presented below (Figure 2), coined GoodWeb, and applied to an evaluation of e-Bug, an international educational health website [14]. Allison at al [86] show the full details of how GoodWeb has been applied and outcomes of the e-Bug website evaluation.

Figure 2. Framework for website evaluation.



Step 1: What are the important website attributes that affect user's experience of the chosen website? (Multimedia Appendix 4 shows the full breakdown of the most used website attributes)

1. Usability/ease of use is the degree to which a website can be used to achieve given goals. It includes navigation such as intuitiveness, learnability, memorability and information architecture; effectiveness such as errors; and efficiency.
2. Content: completeness, accuracy, relevancy, timeliness and understandability of the information.
3. Web design criteria: use of media, search engines, help resources, originality of the website, site map, user interface, multilanguage and maintainability.
4. Functionality: links, website speed, security and compatibility with devices and browsers.
5. Appearance: layout, font, colors and page length.
6. Interactivity: sense of community, such as ability to leave feedback and comments, email or share with a friend option or forum discussion boards; personalization; help options such as FAQs or customer services; and background music.
7. Satisfaction: usefulness, entertainment, look and feel and pleasure.
8. Loyalty: first impression of the website.

Step 2: What is the best way to evaluate these attributes? (Multimedia Appendix 3)

1. Questionnaire: ranking and rating; open questions; at the beginning; at the end.
2. Browsing the website: task completion; cognitive walkthrough; screen capture; recording the time to complete tasks; mouse movement/clicks; eye tracking; think aloud protocol.
3. Qualitative data collection: interviews; focus groups; observations.
4. Automated usability evaluation software
5. Card sorting
6. Web usage analysis

Step 3: Who should evaluate the website? (See Multimedia Appendix 1 for a breakdown of the participants used)

1. Users/potential users of the website
2. Experts in the field of website design

Step 4: What setting should be used for the chosen methods and to maximise sample size?

1. Face-to-face/controlled
2. Remotely

Step 1. What Are the Important Website Attributes That Affect User's Experience of the Chosen Website?

Usability or ease of use, content, Web design criteria, functionality, appearance, interactivity, satisfaction, and loyalty were the umbrella terms that encompassed the website attributes identified or evaluated in the 69 studies in this scoping study. Multimedia Appendix 4 contains a summary of the most used website attributes that have been assessed. Recent website evaluations have shifted focus from usability of websites to an overall user's experience of website use. A decision on which website attributes to evaluate for specific websites could come from interviews or focus groups with users or experts or a literature search of attributes used in similar evaluations.

Application

In the scenario of evaluating e-Bug or similar educational health websites, the attributes chosen to assess could be the following:

1. Appearance: colors, fonts, media or graphics, page length, style consistency, and first impression
2. Content: clarity, completeness, current and timely information, relevance, reliability, and uniqueness
3. Interactivity: sense of community and modern features

4. Ease of use: home page indication, navigation, guidance, and multilanguage support
5. Technical adequacy: compatibility with other devices, load time, valid links, and limited use of special plug-ins
6. Satisfaction: loyalty

These cover the main website attributes appropriate for an educational health website. If the website did not currently have features such as search engines, site map, background music, it may not be appropriate to evaluate these, but may be better suited to question whether they would be suitable additions to the website; or these could be combined under the heading *modern features*. Furthermore, security may not be a necessary attribute to evaluate if participant identifiable information or bank details are not needed to use the website.

Step 2. What Is the Best Way to Evaluate These Attributes?

Often, a combination of methods is suitable to evaluate a website, as 1 method may not be appropriate to assess all attributes of interest [29] (see Multimedia Appendix 3 for a summary of the most used methods for evaluating websites). For example, screen capture of task completion may be appropriate to assess the efficiency of a website but would not be the chosen method to assess loyalty. A questionnaire or qualitative interview may be more appropriate for this attribute.

Application

In the scenario of evaluating e-Bug, a questionnaire before browsing the website would be appropriate to rank the importance of the selected website attributes, chosen in step 1. It would then be appropriate to observe browsing of the website, collecting data on completion of typical task scenarios, using the screen capture function for future reference. This method could be used to evaluate the effectiveness (number of tasks successfully completed), efficiency (whether the most direct route through the website was used to complete the task), and learnability (whether task completion is more efficient or effective second time of trying). It may then be suitable to use a follow-up questionnaire to rate e-Bug against the website attributes previously ranked. The attribute ranking and rating could then be combined to indicate where the website performs well and areas for improvement.

Step 3: Who Should Evaluate the Website?

Both users and experts can be used to evaluate websites. Experts are able to identify areas for improvements, in relation to usability; whereas, users are able to appraise quality as well as identify areas for improvement. In this respect, users are able to fully evaluate user’s experience, where experts may not be able to.

Application

For this reason, it may be more appropriate to use current or potential users of the website for the scenario of evaluating e-Bug.

Step 4: What Setting Should Be Used?

A combination of controlled and remote settings can be used, depending on the methods chosen. For example, it may be appropriate to collect data via a questionnaire, remotely, to increase sample size and reach a more diverse audience, whereas a controlled setting may be more appropriate for task completion using eye-tracking methods.

Strengths and Limitations

A scoping study differs from a systematic review, in that it does not critically appraise the quality of the studies before extracting or *charting* the data. Therefore, this study cannot compare the effectiveness of the different methods or methodologies in evaluating the website attributes. However, what it does do is review and summarize a huge amount of literature, from different sources, in a format that is understandable and informative for future designs of website evaluations.

Furthermore, studies that evaluate banking, e-commerce, or online libraries’ websites and do not have transferrable measures to a range of other websites were excluded from this study. This decision was made to limit the number of studies that met the remaining inclusion criteria, and it was deemed that the website attributes for these websites would be too specialist and not necessarily transferable to a range of websites. Therefore, the findings of this study may not be generalizable to all types of website. However, Multimedia Appendix 1 shows that data were extracted from a very broad range of websites when it was deemed that the information was transferrable to a range of other websites.

A robust website evaluation can identify areas for improvement to both fulfill the goals and desires of its users [62] and influence their perception of the organization and overall quality of resources [48]. An improved website could attract and retain more online users; therefore, an evidence-based website evaluation guide is essential.

Conclusions

This scoping study emphasizes the fact that the debate about how to define the quality of websites remains open, and there are numerous approaches and models to evaluate it. Multimedia Appendix 2 shows existing methodologies or tools that can be used to evaluate websites. Many of these are variations of similar approaches; therefore, it is not strictly necessary to use these tools at face value; however, some could be used to guide analysis, following data collection. By following steps 1 to 4 of GoodWeb, the framework suggested in this study, taking into account the desired participants and setting and website evaluation methods, can be tailored to the needs of specific websites and individual aims of evaluations.

Acknowledgments

This work was supported by the Primary Care Unit, Public Health England. This study is not applicable as secondary research.

Authors' Contributions

RA wrote the protocol with input from CH, CM, and VY. RA and CH conducted the scoping review. RA wrote the final manuscript with input from CH, CM, and VY. All authors reviewed and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Summary of included studies, including information on the participant.
[PDF File (Adobe PDF File), 133 KB - formative_v3i4e14372_app1.pdf]

Multimedia Appendix 2
Interventions: methodologies and tools to evaluate websites.
[PDF File (Adobe PDF File), 85 KB - formative_v3i4e14372_app2.pdf]

Multimedia Appendix 3
Methods used or described in each study.
[XLSX File (Microsoft Excel File), 20 KB - formative_v3i4e14372_app3.xlsx]

Multimedia Appendix 4
Summary of the most used website attributes evaluated.
[XLSX File (Microsoft Excel File), 26 KB - formative_v3i4e14372_app4.xlsx]

References

1. Straub DW, Watson RT. Research Commentary: Transformational Issues in Researching IS and Net-Enabled Organizations. Info Syst Res 2001;12(4):337-345. [doi: 10.1287/isre.12.4.337.9706]
2. Bairamzadeh S, Bolhari A. Investigating factors affecting students' satisfaction of university websites. In: 2010 3rd International Conference on Computer Science and Information Technology. 2010 Presented at: ICCSIT'10; July 9-11, 2010; Chengdu, China p. 469-473. [doi: 10.1109/iccsit.2010.5565050]
3. Fink D, Nyaga C. Evaluating web site quality: the value of a multi paradigm approach. Benchmarking 2009;16(2):259-273. [doi: 10.1108/14635770910948259]
4. Markaki OI, Charilas DE, Askounis D. Application of Fuzzy Analytic Hierarchy Process to Evaluate the Quality of E-Government Web Sites. In: Proceedings of the 2010 Developments in E-systems Engineering. 2010 Presented at: DeSE'10; September 6-8, 2010; London, UK p. 219-224. [doi: 10.1109/dese.2010.42]
5. Eysenbach G, Powell J, Kuss O, Sa E. Empirical studies assessing the quality of health information for consumers on the world wide web: a systematic review. J Am Med Assoc 2002;287(20):2691-2700. [doi: 10.1001/jama.287.20.2691] [Medline: 12020305]
6. International Organization for Standardization. ISO 9241-11: Ergonomic Requirements for Office Work with Visual Display Terminals (VDTs): Part 11: Guidance on Usability. Switzerland: International Organization for Standardization; 1998.
7. Hartmann J, Sutcliffe A, Angeli AD. Towards a theory of user judgment of aesthetics and user interface quality. ACM Trans Comput-Hum Interact 2008;15(4):1-30. [doi: 10.1145/1460355.1460357]
8. Bargas-Avila JA, Hornbæk K. Old Wine in New Bottles or Novel Challenges: A Critical Analysis of Empirical Studies of User Experience. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2011 Presented at: CHI'11; May 7-12, 2011; Vancouver, BC, Canada p. 2689-2698. [doi: 10.1145/1978942.1979336]
9. Hassenzahl M, Tractinsky N. User experience - a research agenda. Behav Info Technol 2006;25(2):91-97. [doi: 10.1080/01449290500330331]
10. Aranyi G, van Schaik P. Testing a model of user-experience with news websites. J Assoc Soc Inf Sci Technol 2016;67(7):1555-1575. [doi: 10.1002/asi.23462]
11. Tsai W, Chou W, Lai C. An effective evaluation model and improvement analysis for national park websites: a case study of Taiwan. Tour Manag 2010;31(6):936-952. [doi: 10.1016/j.tourman.2010.01.016]
12. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Method 2005;8(1):19-32. [doi: 10.1080/1364557032000119616]

13. Anderson S, Allen P, Peckham S, Goodwin N. Asking the right questions: scoping studies in the commissioning of research on the organisation and delivery of health services. Health Res Policy Syst 2008;6:7 [FREE Full text] [doi: 10.1186/1478-4505-6-7] [Medline: 18613961]
14. e-Bug. 2018. Welcome to the e-Bug Teachers Area! URL: https://e-bug.eu/eng_home.aspx?cc=eng&ss=1&t=Welcome%20to%20e-Bug [accessed 2019-08-23]
15. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. PLoS Med 2009 Jul 21;6(7):e1000100 [FREE Full text] [doi: 10.1371/journal.pmed.1000100] [Medline: 19621070]
16. Badger D, Nursten J, Williams P, Woodward M. Should All Literature Reviews be Systematic? Eval Res Educ 2000;14(3-4):220-230. [doi: 10.1080/09500790008666974]
17. Ritchie J, Spencer L. Qualitative data analysis for applied policy research. In: Bryman A, Burgess B, editors. Analyzing Qualitative Data. Abingdon-on-Thames: Routledge; 2002:187-208.
18. Thomsett-Scott BC. Web site usability with remote users: Formal usability studies and focus groups. J Libr Adm 2006;45(3-4):517-547. [doi: 10.1300/j111v45n03_14]
19. Moreno JM, Morales del Castillo JM, Porcel C, Herrera-Viedma E. A quality evaluation methodology for health-related websites based on a 2-tuple fuzzy linguistic approach. Soft Comput 2010;14(8):887-897. [doi: 10.1007/s00500-009-0472-7]
20. Alva M, Martínez A, Labra Gayo J, Del Carmen Suárez M, Cueva J, Sagástegui H. Proposal of a tool of support to the evaluation of user in educative web sites. 2008 Presented at: 1st World Summit on the Knowledge Society, WSKS 2008; 2008; Athens p. 149-157. [doi: 10.1007/978-3-540-87781-3_17]
21. Usability: Home. Usability Evaluation Basics URL: https://www.usability.gov/what-and-why/usability-evaluation.html [accessed 2019-08-24]
22. AddThis: Get more likes, shares and follows with smart. 10 Criteria for Better Website Usability: Heuristics Cheat Sheet URL: http://www.addthis.com/blog/2015/02/17/10-criteria-for-better-website-usability-heuristics-cheat-sheet/#.V712QfkrJD8 [accessed 2019-08-24]
23. Akgül Y. Quality evaluation of E-government websites of Turkey. In: Proceedings of the 2016 11th Iberian Conference on Information Systems and Technologies. 2016 Presented at: CISTI'16; June 15-18 2016; Las Palmas, Spain p. 1-7. [doi: 10.1109/cisti.2016.7521567]
24. Al Zaghoul FA, Al Nsour AJ, Rababah OM. Ranking Quality Factors for Measuring Web Service Quality. In: Proceedings of the 1st International Conference on Intelligent Semantic Web-Services and Applications. 2010 Presented at: 1st ACM Jordan Professional Chapter ISWSA Annual - International Conference on Intelligent Semantic Web-Services and Applications, ISWSA'10; June 14-16, 2010; Amman, Jordan. [doi: 10.1145/1874590.1874608]
25. Alharbi A, Mayhew P. Users' Performance in Lab and Non-lab Enviornments Through Online Usability Testing: A Case of Evaluating the Usability of Digital Academic Libraries' Websites. In: 2015 Science and Information Conference. 2015 Presented at: SAI'15; July 28-30, 2015; London, UK p. 151-161. [doi: 10.1109/sai.2015.7237139]
26. Aliyu M, Mahmud M, Md Tap AO. Preliminary Investigation of Islamic Websites Design & Content Feature: A Heuristic Evaluation From User Perspective. In: Proceedings of the 2010 International Conference on User Science and Engineering. 2010 Presented at: iUSEr'10; December 13-15, 2010; Shah Alam, Malaysia p. 262-267. [doi: 10.1109/iuser.2010.5716763]
27. Aliyu M, Mahmud M, Tap AO, Nassr RM. Evaluating Design Features of Islamic Websites: A Muslim User Perception. In: Proceedings of the 2013 5th International Conference on Information and Communication Technology for the Muslim World. 2013 Presented at: ICT4M'13; March 26-27, 2013; Rabat, Morocco. [doi: 10.1109/ict4m.2013.6518878]
28. Al-Radaideh QA, Abu-Shanab E, Hamam S, Abu-Salem H. Usability evaluation of online news websites: a user perspective approach. World Acad Sci Eng Technol 2011;74:1058-1066 [FREE Full text]
29. Aranyi G, van Schaik P, Barker P. Using think-aloud and psychometrics to explore users' experience with a news web site. Interact Comput 2012;24(2):69-77. [doi: 10.1016/j.intcom.2012.01.001]
30. Arrue M, Fajardo I, Lopez JM, Vigo M. Interdependence between technical web accessibility and usability: its influence on web quality models. Int J Web Eng Technol 2007;3(3):307-328. [doi: 10.1504/ijwet.2007.012059]
31. Arrue M, Vigo M, Abascal J. Quantitative metrics for web accessibility evaluation. 2005 Presented at: Proceedings of the ICWE 2005 Workshop on Web Metrics and Measurement; 2005; Sydney.
32. Atterer R, Wnuk M, Schmidt A. Knowing the User's Every Move: User Activity Tracking for Website Usability Evaluation and Implicit Interaction. In: Proceedings of the 15th international conference on World Wide. 2006 Presented at: WWW'06; May 23-26, 2006; Edinburgh, Scotland p. 203-212. [doi: 10.1145/1135777.1135811]
33. Bahry F, Masrom M, Masrek M. Website evaluation measures, website user engagement and website credibility for municipal website. ARPN J Eng Appl Sci 2015;10(23):18228-18238. [doi: 10.1109/iuser.2016.7857963]
34. Bañón-Gomis A, Tomás-Miquel JV, Expósito-Langa M. Improving user experience: a methodology proposal for web usability measurement. In: Strategies in E-Business: Positioning and Social Networking in Online Markets. New York City: Springer US; 2014:123-145.
35. Barnes SJ, Vidgen RT. Data triangulation and web quality metrics: a case study in e-government. Inform Manag 2006;43(6):767-777. [doi: 10.1016/j.im.2006.06.001]

36. Bolchini D, Garzotto F. Quality of Web Usability Evaluation Methods: An Empirical Study on MiLE+. In: Proceedings of the 2007 international conference on Web information systems engineering. 2007 Presented at: WISE'07; December 3-3, 2007; Nancy, France p. 481-492. [doi: 10.1007/978-3-540-77010-7_47]
37. Chen FH, Tzeng G, Chang CC. Evaluating the enhancement of corporate social responsibility websites quality based on a new hybrid MADM model. Int J Inf Technol Decis Mak 2015;14(03):697-724. [doi: 10.1142/s0219622015500121]
38. Cherfi SS, Tuan AD, Comyn-Wattiau I. An Exploratory Study on Websites Quality Assessment. In: Proceedings of the 32nd International Conference on Conceptual Modeling Workshops. 2013 Presented at: ER'13; November 11-13, 2014; Hong Kong, China p. 170-179. [doi: 10.1007/978-3-319-14139-8_19]
39. Chou W, Cheng Y. A hybrid fuzzy MCDM approach for evaluating website quality of professional accounting firms. Expert Sys Appl 2012;39(3):2783-2793. [doi: 10.1016/j.eswa.2011.08.138]
40. Churm T. Usability Geek. 2012 Jul 9. An Introduction To Website Usability Testing URL: http://usabilitygeek.com/an-introduction-to-website-usability-testing/ [accessed 2019-08-24]
41. Demir Y, Gozum S. Evaluation of quality, content, and use of the web site prepared for family members giving care to stroke patients. Comput Inform Nurs 2015 Sep;33(9):396-403. [doi: 10.1097/CIN.0000000000000165] [Medline: 26125455]
42. Dominic P, Jati H, Hanim S. University website quality comparison by using non-parametric statistical test: a case study from Malaysia. Int J Oper Res 2013;16(3):349-374. [doi: 10.1504/ijor.2013.052335]
43. Elling S, Lentz L, de Jong M, van den Bergh H. Measuring the quality of governmental websites in a controlled versus an online setting with the ‘Website Evaluation Questionnaire’. Gov Inf Q 2012;29(3):383-393. [doi: 10.1016/j.giq.2011.11.004]
44. Fang X, Holsapple CW. Impacts of navigation structure, task complexity, and users’ domain knowledge on web site usability—an empirical study. Inf Syst Front 2011;13(4):453-469. [doi: 10.1007/s10796-010-9227-3]
45. Fernandez A, Abrahão S, Insfran E. A systematic review on the effectiveness of web usability evaluation methods. In: Proceedings of the 16th International Conference on Evaluation & Assessment in Software Engineering. 2012 Presented at: EASE'12; May 14-15, 2012; Ciudad Real, Spain. [doi: 10.1049/ic.2012.0007]
46. Flavián C, Guinalíu M, Gurrea R. The role played by perceived usability, satisfaction and consumer trust on website loyalty. Inform Manag 2006;43(1):1-14. [doi: 10.1016/j.im.2005.01.002]
47. Flavián C, Guinalíu M, Gurrea R. The influence of familiarity and usability on loyalty to online journalistic services: the role of user experience. J Retail Consum Serv 2006;13(5):363-375. [doi: 10.1016/j.jretconser.2005.11.003]
48. Gonzalez ME, Quesada G, Davis J, Mora-Monge C. Application of quality management tools in the evaluation of websites: the case of sports organizations. Qual Manage J 2015;22(1):30-46. [doi: 10.1080/10686967.2015.11918417]
49. Harrison C, Pétrie H. Deconstructing Web Experience: More Than Just Usability and Good Design. In: Proceedings of the 12th international conference on Human-computer interaction: applications and services. 2007 Presented at: HCI'07; July 22-27, 2007; Beijing, China p. 889-898. [doi: 10.1007/978-3-540-73111-5_98]
50. Hart D, Portwood DM. Usability Testing of Web Sites Designed for Communities of Practice: Tests of the IEEE Professional Communication Society (PCS) Web Site Combining Specialized Heuristic Evaluation and Task-based User Testing. In: Proceedings of the 2009 IEEE International Professional Communication Conference. 2009 Presented at: 2009 IEEE International Professional Communication Conference; July 19-22, 2009; Waikiki, HI, USA. [doi: 10.1109/ipcc.2009.5428218]
51. Hedegaard S, Simonsen JG. Extracting Usability and User Experience Information From Online User Reviews. 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. 2089-2098. [doi: 10.1145/2470654.2481286]
52. Herrera F, Herrera-Viedma E, Martínez L, Ferez LG, López-Herrera AG, Alonso S. A multi-granular linguistic hierarchical model to evaluate the quality of web site services. In: Mathew S, Mordeson JN, Malik DS, editors. Studies in Fuzziness and Soft Computing. New York City: Springer; 2006:247-274.
53. Hinchliffe A, Mummery WK. Applying usability testing techniques to improve a health promotion website. Health Promot J Austr 2008 Apr;19(1):29-35. [doi: 10.1071/HE08029] [Medline: 18481929]
54. Ijaz T, Andlib F. Impact of Usability on Non-technical Users: Usability Testing Through Websites. In: Proceedings of the 2014 National Software Engineering Conference. 2014 Presented at: 2014 National Software Engineering Conference, NSEC 2014; November 11-12, 2014; Rawalpindi, Pakistan. [doi: 10.1109/nsec.2014.6998238]
55. Janiak E, Rhodes E, Foster AM. Translating access into utilization: lessons from the design and evaluation of a health insurance Web site to promote reproductive health care for young women in Massachusetts. Contraception 2013 Dec;88(6):684-690. [doi: 10.1016/j.contraception.2013.09.004] [Medline: 24120249]
56. Kaya T. Multi-attribute evaluation of website quality in e-business using an integrated fuzzy AHP-TOPSIS methodology. Int J Comput Intell Syst 2010;3(3):301-314. [doi: 10.2991/ijcis.2010.3.3.6]
57. Kincl T, Štrach P. Measuring website quality: asymmetric effect of user satisfaction. Behav Inform Technol 2012;31(7):647-657. [doi: 10.1080/0144929x.2010.526150]
58. Koutsabasis P, Istikopoulou TG. Perceived website aesthetics by users and designers: implications for evaluation practice. Int J Technol Human Interact 2014;10(2):21-34. [doi: 10.4018/ijthi.2014040102]

59. Leuthold S, Schmutz P, Bargas-Avila JA, Tuch AN, Opwis K. Vertical versus dynamic menus on the world wide web: eye tracking study measuring the influence of menu design and task complexity on user performance and subjective preference. Comput Human Behav 2011;27(1):459-472. [doi: 10.1016/j.chb.2010.09.009]
60. Longstreet P. Evaluating Website Quality: Applying Cue Utilization Theory to WebQual. In: Proceedings of the 2010 43rd Hawaii International Conference on System Sciences. 2010 Presented at: HICSS'10; January 5-8, 2010; Honolulu, HI, USA. [doi: 10.1109/hicss.2010.191]
61. Manzoor M. Measuring user experience of usability tool, designed for higher educational websites. Middle East J Sci Res 2013;14(3):347-353. [doi: 10.5829/idosi.mejsr.2013.14.3.2053]
62. Mashable India. 22 Essential Tools for Testing Your Website's Usability URL: http://mashable.com/2011/09/30/website-usability-tools/#cNv8ckxZsmqw [accessed 2019-08-16]
63. Matera M, Rizzo F, Carughi GT. Web usability: principles and evaluation methods. In: Web Engineering. New York City: Springer; 2006:143-180.
64. McClellan MA, Karumur RP, Vogel RI, Petzel SV, Cragg J, Chan D, et al. Designing an educational website to improve quality of supportive oncology care for women with ovarian cancer: an expert usability review and analysis. Int J Hum Comput Interact 2016;32(4):297-307 [FREE Full text] [doi: 10.1080/10447318.2016.1140528] [Medline: 27110082]
65. Nakamichi N, Shima K, Sakai M, Matsumoto KI. Detecting Low Usability Web Pages Using Quantitative Data of Users' Behavior. In: Proceedings of the 28th international conference on Software engineering. 2006 Presented at: ICSE'06; May 20-28, 2006; Shanghai, China p. 569-576. [doi: 10.1145/1134285.1134365]
66. Nathan RJ, Yeow PH. An empirical study of factors affecting the perceived usability of websites for student internet users. Univ Access Inf Soc 2009;8(3):165-184. [doi: 10.1007/s10209-008-0138-8]
67. Oliver H, Diallo G, de Quincey E, Alexopoulou D, Habermann B, Kostkova P, et al. A user-centred evaluation framework for the Sealife semantic web browsers. BMC Bioinform 2009;10(S10). [doi: 10.1186/1471-2105-10-s10-s14]
68. Paul A, Yadamsuren B, Erdelez S. An Experience With Measuring Multi-User Online Task Performance. In: Proceedings of the 2012 World Congress on Information and Communication Technologies. 2012 Presented at: WICT'12; October 30-November 2, 2012; Trivandrum, India p. 639-644. [doi: 10.1109/wict.2012.6409154]
69. Petrie H, Power C. What Do Users Really Care About?: A Comparison of Usability Problems Found by Users and Experts on Highly Interactive Websites. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2012 Presented at: CHI'12; May 5-10, 2012; Austin, Texas, USA p. 2107-2116. [doi: 10.1145/2207676.2208363]
70. Rekik R, Kallel I. Fuzzy Reduced Method for Evaluating the Quality of Institutional Web Sites. In: Proceedings of the 2011 7th International Conference on Next Generation Web Services Practices. 2011 Presented at: NWeSP'11; October 19-21, 2011; Salamanca, Spain p. 296-301. [doi: 10.1109/nwesp.2011.6088194]
71. Reynolds E. The secret to patron-centered Web design: cheap, easy, and powerful usability techniques. Comput Librar 2008;28(6):44-47 [FREE Full text]
72. Sheng H, Lockwood NS, Dahal S. Eyes Don't Lie: Understanding Users' First Impressions on Websites Using Eye Tracking. In: Proceedings of the 15th International Conference on Human Interface and the Management of Information: Information and Interaction Design. 2013 Presented at: HCI'13; July 21-26, 2013; Las Vegas, NV, USA p. 635-641. [doi: 10.1007/978-3-642-39209-2_71]
73. Swaid SI, Wigand RT. Measuring Web-based Service Quality: The Online Customer Point of View. In: Proceedings of the 13th Americas Conference on Information Systems. 2007 Presented at: AMCIS'07; August 9-12, 2007; Keystone, Colorado, USA p. 778-790.
74. Tan GW, Wei KK. An empirical study of Web browsing behaviour: towards an effective Website design. Elect Commer Res Appl 2006;5(4):261-271. [doi: 10.1016/j.elerap.2006.04.007]
75. Tan W, Liu D, Bishu R. Web evaluation: heuristic evaluation vs user testing. Int J Ind Ergonom 2009;39(4):621-627. [doi: 10.1016/j.ergon.2008.02.012]
76. Tao D, LeRouge CM, Deckard G, de Leo G. Consumer Perspectives on Quality Attributes in Evaluating Health Websites. In: Proceedings of the 2012 45th Hawaii International Conference on System Sciences. 2012 Presented at: HICSS'12; January 4-7, 2012; Maui, HI, USA. [doi: 10.1109/hicss.2012.180]
77. The Whole Brain Group. 2011. Conducting a Quick & Dirty Evaluation of Your Website's Usability URL: http://blog.thewholebraingroup.com/conducting-quick-dirty-evaluation-websites-usability [accessed 2019-08-24]
78. Thielsch MT, Blotenberg I, Jaron R. User evaluation of websites: From first impression to recommendation. Interact Comput 2014;26(1):89-102. [doi: 10.1093/iwc/iwt033]
79. Tung LL, Xu Y, Tan FB. Attributes of web site usability: a study of web users with the repertory grid technique. Int J Elect Commer 2009;13(4):97-126. [doi: 10.2753/jec1086-4415130405]
80. Agarwal R, Venkatesh V. Assessing a firm's Web presence: a heuristic evaluation procedure for the measurement of usability. Inform Syst Res 2002;13(2):168-186. [doi: 10.1287/isre.13.2.168.84]
81. Venkatesh V, Ramesh V. Web and wireless site usability: understanding differences and modeling use. Manag Inf Syst Q 2006;30(1):181-206. [doi: 10.2307/25148723]

82. Vaananen-Vainio-Mattila K, Wäljas M. Development of Evaluation Heuristics for Web Service User Experience. In: Proceedings of the Extended Abstracts on Human Factors in Computing Systems. 2009 Presented at: CHI'09; April 4-9, 2009; Boston, MA, USA p. 3679-3684. [doi: 10.1145/1520340.1520554]
83. Wang WT, Wang B, Wei YT. Examining the Impacts of Website Complexities on User Satisfaction Based on the Task-technology Fit Model: An Experimental Research Using an Eyetracking Device. In: Proceedings of the 18th Pacific Asia Conference on Information Systems. 2014 Presented at: PACIS'14; June 18-22, 2014; Jeju Island, South Korea.
84. Yen B, Hu PJ, Wang M. Toward an analytical approach for effective web site design: a framework for modeling, evaluation and enhancement. Elect Commer Res Appl 2007;6(2):159-170. [doi: 10.1016/j.elerap.2006.11.004]
85. Yen PY, Bakken S. A comparison of usability evaluation methods: heuristic evaluation versus end-user think-aloud protocol - an example from a web-based communication tool for nurse scheduling. AMIA Annu Symp Proc 2009 Nov 14;2009:714-718 [FREE Full text] [Medline: 20351946]
86. Allison R, Hayes C, Young V, McNulty CAM. Evaluation of an educational health website on infections and antibiotics: a mixed-methods, user-centred approach in England. JMIR Formative Research 2019 (forthcoming).

Abbreviations

E-commerce: electronic commerce

Edited by G Eysenbach; submitted 26.04.19; peer-reviewed by C Eley, C Brown; comments to author 31.05.19; revised version received 24.06.19; accepted 18.08.19; published 24.10.19.

Please cite as:
Allison R, Hayes C, McNulty CAM, Young V
A Comprehensive Framework to Evaluate Websites: Literature Review and Development of GoodWeb
JMIR Form Res 2019;3(4):e14372
URL: http://formative.jmir.org/2019/4/e14372/
doi:10.2196/14372
PMID:31651406

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org