JMIR Formative Research

Impact Factor (2023): 2.0
Volume 4 (2020), Issue 8 ISSN 2561-326X Editor in Chief: Amaryllis Mavragani, PhDc

Contents

Original Papers

Viewpoint

Review

Original Paper

Needs and Requirements in the Designing of Mobile Interventions for Patients With Peripheral Arterial Disease: Questionnaire Study

Julia Lortz[1], MD, PD Dr; Jan Simanovski[2], BSc; Tabea Kuether[3], BSc; Ilonka Kreitschmann-Andermahr[4], MD, Prof Dr; Greta Ullrich[1]; Martin Steinmetz[1], MD, Dr; Christos Rammos[1], MD, PD Dr; Rolf Alexander Jánosi[1], MD; Susanne Moebus[2], PhD, Prof Dr; Tienush Rassaf[1], MD, Prof Dr; Katrin Paldán[2,3], PhD, Dr Phil

[1]Department of Cardiology and Vascular Medicine, West-German Heart and Vascular Center Essen, University of Duisburg-Essen, Essen, Germany

[2]Centre for Urban Epidemiology, Institute for Medical Informatics, Biometry and Epidemiology, University of Duisburg-Essen, Essen, Germany

[3]Centre of Competence Personal Analytics at the University of Duisburg-Essen, Department of Engineering Sciences, University of Duisburg-Essen, Duisburg, Germany

[4]Department of Neurosurgery, University of Duisburg-Essen, Essen, Germany

Corresponding Author:
Julia Lortz, MD, PD Dr
Department of Cardiology and Vascular Medicine
West-German Heart and Vascular Center Essen
University of Duisburg-Essen
Hufelandstraße 55
Essen, 45147
Germany
Phone: 49 201 723 4801
Email: julia.lortz@uk-essen.de

Abstract

Background: The development of mobile interventions for noncommunicable diseases has increased in recent years. However, there is a dearth of apps for patients with peripheral arterial disease (PAD), who frequently have an impaired ability to walk.

Objective: Using a patient-centered approach for the development of mobile interventions, we aim to describe the needs and requirements of patients with PAD regarding the overall care situation and the use of mobile interventions to perform supervised exercise therapy (SET).

Methods: A questionnaire survey was conducted in addition to a clinical examination at the vascular outpatient clinic of the West-German Heart and Vascular Center of the University Clinic Essen in Germany. Patients with diagnosed PAD were asked to answer questions on sociodemographic characteristics, PAD-related need for support, satisfaction with their health care situation, smartphone and app use, and requirements for the design of mobile interventions to support SET.

Results: Overall, a need for better support of patients with diagnosed PAD was identified. In total, 59.2% (n=180) expressed their desire for more support for their disease. Patients (n=304) had a mean age of 67 years and half of them (n=157, 51.6%) were smartphone users. We noted an interest in smartphone-supported SET, even for people who did not currently use a smartphone. “Information,” “feedback,” “choosing goals,” and “interaction with physicians and therapists” were rated the most relevant components of a potential app.

Conclusions: A need for the support of patients with PAD was determined. This was particularly evident with regard to disease literacy and the performance of SET. Based on a detailed description of patient characteristics, proposals for the design of mobile interventions adapted to the needs and requirements of patients can be derived.

(JMIR Form Res 2020;4(8):e15669) doi:10.2196/15669

KEYWORDS

peripheral arterial disease; mHealth; digital intervention; supervised exercise therapy; smartphone usage

Introduction

Circulatory disorders of peripheral arteries due to atherosclerotic lesions are the third most frequent manifestation of atherosclerotic disease after its manifestation in coronary and cerebrovascular arteries [1]. Nevertheless, peripheral artery disease (PAD) causes the highest treatment costs for health care providers of all cardiovascular disorders [2]. The prevalence of PAD increases with age and affects a substantial proportion of the elderly population (>20% in those aged >80 years) [1,3]. Additionally, PAD is linked to higher morbidity and mortality and leads to a significantly reduced quality of life including daily life restrictions [4], ranging from mild impairment in walking distance to limb amputations [5,6].

One recommendation of the current guidelines is supervised exercise therapy (SET) or a supervised exercise program (SEP) [3,7]. For better readability, we refer to both SET and SEP as SET from here on. Regularly reaching the pain threshold leads to better leg perfusion and makes SET one of the most effective (both medically and economically) conservative therapies for extending pain-free walking distance [7-10]. The regular performance of SET has already been proven to be associated with decreased mortality and also results in an improvement in functional health and quality of life [3,10-12].

Recent studies have shown that patient empowerment helps to increase therapy adherence. This is mainly achieved through gaining greater control in health decisions [13-15]. The needs and preferences of the patient have to receive more attention and patients should be involved more closely in the process of care. A deeper exploration of clinical and demographic characteristics may influence the response to SET, help to overcome barriers and allow for the possibility of designing tailor-made solutions to implement SET in a patient's everyday life [10].

Mobile health (mHealth) technologies provide digital solutions to close gaps in care [16,17]. The use of mobile devices (eg, smartwatches and smartphones) permits the monitoring of health data that far exceeds the information gathered in a brief clinical encounter [18]. Based on persuasive design aspects, mobile devices also offer opportunities to support patients' health-related behavior [19]. The development of mHealth interventions for noncommunicable diseases has progressively received attention in recent years. However, patients with PAD and their specific requirements (due to a frequently impaired ability to walk) have been neglected thus far. Current gaps in research arise from the fact that no specific apps are designed for patients with PAD. Studies either focus on the general aspects of cardiovascular health including (remote) counselling [20], or use nonspecific mHealth technologies that aim to raise the level of general activity [21], instead of providing a PAD-specific approach in promoting SET. PAD-specific mHealth solutions are not currently available, but their development should be informed by first identifying the requirements of the PAD-population.

As a first step in a patient-centered approach to develop PAD-specific mobile interventions, we describe the needs and requirements from a patient perspective.

The aim of the study was to determine the needs and requirements of patients with PAD. This included their overall care situation and the potential use of mobile interventions.

In addition to the clinical examination, we answer the following research questions:

1. What is the current perception of medical care in patients with PAD? Can a need for medical support be determined among the study participants?
2. Do patients with PAD currently use smartphones and apps? What are the characteristics of smartphone users and nonusers?
3. What are the requirements for the design of mobile interventions to support patients with PAD in performing supervised exercise therapy?

Methods

Study Design and Patient Recruitment

In addition to the clinical examination, we conducted a questionnaire-based survey at the vascular outpatient clinic of the West-German Heart and Vascular Center Essen of the University Clinic Essen, Germany. This clinic treats more than 1500 patients with PAD annually. Patients were recruited between September and December 2018.

Inclusion and Exclusion Criteria

Consecutively, patients with diagnosed PAD were asked to participate in this study. The inclusion criteria were male or female patients aged 18 or older with PAD. PAD had to be diagnosed at least 3 months prior to the study.

Furthermore, patients were excluded if they were unable to complete the questionnaire themselves (eg, severe dementia or cognitive dysfunction). We also excluded individuals who did not have sufficient knowledge of the German language.

Sample Size and Basic Sociodemographic Characteristics

In total, we surveyed 304 patients with PAD. Two-thirds of the patients were men (n=203, 66.8%; Table 1). The participants were aged between 41 and 90 years (mean 67 years, SD 10.21). In total, 133 (46.5%) of the participants had an upper-medium educational attainment (12 to 13 years of education), and 62 (21.7%) had a lower educational attainment (≤12 years of education). Overall, 24 (7.9%) said they did not have a secondary school graduation certificate.

Table 1. Sociodemographic characteristics of patients with peripheral arterial disease (PAD) divided into all patients, smartphone users, and non–smartphone users.

Sociodemographic characteristics	All patients, n (%)	Smartphone users, n (%)	Non–smartphone users, n (%)	*P* value (χ^2 value)
Sex	n=304	n=157	n=147	
Male	203 (66.8)[a]	106 (52.2)[b]	97 (47.8)[b]	.73 (0.121)
Age (years)	n=301	n=155	n=146	
40-49	18 (6.0)	11 (61.1)	7 (38.9)	.74 (0.113)
50-59	58 (19.3)	34 (58.6)	24 (41.4)	.46 (0.556)
60-69	102 (33.9)	64 (62.7)	38 (37.3)	.09 (2.870)
70-79	85 (28.2)	35 (41.2)	50 (58.8)	.32 (1.002)
≥80	37 (12.3)	11 (29.7)	26 (70.3)	.12 (2.382)
Educational attainment (years)	n=286	n=151	n=135	
<10	40 (14.0)	19 (12.6)	21 (15.6)	.91 (0.012)
10-11	22 (7.7)	13 (8.6)	9 (6.7)	.76 (0.091)
12-13	133 (46.5)	67 (44.4)	66 (48.9)	.09 (0.001)
14-17	66 (23.1)	35 (23.2)	31 (23.0)	.32 (0.030)
>17	25 (8.7)	17 (11.3)	8 (5.9)	.12 (1.013)
Employment status	n=304	n=157	n=147	
Currently employed	77 (25.3)	43 (55.8)	34 (44.2)	.57 (0.319)
Retired	140 (46.1)	86 (42.9)	94 (57.1)	.75 (0.100)
Retired due to illness	40 (13.2)	26 (65.0)	14 (35.0)	.26 (0.258)
Burden of PAD	n=304	n=157	n=147	
Not at all	21 (6.9)	7 (33.3)	14 (66.7)	.43 (0.612)
A little	49 (16.1)	23 (46.9)	26 (53.1)	.92 (0.102)
Average	74 (24.3)	36 (48.6)	38 (51.4)	.97 (0.002)
Fair	98 (32.2)	57 (58.2)	41 (41.8)	.32 (1.007)
Great	62 (20.4)	34 (54.8)	28 (45.2)	.72 (0.129)
Burden of disease	1.36 (0.76)	1.45 (0.73)	1.25 (0.76)	.02 (2.33)
Burden of environmental conditions	0.90 (0.63)	1.02 (0.58)	0.77 (0.64)	<.001 (3.377)

[a]The percentage is based on the number of all responses for the associated sociodemographic characteristic (sex, age, educational attainment, employment status, and burden of PAD).

[b]The percentage is based on the total number of observations within the associated sociodemographic characteristic group.

Ethics

The study was conducted in accordance with the Declaration of Helsinki, and the protocol was approved by the local ethics committee at the Faculty of Medicine of the University Duisburg-Essen. Patient records were deidentified and analyzed anonymously. Written consent was obtained from each patient included in the study.

Measurements: Questionnaire

The questionnaire was prepared specifically for this study and was pretested on 5 PAD patients not included in the study sample. The pretest did not reveal the need for any changes.

The 10-page questionnaire encompasses a total of 31 questions on sociodemographic characteristics; subjective burden of disease and PAD-related care situations; subjective burden of environmental conditions; implementation and feasibility of SET; mobile or app usage; interests and knowledge regarding SET and medication; and the need for support and satisfaction with the health care situation. In the questionnaire, we used the term supervised walking training instead of the technical term supervised exercise therapy (SET) because the German word “Gehtraining” is more established in clinical practice.

The questionnaire included dichotomous and 5-point assessments similar to the Likert scale, adapted response scales, and open-ended questions. The questionnaire (English translation) is provided in Multimedia Appendix 1.

The thematic structure of the questionnaire included the following topics:

- Need for support
- Satisfaction with health care situation
- Sociodemographic characteristics
- Burden of environmental conditions
- Burden of PAD and other diseases
- Pain-free walking distance
- Clinical characteristics
- Preferences regarding offers to support patients with PAD
- Smartphone usage, knowledge about health apps, and health app usage
- Design categories in health apps to support patients with PAD

The detailed classification of the topics is shown in Multimedia Appendix 2. The items were chosen to measure the current level of burden in terms of PAD and other diseases. Relevant characteristics to classify PAD severity were also interrogated. In addition, it should be examined to what extent digital interventions represent a possible approach to support affected patients. Sociodemographic characteristics should provide information about the special requirements of different subgroups.

Analysis

We performed descriptive data analysis using SPSS (Version 23; IBM Corp). Variables are presented as frequencies and percentages or as means and standard deviations. Variables were compared using an unpaired *t* test or chi-square test, and a one-way analysis of variance when more than two groups were compared. Values of $P<.05$ were considered statistically significant. For comparative analyses of normally distributed variables, parametric tests such as the Student *t* test were used to test the assumption of homogeneity or nonhomogeneity of variance.

Results

Research Question 1: What Perceptions Do Patients With PAD Have of Their Medical Care? Do Study Participants Indicate a Need for Medical Support?

Overall, the need for more medical support in patients with PAD was identified (Table 2). Two-thirds (n=180, 59.2%) of the surveyed patients expressed their desire for more support in regard to their disease. More than half (n=162, 53.2%) of the patients were not very satisfied with their health care situation. The question regarding patient knowledge about current medical therapies and recommendations regarding SET indicated both a poor level of patient care and information. Two-thirds (n=198, 65.1%) of the participants stated that they did not know if they were taking medication to treat their PAD. The lack of suitable medication occurred in all Fontaine stages (I: n=101, 33.3%; IIa: n=34, 11.1%; IIb: n=135, 44.4%; and IV: n=34, 11.1%), although medication is recommended for all stages. The vast majority of patients (n=264, 86.8%) reported that their physician did not explain why their prescribed medication was important. Two-thirds of the patients (n=194, 63.8%) answered “no” to the question of whether they had already been recommended to perform walking training for the treatment of PAD. More than half of the patients (n=163, 53.6%) were even not familiar with the term “supervised exercise therapy.” Only 26% (n=79) said they already performed walking training on a regular basis.

Table 2. Current perceptions of central aspects of medical care and need for support in patients with peripheral arterial disease.

Items	Total responses, n (%) N=304	40-49 years, n (%) n=18	50-59 years, n (%) n=58	60-69 years, n (%) n=102	70-79 years, n (%) n=85	>80 years, n (%) n=38
Overall need for support						
Yes	180 (59.2)[a]	8 (44.5)[b]	31 (53.4)[b]	60 (58.8)[b]	49 (57.6)[b]	29 (76.3)
Health care satisfaction						
Completely dissatisfied	50 (16.4)	2 (11.1)	16 (27.6)	12 (11.8)	10 (11.8)	9 (23.7)
Rather dissatisfied	112 (36.8)	8 (44.4)	20 (34.5)	37 (36.3)	36 (42.4)	11 (28.9)
Neither dissatisfied nor satisfied	63 (20.7)	2 (11.1)	12 (20.7)	24 (23.5)	20 (23.5)	5 (13.2)
Rather satisfied	46 (15.1)	2 (11.1)	7 (12.1)	18 (17.6)	12 (14.1)	6 (15.8)
Very satisfied	33 (10.9)	4 (22.2)	3 (5.2)	11 (10.8)	7 (8.2)	7 (18.4)
Medication for peripheral arterial disease						
Yes	77 (25.3)	8 (44.4)	14 (24.1)	24 (23.5)	22 (25.9)	8 (21.1)
Information about medication						
Yes	40 (13.2)	3 (16.7)	8 (13.8)	9 (8.8)	14 (16.5)	6 (15.8)
Recommendation for supervised walking training[c]						
Yes	13 (4.3)	0 (0)	3 (5.2)	4 (3.9%	3 (3.5)	3 (7.9)
Information about supervised walking training[c]						
Yes	36 (11.8)	2 (11.1)	5 (8.6)	15 (14.7)	9 (10.6)	5 (13.2)
Performance of supervised walking training[c]						
Yes	79 (26.0)	3 (16.7)	16 (27.6)	32 (31.4)	21 (24.7)	7 (18.4)

[a]The percentage is based on the total number of responses for the associated item.

[b]The percentage is based on the total number of people in the age group.

[c]In the questionnaire, we used the term supervised walking training instead of the technical term supervised exercise therapy (SET) because the German word “Gehtraining” is more established in clinical practice.

Research Question 2: Do Patients with PAD Currently Use Smartphones and Apps? What are the Characteristics of Smartphone Users and Non–Smartphone Users?

Table 1 presents the following sociodemographic characteristics: (1) the participants who use a smartphone, (2) the participants who do not use a smartphone, and (3) a summary of the entire study sample. In total, 304 patients provided information on whether they use a smartphone.

Half of the patients (n=157, 51.6%) were smartphone users. Health apps were used by only a minority of patients (n=17, 5.7%). However, almost half (n=146, 48%) of all participants said they had already heard about health apps for smartphones that are designed to support health improvement.

The proportion of men and women who used a smartphone was comparable (n=159, 52.2% versus n=154, 50.5%, P=.73, χ^2=0.728). Patients aged between 40 and 69 years were more likely to use a smartphone than not (n=21, 61.2% were active users). This trend changed in patients aged 70 years or older. Two-thirds of the patients aged 70 years or older did not use a smartphone (n=116, 38.1% were active users, P=.07, χ^2=3.262).

Among those who had a low to upper-medium educational attainment (≤17 years of education), we did not see notable differences between users of smartphones and nonusers (P=.83, χ^2=0.048). However, in patients with a high educational attainment (>17 years; equivalent to a university degree), we found a tendency toward higher smartphone use, but this was not statistically significant (n=4, 11.3% versus n=18, 5.9%, P=.31, χ^2=1.013). Three-fourths of the participants (n=234, 76.9%) were not currently employed; of these, 94 (63.9%) were retired. The most frequent reason for retirement was having reached retirement age (n=165, 54.4%). Only 14 patients (9.5%) had retired due to illness.

Overall, patients tended to feel “quite burdened” (n=98, 32.2%) to “very burdened” (n=62, 20.4%) by their PAD. In addition to PAD, patients were mainly affected by diseases of the musculoskeletal system (mean 2.32, SD 1.63), diseases of the cardiovascular system (mean 2.14, SD 1.50) and respiratory diseases (mean 1.54, SD 1.43).

The burden of environmental conditions was indicated by a mean of 0.90 (SD 0.63), which corresponds to a low burden of environmental conditions. Patients were mainly affected by environmental conditions such as "financial worries" (mean 1.32, SD 1.28), followed by "constant responsibility for their family" (mean 1.23, SD 1.22) and "household" (mean 1.23, SD 5.85), but the standard deviation for the "household" item was conspicuously large. The group of patients without a smartphone felt somewhat less burdened, both in terms of burden of disease and burden of environmental conditions (P=.02, t=2.330 and P<.001, t=3.377, respectively).

In Table 3, we report the differences in health status and risk factors between participants with and without smartphone use. The pain-free walking distance is a relevant marker for compensated PAD. One-third (n=101, 33.2%) of participants said that they could walk <200 meters without pain. Additionally, 28% (n=85) of the participants reported they could walk between 200 and 1000 meters, and 29% (n=88) were hardly restricted and reported a pain-free walking distance of more than 1000 meters (n=30 or 9.9% chose the answer option "I do not know"). The two groups (smartphone users and nonusers) were comparable with regard to their pain-free walking distance and did not differ substantially.

Table 3. Health status and risk factors of patients with peripheral arterial disease divided into all patients, smartphone users, and non–smartphone users.

Health status or risk factor	All patients	Smartphone users	Non–smartphone users	P value (χ^2 value)
Pain-free walking distance, mean (SD)	n=304	n=157	n=147	
<200 m	101 (33.2)[a]	53 (52.5)[b]	48 (47.5)[b]	.83 (0.045)
200-1000 m	85 (28.0)	47 (55.3)	38 (44.7)	.59 (0.289)
>1000 m	88 (28.9)	43 (48.9)	45 (51.1)	.97 (0.001)
I do not know	30 (9.9)	14 (46.7)	16 (53.3)	.95 (0.004)
Disease severity according to Fontaine [22], mean (SD)	n=299	n=155	n=144	
Stage I	125 (41.8)	66 (52.8)	59 (47.2)	.75 (0.100)
Stage IIa	44 (14.7)	28 (63.6)	16 (36.4)	.28 (1.158)
Stage IIb	103 (34.4)	52 (50.5)	51 (49.5)	.97 (0.001)
Stage III	7 (2.3)	5 (71.4)	2 (28.6)	.78 (0.075)
Stage IV	20 (6.7)	4 (20.0)	16 (80.0)	.10 (2.747)
BMI, mean (SD)	n=176	n=90	n=86	
Underweight (<18.5 kg/m^2)	7 (4.0)	2 (2.2)	5 (5.8)	.78 (0.075)
Normal (18.5-24.9 kg/m^2)	55 (31.2)	27 (30.0)	28 (32.6)	.96 (0.002)
Overweight (25.0-29.9 kg/m^2)	69 (39.2)	36 (52.2)	33 (38.4)	.93 (0.007)
Obese (>30 kg/m^2)	45 (25.6)	25 (27.8)	20 (23.3)	.75 (0.100)
Currently smoking, mean (SD)	n=275	n=142	n=133	
Yes	86 (31.3)	47 (33.1)	39 (29.3)	.65 (0.210)
Not anymore	39 (14.2)	19 (13.4)	20 (15.0)	.96 (0.003)

[a]Percentage is based on the number of responses for the associated health status or risk factor (pain-free walking distance, disease severity, BMI, smoking).

[b]Percentage is based on the total number of observations within the associated group of health outcomes or risk factors, regardless of smartphone use.

Based on the severity of the disease, 42% (n=128) were in Fontaine Stage I (corresponding to mild PAD), 15% (n=46) were in Stage IIa, 34% (n=103) were in Stage IIb, 2% (n=6) were in Stage III, and 7% (n=21) were in Stage IV (corresponding to very severe PAD). On average, patients in the smartphone group (mean 2.05, SD 1.06) and patients in the non–smartphone group (mean 2.30, SD 1.31) had mild PAD. However, in the non–smartphone group, there were more cases of severe PAD (Stage IV, 20% [n=61] versus 80% [n=243], P=.10).

In total, more than one-third (n=119, 39.2%) of the participants were overweight, and an additional 26% (n=79) were obese. Normal weight was documented in 31% (n=94) of the participants, and 4% (n=12) of the participants were underweight.

Almost half of the participants (n=140, 46%) had smoked at one stage of their life, and of these participants, 31% (n=94) were current smokers and 14% (n=43) had already quit smoking. With regard to smoking behavior, the smartphone users and

nonusers did not show substantial differences (P=.65 and P=.96, respectively).

Research Question 3: What are the Requirements for the Design of Mobile Interventions to Support Patients with PAD?

When asked how likely it was that they would use the listed services, participants indicated that they were most likely to use a "training app" on their smartphone (mean 3.18, SD 1.28), followed by "informational material" (mean 2.83, SD 1.48) and "training groups with instructions" (mean 2.53, SD 1.45). "Online platforms" (mean 1.73, SD 1.10) and "support groups" (mean 1.87, SD 1.87) were the response options that participants indicated they were least likely to use. The probability of making use of the listed options for patients with PAD is summarized in Figure 1.

Figure 2 shows the ranking of the components of health apps that can be used to support patients. The most relevant components were "information" (mean 3.29, SD 1.33), "feedback" (mean 3.27, SD 1.38), "choosing goals" (mean 3.06, SD 1.32) and "interaction with physicians and therapists" (mean 3.05, SD 1.40). The least relevant component was "interaction with other patients" (mean 2.44, SD 1.30).

Figure 1. Descriptive analysis of user preferences in terms of offered app components for patients with peripheral arterial disease (refers to Question 15 of the questionnaire). Note that multiple choices were possible.

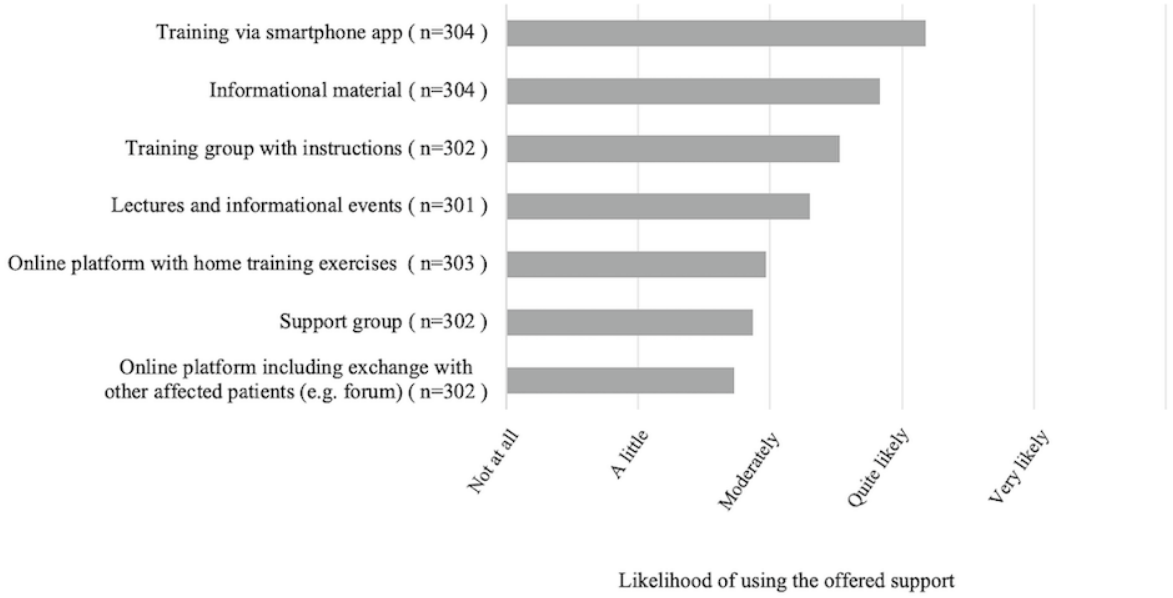


Figure 2. Descriptive analysis of user-reported relevance in terms of health app components that would assist patients with peripheral arterial disease performing supervised walking training. The analysis refers to question 15 of the questionnaire. Note that multiple choices were possible. In the questionnaire, we used the term "supervised walking training" instead of the technical term "supervised exercise therapy" (SET) because the German word "Gehtraining" is more established in clinical practice.

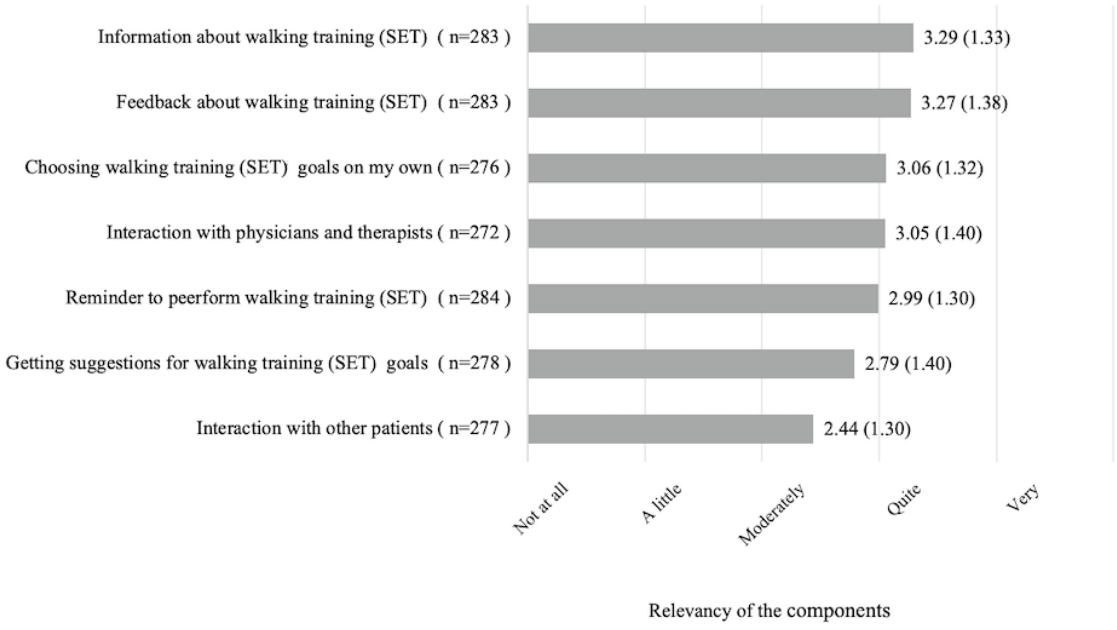


Table 4 shows the ranking of components of health apps that can be used to support patients depending on disease severity according to Fontaine stage [22]. The importance of the components were most likely to be seen in participants with Stage IIa across all components of health apps that support patients with PAD, although the averages did not surpass "moderately" to "fairly important." All components were rated highest and at least moderately important by patients with Stage IIa. Although patients with Stage I were, by definition, not restricted in everyday life, support components such as "information" (mean 3.32, SD 1.30), "feedback" (mean 3.36, SD 1.36), "choosing goals" (mean 3.14, SD 1.36), and "interaction with physicians and therapists" (mean 3.10, SD 1.42) were seen to be moderately to fairly important for patients in Stage I with regard to performing SET using a health app. The performed variance analysis showed a significant difference between Fontaine stages for "interaction with physicians and therapists" (P=.04, F=4.231) and "getting suggestions for walking training (SET) goals" (P=.03, F=5.026). There was a trend toward differences between the Fontaine stages for "reminder to perform walking training (SET)" and "choosing walking training (SET) goals on my own" but these differences were not statistically significant (P=.06, F=3.668 and P=.08, F=3.034, respectively).

Age was only found to have an effect on the answer to "reminder to perform walking training (SET)." The older the participants were, the more they preferred a reminder function of a health app (P=.02, F=5.933). Other effects of age were not found in terms of supporting app components.

The ranking of the individual components differed slightly depending on the severity of the disease. For participants in Stage IIa, the ranking was as follows: (1) interaction with physicians and therapists, (2) information about SET, (3) feedback about SET, (4) choosing goals, (5) suggestions for goal setting, (6) reminders, and (7) interaction with other patients. Interaction with physicians and therapists was less important for patients in Stages I, IIb, and IV.

Information regarding SET and feedback about SET were moderately to fairly important for patients regardless of disease stage. Interactions with other patients were considered least important by participants in Stages I to III. For Stage IV participants, interactions with other participants were considered more important than choosing goals, reminders, and suggestions for goal setting. The core results of the study are summarized in Table 5.

Table 4. Relevance of components of health apps to support patients with peripheral arterial disease by disease severity according to Fontaine stages [22].

Components of health apps to support patients	Stage I		Stage IIa		Stage IIb		Stage III		Stage IV	
	Patients, n	Mean (SD)	Patients, n	Mean (SD)	Patients, n	Mean (SD)	Patients, n	Mean (SD)	Patients, n	Mean (SD)
Information about walking training[a]	113	3.32 (1.30)	42	3.66 (1.26)	97	3.28 (1.37)	6	2.66 (1.21)	20	2.80 (1.15)
Feedback about walking training	111	3.36 (1.36)	42	3.59 (1.30)	99	3.12 (1.40)	7	2.57 (1.27)	19	3.36 (1.38)
Choosing walking training goals on my own	110	3.14 (1.36)	40	3.35 (1.16)	95	3.03 (1.30)	6	2.16 (1.47)	20	2.70 (1.26)
Interaction with physicians and therapists	107	3.10 (1.42)	39	3.71 (1.21)	95	2.91 (1.38)	7	2.42 (1.27)	20	2.75 (1.25)
Reminder to perform walking training	112	3.08 (1.32)	41	3.17 (1.30)	100	2.99 (1.26)	7	2.57 (1.27)	19	2.42 (1.12)
Getting suggestions for walking training goals	110	2.90 (1.41)	40	3.32 (1.36)	98	2.69 (1.40)	7	2.00 (1.15)	19	2.31 (1.00)
Interaction with other patients	111	2.39 (1.28)	39	3.00 (1.31)	97	2.30 (1.27)	7	1.71 (1.11)	19	2.73 (1.19)

[a]In the questionnaire, we used the term "supervised walking training" instead of the technical term "supervised exercise therapy" because the German word "Gehtraining" is more established in clinical practice.

Table 5. Summary of core results.

Research question	Summary of core results
1	A need for support was determined. Receiving more educational health information, increased support in the form of prescribed medication, and help in terms of implementing supervised exercise therapy (SET) are the most desired actions for improving the care of patients with PAD.
2	• Half of the participants use smartphones. For them, mobile interventions to support SET and medication can be a relevant treatment component. • Patients aged >70 years are less likely to use smartphones than younger patients. With regard to characteristics such as sex, education, profession, BMI, smoking behavior, exposure to illness or the environment, or the current state of illness, the data did not reveal any significant differences between smartphone users and nonusers within the patient population.
3	• Interest in smartphone-supported training is present, even for people who do not currently use a smartphone. • Health app components such as "information," "monitoring," and "feedback" were the most relevant for patients with PAD. Other components such as "choosing goals," "interaction with physicians and therapists," "interaction with other patients," and "reminders and suggestions for goal setting" were less relevant for the patients and should be selectable on demand according to patient preference.

Discussion

There Is a Need for Supporting the Care of Patients With PAD

More than half (53.2%) of the participants were less than satisfied or completely unsatisfied with their health care situation. Patients do not feel well-informed enough in terms of SET and their prescribed medication. Since both are cornerstones in the treatment of PAD, this finding is alarming in terms of secondary prevention and long-term outcomes. The lack of educational background is expected to be associated with poor medication and exercise compliance, impeding the successful empowerment of patients. Previous research found that mHealth interventions improve adherence to prescribed medication in patients with cardiovascular disease [23].

Our results show an evident need for action to support patients with PAD in secondary prevention. A major goal should include patient empowerment. The demand for more support was found in all subgroups, independent of age or severity of disease. Institutional barriers in particular (eg, a lack of training groups and primary health care providers providing care to patients with PAD) limit the likelihood of an adequate health care offer for affected patients. Previous studies already reported the undersupply of primary health care for patients with PAD in general as well as those from various sociodemographic backgrounds [24,25].

A Call for Patient-Centered Mobile Interventions

Personal barriers are primarily linked to poor knowledge about the disease and low empowerment. Mobile interventions might play an ever-increasing role, since they are widely accessible and have a low threshold for access. Time resources for consultations between patients and doctors are limited. In clinical practice, lifestyle recommendations are made within a few minutes. To increase the probability of patients' adherence and their empowerment to take responsibility for their own health, personalized approaches are promising [26]. However, these require the involvement of the patient. The analysis of patient characteristics, smartphone use, and requirements for support measures is a first step to identify patient needs. On the basis of surveys focusing on patients' specific needs and requirements, patient-centered interventions can be developed; at the same time, deficiencies in the current health care situation can be identified and potentially improved.

The use of patient-centered methods to develop persuasive strategies for mHealth interventions [27], as well as the conception and implementation of analog interventions, can help to take different social groups into account, with respect to their specific and individual needs.

Requirements for the Design of Mobile Interventions

The idea of using a training app was of strong interest, even to participants who currently do not use smartphones. Based on this preference for digital support, the need to design and implement motivating tools that provide educational information was identified. In this setting, the analysis of assessed data regarding usage and user preferences might also be helpful in the feedback process. The current study also demonstrated patients' priorities regarding important features, such as the opportunity to set individual goals or to get in touch with professionals, including physicians or therapists. Conspicuously, the offered support of interactions with other patients tended to perform poorly, both as a proposed digital chat component in an app and as a component of an analog intervention in the sense of a support group. A previous study showed a high acceptance of electronic health information and disease-related community forums in patients with PAD [28]. This discrepancy might result from the sole query regarding the preferences of interaction between patients without further explanation. On the other hand, one might deduce that tools do not primarily have to offer (specific) messenger services to enable and support networking between the patients. However, the opportunity for patients to compare and compete with each other (within an app) was not reviewed. Further research is needed to evaluate this kind of digital support.

Modular Concepts Adapted to the Needs of Patients Appear Promising

Depending on the severity of their disease, the participants' ranking of useful components within a digital intervention app differed slightly. Although "information," "monitoring," and

"feedback" should be fixed components within apps that support patients with PAD, other components, such as "goal selection," "interaction with physicians or therapists," "interaction with other patients," "reminders for structured walking training," and "suggestions for individual goals" can be offered additionally, as a voluntary, selectable feature according to patients' preferences.

In addition to tools for the implementation of SET, supplementary components that support medication use, healthier nutrition, or cessation of smoking appear useful. Considering the BMI of our study population compared to the German population in 2017 (overweight: 35.9%; obese: 18.1%), the sample was above the national average [29]. Therefore, a combination of components that promote more exercise and a healthier lifestyle including nutrition may contribute to weight management. Weight reduction also reduces the risk of multimorbidity, which was a common burden in our cohort.

Challenges to Face

The efficacy of digital interventions is significantly influenced by an individual's engagement with, for example, a specific app [30]. In particular, mHealth technologies face one major limiting factor in terms of long-term engagement: the vast majority do not exceed 6 months of regular app use [31,32]. This phenomenon does not occur only in healthy subjects who aim for a healthier lifestyle, but also in secondary prevention, where long-term behavioral changes toward a more active lifestyle are associated with health benefits [33,34]. Strategies that improve user engagement linked to these technologies may include elements of gamification [35] and devices deeply intertwined with everyday life, such as smartphones or wearables [36], that deliver instant feedback of good behavior. Our results showed a Fontaine stage–dependent decline in interest in interactions with physicians and therapists, and suggested walking training (SET) goals. This effect might be linked to a higher frustration level in patients with an advanced disease stage [37] and emphasizes the importance of early diagnosis and treatment.

Another consideration is the age of the target group, which often includes older patients with noncommunicable diseases. The mean age of participants in this study was 67 years. More than half of the patients aged 70 years or older were not reachable by mobile interventions. This finding has to be taken into consideration when designing digital interventions. Similar results regarding the use of digital interventions in older patients were previously observed [38]. Nevertheless, older patients are not to be neglected in terms of the development of digital interventions as the aging population is expected to become increasingly accustomed to the use of smart technologies [39]. Further physical impairments like diminished eyesight related to diabetes or deteriorated motor skills have to be taken into account, since they occur more often with increasing age. Motivational and cognitive barriers of older adults are major challenges [40,41] and need to be investigated more in detail. In this study, we found that a reminder to perform walking training (SET) would be of special interest in the older population, but other differences between the younger and older patients were not seen. Since this study did not focus on age-related preferences of app development in detail, further research is needed to investigate special needs and requirements in older adults.

Limitations and Future Work

This analysis merely serves as an empirically sound description of the addressed problem and identifies approaches to improve the care of patients with PAD. This study included only a small sample; thus, the results cannot be generalized.

The study focused on a selection of personal characteristics to avoid time-consuming interviews before starting the actual clinical examination. Other characteristics of patients with PAD that may affect the need for and responsiveness to interventions supporting SET in daily living (eg, self-efficacy, motivation to change, race/ethnicity, income, social capital) were not examined. Based on the present study findings, we developed an app to support SET for patients with PAD [42], but the results are pending.

This study also did not address environmental factors. This is a potential point of criticism. With regard to health, in addition to personal characteristics, environmental characteristics play an important role in the implementation of health-promoting and therapy-compliant behaviors [43]. To obtain health-related environmental data (eg, neighborhood-related resources and walkability), the recorded postcode can be used in further studies.

Additionally, we offered only an abridged list of design components for an app, rather than all components that are conceivable in principle. The additional demonstration of mock-ups and prototypes to determine the preferences and desires of the participants might be useful in future surveys. Although user-centered methods for app design that combine different methods (eg, design thinking research) are time-consuming, they may improve the effectiveness of behavior-change support systems [44].

The description of the sociodemographic characteristics of our participants, grouped into smartphone users and non–smartphone users, showed that participants younger than the age of 70 years used smartphones much more often than older participants. Hence, the latter group of patients is hard to reach with mobile interventions. To improve the success of therapy for non–smartphone users, analog interventions (supporting medication use and the implementation of SET for older patients) should also be offered. Except for age, we found no noticeable differences between smartphone and non–smartphone users. The analysis of patient characteristics (ie, sex, education, burden, and health status) with respect to smartphone use, did not reveal any other significant differences between smartphone users and nonusers.

Conclusion

This survey of patients with PAD indicates the necessity of improving the care situation of these patients. A need for support can be determined and identified with regard to educational and general support deficiencies. This need includes a better understanding of the prescribed medication and the necessary

implementation of SET as a central pillar of the guideline-oriented care of patients with PAD.

There also exists a great interest in mobile support services. To improve the care situation of these patients, mobile interventions are promising. The large reach and wide availability of these interventions are major advantages.

Authors' Contributions

JL contributed to study design, data collection, data analysis, data interpretation, manuscript writing, and final approval. JS contributed to the literature search, figures, data collection, manuscript writing, and final approval. TK contributed to the literature search, figures, tables, and final approval. IKA contributed to study design, manuscript writing, and final approval. GU contributed to the literature search, figures, data collection, and final approval. MS contributed to the data collection, data analysis, manuscript writing, and final approval. CR contributed to the data collection, data analysis, manuscript writing, and final approval. RAJ contributed to the literature search, data analysis, manuscript writing, and final approval. SM contributed administrative support, data analysis, data interpretation, manuscript writing, and final approval. TR contributed administrative support, data interpretation, manuscript writing, and final approval. KP contributed to study design, data collection, data analysis, data interpretation, manuscript writing, and final approval.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Questionnaire.
[PDF File (Adobe PDF File), 973 KB - formative_v4i8e15669_app1.pdf]

Multimedia Appendix 2
Classification of topics.
[DOCX File , 101 KB - formative_v4i8e15669_app2.docx]

References

1. Fowkes FGR, Rudan D, Rudan I, Aboyans V, Denenberg JO, McDermott MM, et al. Comparison of global estimates of prevalence and risk factors for peripheral artery disease in 2000 and 2010: a systematic review and analysis. Lancet 2013 Oct 19;382(9901):1329-1340. [doi: 10.1016/S0140-6736(13)61249-0] [Medline: 23915883]
2. Smolderen KG, Wang K, de Pouvourville G, Brüggenjürgen B, Röther J, Zeymer U, REACH Registry Investigators. Two-year vascular hospitalisation rates and associated costs in patients at risk of atherothrombosis in France and Germany: highest burden for peripheral arterial disease. Eur J Vasc Endovasc Surg 2012 Feb;43(2):198-207 [FREE Full text] [doi: 10.1016/j.ejvs.2011.09.016] [Medline: 22001145]
3. Harwood AE, Totty JP, Pymer S, Huang C, Hitchman L, Carradice D, et al. Cardiovascular and musculoskeletal response to supervised exercise in patients with intermittent claudication. J Vasc Surg 2019 Jun;69(6):1899-1908.e1. [doi: 10.1016/j.jvs.2018.10.065] [Medline: 30583899]
4. Pande RL, Perlstein TS, Beckman JA, Creager MA. Secondary prevention and mortality in peripheral artery disease: National Health and Nutrition Examination Study, 1999 to 2004. Circulation 2011 Jul 05;124(1):17-23 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.110.003954] [Medline: 21690489]
5. Letterstål A, Forsberg C, Olofsson P, Wahlberg E. Risk attitudes to treatment among patients with severe intermittent claudication. J Vasc Surg 2008 May;47(5):988-994 [FREE Full text] [doi: 10.1016/j.jvs.2007.12.055] [Medline: 18455642]
6. Maksimovic M, Vlajinac H, Marinkovic J, Kocev N, Voskresenski T, Radak D. Health-related quality of life among patients with peripheral arterial disease. Angiology 2014 Jul;65(6):501-506. [doi: 10.1177/0003319713488640] [Medline: 23657177]
7. Bouwens E, Klaphake S, Weststrate KJ, Teijink JA, Verhagen HJ, Hoeks SE, et al. Supervised exercise therapy and revascularization: Single-center experience of intermittent claudication management. Vasc Med 2019 Jun;24(3):208-215 [FREE Full text] [doi: 10.1177/1358863X18821175] [Medline: 30795714]
8. Lane R, Harwood A, Watson L, Leng GC. Exercise for intermittent claudication. Cochrane Database Syst Rev 2017 Dec 26;12:CD000990. [doi: 10.1002/14651858.CD000990.pub4] [Medline: 29278423]
9. Lee HLD, Mehta T, Ray B, Heng MST, McCollum PT, Chetter IC. A non-randomised controlled trial of the clinical and cost effectiveness of a Supervised Exercise Programme for claudication. Eur J Vasc Endovasc Surg 2007 Feb;33(2):202-207 [FREE Full text] [doi: 10.1016/j.ejvs.2006.08.005] [Medline: 17142065]
10. Treat-Jacobson D, McDermott MM, Bronas UG, Campia U, Collins TC, Criqui MH, American Heart Association Council on Peripheral Vascular Disease; Council on Quality of CareOutcomes Research;Council on CardiovascularStroke Nursing. Optimal Exercise Programs for Patients With Peripheral Artery Disease: A Scientific Statement From the American Heart Association. Circulation 2019 Jan 22;139(4):e10-e33. [doi: 10.1161/CIR.0000000000000623] [Medline: 30586765]

11. Fakhry F, Rouwet EV, den Hoed PT, Hunink MGM, Spronk S. Long-term clinical effectiveness of supervised exercise therapy versus endovascular revascularization for intermittent claudication from a randomized clinical trial. Br J Surg 2013 Aug;100(9):1164-1171. [doi: 10.1002/bjs.9207] [Medline: 23842830]
12. McDermott MM. Exercise training for intermittent claudication. J Vasc Surg 2017 Dec;66(5):1612-1620 [FREE Full text] [doi: 10.1016/j.jvs.2017.05.111] [Medline: 28874320]
13. Edwards M, Davies M, Edwards A. What are the external influences on information exchange and shared decision-making in healthcare consultations: a meta-synthesis of the literature. Patient Educ Couns 2009 Apr;75(1):37-52. [doi: 10.1016/j.pec.2008.09.025] [Medline: 19036550]
14. Kambhampati S, Ashvetiya T, Stone NJ, Blumenthal RS, Martin SS. Shared Decision-Making and Patient Empowerment in Preventive Cardiology. Curr Cardiol Rep 2016 May;18(5):49. [doi: 10.1007/s11886-016-0729-6] [Medline: 27098670]
15. Wildenbos GA, Horenberg F, Jaspers M, Peute L, Sent D. How do patients value and prioritize patient portal functionalities and usage factors? A conjoint analysis study with chronically ill patients. BMC Med Inform Decis Mak 2018 Nov 21;18(1):108 [FREE Full text] [doi: 10.1186/s12911-018-0708-5] [Medline: 30463613]
16. Argent R, Daly A, Caulfield B. Patient Involvement With Home-Based Exercise Programs: Can Connected Health Interventions Influence Adherence? JMIR mHealth uHealth 2018 Mar 01;6(3):e47 [FREE Full text] [doi: 10.2196/mhealth.8518] [Medline: 29496655]
17. Kostkova P. Grand challenges in digital health. Front Public Health 2015;3:134 [FREE Full text] [doi: 10.3389/fpubh.2015.00134] [Medline: 26000272]
18. Burke LE, Ma J, Azar KMJ, Bennett GG, Peterson ED, Zheng Y, American Heart Association Publications Committee of the Council on EpidemiologyPrevention, Behavior Change Committee of the Council on Cardiometabolic Health, Council on CardiovascularStroke Nursing, Council on Functional GenomicsTranslational Biology, Council on Quality of CareOutcomes Research,Stroke Council. Current Science on Consumer Use of Mobile Health for Cardiovascular Disease Prevention: A Scientific Statement From the American Heart Association. Circulation 2015 Sep 22;132(12):1157-1213 [FREE Full text] [doi: 10.1161/CIR.0000000000000232] [Medline: 26271892]
19. Bendig E, Bauereiß N, Ebert DD, Snoek F, Andersson G, Baumeister H. Internet- Based Interventions in Chronic Somatic Disease. Dtsch Arztebl Int 2018 Nov 05;115(40):659-665 [FREE Full text] [doi: 10.3238/arztebl.2018.0659] [Medline: 30381130]
20. McDermott MM, Spring B, Berger JS, Treat-Jacobson D, Conte MS, Creager MA, et al. Effect of a Home-Based Exercise Intervention of Wearable Technology and Telephone Coaching on Walking Performance in Peripheral Artery Disease: The HONOR Randomized Clinical Trial. JAMA 2018 Dec 24;319(16):1665-1676 [FREE Full text] [doi: 10.1001/jama.2018.3275] [Medline: 29710165]
21. Shalan A, Abdulrahman A, Habli I, Tew G, Thompson A. YORwalK: Desiging a Smartphone Exercise Application for People with Intermittent Claudication. Stud Health Technol Inform 2018;247:311-315. [Medline: 29677973]
22. Aboyans V, Ricco J, Bartelink MEL, Björck M, Brodmann M, Cohnert T, ESC Scientific Document Group. 2017 ESC Guidelines on the Diagnosis and Treatment of Peripheral Arterial Diseases, in collaboration with the European Society for Vascular Surgery (ESVS): Document covering atherosclerotic disease of extracranial carotid and vertebral, mesenteric, renal, upper and lower extremity arteriesEndorsed by: the European Stroke Organization (ESO)The Task Force for the Diagnosis and Treatment of Peripheral Arterial Diseases of the European Society of Cardiology (ESC) and of the European Society for Vascular Surgery (ESVS). Eur Heart J 2018 Mar 01;39(9):763-816. [doi: 10.1093/eurheartj/ehx095] [Medline: 28886620]
23. Gandapur Y, Kianoush S, Kelli HM, Misra S, Urrea B, Blaha MJ, et al. The role of mHealth for improving medication adherence in patients with cardiovascular disease: a systematic review. Eur Heart J Qual Care Clin Outcomes 2016 Oct 01;2(4):237-244 [FREE Full text] [doi: 10.1093/ehjqcco/qcw018] [Medline: 29474713]
24. Lecouturier J, Scott J, Rousseau N, Stansby G, Sims A, Allen J. Peripheral arterial disease diagnosis and management in primary care: a qualitative study. BJGP Open 2019 Aug 20;29(3):1. [doi: 10.3399/bjgpopen19x101659]
25. Jelani Q, Jhamnani S, Spatz ES, Spertus J, Smolderen KG, Wang J, et al. Financial barriers in accessing medical care for peripheral artery disease are associated with delay of presentation and adverse health status outcomes in the United States. Vasc Med 2019 Oct 11;25(1):13-24. [doi: 10.1177/1358863x19872542] [Medline: 31603393]
26. Bradbury K, Morton K, Band R, van Woezik A, Grist R, McManus RJ, et al. Using the Person-Based Approach to optimise a digital intervention for the management of hypertension. PLoS One 2018;13(5):e0196868 [FREE Full text] [doi: 10.1371/journal.pone.0196868] [Medline: 29723262]
27. Asbjørnsen RA, Smedsrød ML, Solberg Nes L, Wentzel J, Varsi C, Hjelmesæth J, et al. Persuasive System Design Principles and Behavior Change Techniques to Stimulate Motivation and Adherence in Electronic Health Interventions to Support Weight Loss Maintenance: Scoping Review. J Med Internet Res 2019 Jun 21;21(6):e14265 [FREE Full text] [doi: 10.2196/14265] [Medline: 31228174]
28. Castaneda P, Sales A, Osborne NH, Corriere MA. Scope, Themes, and Medical Accuracy of eHealth Peripheral Artery Disease Community Forums. Ann Vasc Surg 2019 Jan;54:92-102. [doi: 10.1016/j.avsg.2018.09.004] [Medline: 30267913]

29. Yates N, Teuner CM, Hunger M, Holle R, Stark R, Laxy M, et al. The Economic Burden of Obesity in Germany: Results from the Population-Based KORA Studies. Obes Facts 2016;9(6):397-409 [FREE Full text] [doi: 10.1159/000452248] [Medline: 27951530]
30. Elliott M, Eck F, Khmelev E, Derlyatka A, Fomenko O. Physical Activity Behavior Change Driven by Engagement With an Incentive-Based App: Evaluating the Impact of Sweatcoin. JMIR mHealth uHealth 2019 Jul 08;7(7):e12445. [doi: 10.2196/12445] [Medline: 31287064]
31. Grady A, Yoong S, Sutherland R, Lee H, Nathan N, Wolfenden L. Improving the public health impact of eHealth and mHealth interventions. Australian and New Zealand Journal of Public Health 2018 Jan 31;42(2):118-119. [doi: 10.1111/1753-6405.12771]
32. Jee H. Review of researches on smartphone applications for physical activity promotion in healthy adults. J Exerc Rehabil 2017 Feb 27;13(1):3-11. [doi: 10.12965/jer.1732928.464]
33. Lear SA, Hu W, Rangarajan S, Gasevic D, Leong D, Iqbal R, et al. The effect of physical activity on mortality and cardiovascular disease in 130 000 people from 17 high-income, middle-income, and low-income countries: the PURE study. The Lancet 2017 Dec;390(10113):2643-2654. [doi: 10.1016/S0140-6736(17)31634-3]
34. Kononova A, Li L, Kamp K, Bowen M, Rikard RV, Cotten S, et al. The Use of Wearable Activity Trackers Among Older Adults: Focus Group Study of Tracker Perceptions, Motivators, and Barriers in the Maintenance Stage of Behavior Change. JMIR mHealth uHealth 2019 Apr 05;7(4):e9832 [FREE Full text] [doi: 10.2196/mhealth.9832] [Medline: 30950807]
35. Kamel BMN, Gammon S, Dixon MC, MacRury SM, Fergusson MJ, Miranda RF, et al. Digital games for type 1 and type 2 diabetes: underpinning theory with three illustrative examples. JMIR Serious Games 2015;3(1):e3 [FREE Full text] [doi: 10.2196/games.3930] [Medline: 25791276]
36. Bennett J, Rokas O, Chen L. Healthcare in the Smart Home: A Study of Past, Present and Future. Sustainability 2017 May 17;9(5):840. [doi: 10.3390/su9050840]
37. Davie-Smith F, Coulter E, Kennon B, Wyke S, Paul L. Factors influencing quality of life following lower limb amputation for peripheral arterial occlusive disease: A systematic review of the literature. Prosthet Orthot Int 2017 Feb 02;41(6):537-547. [doi: 10.1177/0309364617690394]
38. Sanders AB, Conroy DE, Schmitz KH, Gusani N. Physical Activity and Sedentary Behavior in Older Gastrointestinal Cancer Survivors: Need and Acceptability of Digital Health Interventions. J Gastrointest Cancer 2018 Jun 18. [doi: 10.1007/s12029-018-0128-x] [Medline: 29911290]
39. Manini TM, Mendoza T, Battula M, Davoudi A, Kheirkhahan M, Young ME, et al. Perception of Older Adults Toward Smartwatch Technology for Assessing Pain and Related Patient-Reported Outcomes: Pilot Study. JMIR Mhealth Uhealth 2019 Mar 26;7(3):e10044 [FREE Full text] [doi: 10.2196/10044] [Medline: 30912756]
40. Theis S, Schäfer D, Bröhl C, Schäfer K, Rasche P, Wille M, et al. Predicting technology usage by health information need of older adults: Implications for eHealth technology. WOR 2019 Mar 19;62(3):443-457. [doi: 10.3233/WOR-192878]
41. Wildenbos GA, Jaspers MWM, Schijven MP, Dusseljee-Peute LW. Mobile health for older adult patients: Using an aging barriers framework to classify usability problems. Int J Med Inform 2019 Dec;124:68-77 [FREE Full text] [doi: 10.1016/j.ijmedinf.2019.01.006] [Medline: 30784429]
42. Paldán K, Simanovski J, Ullrich G, Steinmetz M, Rammos C, Jánosi RA, et al. Feasibility and Clinical Relevance of a Mobile Intervention Using TrackPAD to Support Supervised Exercise Therapy in Patients With Peripheral Arterial Disease: Study Protocol for a Randomized Controlled Pilot Trial. JMIR Res Protoc 2019 Jun 26;8(6):e13651 [FREE Full text] [doi: 10.2196/13651] [Medline: 31244477]
43. Graham H, White PCL. Social determinants and lifestyles: integrating environmental and public health perspectives. Public Health 2016 Dec;141:270-278 [FREE Full text] [doi: 10.1016/j.puhe.2016.09.019] [Medline: 27814893]
44. Herrmanny K, Schwarz M, Paldán K, Beckmann N, Sell J, Wagner NF, et al. Designing a Holistic Behavior Change Support System for Healthy Aging. i-com 2017;16(2):99-111. [doi: 10.1515/icom-2017-0008]

Abbreviations

mHealth: mobile Health
PAD: peripheral arterial disease
SEP: supervised exercise program
SET: supervised exercise therapy

Edited by G Eysenbach; submitted 27.07.19; peer-reviewed by A Hoffmann, C Faust-Christmann; comments to author 28.09.19; revised version received 11.11.19; accepted 29.03.20; published 04.08.20.

Please cite as:
Lortz J, Simanovski J, Kuether T, Kreitschmann-Andermahr I, Ullrich G, Steinmetz M, Rammos C, Jánosi RA, Moebus S, Rassaf T, Paldán K
Needs and Requirements in the Designing of Mobile Interventions for Patients With Peripheral Arterial Disease: Questionnaire Study
JMIR Form Res 2020;4(8):e15669
URL: https://formative.jmir.org/2020/8/e15669
doi:10.2196/15669
PMID:32663154

Original Paper

Clinician Perspectives and Design Implications in Using Patient-Generated Health Data to Improve Mental Health Practices: Mixed Methods Study

Danny T Y Wu[1,2], PhD, MSI; Chen Xin[1,3], MDes; Shwetha Bindhu[1,4]; Catherine Xu[1,4]; Jyoti Sachdeva[5], MD; Jennifer L Brown[5], PhD; Heekyoung Jung[3], PhD

[1]Department of Biomedical Informatics, College of Medicine, University of Cincinnati, Cincinnati, OH, United States

[2]Department of Pediatrics, College of Medicine, University of Cincinnati, Cincinnati, OH, United States

[3]School of Design, College of Design, Architecture, Art, and Planning, University of Cincinnati, Cincinnati, OH, United States

[4]Medical Sciences Baccalaureate Program, College of Medicine, University of Cincinnati, Cincinnati, OH, United States

[5]Department of Psychiatry and Behavioral Neuroscience, College of Medicine, University of Cincinnati, Cincinnati, OH, United States

Corresponding Author:
Danny T Y Wu, PhD, MSI
Department of Biomedical Informatics
College of Medicine
University of Cincinnati
231 Albert Sabin Way, ML0840
Cincinnati, OH, 45229
United States
Phone: 1 5135586464
Email: wutz@ucmail.uc.edu

Abstract

Background: Patient-generated health data (PGHD) have been largely collected through mobile health (mHealth) apps and wearable devices. PGHD can be especially helpful in mental health, as patients' illness history and symptom narratives are vital to developing diagnoses and treatment plans. However, the extent to which clinicians use mental health–related PGHD is unknown.

Objective: A mixed methods study was conducted to understand clinicians' perspectives on PGHD and current mental health apps. This approach uses information gathered from semistructured interviews, workflow analysis, and user-written mental health app reviews to answer the following research questions: (1) What is the current workflow of mental health practice and how are PGHD integrated into this workflow, (2) what are clinicians' perspectives on PGHD and how do they choose mobile apps for their patients, (3) and what are the features of current mobile apps in terms of interpreting and sharing PGHD?

Methods: The study consists of semistructured interviews with 12 psychiatrists and clinical psychologists from a large academic hospital. These interviews were thematically and qualitatively analyzed for common themes and workflow elements. User-posted reviews of 56 sleep and mood tracking apps were analyzed to understand app features in comparison with the information gathered from interviews.

Results: The results showed that PGHD have been part of the workflow, but its integration and use are not optimized. Mental health clinicians supported the use of PGHD but had concerns regarding data reliability and accuracy. They also identified challenges in selecting suitable apps for their patients. From the app review, it was discovered that mHealth apps had limited features to support personalization and collaborative care as well as data interpretation and sharing.

Conclusions: This study investigates clinicians' perspectives on PGHD use and explored existing app features using the app review data in the mental health setting. A total of 3 design guidelines were generated: (1) improve data interpretation and sharing mechanisms, (2) consider clinical workflow and electronic health record integration, and (3) support personalized and collaborative care. More research is needed to demonstrate the best practices of PGHD use and to evaluate their effectiveness in improving patient outcomes.

(JMIR Form Res 2020;4(8):e18123) doi:10.2196/18123

KEYWORDS

patient-generated health data; mental health; workflow; mobile application; interview

Introduction

Background and Significance

With advances in mobile technology and the pervasive use of wearable devices, a large amount of digital health data have been generated by patients. Patient-generated health data (PGHD) refer to "health-related data created and recorded by or from patients outside of the clinical setting to help address a health concern," as defined by the Office of the National Coordinator for Health Information Technology (ONC) [1]. According to the ONC, the adoption of PGHD can have several benefits, including, but not limited to, enhancing patient experience, alerting care teams for early intervention, and improving patient health outcomes [1]. Several projects have piloted these ideas and implemented informatics solutions to collect, use, and share PGHD [2], such as in postsurgical surveillance [3,4]. Patients seem to have positive attitudes toward PGHD and are willing to share their data with the care team to support long-term health management. Studies have shown the feasibility of using PGHD to support personalized and effective care management [5]. Using PGHD can improve data work for both clinicians and patients. A case study at a cancer rehabilitation clinic showed that the data collection became more distributed, the nurses asked more focused questions during the consultations, and the patients gradually developed competence in managing their own health [6].

Despite the potential benefits of PGHD, it currently remains limited with a focus on health history, verified surveys, and biometric activities [7]. Ultimately, PGHD should be seamlessly integrated with electronic health records (EHRs) to support the clinical decision-making process [8]. However, several challenges need to be addressed to maximize the benefits of PGHD, as many pertain to its use at the point of care. First, integrating PGHD into a clinical setting necessarily involves considering clinical workflow redesign, data management concerns, patient privacy protections, and ease of PGHD use [9]. This is especially important because clinicians spend a significant amount of time (25%-50%) on documentation tasks in their daily work [10-12]. Second, clinicians may have concerns about the impact of PGHD on reimbursement and data reliability, as PGHD are generated by patients in their daily lives and require extra effort to be consumed in clinics [13,14]. Third, PGHD use impacts the relationship between both patients and clinicians. Therefore, a collaborative model may be developed to fulfill the desire of both parties: patients' desire to know more about their health and clinicians' desire to have better practices [15]. With the growing interest in patient-centered health care, there have been studies that investigate patients' motivation and attitudes toward tracking and sharing personal data [16,17]. However, the extent and methodology behind the use of PGHD and its integration into clinicians' workflow remain unknown and, therefore, require further investigation into the demonstration of best practices.

Objectives

To provide empirical evidence to address these challenges, this study investigated the current use of PGHD within mental health practices, with a focus on workflow, clinicians' perspectives, and data interpretation and sharing. The main reason for choosing a mental health practice setting was that clinicians often rely on patient narratives, observations, and patient-reported outcomes (PROs) to assess the efficacy of psychiatric treatment [18]. PROs refer to "a measurement of any aspect of a patient's health status that comes directly from the patient" [19]. As the purpose of both PROs and PGHD is to collect data from the patients' perspective, clinicians from a mental health practice are more likely to adopt and use PGHD. A total of 3 research questions arise from this: (1) What is the current workflow of mental health practice and how are PGHD integrated into this workflow, (2) what are the clinicians' perspectives on PGHD and how do they choose a mobile-based app (mobile app) for their patients, and (3) what are the features of current mobile apps in terms of interpreting and sharing PGHD?

Literature Review

PGHD

Traditionally, clinicians focus on collecting one-time snapshots of patient information in a clinical setting and making decisions based upon them, thereby losing an opportunity to create a thorough understanding of the patient's health status [20]. In these situations, PGHD are a useful tool for continuous monitoring, especially for patients with chronic conditions that require daily management and benefit from effective tracking [21-23]. PGHD can also improve disease surveillance by more accurately assigning patients to disease categories rather than solely using national and regional data [24].

PGHD can come from multiple data sources, such as family history, medication, and physiological data sensing [25]. PGHD, along with other health data, can ultimately form a repository of patient-centered personal health records, which can then be used to store and manage PGHD [26]. Clinicians see the potential of PGHD but raise some concerns about its use, such as difficulties in summarizing PGHD patterns across different clinical specialties and concerns regarding data management, patient privacy, and ethical challenges of PGHD generation [9,27,28].

Opportunities and Challenges of Mobile Technology

Internet use has been growing steadily since 1990, with approximately 4 billion internet users worldwide, with 15- to 24-year-olds leading the frontier of internet adoption [29,30]. The internet has also become increasingly mobile. An unsurprising result of this trend has been increasingly sophisticated mobile technology and wearable devices, including the development of numerous mobile health (mHealth) apps, all of which have helped generate large amounts of PGHD.

Wearable devices and mHealth apps can facilitate health behavior improvements if designed with proper engagement strategies and data collection methods [31,32]. Effective mHealth app design has helped users monitor and promote positive health habits, such as physical activity and eating behaviors [33,34]. Furthermore, PGHD can help improve clinical trials' efficiency and output [35]. However, these opportunities come with a number of challenges [2]. First, the management and interoperability of massive PGHD require standard vocabularies and data models. Consumer health vocabulary (CHV) is one such effort that has been developed since 2007 [36,37] and enhanced by several text mining–based projects [38-41]. In addition, frameworks and data models have been proposed to incorporate PGHD into EHRs [42,43]. However, very few projects use both CHV and common data models to facilitate PGHD comprehension. Indeed, to our knowledge, only one position paper proposes an interpretability-aware framework to systematically understand PGHD [44]. The second challenge is the lack of guidelines and best practices for integrating PGHD into the clinical workflow [15]. In addition, most PGHD are collected through mHealth apps, which should be carefully evaluated in a standardized manner for efficacy and health outcome improvement [45]. Finally, there are concerns about the quality and ownership of PGHD gathered by mHealth apps and wearable devices [28,46]. App developers should develop more transparent data ownership policies so that users can make informed decisions regarding their PGHD [47]. Guidelines should also be developed to ensure that high-quality data are gathered by users in their daily routines [48].

PGHD Use in Psychiatry

Psychiatry is a medical specialty focused on "the diagnosis, treatment, and prevention of mental, emotional, and behavioral disorders" [49]. Clinicians use a variety of data from patients to determine psychiatric diagnoses. However, these data are often collected solely in clinical settings. Since the 1960s, mental health clinicians have begun to pay more attention to patients' personal perspectives, including health-related quality of life (HRQoL) outcomes. HRQoL is a type of PRO that includes "symptoms of disease or health condition, treatment side effects, and functional status across physical, social, and mental health life domains" [50]. PROs are derived from patients completing standardized questionnaires and cannot guarantee large-scale, continuous data collection [51,52]. However, with the help of mobile technology and wearable devices, PGHD can collect large amounts of patient health data unobtrusively and continuously. Furthermore, because psychiatry has historically depended on PROs and HRQoLs, clinicians in this field should be able to use PGHD without many conceptual barriers.

For instance, ecological momentary assessment (EMA) was proposed to assist clinical psychologists in monitoring HRQoL changes by tracking patient behaviors in real time and in their natural environment. EMA uses various data collection tools, such as written diaries and telephones, and PGHD gathered by mHealth apps and wearables can further support this approach [53]. PGHD can facilitate large-scale environmental psychiatric research in naturalistic settings and create *digital footprints* to measure patients' health status (eg, mood and sleep) in an unobtrusive and longitudinal fashion [54].

Although studies have shown that wearable devices and mHealth apps help in the treatment of mental health by increasing awareness and giving reinforcement, such as in the study by Ng et al [55], research regarding the implementation of mobile technology into the care process is lacking. Studies have only just started to design and develop apps that focus on interface usability and workflow integration. Notably, Bauer et al [56] applied Principles for Digital Development to develop a highly usable mHealth app to support collaborative care for patients and generated 4 more principles based on user feedback. Recently, mental health studies have focused on PGHD, showing a transition to participatory and personalized medicine [57]. However, other stakeholders, such as consultants, policy makers, and vendors, should also be considered when integrating PGHD into mental health [58].

Clearly, more PGHD will be generated in the coming future through mobile technology and hold the potential to improve mental health practice, which has long relied on patient-reported data. Therefore, we follow this trend and aim to provide empirical evidence regarding the current use of PGHD in mental health practices, clinicians' attitudes toward PGHD, and the mHealth app features and selection criteria considered by mental health practitioners.

Methods

Clinical Setting

This study was conducted at the Department of Psychiatry and Behavioral Neuroscience in a large academic hospital in the Midwest United States. The department is a nationally recognized leader in advancing the diagnosis and treatment of mental and behavioral disorders. The department has more than 90 faculty members, with half of them trained as psychiatrists with a medical degree (Doctor of Medicine or Doctor of Osteopathic Medicine). Our study targets were faculty members who are actively seeing patients. As there was a significant portion of faculty trained as psychologists (Doctor of Philosophy or Doctor of Psychology), our study participants included clinicians from both groups.

Study Design

This study contains a set of semistructured interviews and an mHealth app review. The interviews collected qualitative data from both psychiatrists and clinical psychologists to understand their clinical workflow, attitudes toward PGHD, and the use and sharing of PGHD in clinics. In the app review, a set of mHealth apps was selected and reviewed systematically. The main objective of this review was to understand patients' experiences of using mental health–related apps and their opinions around it. The review comments were also downloaded from the Google Play Store and the Apple Store using existing application programming interfaces (APIs) [59,60]. The app review data were summarized in terms of data interpretation and shared features. Data interpretation here is defined as the manner in which PGHD are collected and presented, either qualitatively or quantitatively assessed, and may be subsequently

visualized. Qualitative assessment referred to users' self-reported data on their sleep quality or mood status. Quantitative assessment, on the other hand, relied on the automatically recorded data by an app and its sensors. The clinicians' perspectives and the app features were then synthesized to generate design recommendations for mHealth app designers and developers. This study was reviewed and approved by the institutional review board of the study site (IRB# 2018-6453).

Participant Recruitment

A total of 12 clinicians, including 7 psychiatrists and 5 clinical psychologists, were recruited using convenience sampling and snowball sampling. Two coauthors facilitated the interview invitation to their colleagues; each participant was asked to provide a few names to reach out at the end of the interview. Table 1 shows the participant subgroups and their average professional experience (number of years). Male psychiatrists were the largest group among the participants (n=5), followed by slightly more female clinical psychologists.

Table 1. Participant subgroups.

Characteristics of interviewees	Male	Female	Years of experience, mean (SD)
Psychiatrist, n (%)	5 (41)	2 (17)	13.9 (6.7)
Clinical psychologist, n (%)	2 (17)	3 (25)	12.0 (10.3)
Years of experience, mean (SD)	14.5 (9.3)	11.1 (6.2)	13.1 (8.0)

Qualitative Interviews and Analysis

Each semistructured interview was conducted by 2 members in the research team and lasted for 30 to 45 min. The interview questions were organized in 5 areas: (1) job title and responsibility, (2) clinical workflow, (3) PGHD attitudes, (4) the selection and use of mHealth apps, and (5) the use and sharing of PGHD. Although participants' attitudes toward PGHD were specifically asked, their attitudes toward EHR integration was not prompted. The interviews were audio recorded and transcribed verbatim in 2 steps. First, the audio recordings were transcribed by the Google Cloud Speech-to-Text API. Second, the transcription drafts were reviewed by the research team to ensure high quality and deidentification of the data. The participants were coded from P01 to P12, with the first 7 participants being psychiatrists.

Interview data were analyzed in multiple rounds. Specifically, the transcriptions were coded using the Work Elements Model with a focus on actors, actions, and artifacts [61]. In the first round, a set of swim lane workflow diagrams was generated. In this swim lane chart, the columns (lanes) represent the actors, and the rectangles represent the actions. A workflow diagram was drawn for each psychiatrist. All psychiatric workflow charts were then consolidated into one workflow chart. The same process was repeated for the psychologists. It is worth noting that 3 participants (P09, P10, and P12) had a portion of the workflow that significantly deviated from others because of their job responsibilities. This reflects the variety of mental health practices. The deviation was due to our sampling methods and the relatively small sample size. These portions of workflow were excluded from the consolidated workflow because of their uniqueness.

Following the steps in thematic analysis [62], the transcriptions were analyzed to understand clinicians' attitudes toward PGHD and the interpretation and sharing of PGHD in clinics. To ensure the coding quality, one researcher independently coded all the transcriptions and extracted themes, which were then reviewed by another researcher. The 2 researchers met and resolved any disagreements.

Sleep and Mood App Tracking Review

A total of 31 sleep tracking apps and 25 mood tracking apps in current mobile app markets (App Store and Google Play) were selected and reviewed. Instead of developing our own search keywords, the results of 2 published systematic review papers were used [63,64], providing a list of 73 and 32 tracking apps for sleep and mood, respectively. The papers were published in 2018 and 2019. The details of the sleep and mood app selection are described in Multimedia Appendix 1.

The 2 types of mHealth apps were selected because of their significant role in tracking and understanding patients' status and treatment effects. This was indicated through the prevalent use of these 2 types of apps by the participating psychiatrists and clinical psychologists, as shown in our interviews with them. In total, 9 of 12 participants mentioned these apps. Specifically, 3 of them reported using both sleep tracking and mood tracking mobile apps, 3 of them reported using only mood tracking apps, and 3 of them reported using only sleep tracking apps. Biometric or fitness tracking apps were also mentioned but less frequently; therefore, they were not our investigation focus.

The information of the selected apps was extracted and organized using a spreadsheet in the following 10 columns: (1) app type (either sleep or mood), (2) app ID, (3) source (Apple or Google), (4) app name, (5) rating, (6) advantages from review, (7) hindrance from review, (8) data interpretation features, (9) data sharing features, and (10) review quotes. The app review data were extracted independently by 2 researchers, one of whom reviewed sleep apps and the other who reviewed mood apps. The extracted data were then reviewed by another researcher for quality checks. Information was collected solely from users' comments. Advantages and hindrances were defined as qualities that users predominantly classified as benefiting the user experience or detracting from it, respectively. The details regarding app sharing capabilities (such as social media and exports to other file types) and methods of data display and interpretation were also included. Selected quotes were compiled from users who offered a holistic perspective on the app. It is worth noting that app review data mostly reflected end user

experience and highlighted general user satisfaction or complaints with the apps, complementing our interviews with the clinicians.

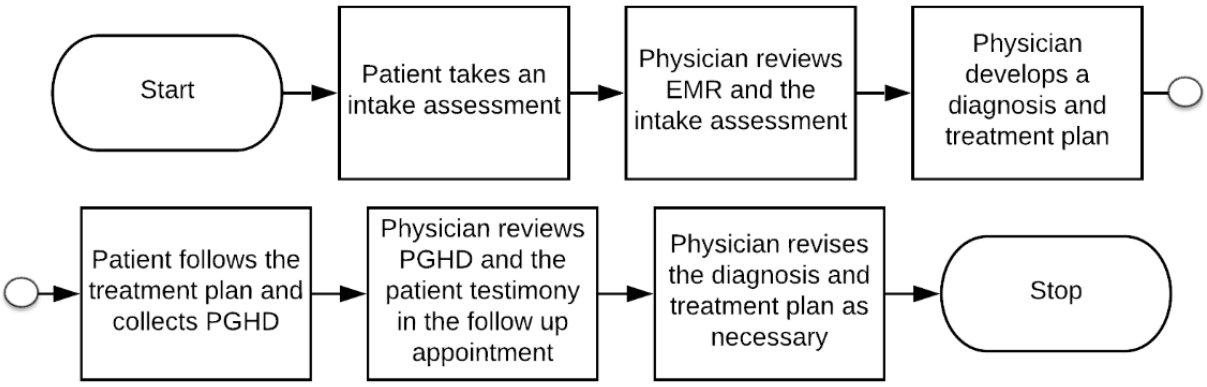
Results

Qualitative Analysis: Clinical Workflow

Consolidated Workflow and PGHD Use

Figure 1 shows a simplified version of the consolidated workflow diagram, which includes step-by-step movement from a patient taking an initial assessment, development of diagnosis and treatment plan, meeting with a physician, the influence of PGHD, and subsequent adjustments in the diagnosis and treatment plan. The detailed workflow diagram is included in Multimedia Appendix 2.

Figure 1. Simplified consolidate workflow diagram of psychiatrists and psychologists based on semistructured interviews. EMR: electronic medical record; PGHD: patient-generated health data.



As shown in Figure 1, a typical clinician workflow in mental health started from gathering patient information and PROs using an intake assessment. Mental health clinicians then reviewed the form along with EHR information, if any, and conducted a clinical interview to establish psychiatric diagnoses and a corresponding treatment plan. Next, patients participated in the treatment (eg, practice assignments and collection of PGHD), following their clinicians' requirements. The ways to collect PGHD varied depending on the clinicians' preferences and patient situations. Next, in subsequent clinical encounters, clinicians reviewed the PGHD and assessed the changes in their patients' symptoms and functioning since the previous session. Clinicians may revise the treatment plan based on the updated patient health status.

Many participants do not consider themselves as the sole decision maker in developing the treatment plan. Instead, they worked with patients to align their goals with the treatment. Homework is a common term used by clinicians to discuss patients' efforts to improve their mental health status between clinic visits. PROs and PGHD can assist patients in their homework, demonstrate their achievement, and provide information to clinicians to make evidence-based decisions on the treatment plan and assess treatment progress. However, participants reported no standardized way to collect and manage PGHD.

Although collecting PROs and PGHD has been part of mental health practices, its use is not optimized. Taking the intake forms as an example, one participant talked about troubles using survey scales when patients wait in the waiting room. This suboptimal data collection may slow down the clinics and reduce the efficiency and quality of care:

> *As I mentioned the scales aren't always filled out and they may not have enough time to fill out the scales if they're call[ed] back right away.* [P06]

Sometimes, PROs can be ambiguous and confusing and require more data to understand patients' health status change and nuanced differences. In this case, PGHD can be complementary and provide more detailed behavioral data to inform the shared clinical decisions on the treatment plans. One participant explained how PROs may be confusing:

> *So, for example, like number three on the PHQ9, in one item, it's assessing "trouble falling asleep," "staying asleep" or "sleeping too much," which could mean very different things in terms of planning to do for treatment. And so, I will always ask a follow-up and then I will manually circle the ones that applied for the patient.* [P10]

Workflow Comparison Between Psychiatrists and Clinical Psychologists

All participants followed the simplified consolidated workflow in their practices. However, there were some noticeable differences between psychiatrists and clinical psychologists. Psychiatric appointments, in general, were shorter (30 min), followed by appointments with clinical psychologists (45-60 min). As psychiatrists can prescribe medication, they check medication use and effects in every patient visit. Psychiatrists also conduct psychotherapy and value the therapeutic

relationship between patients and themselves. On the other hand, clinical psychologists cannot prescribe medication, so they focus solely on conducting psychological assessments and delivering psychotherapy. Therefore, the workflow for clinical psychologists can be very dynamic and conversational. Clinical psychologists pay significant attention to patients' narratives of experience to understand their unique health status and changes. The workflow would be more standardized, however, if a clinical psychologist only focuses on psychological assessments, for example, a cognitive evaluation, as the assessments have a validated procedure to follow and tools to use.

Qualitative Analysis: Clinician Perspectives

Dual Attitudes Toward PGHD

All participants had worked with their patients to track their sleep and/or mood behaviors through an mHealth app and/or a wearable device. Overall, 6 participants held a dual attitude toward PGHD use. On the one hand, clinicians had seen the potential of PGHD and looked forward to taking advantage of them, especially the ability to track their patients' activities between consultations better:

> *There is a long history of using mood scales, potentially longitudinal mood scales, basic tracking charts for depression, bipolar disorder—can be really helpful, also can be somewhat tedious. And there are some apps now that do that very well...*" [P05]

On the other hand, some concerns were raised about data validity and reliability. The participants were being cautious because they identified the need for reliable PGHD to inform evidence-based treatment implementation and evaluation:

> *When you're saying hard data from a device that measures sleep, I would need to know for myself how it's measuring sleep... I think I would tend to question the specificity and accuracy of those for actual sleep... I find that patient self-reports of sleep are unreliable.* [P04]

Concerns About Integrating PGHD Into Workflow and EHR

Clinicians' concerns about integrating PGHD into workflow and EHR systems were a recurring theme in the interviews. One participant indicated that there was no app that seamlessly integrates its data into EHR:

> *So... the way things work now [is] very much sort of pen and paper. You show up at a doctor's office. They get handed some of these screeners and somebody must manually enter it into Epic, which is kind of a pain. We would love to be able to send patients a MyChart message or something and say fill this out and send it back to us and have it automatically go into a flow sheet in Epic that we then track over time with the patient. That would be amazing.* [P03]

A total of 7 participants preferred using paper-based PGHD in their practice because of the patients' preference or the lack of data sharing mechanisms in the apps. In this case, the data flow is deemed indirect. Clinicians would review the PGHD in the session and put the interpretations in the clinical notes, which can slow down the clinical workflow. Generating data visualization of PGHD to facilitate the interpretation of the data seems to be a preferable method for both clinicians and patients:

> *We also graph patient data over time. I do this with all my patients... It would take 20 minutes of the patient session—patient wouldn't get care for that 20 minutes and it's literally I'll have a calculator out, an actual calculator, my graphing calculator, and I'll sit there and calculate the data and then manually make the graph. So having programs built in that aggregate data and automatically populate graphs are great; and patients like to see the visual, the graph—they do, overall. And generally it's good discussion even if the graph is not great in terms of what it's explaining.* [P10]

Manually transcribing data into electronic medical records is not only time consuming but also interferes with the patient-clinician interaction because the clinicians are distracted by the data entry tasks on their computer. Common concerns found included the difficulties of having to face the computer and type in data during sessions:

> *I'm put[ting] in the data in the computer and looking at the questionnaires and if there was a different way of doing that, you know, it would actually make me take my life away from the computer screen and interact with my patient.* [P01]

Lack of Information for App Selection

In addition to the challenges in using data from mHealth apps, it can be challenging to identify what apps to use in the first place. In all, 3 participants found that it was difficult to find an app that met all the requirements. Others sometimes found useful apps through patient recommendations. The following 2 quotes exemplify this situation:

> *I've not been able to figure out an app that I could actually use [and] that was specific to be able to individualized for patients; that's been a challenge.* [P11]

> *There are lots of great apps out there for sure...and sometimes they'll bring stuff to me that I'm unaware of.* [P08]

When asked about using apps for collecting data, some clinicians responded that they did not know that mobile apps could perform certain data tracking tasks:

> *So, we look at this daily measure sheet, which would be perfect for an app that looks at what their moods have been, what kind of sleep they had.* [P11]

In contrast, the app review (see details given in the section *Sleep and Mood Tracking App Reviews*) showed that mood apps in the current market can offer certain functions, which means there is a lack of information to increase app awareness and support app selection process, especially for clinicians. This suggests that an app recommendation system for the clinician to use would be very helpful.

Limited App Features to Support Personalization and Collaborative Care

One factor contributing to the difficulty of finding the right app is the variety in patients' health statuses and conditions. Clinicians would prefer to personalize treatment and collect data in a personalized manner. However, current apps were not designed with features to support personalization in data collection. One participant elaborated that:

> *A psychiatric exam, ugh, depends on a lot offactors actually...Let's say I have a 60-year-old sort of borderline intellectual functioning person, maybe 8th grade education with never seen a psychiatrist and coming in first time... So that exam...maybe take more time there, more explanation, more education, more time to elicit... I may have to offer something, a different questionnaire other than the standard four to five that I sent to every patient.* [P01]

Another type of feature that current apps may be missing is to support collaborative care. It is beneficial to support collaborative care so that patients can be more engaged in their health and take control of their care, and clinicians can create a treatment plan that is most suitable and effective for the patient's situation. One clinician further talked about the difficulty in finding an app in current app markets to address his needs of data collection, although they do have data collection features:

> *The app would have to collect data on several different domains like nutrition, like sleep, like physical activity, and then maybe sampling mood multiple times a day. So there are apps that look at the individual factors, but I don't know, I'm not aware there may be an app [that can offer all above].* [P09]

Sleep and Mood Tracking App Reviews

Statistical Summary

Table 2 shows the statistical summary of the selected apps. The data were retrieved on September 8, 2019. The average number of reviews was around 400, and the average user rating was around 4.0 for both sleep and mood tracking apps.

Table 2. Statistical summary of the selected sleep and mood tracking apps.

App characteristics	Sleep	Mood
Number of apps, n (%)	31 (55)	25 (45)
Number of reviews, mean (SD)	346.1 (896.8)	387.8 (912.5)
User rating, mean (SD)	3.88 (0.78)	3.99 (0.54)
Had a data interpretation feature, n (%)	31 (55)	21 (38)
Had a data sharing feature, n (%)	18 (32)	16 (29)

Data Interpretation and Sharing Features

The data interpretation, which includes collection and visualization, and sharing features in the selected apps varied. All 31 sleep tracking apps offered at least one feature for data interpretation, including sleep quantity statistics (n=27) and sleep quality analysis (n=5). Of the 31 sleep tracking apps, 18 (58%) support data sharing in various means, including direct sharing with other people (n=8) and integration with other apps (n=7).

The major source of data interpretation is through collecting statistics, through recording several days' worth of accumulated data points. These data points are either manually inputted by users or automatically recorded by the app. Overall, 5 apps also offer qualitative assessments of sleep patterns by presenting the sleep cycles that users experience, recording sound files, and/or providing descriptive sleep analyses.

Many sleep tracking apps also support sharing data. This is usually done by exporting and downloading the data as a comma-separated values (CSV), PDF, etc file (N=8) or via integration with an alternate app, such as Apple Health (n=7). Other sources include social media (n=3) and email (n=4). Many apps also support multiple forms of sharing.

Many users find accurate sleep tracking to be helpful in improving their sleep quality and daily life:

> *I love this app. Pleasing to the eye and so many great features! I love that it keeps stats on your sleep cycles and can be set to accommodate how your sleep may be affected by working out, caffeine, or other factors.*

On the other hand, common complaints of sleep apps included inaccurate tracking, failures in data collection, and difficulty in use due to technical issues. For example, one user expressed frustration in inaccurate tracking:

> *This app has been super frustrating. I use it while wearing my watch and it has recorded me in a deep sleep while I was making food.*

The analysis of mood tracking apps shows similar results. Of the 25 mood tracking apps, 21 (84%) offered either a qualitative (n=4) and/or quantitative (n=22) data interpretation feature. Eight mood apps had more than one mode of quantitative data interpretation. Of the 25 mood tracking apps, 16 (64%) had data sharing capabilities, mainly through direct sharing with other users (n=13) or social media (n=5). Three mood apps had more than one mode of data sharing.

Many apps offer multiple means of data interpretation. The forms of qualitative data interpretation include monthly reports of the moods logged. Forms of quantitative data interpretation include daily, weekly, or monthly graphs or charts of the frequency of moods logged by the user. The modes of sharing data are exporting the data as a spreadsheet, CSV, PDF, etc file (n=6); via email (n=3); or via cloud sharing (n=3).

Many users find that the ability of the apps to track and reflect on one's mood trends was very helpful:

> *I am not exaggerating to say this app has been life changing. This app allows me to customize the settings so I can simultaneously track my moods, medications, activity level, wellness tools, triggers and alternative treatments. All of this info is so important. Daily logging with an easy interface means I can track LOTS OF DATA in an organized way. The visual graphs show the relationships between these various data points. The info gathered is a valuable tool in my wellness and recovery.*

The most common complaint was the limited options of mood offered, which caused difficulties for users in tracking their real mood, as indicated in the following review comments:

> *More moods. Exactly the kind of app I need. I just wish there were more moods like: anxious, on edge, tearful, sensitive, irritable, exhausted.*

With both mood and sleep apps, the primary method of data visualization was through graphs and charts (n=37). Sleep apps, which recorded quantitative data more frequently than qualitative data, used statistics (such as the length of time slept) to present graphical summaries of sleep over time. These graphs were then analyzed to show the quality of sleep over time, track trends, inconsistencies, and improvements. Similarly, mood apps also used graphical representations to present data summaries. However, these apps typically used qualitative data entries (such as recorded moods on certain days) to provide graphs and charts that showed the number of times certain moods were recorded and how moods fluctuated over the course of weeks or months. The popular usage of graphs across both categories suggests that regardless of the nature of data collection (either quantitative or qualitative), users prefer apps that visualize data through digestible and succinct representations that make trends identifiable and trackable over time.

Discussion

Principal Findings

We conducted a mixed methods study with 12 clinicians in mental health practice to understand their perspectives on PGHD and the current use and to share features of the mHealth apps on the market. The results show that mental health clinicians had a dual attitude toward PGHD. The advantages and concerns of PGHD use were aligned with those in the literature. It is not surprising that mental health clinicians have seen the potential of PGHD because they have been largely relying on PROs to develop treatment plans [18]. Our results also confirmed that mental health clinicians have concerns about data validity and reliability similar to clinicians in other specialties [13,14]. Although mental health practices have started to use PGHD, their use has not been optimized in clinical workflow and integrated into EHR. However, limited PGHD integration with EHRs may not be totally native. Clinicians who have concerns with data reliability and ability may prefer to review PGHD before putting them into EHR, rather than including them directly from mHealth apps and/or wearables. Moreover, our findings revealed that there are different ways to make use of PGHD without integrating it into EHR systems, such as using it to check patient conditions and homework in between clinical sessions. However, personalized data tracking and visualization are critical factors in the successful use of PGHD for both patients and clinicians.

In addition to using PGHD in clinics, we found that mental health clinicians may have a hard time finding the right mHealth apps for their patients to collect PGHD in the first place. There was a lack of information to help them choose the most suitable apps for their patients to use. Part of the reason is that each patient has a unique mental health status and condition, and mHealth apps do not support much personalization. Moreover, most of the mHealth apps were patient centered but may not support collaborative care. As clinicians and patients frequently make shared decisions for the treatment plan, mHealth apps without features to support collaborative care could reduce clinicians' willingness to adopt them or introduce barriers in clinical workflow. In addition, our review on sleep and mood tracking apps confirmed that the current mHealth apps on the market had limited features in data interpretation (eg, visualization) and limited mechanism to share PGHD with other people and EHR systems.

Design Implications

Improve Data Interpretation and Sharing Mechanisms

Current mood or sleep tracking apps are focused on collecting PGHD in a patient-centered manner, which is a critical first step. However, to maximize the value of PGHD, these apps should improve their mechanisms in data interpretation and sharing. Specifically, data visualization can be a viable way to help both mental health clinicians and patients interpret much PGHD and identify patterns and trends regardless of their integration into EHR. On the other hand, mHealth apps should enable data sharing mechanisms with different parties, including, but not limited to, clinicians, families, friends, and other practitioners, as this was a concern noted by both clinicians during interviews and users in the app review [58]. It is worth emphasizing the importance of information confidentiality when designing a sharing mechanism. Psychology clinics are considered as a *safe bed* for patients to discuss their mental health status and conditions with clinicians. Hence, data sharing mechanisms should not be one-size-fit-all; they should be designed to allow patients to select which part of PGHD to share and how to share to keep highest data confidentiality based on their psychiatric conditions.

Consider Clinical Workflow and EHR Integration

Technology-enabled clinical data capture and documentation should consider the clinical workflow [65]. Similarly, we suggest that mHealth apps designed to gather PGHD should consider clinical workflow to improve the quality of patient experience. Although PROs and PGHD have been used in mental health clinics, there are no guidelines for data collection and use. It would be beneficial to conduct observational workflow analysis, such as time and motion studies, to understand when and where PGHD are used and identified

bottlenecks. In fact, it is common to conduct such workflow analysis to design and develop any health information technology that will be used in clinics (eg, clinical decision support tools) [66].

On the other hand, mHealth apps should also consider EHR integration. We have discussed the importance of sharing PGHD with various parties in a way to protect information confidentiality. PGHD may be shared with clinicians directly through EHR integration or indirectly through data summarization and electronic reports. We have also seen a working example in our interviews to integrate PGHD with EHR using a Research Electronic Data Capture database [67]. As clinicians may have concerns on data reliability and accuracy, having an indirect EHR integration (eg, a dashboard or PDF report demonstrating patterns in PGHD for review) may be a viable way to reduce clinicians' concerns and increase their PGHD adoption in clinics. As many apps are coming into the market and are also going obsolete at a fast pace, a standardized data management and export system could be proposed to better integrate PGHD into clinical practice regardless of specific kinds of apps.

Support Personalized and Collaborative Care

Our results showed that mHealth apps must support both personalization in data collection and collaboration between patients and clinicians during clinics. In terms of personalization, as each patient has a unique mental health status and conditions in various social contexts, it may be difficult to find an app that covers all kinds of needs of PGHD collection. There are 2 ways to approach this issue from a design perspective. First, mHealth apps should maximize their ability to personalize data collection methods to fit different patient needs. Participatory design methods may be helpful in identifying such needs and incorporating them into app features. Second, as clinicians do not always have the information of some existing apps that are potentially useful, an app recommendation system may be developed to assist clinicians in choosing which apps to use to collect PGHD. Currently, some clinicians rely on patients' recommendations. This app recommendation system may be maintained by both clinicians and patients.

In addition, mHealth apps should be designed to support collaborative care. It is critical to ensure that PGHD are used effectively and efficiently during clinics. Ethnographic observations may be needed to systematically document the behaviors and interactions between patients and clinicians. The results can help researchers better understand the role that PGHD and mHealth apps play during clinic visits and generate guidelines to redesign mHealth apps and improve the use of PGHD to support shared decision-making and collaborative care.

Limitations

This study has a few limitations. First, it was conducted in one institution using convenience sampling to recruit participants, thereby limiting generalizability. Second, the study only included clinicians, whereas patients' perspectives were only indirectly approached through our research on app reviews. However, we believe that the clinician perspective on PGHD was sufficiently understood because we continued recruiting the participants until the data reached saturation. Moreover, we compared the interview data with the app review data, which were complementary, to generate the design implications. Third, the app review was conducted using a snapshot of mHealth apps investigated in the past 2 years. The apps may have since then been updated, resulting in fluctuating ratings and different features. However, the efficacy of the app review remains the same, as it helps identify the keep features that users look for in mHealth apps and provides sufficient data to inform app designers. Finally, clinicians can have several subspecialties in mental health, which affects their workflow and how they use PGHD. We were not able to recruit a diverse sample to include all the opinions from the mental health clinicians. However, this study focused on the common workflow components in mental health practices and served as a pilot study to understand clinician perspectives.

Future Directions

We will continue investigating the best practices for using PGHD at the point of care, considering clinical workflow and developing informatics solutions to facilitate the development of a collaborative model to make sense of PGHD to inform shared decisions. The PGHD here will not be limited to data gathered from mHealth apps and wearable devices. They can include data from social media (eg, Twitter and forums) to synthesize more information about patients' opinions on their health [68]. Interactive data visualization may be a viable way to achieve the common goals of clinicians and patients. Moreover, we will pay specific attention to the integration between PGHD and EHR and further develop clinical decision support tools with machine intelligence to use this new and valuable dataset to improve patient outcomes.

Conclusions

This study demonstrated the clinicians' perspectives on PGHD and the current features of mHealth apps. The results showed that PGHD have been used in mental health practices but in a suboptimal way without guidelines. Clinicians look forward to the potential benefit of using PGHD but have dual attitudes toward PGHD. That is, clinicians see the potential of PGHD but hesitate to embrace them mainly because of data validity and reliability concerns. Other concerns about workflow and EHR integration exist prevalently. Moreover, clinicians experienced challenges in selecting suitable apps for their patients, partly because of the limited features of mHealth apps in supporting personalized and collaborative care. We identified 3 design implications: (1) improve data interpretation and sharing mechanisms, (2) consider clinical workflow and EHR integration, and (3) support personalized and collaborative care. We will continue our research with a focus on designing and developing informatics solutions to demonstrate the best practices of PGHD use and evaluate their effectiveness in improving patient outcomes.

Acknowledgments

This study was conducted using the first (corresponding) author's start-up funds. The author led and coordinated the effort in study design, data collection and analysis, results interpretation, manuscript writing, and decisions to submit for publication. The authors would like to thank all the participating mental health clinicians in this study. The authors would also like to thank Mr Jason Revalee, Mr Ruthvik Abbu, Ms Shuai Mu, Ms Longwei Li, and Ms Qiyang Zhou for their help in interview data collection and analysis, literature review, and manuscript proofread.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Sleep and mood app selection process.
[DOCX File , 67 KB - formative_v4i8e18123_app1.docx]

Multimedia Appendix 2
Detailed consolidated workflow diagram.
[DOCX File , 3820 KB - formative_v4i8e18123_app2.docx]

References

1. Cortez A, Hsii P, Mitchell E, Riehl V, Smith P. Conceptualizing a Data Infrastructure for the Capture, Use, and Sharing of Patient-Generated Health Data in Care Delivery and Research through 2024. Office of the National Coordinator for Health Information. 2018. URL: https://www.healthit.gov/sites/default/files/onc_pghd_final_white_paper.pdf [accessed 2020-07-09]
2. Lai AM, Hsueh PS, Choi YK, Austin RR. Present and future trends in consumer health informatics and patient-generated health data. Yearb Med Inform 2017 Aug;26(1):152-159 [FREE Full text] [doi: 10.15265/IY-2017-016] [Medline: 29063559]
3. Lober WB, Evans HL. Patient-generated health data in surgical site infection: changing clinical workflow and care delivery. Surg Infect (Larchmt) 2019 Oct;20(7):571-576. [doi: 10.1089/sur.2019.195] [Medline: 31397635]
4. Semple JL, Evans HL, Lober WB, Lavallee DC. Implementing mobile health interventions to capture post-operative patient-generated health data. Surg Infect (Larchmt) 2019 Oct;20(7):566-570. [doi: 10.1089/sur.2019.151] [Medline: 31429637]
5. Lv N, Xiao L, Simmons ML, Rosas LG, Chan A, Entwistle M. Personalized hypertension management using patient-generated health data integrated with electronic health records (EMPOWER-H): six-month pre-post study. J Med Internet Res 2017 Sep 19;19(9):e311 [FREE Full text] [doi: 10.2196/jmir.7831] [Medline: 28928111]
6. Islind AS, Lindroth T, Lundin J, Steineck G. Shift in translations: data work with patient-generated health data in clinical practice. Health Informatics J 2019 Sep;25(3):577-586. [doi: 10.1177/1460458219833097] [Medline: 30866707]
7. Adler-Milstein J, Nong P. Early experiences with patient generated health data: health system and patient perspectives. J Am Med Inform Assoc 2019 Oct 1;26(10):952-959. [doi: 10.1093/jamia/ocz045] [Medline: 31329886]
8. Demiris G, Iribarren SJ, Sward K, Lee S, Yang R. Patient generated health data use in clinical practice: a systematic review. Nurs Outlook 2019;67(4):311-330. [doi: 10.1016/j.outlook.2019.04.005] [Medline: 31277895]
9. Cohen DJ, Keller SR, Hayes GR, Dorr DA, Ash JS, Sittig DF. Integrating patient-generated health data into clinical care settings or clinical decision-making: lessons learned from project healthdesign. JMIR Hum Factors 2016 Oct 19;3(2):e26 [FREE Full text] [doi: 10.2196/humanfactors.5919] [Medline: 27760726]
10. Ammenwerth E, Spötl HP. The time needed for clinical documentation versus direct patient care. A work-sampling analysis of physicians' activities. Methods Inf Med 2009;48(1):84-91. [Medline: 19151888]
11. Poissant L, Pereira J, Tamblyn R, Kawasumi Y. The impact of electronic health records on time efficiency of physicians and nurses: a systematic review. J Am Med Inform Assoc 2005;12(5):505-516 [FREE Full text] [doi: 10.1197/jamia.M1700] [Medline: 15905487]
12. Sinsky C, Colligan L, Li L, Prgomet M, Reynolds S, Goeders L, et al. Allocation of physician time in ambulatory practice: a time and motion study in 4 specialties. Ann Intern Med 2016 Dec 6;165(11):753-760. [doi: 10.7326/M16-0961] [Medline: 27595430]
13. Chung AE, Basch EM. Potential and challenges of patient-generated health data for high-quality cancer care. J Oncol Pract 2015 May;11(3):195-197 [FREE Full text] [doi: 10.1200/JOP.2015.003715] [Medline: 25852139]
14. Nundy S, Lu CE, Hogan P, Mishra A, Peek ME. Using patient-generated health data from mobile technologies for diabetes self-management support: provider perspectives from an academic medical center. J Diabetes Sci Technol 2014 Jan;8(1):74-82 [FREE Full text] [doi: 10.1177/1932296813511727] [Medline: 24876541]
15. Reading MJ, Merrill JA. Converging and diverging needs between patients and providers who are collecting and using patient-generated health data: an integrative review. J Am Med Inform Assoc 2018 Jun 1;25(6):759-771 [FREE Full text] [doi: 10.1093/jamia/ocy006] [Medline: 29471330]

16. Kelley C, Lee B, Wilcox L. Self-Tracking for Mental Wellness: Understanding Expert Perspectives and Student Experiences. In: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. 2017 Presented at: CHI'17; May 6-11, 2017; Denver, USA. [doi: 10.1145/3025453.3025750]
17. Lee K, Hong H. MindNavigator: Exploring the Stress and Self-Interventions for Mental Wellness. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. 2018 Presented at: CHI'18; April 21-26, 2018; Montreal, QC, Canada. [doi: 10.1145/3173574.3174146]
18. Sartorius N. Patient-reported outcomes in psychiatry. Dialogues Clin Neurosci 2014 Jun;16(2):123-124 [FREE Full text] [Medline: 25152651]
19. US Department of Health and Human Services FDA Center for Drug Evaluation and Research, US Department of Health and Human Services FDA Center for Biologics, Evaluation and Research, US Department of Health and Human Services FDA Center for Devices and Radiological Health. Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims: draft guidance. Health Qual Life Outcomes 2006 Oct 11;4:79 [FREE Full text] [doi: 10.1186/1477-7525-4-79] [Medline: 17034633]
20. Deering MJ. Issue Brief: Patient-Generated Health Data and Health IT. Office of the National Coordinator for Health Information. 2013. URL: https://docs.google.com/viewer?url=http%3A%2F%2Fwww.healthit.gov%2Fsites%2Fdefault%2Ffiles%2Fpghd_brief_final122013.pdf. [accessed 2020-07-09]
21. Ancker JS, Mauer E, Kalish RB, Vest JR, Gossey JT. Early adopters of patient-generated health data upload in an electronic patient portal. Appl Clin Inform 2019 Mar;10(2):254-260 [FREE Full text] [doi: 10.1055/s-0039-1683987] [Medline: 30970383]
22. Reading M, Baik D, Beauchemin M, Hickey KT, Merrill JA. Factors influencing sustained engagement with ECG self-monitoring: perspectives from patients and health care providers. Appl Clin Inform 2018 Oct;9(4):772-781 [FREE Full text] [doi: 10.1055/s-0038-1672138] [Medline: 30304745]
23. Vizer LM, Eschler J, Koo BM, Ralston J, Pratt W, Munson S. 'It's not just technology, it's people': constructing a conceptual model of shared health informatics for tracking in chronic illness management. J Med Internet Res 2019 Apr 29;21(4):e10830 [FREE Full text] [doi: 10.2196/10830] [Medline: 31033452]
24. Bourgeois FT, Porter SC, Valim C, Jackson T, Cook EF, Mandl KD. The value of patient self-report for disease surveillance. J Am Med Inform Assoc 2007;14(6):765-771 [FREE Full text] [doi: 10.1197/jamia.M2134] [Medline: 17712092]
25. Woods SS, Evans NC, Frisbee KL. Integrating patient voices into health information for self-care and patient-clinician partnerships: Veterans Affairs design recommendations for patient-generated data applications. J Am Med Inform Assoc 2016 May;23(3):491-495. [doi: 10.1093/jamia/ocv199] [Medline: 26911810]
26. Park YR, Lee Y, Kim JY, Kim J, Kim HR, Kim Y, et al. Managing patient-generated health data through mobile personal health records: analysis of usage data. JMIR Mhealth Uhealth 2018 Apr 9;6(4):e89 [FREE Full text] [doi: 10.2196/mhealth.9620] [Medline: 29631989]
27. Huba N, Zhang Y. Designing patient-centered personal health records (PHRs): health care professionals' perspective on patient-generated data. J Med Syst 2012 Dec;36(6):3893-3905. [doi: 10.1007/s10916-012-9861-z] [Medline: 22644130]
28. Nittas V, Lun P, Ehrler F, Puhan MA, Mütsch M. Electronic patient-generated health data to facilitate disease prevention and health promotion: scoping review. J Med Internet Res 2019 Oct 14;21(10):e13320 [FREE Full text] [doi: 10.2196/13320] [Medline: 31613225]
29. Number of Internet Users Worldwide From 2005 to 2019 (in Millions). Statista. URL: https://www.statista.com/statistics/273018/number-of-internet-users-worldwide/ [accessed 2019-09-09]
30. ICT Fats and Figures. Google Docs. 2017. URL: https://docs.google.com/viewer?url=https%3A%2F%2Fwww.itu.int%2Fen%2FITU-D%2FStatistics%2FDocuments%2Ffacts%2FICTFactsFigures2017.pdf [accessed 2019-09-09]
31. Burns K, McBride CA, Patel B, FitzGerald G, Mathews S, Drennan J. Creating consumer-generated health data: interviews and a pilot trial exploring how and why patients engage. J Med Internet Res 2019 Jun 13;21(6):e12367 [FREE Full text] [doi: 10.2196/12367] [Medline: 31199312]
32. Patel MS, Asch DA, Volpp KG. Wearable devices as facilitators, not drivers, of health behavior change. J Am Med Assoc 2015 Feb 3;313(5):459-460. [doi: 10.1001/jama.2014.14781] [Medline: 25569175]
33. Al Ayubi SU, Parmanto B, Branch R, Ding D. A persuasive and social mhealth application for physical activity: a usability and feasibility study. JMIR Mhealth Uhealth 2014 May 22;2(2):e25 [FREE Full text] [doi: 10.2196/mhealth.2902] [Medline: 25099928]
34. Du H, Venkatakrishnan A, Youngblood GM, Ram A, Pirolli P. A group-based mobile application to increase adherence in exercise and nutrition programs: a factorial design feasibility study. JMIR Mhealth Uhealth 2016 Jan 15;4(1):e4 [FREE Full text] [doi: 10.2196/mhealth.4900] [Medline: 26772910]
35. Wood WA, Bennett AV, Basch E. Emerging uses of patient generated health data in clinical research. Mol Oncol 2015 May;9(5):1018-1024 [FREE Full text] [doi: 10.1016/j.molonc.2014.08.006] [Medline: 25248998]
36. Keselman A, Tse T, Crowell J, Browne A, Ngo L, Zeng Q. Assessing consumer health vocabulary familiarity: an exploratory study. J Med Internet Res 2007 Mar 14;9(1):e5 [FREE Full text] [doi: 10.2196/jmir.9.1.e5] [Medline: 17478414]

37. Zeng QT, Tse T, Divita G, Keselman A, Crowell J, Browne AC, et al. Term identification methods for consumer health vocabulary development. J Med Internet Res 2007 Feb 28;9(1):e4 [FREE Full text] [doi: 10.2196/jmir.9.1.e4] [Medline: 17478413]
38. Doing-Harris KM, Zeng-Treitler Q. Computer-assisted update of a consumer health vocabulary through mining of social network data. J Med Internet Res 2011 May 17;13(2):e37 [FREE Full text] [doi: 10.2196/jmir.1636] [Medline: 21586386]
39. Gu G, Zhang X, Zhu X, Jian Z, Chen K, Wen D, et al. Development of a consumer health vocabulary by mining health forum texts based on word embedding: semiautomatic approach. JMIR Med Inform 2019 May 23;7(2):e12704 [FREE Full text] [doi: 10.2196/12704] [Medline: 31124461]
40. He Z, Chen Z, Oh S, Hou J, Bian J. Enriching consumer health vocabulary through mining a social Q&A site: a similarity-based approach. J Biomed Inform 2017 May;69:75-85 [FREE Full text] [doi: 10.1016/j.jbi.2017.03.016] [Medline: 28359728]
41. Vydiswaran VG, Mei Q, Hanauer DA, Zheng K. Mining consumer health vocabulary from community-generated text. AMIA Annu Symp Proc 2014;2014:1150-1159 [FREE Full text] [Medline: 25954426]
42. Gollamudi SS, Topol EJ, Wineinger NE. A framework for smartphone-enabled, patient-generated health data analysis. PeerJ 2016;4:e2284 [FREE Full text] [doi: 10.7717/peerj.2284] [Medline: 27547580]
43. Plastiras P, O'Sullivan D. Exchanging personal health data with electronic health records: a standardized information model for patient generated health data and observations of daily living. Int J Med Inform 2018 Dec;120:116-125. [doi: 10.1016/j.ijmedinf.2018.10.006] [Medline: 30409336]
44. Hsueh PS, Dey S, Das S, Wetter T. Making sense of patient-generated health data for interpretable patient-centered care: the transition from 'more' to 'better'. Stud Health Technol Inform 2017;245:113-117. [Medline: 29295063]
45. Kumar S, Nilsen WJ, Abernethy A, Atienza A, Patrick K, Pavel M, et al. Mobile health technology evaluation: the mHealth evidence workshop. Am J Prev Med 2013 Aug;45(2):228-236 [FREE Full text] [doi: 10.1016/j.amepre.2013.03.017] [Medline: 23867031]
46. Abdolkhani R, Gray K, Borda A, de Souza R. Patient-generated health data quality for clinical use: human and technology factors. Iproc 2018 Sep 17;4(2):e11703. [doi: 10.2196/11703]
47. Petersen C, DeMuro P. Legal and regulatory considerations associated with use of patient-generated health data from social media and mobile health (mHealth) devices. Appl Clin Inform 2015;6(1):16-26 [FREE Full text] [doi: 10.4338/ACI-2014-09-R-0082] [Medline: 25848410]
48. Abdolkhani R, Borda A, Gray K. Quality management of patient generated health data in remote patient monitoring using medical wearables - a systematic review. Stud Health Technol Inform 2018;252:1-7. [Medline: 30040674]
49. What Is Psychiatry? American Psychiatric Association. URL: https://www.psychiatry.org/patients-families/what-is-psychiatry [accessed 2020-06-29]
50. Revicki DA, Kleinman L, Cella D. A history of health-related quality of life outcomes in psychiatry. Dialogues Clin Neurosci 2014 Jun;16(2):127-135 [FREE Full text] [Medline: 25152652]
51. Bendich I, Chung C, Hwang K, Patterson J, Mulvihill J, Barry J, et al. Changes in prospectively collected longitudinal patient-generated health data are associated with short-term patient-reported outcomes after total joint arthroplasty: a pilot study. Arthroplast Today 2019 Mar;5(1):61-63 [FREE Full text] [doi: 10.1016/j.artd.2019.01.005] [Medline: 31020024]
52. Cella D, Hahn EA, Jensen SE, Butt Z, Nowinski CJ, Rothrock N, et al. Patient-reported outcomes in performance measurement. Research Triangle Park 2015:-. [Medline: 28211667]
53. Shiffman S, Stone AA, Hufford MR. Ecological momentary assessment. Annu Rev Clin Psychol 2008;4:1-32. [doi: 10.1146/annurev.clinpsy.3.022806.091415] [Medline: 18509902]
54. Bidargaddi N, Musiat P, Makinen V, Ermes M, Schrader G, Licinio J. Digital footprints: facilitating large-scale environmental psychiatric research in naturalistic settings through data from everyday technologies. Mol Psychiatry 2017 Feb;22(2):164-169 [FREE Full text] [doi: 10.1038/mp.2016.224] [Medline: 27922603]
55. Ng A, Reddy M, Zalta AK, Schueller SM. Veterans' perspectives on FitBit use in treatment for post-traumatic stress disorder: an interview study. JMIR Ment Health 2018 Jun 15;5(2):e10415 [FREE Full text] [doi: 10.2196/10415] [Medline: 29907556]
56. Bauer AM, Hodsdon S, Bechtel JM, Fortney JC. Applying the principles for digital development: case study of a smartphone app to support collaborative care for rural patients with posttraumatic stress disorder or bipolar disorder. J Med Internet Res 2018 Jun 6;20(6):e10048 [FREE Full text] [doi: 10.2196/10048] [Medline: 29875085]
57. Berrouiguet S, Perez-Rodriguez MM, Larsen M, Baca-García E, Courtet P, Oquendo M. From ehealth to ihealth: transition to participatory and personalized medicine in mental health. J Med Internet Res 2018 Jan 3;20(1):e2 [FREE Full text] [doi: 10.2196/jmir.7412] [Medline: 29298748]
58. Petersen C, Adams SA, DeMuro PR. mHealth: don't forget all the stakeholders in the business case. Med 2 0 2015 Dec 31;4(2):e4 [FREE Full text] [doi: 10.2196/med20.4349] [Medline: 26720310]
59. Olano F. Google-Play-Scraper. GitHub. URL: https://github.com/facundoolano/google-play-scraper [accessed 2020-06-29]
60. Olano F. App-Store-Scraper. GitHub. URL: https://github.com/facundoolano/app-store-scraper [accessed 2020-06-29]
61. Unertl KM, Novak LL, Johnson KB, Lorenzi NM. Traversing the many paths of workflow research: developing a conceptual framework of workflow terminology through a systematic literature review. J Am Med Inform Assoc 2010;17(3):265-273 [FREE Full text] [doi: 10.1136/jamia.2010.004333] [Medline: 20442143]

62. Maguire M, Delahunt B. Doing a thematic analysis: A practical, step-by-step guide for learning and teaching scholars. All Ireland Journal of Higher Education [FREE Full text]
63. Caldeira C, Chen Y, Chan L, Pham V, Chen Y, Zheng K. Mobile apps for mood tracking: an analysis of features and user reviews. AMIA Annu Symp Proc 2017;2017:495-504 [FREE Full text] [Medline: 29854114]
64. Choi YK, Demiris G, Lin S, Iribarren SJ, Landis CA, Thompson HJ, et al. Smartphone applications to support sleep self-management: review and evaluation. J Clin Sleep Med 2018 Oct 15;14(10):1783-1790 [FREE Full text] [doi: 10.5664/jcsm.7396] [Medline: 30353814]
65. Cusack CM, Hripcsak G, Bloomrosen M, Rosenbloom ST, Weaver CA, Wright A, et al. The future state of clinical data capture and documentation: a report from AMIA's 2011 policy meeting. J Am Med Inform Assoc 2013 Jan 1;20(1):134-140 [FREE Full text] [doi: 10.1136/amiajnl-2012-001093] [Medline: 22962195]
66. Ozkaynak M, Wu DT, Hannah K, Dayan PS, Mistry RD. Examining workflow in a pediatric emergency department to develop a clinical decision support for an antimicrobial stewardship program. Appl Clin Inform 2018 Apr;9(2):248-260 [FREE Full text] [doi: 10.1055/s-0038-1641594] [Medline: 29642247]
67. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
68. Golder SA, Macy MW. Digital footprints: opportunities and challenges for online social research. Annu Rev Sociol 2014 Jul 30;40(1):129-152. [doi: 10.1146/annurev-soc-071913-043145]

Abbreviations

API: application programming interface
CHV: consumer health vocabulary
CSV: comma-separated values
EHR: electronic health record
EMA: ecological momentary assessment
HRQoL: health-related quality of life
mHealth: mobile health
ONC: Office of the National Coordinator for Health Information Technology
PGHD: patient-generated health data
PRO: patient-reported outcome

Edited by G Eysenbach; submitted 04.02.20; peer-reviewed by MY Chih, K Blondon; comments to author 30.03.20; revised version received 25.05.20; accepted 15.06.20; published 07.08.20.

Please cite as:
Wu DTY, Xin C, Bindhu S, Xu C, Sachdeva J, Brown JL, Jung H
Clinician Perspectives and Design Implications in Using Patient-Generated Health Data to Improve Mental Health Practices: Mixed Methods Study
JMIR Form Res 2020;4(8):e18123
URL: https://formative.jmir.org/2020/8/e18123
doi:10.2196/18123
PMID:32763884

Original Paper

Perceptions and Attitudes Toward a Mobile Phone App for Mental Health for College Students: Qualitative Focus Group Study

Bree E Holtz[1], PhD; Alexis M McCarroll[1], BA; Katharine M Mitchell[1], MA

Michigan State University, East Lansing, MI, United States

Corresponding Author:
Bree E Holtz, PhD
Michigan State University
404 Wilson Road, Room 309
East Lansing, MI
United States
Phone: 1 5178844537
Email: bholtz@msu.edu

Abstract

Background: Many college students who have mental health issues do not receive professional care for various reasons. Students who do not receive help often have both short- and long-term adverse health outcomes. Mobile apps for mental health services such as MySSP, a service provided to college students through their university, may help eliminate barriers to seeking mental health care and result in more positive outcomes for college students.

Objective: This qualitative study aims to better understand college students' perceptions and attitudes toward the adoption and use of a mobile phone app for mental health, MySSP, using the technology acceptance model (TAM).

Methods: A series of nine focus groups were conducted with college students (N=30) between February and May 2019 at a large, public Midwestern university. The moderator's guide was based on the TAM, and focus group sessions primarily focused on the use and knowledge of apps for mental health, specifically, MySSP. The focus group transcriptions were hand-coded to develop a set of themes that encompassed students' perceptions and attitudes toward MySSP.

Results: The analysis of the focus groups suggested the following themes: (1) existing awareness of the app, (2) perceived usefulness, (3) perceived ease of use, (4) attitudes toward apps for mental health and MySSP, and (5) social influence.

Conclusions: The results of this study provide deeper insights into the perceptions of a mobile app for mental health among college students. Future research should explore the specific contexts in which an app for mental health will be most effective for college students.

(JMIR Form Res 2020;4(8):e18347) doi:10.2196/18347

KEYWORDS

mental health; mobile phone; mHealth

Introduction

Background

The prevalence and severity of mental health disorders among college students has been steadily rising [1]. Data from the American College Health Association showed that 45.1% of college students reported feeling so depressed that it was difficult to function and 65.7% felt overwhelmed with anxiety over the last 12 months [2]. In addition, from 2015 to 2019, there has been a 4% (from 9.3% to 13.3%) increase in students reporting that they have seriously considered suicide in the last 12 months [2,3]. However, the same report showed that only 24.3% of students had sought help for anxiety over the past 12 months and even fewer had sought help for depression (20%). Many college students who have mental health issues do not receive professional help because of the limited resources on college campuses, the time required to receive help, the lack of awareness of college mental health resources, and the stigma around receiving care [4-6]. Students who do not receive help have a higher incidence of dropping out before completing a degree [7,8] and experience long-term adverse outcomes, including low employment [9], perpetual emotional and physical health problems [10], and relationship dysfunction [11]. Consequently, many colleges and universities are seeking innovative ways to help students receive the help they need.

Mobile phones may prove to be a beneficial tool for providing mental health access and to overcome the barriers associated with receiving treatment. Mobile phones are ubiquitous among college students, with 95% of college students owning a smartphone [12]. College students and young adults are rarely without their phones, using them for several facets of their lives, generally through mobile apps. Some research suggests that college students would rather use in-person resources than web-based resources for mental health; however, college students are likely to use web-based resources because of their availability, convenience, and confidentiality [13-15]. Mobile apps for mental health services can provide users with additional benefits that seeking face-to-face help does not provide, such as its relatively low cost for care and the lack of stigma from seeking in-person treatment. In addition, mental health apps could be a helpful supplement to in-person care or as a first step in seeking in-person care.

Although there has been much research on mental health apps for college students, this qualitative study uses the technology acceptance model (TAM) as a framework to better understand college students' perceptions and attitudes toward the adoption and use of a mobile phone app, MySSP (Morneau Shepell), for mental health [16,17]. MySSP is a mobile app service purchased by universities and is a resource for college students to receive mental health help, including support from professional counselors. As MySSP is readily available to participants through their university, this study focuses on perceptions and attitudes toward the adoption of this app specifically. Focus groups were conducted with both undergraduate and graduate students to explore their perceptions and use of mental health apps, specifically MySSP. This study can also provide insights into this population's attitude of other apps that are developed to support mental health. In the next section *Literature Review*, we review the research related to mental health services via information and communications technologies (ICTs), including mobile phone apps for mental health for college students, followed by a brief review of the TAM. Finally, our research questions are presented.

Literature Review

Mobile Health for Mental Health

There has been an increasing number of ICTs that seek to help people with mental health issues, such as web-based psychological therapy, remote video consultation, and social media platforms. The use of these technologies "represent a cultural change in mental health care by empowering patients to exercise greater choice and control" [18]. These apps aim to help with a variety of mental health conditions such as anxiety, depression, and obsessive-compulsive disorder, to name a few. Many of these apps attempt to provide different functionalities, such as controlled breathing, positive thinking, and meditation [19,20]. A vast majority of these apps remain unstudied and have not been thoroughly tested, leaving little evidence for their proposed benefits [20-22]. However, a handful of apps that have been studied provide valuable insights into their usefulness. Clinical trials that have been conducted sought to test the effectiveness of mobile apps for mental health, and the findings suggest that the apps are more beneficial for those with low or moderate levels of depression but can improve symptoms of depression and anxiety [23-25]. Further research has demonstrated that many of these apps had limited and short-term use because of some acceptance barriers; however, the results suggest that some treatment is better than no treatment [22,26-29].

Mental Health Apps and College Students

When examining college students' openness and attitudes toward using mental health apps in a college setting, researchers found that 26.1% of college students were open to using mental health apps; however, only 7.3% had ever used such an app before [13]. The relatively low rate of mental health app adoption was attributed to participants having no current mental health needs, perceptions that mental health apps felt too impersonal, confidentiality concerns, and the utility of the apps [13,30]. Similar to the general population, college students have also tended to use mobile apps for mental health for short periods (≤4 weeks), although it is unknown if this short-term adoption can be attributed to feeling better, poor app design, student workload, or something else [13,14].

TAM

The TAM is one of the most widely used frameworks to examine the adoption of technology in a multitude of settings [16]. An extension of the theory of reasoned action [31], the TAM suggests that there are 2 key factors that predict the acceptance and use of a new technology, including perceived usefulness and perceived ease of use, both of which impact the attitude toward using a technology. Perceived usefulness is the users' perceptions of how well a technology will improve their current practices. The ease of use is defined as how easy the technology is to learn and use. Attitude refers to the general feeling of a user when implementing a technology into their everyday routines. The TAM has been used in a variety of health contexts, including among college students and in mental health mobile apps [32]. Furthermore, the TAM has been used in the analysis of qualitative data using a grounded theory approach [33], but the TAM has also been used to deductively inform the development of semistructured interview protocols [34,35].

In addition, social factors deeply affect an individual's attitude, and a multitude of theories posit that social influence is a key consideration in understanding adoption behaviors [31,36,37]. In the context of technology, the perception that the technology is accepted by one's peers has been demonstrated to be an important factor of adoption [38]. Therefore, the individual's attitudes, informed partially by their peers and partially by their evaluation of the technology (ie, ease of use and usefulness), will influence their overall intentions of acceptance of an app for mental health issues [39]. As college-aged adults have a greater communication with their peers about topics of personal importance such as health [38], we believe that social influence will be a salient factor for our population. Therefore, this study uses the TAM to deductively explore college students' perspectives of perceived usefulness and ease of use, related to MySSP, which are key to understanding their intentions to use the app.

MySSP App Description

MySSP is a mental health services app, developed by Morneau Shepell, which is purchased by universities for their students. In the summer of 2018, the student health system of the university where this study took place made a contract with the company to provide mental health services to the university's students. Once a contract is signed, the app is freely available to the university's students providing immediate and confidential support from professional counselors through Morneau Shepell through chat (texting), voice, or video. The counselors are trained about the university and are connected to the university's mental health staff. The app also provides text tips about mental health and well-being and provides a variety of informational articles that are relevant to the mental health and well-being of college students (eg, roommate problems, homesickness, stress around exams, etc). The app is available on the Google Play store and the Apple App Store. Figures 1-3 show app screenshots.

Figure 1. Student life screen of the MySSP app.

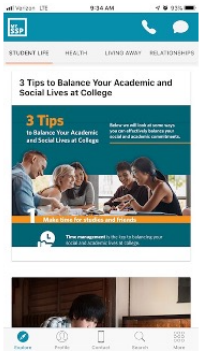


Figure 2. Health screen of the MySSP app.

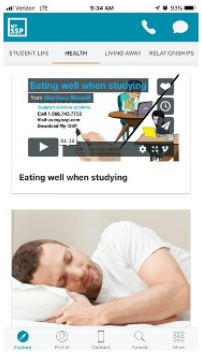
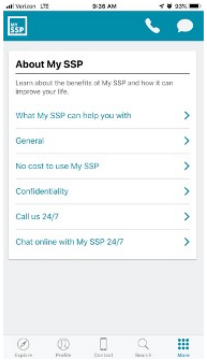


Figure 3. About screen for the MySSP app.



To promote the use of the MySSP app, the university sent out all-campus emails, promoted the service through several social media channels throughout the different departments and colleges, and posted flyers throughout the campus. Faculty members and academic advisors were given information about the resources and asked to share it with their students. In addition, students were told about the app if they called the campus counseling center.

This qualitative study aimed to further our understanding of college students' attitudes and perceptions of mental health apps, specifically MySSP. The TAM guided the framework of the focus groups to better understand students' intention to use the app. In addition, this formative research is a necessary first step in understanding how a campaign could be developed to further promote mental health awareness of campus resources. The research questions that guided this work included understanding the students' general awareness of apps for mental health, specifically, the MySSP app. Once the MySSP app was demonstrated, we sought to understand their overall perceptions of the app.

Methods

Data Source and Participants

A series of 9 focus groups were conducted with college students (N=30) at a large, public, Midwestern land-grant university. The focus groups took place between February and May 2019. To be eligible, participants had to be currently enrolled at the university and be aged older than 18 years. The university's institutional review board approved all portions of this study.

Participant Demographics

Most participants were aged between 18 and 22 years (27/30, 90%) and were identified as female (25/30, 83%). Most participants were undergraduates, and seniors (fourth and fifth year) made up 27% (8/30) of the participants. Most participants indicated their race as white (19/30, 63%), followed by Asian or Pacific Islander (9/30, 30%). Participants were also asked in the demographic survey if they had ever used mental health services; the majority (16/30, 55%) stated that they had never used mental health services. They were also asked if they had used mental health resources in the last 12 months, and 7 students reported that they had used mental health resources. Table 1 provides participant demographics.

Table 1. Participant demographics (N=30).

Variables	Frequency, n (%)
Age (years)	
18	1 (3)
19	9 (30)
20	6 (20)
21	7 (23)
22	4 (13)
30	1 (3)
33	1 (3)
41	1 (3)
Year in school	
First-year undergraduate	4 (13)
Second-year undergraduate	7 (23)
Third-year undergraduate	7 (23)
Fourth-year undergraduate	8 (27)
Graduate or professional student	4 (13)
Gender	
Male	5 (17)
Female	25 (83)
Ethnicity origin (or race)	
White	19 (63)
Black	1 (3)
Asian or Pacific Islander	9 (30)
American Indian	1 (3)
Grade point average	
4.0	6 (20)
3.5	14 (48)
3.0	9 (31)
Use of mental health services	
Yes	13 (45)
No	16 (55)
Use of mental health services (last 12 months)	
Yes	7 (24)
No	22 (76)

Procedure

Participants were recruited through the communication college's SONA system, a subject pool software. The participants were paid $15 to complete the focus group. Participants sat around a table equipped with a recording device, and there was a notetaker present. Sessions were audiotaped and lasted approximately 60 to 90 min. The moderator's guide was developed based on the TAM and included questions such as "Do you think that this app could provide mental health help/care?" and "What must this app have to be considered useful?" The session began with a short written demographic survey. Then, the focus group started with a brief discussion of general app use and was followed by a discussion of apps used for general health. This was done as an icebreaker for the sessions and to understand how general app use is similar to and different from apps for mental health. The majority of the session focused on apps used specifically for mental health, including their use and knowledge of the MySSP app. Following the discussion of apps for mental health, the moderator provided a brief overview of the MySSP app whereas the participants were given flyers, the informational email that had been sent

to all students, and several screenshots of the MySSP app to review. After approximately 3 to 5 min of reviewing the materials, the participants were asked about their perceptions of usefulness, ease of use, and attitudes toward the MySSP app.

Analysis

After each focus group, the notetaker and the moderator debriefed to develop a sense of several overarching themes. Data were analyzed using a descriptive analysis approach [40]. First, after the recordings of the focus groups were transcribed verbatim and returned to the research team, 3 members of the research team familiarized themselves with the data, reading, rereading, and making notes with initial ideas. This led the team to generate initial codes. We then iteratively developed a set of themes that captured the focus group dialog [41]. After the ninth focus group, the team concluded that saturation was reached. The themes were then presented to the whole team for clarification and feedback. Then, 2 members of the team coded a random selection (10%) of the transcripts to ensure intercoder reliability until reliability was reached (Cohen κ>0.8). Then, each coder independently coded half of the transcripts. Disagreements were identified and resolved by primary researchers. As a result of using the TAM to design the focus group protocol, the themes that emerged from the data were aligned with the TAM constructs. However, additional themes emerged from the analysis and are discussed in the Results section.

Results

The results are organized based on themes that emerged from the descriptive analysis of the focus group, which included the constructs of the TAM and the additions of awareness of the app and social influence.

Existing Awareness of the App

Although the app has been available to all students of the university and was announced through emails and flyers around campus starting in the fall of 2018 and continued through the spring of 2019 (when this study took place), very few students had even heard of the MySSP app. Of 30 participants, 7 (23%) had downloaded the MySSP app before attending the focus group. The remaining 23 participants (77%) had never heard of the MySSP app. During the focus group sessions, participants shared similar feedback about MySSP regardless of whether or not they had previously used the app.

Perceived Usefulness

Most participants perceived that the MySSP app would be a useful tool for themselves, their friends, and other students at the university. Participants noted that MySSP would be useful for themselves, particularly in situations related to academic stress. For example, one participant said:

> *I know my friends, none of them have the same major as me so they don't really get my stress that I get from my classes. So maybe in that case if...a student was very stressed out, they could go talk to somebody about it* [in the app]

Another participant felt that MySSP would be useful among their friend group, particularly the instant messaging feature that allows users to *text* a counselor. The participant stated:

> *I like that you can text because a lot of my friends...we're especially anxious about phone calls... texting is kind of stress free for us.*

The participants also described several situations in which the MySSP app would be useful for other students at the university. One participant commented that the app would be particularly helpful to new students who may be having trouble adjusting to living in a new place. The participant said:

> *I think it would be a good resource for freshmen, especially if you're from out of state or you're [an] international [student] and you don't really know anybody. Or if you're having problems making friends, if you're stressed out, or homesick.*

Participants also described scenarios in which a college student might need help navigating. For example, one participant said:

> *I feel like [MySSP] would also work out for people who are having relationship struggles...Maybe someone had a dramatic event happen...maybe they got a little too drunk at a party and they are trying to figure out and recollect what happened...*

There was, however, some skepticism regarding the long-term usefulness of MySSP. One participant felt that the app was not equipped to deal with more serious mental health issues, saying:

> *I feel like it can help in maybe a crisis or if you're feeling lost, but I don't [think]...in the long term, it can do a ton...I think this is way too general for if someone has clinical depression, anxiety, specific OCD...it's more for people who are realizing that they might need help.*

Another participant felt that MySSP would not have the same impact as in-person care in the long run but would still be useful in certain scenarios. This participant stated:

> *I think this mostly seems to be a gateway to something, I don't think it's going to be the same as a therapist or anything but for a lot of problems that students face around here like stress, breakups, relationship issues...this would really work well.*

Perceived Ease of Use

In general, health app use and retention varied among participants. Much of the feedback around short-term use is related to a lack of ease of use. For example, one participant noted how an app they used was too tedious to keep up with, saying:

> *I have this app downloaded but I haven't been keeping up with it...it's just really tedious because you have to really say how has your day been...you have to get all introspective.*

Other participants mentioned that their short-term use was because they were focused on a certain goal (ie, losing weight), and once they achieved the goal, they no longer found a need for that particular app.

After the MySSP app presentation, all participants perceived the app to be easy to use. One participant specifically mentioned the ease of communicating:

> *I think it sounds like it's quite easy...to get in touch with someone and there's a lot of different methods of communication you can use if phone calling is just not your thing, for example.*

Another participant echoed that perspective, saying:

> *I'd say that it's free to all students...and that it's anytime, anywhere that you can be in contact with someone [a counselor], so accessible.*

In addition, participants appreciated that the app was available in multiple languages. One participant said:

> *I like the fact it says you can access an advisor or counselor who speaks your preferred language and understands your culture. That's helpful for an international student.*

Attitudes Toward Apps for Mental Health and MySSP

As expected, the majority of the student participants used social media apps such as Facebook, Instagram, and Snapchat the most. In addition, many participants had experience using general health apps. However, very few of the participants used or had ever used an app for mental health. For example, one participant reported:

> *I haven't really tried them [mental health apps]. I don't know, sometimes my phone stresses me out so I don't want to be on that to try to relax.*

Many of the participants perceived that the app could be beneficial. For example, one participant appreciated that MySSP could provide her quick access to care, saying:

> *I've had interest in going out to find a therapist...but I feel like because we're all so busy all the time it's hard to go out of your way to just talk to somebody. I feel like having access through your phone is something that's really cool.*

Another participant noted that as she would not personally use MySSP for *her* mental health, others in more serious situations might use MySSP. She went on to say:

> *I feel like it just depends where someone is at the stage in their mental health...I wouldn't use it for my mental health. But then other people may be need it in a crisis or just having that person to talk to, so I feel it just depends or where you're at or what you need.*

Some participants felt that MySSP could address a spectrum of mental health needs. One participant said:

> *I feel like every student could benefit from it. People that maybe have several problems, or somebody that is just having a bad day and needs somebody to talk to.*

Participants also approved the appearance of the app, which impacted their perception of the quality of care. Regarding the quality, one participant shared:

> *I'd say it looks a lot better than I had imagined it. So, I think it looks very professional. So, I think the quality of help you would be getting would be more professional* [as well]

Overall, the majority of the participants had positive feelings about the app and that the university was looking for different ways to help service students' mental health issues.

However, some of the participants did not believe that an app was the best way to receive mental health services. For example, a handful of participants felt that they would turn to family or friends for mental health support before turning to an app. One participant said:

> *I can't really imagine that, because I think whatever problems I have, I can talk to my boyfriend first, or my sister, and also my friends.*

Some other participants felt that they did not have any mental health concerns that were *serious* enough to justify using a mental health app. When asked about their attitudes toward using a mental health app for themselves, one participant shared:

> *I just feel that I don't really need to [use a mental health app], I'm not in the state where I [would] download an app to take care of [my mental health]. Most of it is just like situational stress with test or quizzes or things like that. It's not reoccurring or continuing.*

Another participant expressed a similar sentiment saying:

> *I don't have [a] mental health disorder, but everyone gets stressed and I like the way most people cope with it is trying to take their mind off of it. For me, that's just listening to music or something. So, I don't need to go to the extent of doing that* [using a mental health app]

Social Influence

Participants were asked if they had ever downloaded an app because a friend or family member recommended it. Many participants had downloaded an app based on others' recommendations. One participant said:

> *If I'm talking about something interesting and they mention something [about an app], I don't really take a second thought, I'll say, "Oh yeah that sounds good." I'll add it to my phone without thinking about it.*

However, the apps that were recommended to participants were not mental health apps, rather games, photo editing, music streaming, or networking apps.

Researchers asked participants if they thought their friends would use the MySSP app. The majority of participants thought that if their friend needed mental health help and was aware of MySSP, then they would use it. One participant reiterated that MySSP is particularly useful in that their friend would be more likely to use the app than see a counselor face-to-face, saying:

> *I know my one friend struggled with depression and she's always just like, she wants help, but she doesn't want to walk. She's like I don't have time to walk to*

this place, or whatever, but since that's available 24/7, and it's just on your phone. You always have your phone on you, that would be really good.

Participants also mentioned their thoughts if they found that their good friend was using MySSP. Some responses included "Good that they're getting help" and "I'd feel happy for them that they were doing something about the problem they had."

Discussion

Principal Findings

Using the TAM, this study was an initial exploration of college students' perceptions, attitudes, and intentions of using a mobile app for mental health. The results indicate that most students were not aware that MySSP is an available campus resource for mental health issues. Although the app was perceived by the participants, on initial judgment, it was useful and easy to use. Overall, they had positive attitudes toward the app; however, most participants did not perceive that they needed MySSP because they did not have any mental health issues. In addition, social influence appears to be a key component in college students' use of apps. This study furthered our understanding of how TAM can be applied in a university setting for perceptions of mobile health for mental health, using a descriptive analysis to explain reasons for app adoption for mental health by college students.

As the results suggest, students were not aware of the app, which hinders the use and adoption of any intervention, not just this particular one. It is key that when an organization rolls out a resource, it effectively promotes the resource and finds champions who can share the information [42]. Past research shows that the students most likely to use these types of apps are generally women and those with lower levels of depression and anxiety [23,41]. However, male college students also face elevated rates of anxiety and depression. Feedback from the participants will be used to help inform the university on how to better promote mental health resources on campus. In addition, future research should consider how to connect with students who may need the most help.

The multiple modes available for communicating with a counselor, particularly through texting, were seen as key usefulness attributes of MySSP. Physically attending mental health counseling is often time consuming and can leave the student feeling stigmatized [43]. The more anonymous nature of the mediated communication via the app appeared to increase the perceived usefulness for many of the college student participants. In addition, communication mediated through an app is likely to increase rates of self-disclosure, which may positively impact long-term mental health outcomes [44]. Conversely, and similar to past research, participants indicated that the app may not be effective for longer-term use but rather as a tool to overcome the barriers to initial face-to-face counseling. The specific context in which an app for mental health is used is essential to college students' perceived usefulness. Further research should consider the context in which an app, such as MySSP, is most useful for students seeking mental health help.

Although some mental health apps offer a wider variety of features to users, this can increase the complexity of the app. The TAM suggests that apps that are perceived to be easier to use are more likely to be accepted by users [45]. Previous research has found that perceived usefulness and perceived ease of use affect young people's health app acceptance and effectiveness [17]. Participants noted that MySSP appeared to be easy, understandable, and accessible, which are all indications of an overall positive perception of ease of use. Goal setting was mentioned as a context in which long-term use may not apply; however, general health literature suggests that small, specific, and sustainable goals set and achieved over a period of time lead to positive long-term changes [46]. Therefore, with the help of a counselor, the app could implement goal setting in an effective way that actually increases use.

Many participants indicated that they did not feel they need an app for mental health, although the rates of college students who indicate facing issues of mental health are at historic highs [2]. The majority of college students reported feeling stressed on a regular basis. Previous research has indicated that this reoccurring stress can lead to anxiety and depression, in which college students are particularly susceptible. Furthermore, there appears to be a norm in which this stress, anxiety, and depression are accepted by college students. It is possible that the largest barrier to seeking help for mental health, whether face-to-face or via a mobile app, is because of the normative nature of these mental health issues during college. In addition, it is critical to note that when the students were discussing that they would likely not use MySSP, many of the comments were related to feeling like they did not need any professional mental health resources in any form. However, the students in this study did indicate that MySSP would likely help other students overcome the barriers frequently cited for not receiving face-to-face help. This indicates that factors other than barriers to seeking mental health help should be considered in future research.

Social influence is another key factor in establishing intention to use and the effectiveness of health apps [17]. The participants recalled many instances in which they downloaded an app because a friend or family member recommended it. Although none of the participants had experience with a mental health app being recommended to them, the power of peer influence in the uptake of mental health apps may be an avenue for increased use. In addition, participants reflected on how they would feel if their friend used an app for mental health. This feedback further confirmed that college students could benefit from an app for mental health and that the social attitudes around the use of mental health resources are generally positive.

Limitations

As with most research, this study also has some limitations. First, the population of the focus groups was mostly white women, which does not reflect the university's overall population. In addition, all of the participants had a grade point average of 3.0 or higher, indicating that they may be more highly motivated than the general student body. This is also indicated by their willingness to participate in the focus group. The students were also recruited through a communication college SONA system, which may not be representative of the student

body. In addition, because mental health issues are often stigmatized, we might not have been able to get everyone's true opinions regarding the issue. However, the potential use of the MySSP app was able to draw out more neutral experiences. Therefore, these results are still useful when developing a campaign promoting mental health resources, especially at larger public universities.

Our formative research into students' perceptions of app use for mental health provides a path to further explore this issue. A strength of this study is that there were screenshots of an existing app for students to look at, and they could download the app to use during the session. In addition, some students in the focus groups had prior experience using the app. This provided concrete examples rather than an *idea* or concept of what a mobile app for mental health might look like. Using our findings, we have developed and are currently surveying the population about their mental health and intentions of using this app as a resource. This should provide us with more rigorous data and additional results.

Conclusions

This study provides a deeper understanding of the perceptions of college students regarding a mobile app for mental health. The feedback from students will help the student health center promote the university's resources for mental health help and encourage their use. Although this study used the TAM to understand students' perceptions and attitudes toward a mental health app, further research is needed into the specific contexts in which an app for mental health will be most effective for college students. Furthermore, future work must identify why there is a gap between rates of depression and anxiety among college students and their intention to use mental health services.

Conflicts of Interest

None declared.

References

1. Lipson SK, Lattie EG, Eisenberg D. Increased rates of mental health service utilization by US college students: 10-year population-level trends (2007-2017). Psychiatr Serv 2019 Jan 1;70(1):60-63 [FREE Full text] [doi: 10.1176/appi.ps.201800332] [Medline: 30394183]
2. American College Health Association-National College Health Assessment II: Reference Group Executive Summary Spring 2019. American College Health Association (ACHA). 2019. URL: https://www.acha.org/documents/ncha/NCHA-II_SPRING_2019_US_REFERENCE_GROUP_EXECUTIVE_SUMMARY.pdf [accessed 2020-07-27]
3. American College Health Association-National College Health Assessment II: Undergraduate Student Reference Group Data Report Fall 2016. American College Health Association (ACHA). 2017. URL: https://www.acha.org/documents/ncha/NCHA-II_FALL_2016_UNDERGRADUATE_REFERENCE_GROUP_EXECUTIVE_SUMMARY.pdf [accessed 2020-07-27]
4. Pace K, Silk K, Nazione S, Fournier L, Collins-Eaglin J. Promoting mental health help-seeking behavior among first-year college students. Health Commun 2018 Feb;33(2):102-110. [doi: 10.1080/10410236.2016.1250065] [Medline: 27976923]
5. Chekroud SR, Gueorguieva R, Zheutlin AB, Paulus M, Krumholz HM, Krystal JH, et al. Association between physical exercise and mental health in 1·2 million individuals in the USA between 2011 and 2015: a cross-sectional study. Lancet Psychiatry 2018 Sep;5(9):739-746. [doi: 10.1016/S2215-0366(18)30227-X] [Medline: 30099000]
6. Eisenberg D, Gollust SE, Golberstein E, Hefner JL. Prevalence and correlates of depression, anxiety, and suicidality among university students. Am J Orthopsychiatry 2007 Oct;77(4):534-542. [doi: 10.1037/0002-9432.77.4.534] [Medline: 18194033]
7. Hartley MT. Examining the relationships between resilience, mental health, and academic persistence in undergraduate college students. J Am Coll Health 2011;59(7):596-604. [doi: 10.1080/07448481.2010.515632] [Medline: 21823954]
8. Kessler RC, Foster CL, Saunders WB, Stang PE. Social consequences of psychiatric disorders, I: educational attainment. Am J Psychiatry 1995 Jul;152(7):1026-1032. [doi: 10.1176/ajp.152.7.1026] [Medline: 7793438]
9. Bruffaerts R, Mortier P, Kiekens G, Auerbach RP, Cuijpers P, Demyttenaere K, et al. Mental health problems in college freshmen: prevalence and academic functioning. J Affect Disord 2018 Jan 1;225:97-103 [FREE Full text] [doi: 10.1016/j.jad.2017.07.044] [Medline: 28802728]
10. Scott KM, Lim C, Al-Hamzawi A, Alonso J, Bruffaerts R, Caldas-de-Almeida JM, et al. Association of mental disorders with subsequent chronic physical conditions: world mental health surveys from 17 countries. JAMA Psychiatry 2016 Feb;73(2):150-158 [FREE Full text] [doi: 10.1001/jamapsychiatry.2015.2688] [Medline: 26719969]
11. Kerr DC, Capaldi DM. Young men's intimate partner violence and relationship functioning: long-term outcomes associated with suicide attempt and aggression in adolescence. Psychol Med 2011 Apr;41(4):759-769 [FREE Full text] [doi: 10.1017/S0033291710001182] [Medline: 20540815]
12. Galanek JD, Gierdowski DC, Brooks DC. ECAR Study of Undergraduate Students and Information Technology, 2018. Educause Center for Analysis and Research. 2018. URL: https://library.educause.edu/~/media/files/library/2018/10/studentitstudy2018.pdf?la=en [accessed 2020-07-24]
13. Kern A, Hong V, Song J, Lipson SK, Eisenberg D. Mental health apps in a college setting: openness, usage, and attitudes. Mhealth 2018;4:20 [FREE Full text] [doi: 10.21037/mhealth.2018.06.01] [Medline: 30050916]

14. Levin ME, Stocke K, Pierce B, Levin C. Do college students use online self-help? A survey of intentions and use of mental health resources. J College Stud Psychother 2017 Oct 9;32(3):181-198. [doi: 10.1080/87568225.2017.1366283]
15. Horgan A, Sweeney J. Young students' use of the Internet for mental health information and support. J Psychiatr Ment Health Nurs 2010 Mar;17(2):117-123. [doi: 10.1111/j.1365-2850.2009.01497.x] [Medline: 20465756]
16. 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]
17. Russell E, Lloyd-Houldey A, Memon A, Yarker J. Factors influencing uptake and use of a new health information app for young people. Inf Technol J 2019 Jan 28;36(4):222-240. [doi: 10.1080/15228835.2018.1536911]
18. Hollis C, Morriss R, Martin J, Amani S, Cotton R, Denis M, et al. Technological innovations in mental healthcare: harnessing the digital revolution. Br J Psychiatry 2015 Apr;206(4):263-265. [doi: 10.1192/bjp.bp.113.142612] [Medline: 25833865]
19. ADAA Reviewed Mental Health Apps. Anxiety and Depression Association of America: ADAA. 2018. URL: https://adaa.org/finding-help/mobile-apps [accessed 2020-07-22]
20. 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]
21. Anthes E. Mental health: there's an app for that. Nature 2016 Apr 7;532(7597):20-23. [doi: 10.1038/532020a] [Medline: 27078548]
22. Seppälä J, de Vita I, Jämsä T, Miettunen J, Isohanni M, Rubinstein K, M-RESIST Group, et al. Mobile phone and wearable sensor-based mhealth approaches for psychiatric disorders and symptoms: systematic review. JMIR Ment Health 2019 Feb 20;6(2):e9819 [FREE Full text] [doi: 10.2196/mental.9819] [Medline: 30785404]
23. Areàn PA, Ly KH, Andersson G. Mobile technology for mental health assessment. Dialogues Clin Neurosci 2016 Jun;18(2):163-169 [FREE Full text] [Medline: 27489456]
24. Litvin M. How mHealth programmes can treat depression: a randomized controlled trial. Psychreg J Psychol 2019;3(2):65-78 [FREE Full text]
25. Bendtsen M, Müssener U, Linderoth C, Thomas K. A mobile health intervention for mental health promotion among university students: randomized controlled trial. JMIR Mhealth Uhealth 2020 Mar 20;8(3):e17208 [FREE Full text] [doi: 10.2196/17208] [Medline: 32196462]
26. Anastasiadou D, Folkvord F, Serrano-Troncoso E, Lupiañez-Villanueva F. Mobile health adoption in mental health: user experience of a mobile health app for patients with an eating disorder. JMIR Mhealth Uhealth 2019 May 31;7(6):e12920 [FREE Full text] [doi: 10.2196/12920] [Medline: 31199329]
27. Ben-Zeev D, Brenner CJ, Begale M, Duffecy J, Mohr DC, Mueser KT. Feasibility, acceptability, and preliminary efficacy of a smartphone intervention for schizophrenia. Schizophr Bull 2014 Nov;40(6):1244-1253 [FREE Full text] [doi: 10.1093/schbul/sbu033] [Medline: 24609454]
28. Saeb S, Zhang M, Karr CJ, Schueller SM, Corden ME, Kording KP, et al. Mobile phone sensor correlates of depressive symptom severity in daily-life behavior: an exploratory study. J Med Internet Res 2015 Jul 15;17(7):e175. [doi: 10.2196/jmir.4273]
29. Kumar S, Nilsen WJ, Abernethy A, Atienza A, Patrick K, Pavel M, et al. Mobile health technology evaluation: the mHealth evidence workshop. Am J Prev Med 2013 Aug;45(2):228-236 [FREE Full text] [doi: 10.1016/j.amepre.2013.03.017] [Medline: 23867031]
30. Kenny R, Dooley B, Fitzgerald A. Developing mental health mobile apps: exploring adolescents' perspectives. Health Informatics J 2016 Jun;22(2):265-275. [doi: 10.1177/1460458214555041] [Medline: 25385165]
31. Fishbein M, Ajzen I. Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research. Boston, USA: Addison-Wesley; 1977.
32. Anderson K, Burford O, Emmerton L. Mobile health apps to facilitate self-care: a qualitative study of user experiences. PLoS One 2016;11(5):e0156164 [FREE Full text] [doi: 10.1371/journal.pone.0156164] [Medline: 27214203]
33. Campbell JI, Aturinda I, Mwesigwa E, Burns B, Santorino D, Haberer JE, et al. The technology acceptance model for resource-limited settings (TAM-RIS): a novel framework for mobile health interventions targeted to low-literacy end-users in resource-limited settings. AIDS Behav 2017 Nov;21(11):3129-3140 [FREE Full text] [doi: 10.1007/s10461-017-1765-y] [Medline: 28421356]
34. Rogers EM. Diffusion of Innovations. Fifth Edition. New York, USA: Free Press; 2003.
35. Venkatesh V, Morris MG, Davis GB, Davis FD. User acceptance of information technology: toward a unified view. MIS Q 2003;27(3):425. [doi: 10.2307/30036540]
36. Hamari J, Koivisto J. Why do people use gamification services? Int J Inf Manag Sci 2015 Aug;35(4):419-431. [doi: 10.1016/j.ijinfomgt.2015.04.006]
37. 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]
38. Arnett J. Socialization in emerging adulthood: from the family to the wider world, from socialization to self-socialization. In: Grusec JE, Hastings PD, editors. Handbook of Socialization: Theory and Research. New York, USA: The Guilford Press; 2015.

39. Smith J. Qualitative Psychology: A Practical Guide to Research Methods. Thousand Oaks, CA: Sage Publications; 2015.
40. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
41. Toscos T, Carpenter M, Flanagan M, Kunjan K, Doebbeling BN. Identifying successful practices to overcome access to care challenges in community health centers: a 'positive deviance' approach. Health Serv Res Manag Epidemiol 2018;5:2333392817743406 [FREE Full text] [doi: 10.1177/2333392817743406] [Medline: 29552599]
42. Holtz B, Krein S. Understanding nurse perceptions of a newly implemented electronic medical record system. Inf Technol J 2011 Oct;29(4):247-262. [doi: 10.1080/15228835.2011.639931]
43. Corrigan PW, Druss BG, Perlick DA. The impact of mental illness stigma on seeking and participating in mental health care. Psychol Sci Public Interest 2014 Oct;15(2):37-70. [doi: 10.1177/1529100614531398] [Medline: 26171956]
44. WALTHER JB. Computer-mediated communication. Commun Res 2016 Jun 29;23(1):3-43. [doi: 10.1177/009365096023001001]
45. 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]
46. Strecher VJ, Seijts GH, Kok GJ, Latham GP, Glasgow R, DeVellis B, et al. Goal setting as a strategy for health behavior change. Health Educ Q 1995 May;22(2):190-200. [doi: 10.1177/109019819502200207] [Medline: 7622387]

Abbreviations

ICT: information and communications technology
TAM: technology acceptance model

Edited by J Torous; submitted 20.02.20; peer-reviewed by J Melcher, K Cohen; comments to author 26.03.20; revised version received 07.05.20; accepted 22.06.20; published 07.08.20.

Please cite as:
Holtz BE, McCarroll AM, Mitchell KM
Perceptions and Attitudes Toward a Mobile Phone App for Mental Health for College Students: Qualitative Focus Group Study
JMIR Form Res 2020;4(8):e18347
URL: https://formative.jmir.org/2020/8/e18347
doi:10.2196/18347
PMID:32667892

Original Paper

A Web-Delivered Acceptance and Commitment Therapy Intervention With Email Reminders to Enhance Subjective Well-Being and Encourage Engagement With Lifestyle Behavior Change in Health Care Staff: Randomized Cluster Feasibility Stud

Menna Brown[1], BSc, MSc; Nic Hooper[2], BSc, PhD; Phillip James[3], BSc, MRES, PhD; Darren Scott[3], BSc, MSc; Owen Bodger[1], BSc, PhD; Ann John[1], MD, MBBS

[1]Swansea University Medical School, Swansea University, Swansea, United Kingdom

[2]Department of Health and Social Sciences, University of the West of England, Bristol, United Kingdom

[3]Department of Computer Science, Swansea University, Swansea, United Kingdom

Corresponding Author:
Menna Brown, BSc, MSc
Swansea University Medical School
Swansea University
Singleton Park, Data Science Building, Floor 3
Swansea, SA2 8PP
United Kingdom
Phone: 44 1792 606312
Fax: 44 1792 606312
Email: menna.brown@swansea.ac.uk

Abstract

Background: Poor mental health and emotional well-being can negatively impact ability to engage in healthy lifestyle behavior change. Health care staff have higher rates of sickness and absence than other public sector staff, which has implications at both individual and societal levels. Individual efforts to self-manage health and well-being which add to the UK mental health prevention agenda need to be supported.

Objective: The objective of this study was to establish the feasibility and acceptability of the inclusion of a self-guided, automated, web-based acceptance and commitment therapy intervention in an existing health promotion program, to improve subjective well-being and encourage engagement with lifestyle behavior change.

Methods: For this 12-week, 4-armed, randomized controlled cluster feasibility study, we recruited participants offline and randomly allocated them to 1 of 3 intervention arms or control (no well-being intervention) using an automated web-based allocation procedure. Eligibility criteria were current health care staff in 1 Welsh health board, age≥18 years, ability to read English, and ability to provide consent. The primary researcher was blinded to cluster allocation. Feasibility outcomes were randomization procedure, acceptance of intervention, and adherence to and engagement with the wider program. We evaluated health and well-being data via self-assessment at 2 time points, registration and postintervention, using the 14-item Warwick-Edinburgh Mental Well-Being Scale, the 4-item Patient Health Questionnaire, and the 7-item Acceptance and Action Questionnaire—Revised.

Results: Of 124 participants who provided consent and were randomly allocated, 103 completed full registration and engaged with the program. Most participants (76/103) enrolled in at least one health behavior change module, and 43% (41/96) of those randomly allocated to an intervention arm enrolled in the well-being module. Adherence and engagement was low (7/103, 6.8%), but qualitative feedback was positive.

Conclusions: The procedure and randomization process proved feasible, and the addition of the well-being module proved acceptable to health care staff. However, participant engagement was limited, and no one completed the full 12-week program. User feedback should be used to develop the intervention to address poor engagement. Effectiveness should then be evaluated in a full-scale randomized controlled trial, which would be feasible with additional recruitment.

Trial Registration: International Standard Randomised Controlled Trial Number (ISRCTN) 50074817; http://www.isrctn.com/ISRCTN50074817

(JMIR Form Res 2020;4(8):e18586) doi:10.2196/18586

KEYWORDS

well-being; mental health; behavior change; acceptance and commitment therapy; web-delivered intervention; gamification; adherence and engagement; internet-based intervention

Introduction

Background

Poor mental health and emotional well-being underpin many physical diseases and unhealthy lifestyles and can negatively impact an individual's ability to engage in healthy lifestyle behavior change [1-3]. Recent recognition that positive mental health and emotional well-being are fundamental components of good health has given rise to a strong UK mental health prevention agenda. For example, the *Five Year Forward view for Mental Health* [4], the *Prevention Concordat for Better Mental Health* [5] and the National Health Service (NHS) *Long Term Plan* [6] saw local and national services come together (in 2019) at an unprecedented level in their commitment to addressing the mental health crisis. This development built on earlier publications from the UK Department of Health and Social Care, which outlined the importance of well-being and its role in health outcomes [7-9]. Economic analysis has further supported the case for greater investment in mental well-being [10]. This recent and sustained UK focus on well-being builds on global recognition that mental health and emotional well-being are a fundamental component of good health [11].

Mental Health: United Kingdom Picture

Staff sickness and absenteeism in the public sector is high, more so than that of the private sector [12]. Stress is the most commonly cited reason for absenteeism [13], and the associated economic cost is significant at an estimated £105.2 billion a year [14].

Further to their own mental health and well-being needs, health care staff are well placed to promote positive lifestyle behaviors to others, through effective role modelling [15]. Personal health behaviors are critical in establishing effective and confident health behaviors [16,17]. Frontline staff have daily contact with patients and the public and can exhibit health behaviors to be emulated by others. In the United Kingdom, the Nursing and Midwifery Council identified role modelling as a statutory requirement, stating that nurses must "take every opportunity to encourage health-promoting behaviour through education, role modelling and effective communication" [18]. Likewise, for UK medical professionals, the General Medical Council include clear expectations in *Outcomes for Graduates* (pg 23) [19]. Research findings have highlighted the need for ongoing support for health care workers to improve their own health and to fully realize their potential as credible role models and healthy-living advocates for the populations they serve [20-24]. Indeed, Public Health England recently launched Every Mind Matters [25], a website to support well-being and physical health. The proliferation of interest in this area serves to highlight the need for formal evaluation of these approaches.

Web-Based Approach

Global mental health prevention has incorporated diverse initiatives directed at different levels in society, such as individual, community, and societal level regulations. One key area, which has seen exponential growth, is that of web-delivered interventions. The cost-effective benefits of evidence-based, web-delivered therapies that improve mental health are well established [26-28]; however, poor adherence and engagement remain a significant factor that limits effectiveness [29,30] and application.

To address poor adherence and engagement, this study used a therapeutic approach associated with positive adherence: acceptance and commitment therapy (ACT) [31,32]. ACT is based on the principle of psychological flexibility, which involves accepting one's unwanted thoughts and feelings while moving toward personal values. Several successful web-based ACT intervention studies have been reported [33-40], for example, in the treatment of depression among smokers [37] and in the prevention of mental health problems among university students [38]. Earlier systematic reviews and meta-analyses have found web-delivered ACT to be effective for the management of depression [41], and others report its effectiveness in both group and individual settings [42-45].

Randomized Controlled Feasibility Study

This study will provide initial insight into the impact, acceptance, and feasibility of a web-delivered, multifaceted workplace lifestyle behavior change program, which incorporates an ACT-based well-being intervention to support staff mental health in the context of a preventive approach.

Objectives

The study objectives were to (1) determine whether the randomization procedure was feasible, (2) determine whether the inclusion of an ACT-based well-being intervention, within a web-based lifestyle behavior change program, was acceptable to health care staff, (3) determine whether the well-being intervention positively affected adherence and engagement to the wider program, and (4) explore the impact of additional intervention elements.

Methods

Trial Design

This was a 4-armed, cluster randomized controlled feasibility trial.

Trial consent, registration, and assignment to trial arm were automated. The principal researcher (MB) was blinded to cluster allocation throughout the trial; participants were not. A computer code written in Python randomly allocated each cluster to a trial arm using a built-in randomization function. No changes to the program were made during the trial period.

Trial Arms

The control group used Champions for Health, which consisted of 5 lifestyle behavior change modules: Quit Smoking, Drink Responsibly, Weight Optimization, Regular Exercise, and Eat Healthily.

Intervention 1 group used Champions for Health plus the ACT-based well-being module (ACTivate your Well-being).

Intervention 2 group used Champions for Health and ACTivate your Well-being, plus 5 premade well-being films (see Multimedia Appendix 1).

Intervention 3 group used Champions for Health and ACTivate your Well-being, plus a static social norm message (eg, "Other users like you have reduced their weight, on average to 75 kg.").

Ethics

The study received ethical approval from the College of the Human and Health Sciences and the College of Medicine Research Ethics Committee, Swansea University, Swansea, UK, and research and development approval from Abertawe Bro Morgannwg University Health Board Joint Study Review Committee 2017 as service evaluation. The trial is registered with the ISRCTN registry (50074817; Multimedia Appendix 2 [46])

Clusters

Staff from 1 health board in Wales, UK participated. A health board is an organizational and administrative unit consisting of hospitals, community clinics, and general practices (primary care). There are 7 in Wales.

We created 4 clusters based on key hospital and community sites within the participating health board: 3 clusters received the intervention and 1 did not (the control). Use of this trial design is common in health care contexts where cluster trials are an important methodology used to compare different ways of encouraging health behavior change [47,48].

Clusters. We selected this design for pragmatic reasons. The reasons were 3-fold. First, focus group discussions identified that participants who had taken part in earlier releases of the website had discussed its content with colleagues. As such it became evident that if we allocated participants at the individual level, they may discuss and share the content of the intervention with those not allocated to that trial arm. This approach is reported elsewhere [49,50]. Second, the clusters are natural groups of people, determined by their place of employment. Outcomes within naturally occurring clusters may tend to be more correlated than those across clusters; this is because individuals within a health board may have similar practices, arising from organizational culture and shared environment or demographic features that might influence the outcome [48]. Third, allocation by site location may support recruitment. Undertaking randomization after consent and baseline would introduce significant delay, which might have a negative effect on enrollment and engagement.

The Program: Champions for Health

Champions for Health, developed by Public Health Wales, comprised 5 lifestyle behavior change modules. Each included text and images and they were based on the health belief model [51], the theory of planned behavior [52], plan, do, study, act [53], and the self-regulatory model [54].

The Intervention: ACTivate your Well-being

Following a participatory design process [55-58], NHS staff (n=39), researchers, ACT practitioners, mental health experts, and computer scientists worked together to co-design the website and intervention through a series of exploratory interviews, focus groups, usability sessions, and a pilot evaluation.

We developed an automated, interactive, 12-week, self-guided intervention called ACTivate your Well-being for predetermined sequential release. Recommended time spent on each week was 20 minutes per day, 3 days a week. In addition to the structured modules, 3 pop-ups were available: Green Space gallery, Sleep, and Relaxation. These could be accessed freely.

Recruitment and Eligibility

All staff employed by the health board at the time of the study were eligible to participate. Eligibility criteria were current health care staff in 1 Welsh health board, age≥18 years, ability to read English, and ability to provide consent.

We recruited participants between January 28 and February 7, 2019 at 4 hospital sites. Electronic invitation and advertisements were displayed on the hospital intranet, electronic, and physical notice boards. Presentations were made at 4 site locations during induction training to a voluntary staff Well-Being Champions scheme.

Procedure

Participants registered individually via the website. All participants were required to provide consent using a checkbox process prior to completing a registration form, which asked for username, password, gender, age, location, self-rated health, self-rated work performance, sickness leave in the past 6 months, quantity of 1-week absences in the past 30 days, and self-assessed primary outcome measures. Users could opt to receive a semiautomated weekly email reminder, which included the website link and a motivational quote.

Once registered, participants accessed the website freely by logging in to their account. At this point, participants found out whether they had access to the well-being intervention. A personalized dashboard enabled participants to enroll in 1 or more modules and track their progress. We incorporated 2 gamification features to support sustained engagement [59]: Rewards and Feedback. Gamification is the use of game elements in non–game contexts [60]. Health points were rewarded for website engagement and converted into trophies at predetermined thresholds, and a bar showed progress toward each trophy. Feedback graphs were generated to show progress (Multimedia Appendix 3). In week 12 participants were reminded, via the home page and weekly email, to complete the time 2 outcome measures and provide feedback. At this point they were also invited to take part in a focus group. We sent 3 email reminders.

Outcome Measures

Primary Outcome Measure

The 14-item Warwick-Edinburgh Mental Well-Being Scale (WEMWBS) is a validated measure of mental well-being in the general population, responsive to change at both the individual and group levels [61-63]. A higher score represents more positive well-being. A score of 43.5 or less is considered a screening threshold for depression [62]. An increase of 2.77 or greater indicates statistically significant improvement [63].

The 4-item Patient Health Questionnaire (PHQ-4) Anxiety and Depression Scale [64] screened for anxiety and depression. On each subscale, a score of 3 (range 0-6) or greater is considered positive for screening purposes for anxiety and depression [64].

Process Measures

The 7-item Acceptance and Action Questionnaire—Revised (AAQ-II) is a validated, 1-factor measure of psychological inflexibility [65]. Higher scores (range 7-49) indicate greater levels of psychological inflexibility [65]. Cutoff points are not published for this measurement tool; however, a score of 17.5 and greater has been identified to indicate significant psychological inflexibility [66].

Sample Size Considerations

Feasibility trial designs do not commonly employ formal power calculations [67]. As such we aimed to recruit 100 participants, 25 in each trial arm, to allow comparison across groups and to explore the study objectives.

Data Analysis

Randomization Procedure

Functionality was assessed by the web developer during week 1 to ensure that self-reported location data were used accurately to populate trial arms.

Acceptability, Adherence, and Engagement

Descriptive statistics reported website registration and participant characteristics. We conducted statistical analysis using nonparametric methods in IBM SPSS version 26 (IBM Corporation) to explore baseline characteristics of each trial arm. Adherence was measured by completion of outcome measure at baseline and postintervention. Qualitative feedback, collected at the end of the program via a structured feedback form, interview, and focus group discussion, was audio recorded and transcribed verbatim. Data were analyzed using inductive thematic analysis [68].

Primary Outcome Effect

The feasibility study was not powered to test statistical significance, but merely to explore the impact and effectiveness of the intervention on lifestyle behavior change and well-being across trial arms.

Results

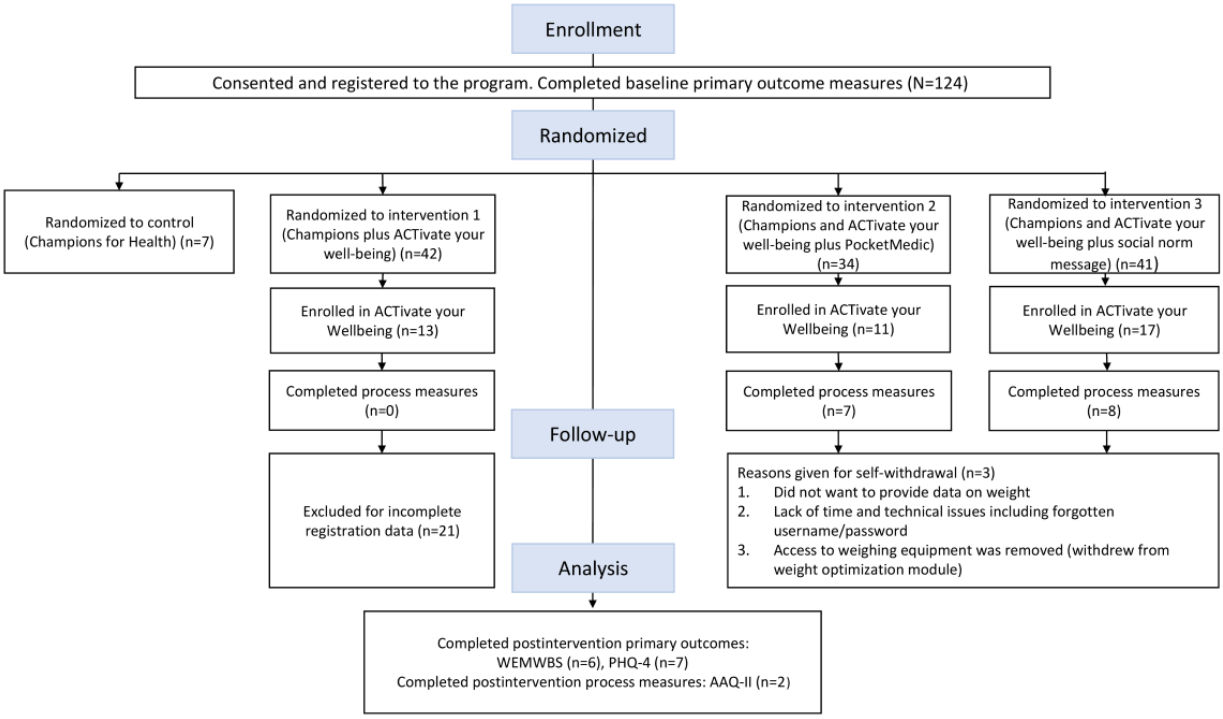
Randomization

The automated randomization procedure proved effective, and participants were allocated as expected.

Recruitment and Participant Summary

A total of 124 users consented to participate and were randomly allocated to a trial arm. Of these, 103 users provided baseline data for the primary outcome measures and were analyzed (Figure 1).

Figure 1. Consolidated Standards of Reporting Trials flow diagram. AAQ-II: Acceptance and Action Questionnaire—Revised; PHQ-4: Patient Health Questionnaire-4; WEMWBS: Warwick-Edinburgh Mental Well-Being Scale.



Users spanned all age brackets (18-65 years old). Most users (91/103, 88.4%) were female and rated their health as "good" to "excellent" (82/101, 81.2%) on a 5-point Likert scale. Over half (63/103, 61.2%) reported no days off work in the past 6 months and, in line with this, the majority (81/103, 78.6%) did not report any 1-week absences in the past month. On average, users rated their general work performance 8 on a scale of 1 to 10, with 10 being highest. All opted to receive email reminders (Table 1).

Table 1. Participant characteristics (n=103).

Characteristics	Trial arm			
	Control (Champions for Health)	Intervention 1 (Champions plus ACTivate your Well-being[a])	Intervention 2 (Champions and ACTivate plus PocketMedic)	Intervention 3 (Champions and ACTivate your Well-being plus social norm message)
Randomized, n	7	42	34	41
Analyzed, n	7	21[b]	34	41
Female, n (%)	6 (86)	19 (90)	30 (88)	38 (93)
Age bracket (years), n (%)				
18-25	0	0	4 (12)	2 (5)
26-35	2 (28)	8 (38)	10 (29)	10 (24)
36-45	3 (43)	5 (24)	11 (32)	12 (29)
46-55	1(14)	5 (24)	7 (20)	14 (32)
56-65	1 (14)	3 (14)	2 (6)	4 (10)
Self-reported days off work, mean (SD)	1.4 (2.1)	9.3 (18.4)	1.0 (1.5)	6.9 (25.3)
Self-rated general work performance score[c], mean (SD)	7.3 (1.6)	7.6 (1.7)	7.5 (1.3)	7.5 (1.7)

[a]ACTivate your Well-being intervention based on acceptance and commitment therapy.

[b]21 users were excluded based on incomplete registration.

[c]On a scale of 1-10, with 10 being the highest.

Trial Arms

We detected no significant differences between trial arms, using the Kruskal-Wallis test, at baseline (n=103) for registration week (*P*=.17), age (*P*=.51), self-rated health (*P*=.36), days off work (*P*=.84), absences of 1-week duration (*P*=.09), self-rated work performance (*P*=.94), WEMWBS (*P*=.24), PHQ-4 (*P*=.27), AAQ-II (*P*=.25) or gender (χ^2_3=103; *P*=.56).

We observed no difference for enrollment in module: Weight Optimization, χ^2_2=103, *P*=.38; Regular Exercise, χ^2_2=103, *P*=.25; Drink Responsibly, χ^2_3=103, *P*=.88; Eat Healthily, χ^2_3=103, *P*=.74; Quit Smoking, χ^2_3=103, *P*=.85). However, module engagement differed across intervention groups (χ^2_2=96, *P*=.40).

Enrollment (χ^2_2=96, *P*=.10) or engagement (as a binary variable) in the well-being intervention did not differ significantly between the 3 intervention groups (χ^2_2=96, *P*=.79) (Table 2).

Table 2. Module engagement (no. of participants) per trial arm (n=103).

Module	Trial arm			
	Control (Champions for Health)	Intervention 1 (Champions and ACTivate your Well-being[a])	Intervention 2 (Champions and ACTivate your Well-being plus PocketMedic)	Intervention 3 (Champions and ACTivate your Well-being plus social norm message)
Randomized, n	7	21	34	41
Enrolled, n (%)				
Champions for Health	5 (71)	13 (62)	29 (85)	27 (66)
ACTivate your Well-being	N/A[b]	13 (62)	11 (32)	17 (41)
Engaged, n (%)				
Champions for Health	2 (28)	0	9 (26)	9 (22)
ACTivate your Well-being	N/A	1 (5)	3 (9)	4 (10)

[a]ACTivate your Well-being intervention based on acceptance and commitment therapy.

Outcome Measures

Primary Outcome Measure

At baseline, participant well-being scores, measured using WEMWBS, was mean 46.3 (SD 8.8, range 29-68). This was lower than the general population. Only 6 participants completed WEMWBS postintervention, and this group recorded a higher score (mean 53.8, SD 3.1) and higher minimum score (43). When this subgroup was tested, 5 participants showed raised scores, which presented some evidence of an improvement, but this fell below the level of statistical significance (*P*=.11; Wilcoxon signed rank test).

Baseline anxiety and depression scores, measured using PHQ-4, were within the normal population range, with means of 1.8 (SD 1.6) and 1.3 (SD 1.5), respectively. A small subset of users met the screening criteria (20/103, 19.4%; 11/103, 10.7%, respectively). The combined PHQ-4 scores (n=7) did not change significantly postintervention (*P*=.34; Wilcoxon signed rank test).

Process Measures

The mean AAQ-II score was 21.7 (SD 9.8; n=22). However, no comparison could be drawn, as only 2 participants completed the postintervention questionnaire.

Acceptability, Adherence, and Engagement

Adherence to the study protocol was poor (7/103, 6.8%). However, the majority of participants (76/103, 73.8%) enrolled in at least one module. Almost half (50/103, 48.5%) enrolled in 1 module, 17 (16.5%) enrolled in 2 modules, 7 (6.8%) enrolled in 3 modules, and 2 (1.9%) enrolled in 4 modules. The most popular modules were well-being (41/96, 43%), Regular Exercise (40/103, 38.8%), and Weight Optimization (39/103, 37.9%).

Of the 9 participants who enrolled in Drink Responsibly, only 3 (33%) engaged. However, health outcomes improved for these active users; 1 reported a reduction in days per week that they consumed alcohol from 4 to 2, with a reduction from 20 drinks per week to 8. Another reduced their consumption from 4 to 3 days per week, with a reduction from 11 to 5 drinks, and the third increased their number of days drinking but their overall alcohol consumption reduced (from 18 to 4). Of the 23 participants (22.3%) who enrolled in Eat Healthily, only 6 (26%) engaged, 1 until week 6. User data indicated poor fruit and vegetable consumption, with few meeting recommended guidelines and some never consuming the recommended 5-a-day portion (5/23). Almost all who enrolled in Weight Optimization (37/39, 95%) engaged and provided an initial weight (mean 75.77, SD 15.80 kg). A small subsample (10/39, 26%) provided a second weight (mean 75, SD 10.89 kg). Of the 40 participants who enrolled in Regular Exercise, 33% (13/40) actively engaged. This module saw the longest sustained engagement of the modules, with activity recorded until week 9. Only 2 (2/103, 1.9%) participants enrolled in Quit Smoking and neither entered nor tracked any data. Well-being had the highest enrollment (41/96, 42%). However, of those who enrolled, only 7 (17%) completed any of the weekly “try-now” activities. This module had the longest sustained engagement overall, with 1 user active until week 10.

Qualitative Data

To fully explore acceptability, we collected a range of qualitative data. We facilitated 2 one-to-one interviews (25-minute duration) and 2 focus groups (77- and 73-minute durations, 4 participants each). A total of 15 participants contacted the principal researcher via email and 8 completed the feedback survey.

Qualitative feedback was positive (Table 3), and it was clear that staff welcomed the inclusion of the well-being intervention; indeed, people who had used the prior releases requested its inclusion. Analysis identified recommendations for future development.

Table 3. Sample quotes from participants' feedback and their recommendations.

Theme	Illustrative quotes and recommendations
Feedback on website and well-being module	"I liked that it would tell you what you should be doing to keep up with the NHS[a] recommendations and then how it compared that to Wales and the rest of the population. I liked all that information." [Interview]
	"I would have liked to set weekly goals." [Interview]
	"Easy to use, visually was nice." [Interview]
	"The health and fitness aspects were quite helpful, I was enrolled on the modules, health eating, weight management, and regular exercise. They were quite simple and straightforward and the information there was very useful." [Interview]
	"I really liked the PocketMedic. I looked at all the films on there." [Focus group]
	"I liked some aspects of the ACT[b] therapy and I found some aspects helpful. Certainly, I've suffered from intrusive thoughts and it's helpful to sort of just accept them rather than fighting against them." [Interview]
Nonadherence	"To me it was purely entering my weight, which I understand needed to be done...where I sit in my office everyone can see my screen clearly and that was why I was not happy to enter my weight." [Email]
	"Couldn't log in, then other priorities took over firefighting through work. It's a time factor thing." [Email]
	"I stopped doing the one about weight management, as when I started there were scales up in outpatients that I used to use and then they took them away." [Interview]
Participant recommendations	Provide an option to set weekly goals and the option to report whether this goal was achieved or not.
	Provide the option to return to the previous week to enter progress data.
	Display progress data per activity undertaken; for example, during week 1 you swam for a total of 80 minutes; you did yoga for 60 minutes, or more detailed track-your-progress options to facilitate competitive and personalized elements.
	Provide personalization of the profile area. Provide the option to edit data displayed on the screen specifically in reference to weight due to lack of privacy in the work setting or the option to hide private details such as weight.
	Streamline access to well-being exercises and activities.
	Include additional signposts to alternative sources of help.
	Incorporate opportunities to connect and interact with others.
	Improve the layout of AAQ-II[c].

[a]NHS: National Health Service.

[b]ACT: ACTivate your Well-being intervention based on acceptance and commitment therapy.

[c]AAQ-II: Acceptance and Action Questionnaire—Revised.

Discussion

Principal Findings

This study explored the feasibility and acceptability of the inclusion of an ACT-based well-being intervention, within an existing web-based, workplace lifestyle behavior change program for health care staff. The cluster design and automated randomization procedure proved feasible. Participants were successfully randomly allocated based on self-reported location, and the principal researcher remained blinded until postintervention. The new multifaceted program also proved acceptable to NHS staff. Recruitment was positive and participation rates compared equally with previous releases; for example, the 2015 campaign administered by Public Health Wales recruited 140 staff from 1 health board.

The proportion of users (43%) who selected to enroll in the well-being intervention highlighted that initiatives such as these are desirable and that well-being support is required in the workplace. Indeed, the qualitative interviews, focus groups, and feedback were positive, and this is encouraging. Staff who engaged with the program and well-being module enjoyed the resources and reported personal benefit. Feedback also highlighted specific ways to develop the intervention to address poor engagement from the staff perspective, which will be used

to support future development. However, the lack of engagement is still a concern. Only 7 participants were active in the try-now elements of the intervention, and engagement was equally poor across the modules. It is worth noting that poor engagement may have occurred as an artifact of what was measured. Specifically, the website recorded activity only for "track your progress:" it was not possible to assess log-on rate, time spent on each page, and general website activity. Therefore, it is possible that users engaged with other program elements. Indeed, the focus group and interview data suggested that the site overall was well used. In line with this, it is also encouraging that most participants enrolled in and engaged with at least one module, the most popular being Regular Exercise and Weight Optimization. Indeed, these were often selected in combination. Quit Smoking was the least popular, with only 2 users, despite local rates of smoking remaining at 22% [69].

Looking at the global picture, this study found health care staffs' well-being scores (WEMWBS) to be lower than the general population (data for England). Mean scores for men and women on this measure are 50.1 and 49.6 [61], respectively, compared with 46.3 in our study. In addition to this, a subset of participants had PHQ-4 scores associated with anxiety and depression. These findings add to the global picture, which suggests prevalence rates of common mental disorders, such as anxiety and depression, are high. In the United Kingdom, estimated population incidence rates are 4% to 10% [70]. Elsewhere, similar instances are reported. For example, lifetime disorder rates in Australia are reported to be 45% [71]. The individual and economic cost associated with common mental disorders is significant. Mental health problems constitute the largest single source of world economic burden, with an estimated global cost of £1.6 trillion [72]. In the United Kingdom, the estimated costs of mental health problems are £70 billion to £100 billion each year and account for 4.5% of gross domestic product [73].

Finally, in this study we were also interested in exploring the impact of the additional intervention elements, PocketMedic and social norm message. There was no significant difference in engagement across the 3 interventions. Qualitative feedback indicated that PocketMedic films were appreciated. Future exploration could collapse the additional elements into one. This is line with earlier findings [74].

Limitations

Several limitations must be acknowledged.

First, the small sample recruited for this feasibility study limited the statistical analysis undertaken. This was particularly relevant to the control arm, as participants were randomly allocated in a 1:3 ratio in favor of the intervention. While we undertook the same recruitment process at each site, only 7 members of staff registered from the location randomized to the control. However, the sample size is not dissimilar to other published studies reporting the effects of similar interventions [36]. In future, cluster size should be considered.

Second, intervention 1 had a high number of participants excluded from the analysis as a result of incomplete registration data. This further reduced the sample size and equality between trial arms. The registration process should include mandatory fields to avoid this issue in future iterations.

A third study limitation resulted from low adherence. Few participants completed the postintervention outcome measures. This limited exploration of intervention impact. In response it will be important to explore alternative ways to encourage self-reported completion at postintervention. One option is to track individual nonuse and request feedback within 1 week. This may go some way to improving adherence rates, as reasons for nonuse could be identified and resolved during the study period. Alternatively, the intervention could be shortened to support continued use. Another option might be to redesign engagement data collection points within the website. In this version, engagement monitoring was limited to the track-your-progress or try-now features, both of which are user initiated. No automated data were recorded. Feedback and interview data suggested that participants engaged at many additional time points, which was not captured in our analysis. This should be addressed. Equally, the use of rewards and feedback did not support sustained engagement. Future developments should consider use of additional or different gamification features, for example, avatars or social interaction [75], or guided support and structured feedback, which have been associated with better adherence.

Fourth, due to an error in the database, which has been amended, we were not able to examine health care worker role in relation to self-reported absenteeism.

Comparison With Prior Work

The inclusion of ACTivate your Well-being within the Champions for Health program created a multifaceted program free and easily accessible to a range of health care staff.

Limited research has explored the role of well-being interventions on lifestyle behavior change programs. To our knowledge, this study is one of the first to explicitly explore the additional benefit of an emotional well-being intervention on lifestyle behavior change. We identified 2 prior studies that incorporated a mental health intervention within a physical health promotion program [76,77]; however, neither explicitly evaluated the additional benefit of a well-being intervention on lifestyle behavior choices and neither used ACT.

The web-based interface used in this health care setting offered an opportunity to provide tailored support to public sector staff through convenient and accessible means. The inclusion of the emotional well-being intervention, in combination with the modules, is a significant step forward in terms of prevention and early intervention for self-management of positive health behaviors and builds on the UK mental health prevention agenda. The multifaceted program targeted both physical health behaviors and emotional well-being in 1 integrated program. This is the unique feature of this program.

Conclusion

This feasibility study was not powered to statistically assess the impact on physical health. A full-scale randomized controlled trial with wider-ranging recruitment methods and additional participant groups would likely support this. We are hopeful

the larger-scale trial will answer this question. Study participants who engaged with ACTivate your Well-being and Champions for Health reported positive feedback and made several useful recommendations to take this project forward.

Acknowledgments

This research was conducted independently with the support of a research grant from Healthcare Research Wales (SCS-14-11). The funders had no role in study design, data collection and analyses, decision to publish, or preparation of the manuscript.

Thank you to eHealth Digital Media Ltd, Swansea, UK, who kindly provided the PocketMedic well-being films.

Conflicts of Interest

None declared.

Multimedia Appendix 1
PocketMedic.
[PDF File (Adobe PDF File), 232 KB - formative_v4i8e18586_app1.pdf]

Multimedia Appendix 2
CONSORT-EHEALTH checklist V1.6.1.
[TXT File , 81 KB - formative_v4i8e18586_app2.txt]

Multimedia Appendix 3
Champions for Health website images and screenshots.
[PDF File (Adobe PDF File), 681 KB - formative_v4i8e18586_app3.pdf]

References

1. Mouchacca J, Abbott GR, Ball K. Associations between psychological stress, eating, physical activity, sedentary behaviours and body weight among women: a longitudinal study. BMC Public Health 2013 Sep 11;13:828 [FREE Full text] [doi: 10.1186/1471-2458-13-828] [Medline: 24020677]
2. McEwen BS. Central effects of stress hormones in health and disease: understanding the protective and damaging effects of stress and stress mediators. Eur J Pharmacol 2008 Apr 07;583(2-3):174-185 [FREE Full text] [doi: 10.1016/j.ejphar.2007.11.071] [Medline: 18282566]
3. Schneiderman N, Ironson G, Siegel SD. Stress and health: psychological, behavioral, and biological determinants. Annu Rev Clin Psychol 2005;1:607-628 [FREE Full text] [doi: 10.1146/annurev.clinpsy.1.102803.144141] [Medline: 17716101]
4. The Mental Health Task Force. The Five Year Forward View for Mental Health: A Report From the Independent Mental Health Taskforce to the NHS in England. 2016 Feb 1. URL: https://www.england.nhs.uk/wp-content/uploads/2016/02/Mental-Health-Taskforce-FYFV-final.pdf [accessed 2020-06-15]
5. Public Health England. Prevention Concordat for Better Mental Health: Prevention Planning Resource for Local Areas. 2017 Aug. URL: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/740587/Prevention_Concordat_for_Better_Mental_Health_Prevention_planning.pdf [accessed 2020-06-15]
6. NHS England, Department of health and Social Care. The NHS Long Term Plan. 2019 Jan 07. URL: https://www.longtermplan.nhs.uk/wp-content/uploads/2019/08/nhs-long-term-plan-version-1.2.pdf [accessed 2020-06-15]
7. De Feo D, Barrett J, Edwards J, Hurst M, Green J. Wellbeing: Why it Matters to Health Policy. London, UK: Department of Health; 2014. URL: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/277566/Narrative__January_2014_.pdf [accessed 2020-06-15]
8. Department of Health. A Compendium of Factsheets: Wellbeing Across the Life Course. The Relationship Between Wellbeing and Health. 2014 Jan. URL: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/295474/The_relationship_between_wellbeing_and_health.pdf [accessed 2020-06-15]
9. Department of Health. A Compendium of Factsheets: Wellbeing Across the Life Course. Wellbeing and Longevity. 2014 Jan. URL: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/277588/Wellbeing_and_Longevity.pdf [accessed 2020-06-15]
10. Department of Health and Social Care. No Health Without Mental Health: A Cross-Government Mental Health Outcomes Strategy for People of All Ages. Supporting Document- The Economic Case For Improving Efficiency and Quality in Mental Health. 2011 Feb 02. URL: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/215808/dh_123993.pdf [accessed 2020-06-15]
11. Prince M, Patel V, Saxena S, Maj M, Maselko J, Phillips MR, et al. No health without mental health. Lancet 2007 Sep 08;370(9590):859-877. [doi: 10.1016/S0140-6736(07)61238-0] [Medline: 17804063]

12. Office for National Statistics. Sickness absence falls to the lowest rate on record. 2018 Jul 30. URL: https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/employmentandemployeetypes/articles/sicknessabsencefallstothelowestratein24years/2018-07-30 [accessed 2020-06-22]
13. Carol B, Frost D. Health at Work – An Independent Review of Sickness Absence. 2011 Nov. URL: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/181060/health-at-work.pdf [accessed 2020-06-15]
14. Office for National Statistics. UK labor market: July 2016. Estimates of employment, unemployment, economic inactivity and other employment-related statistics for the UK. 2016 Jul 20. URL: https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/employmentandemployeetypes/bulletins/uklabourmarket/july2016 [accessed 2016-09-16]
15. Boorman S. NHS Health and Well-Being: Final report November 2009. Leeds, UK: The Health and Well-Being Review Team; 2009. URL: https://webarchive.nationalarchives.gov.uk/20130124052412/http:/www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/documents/digitalasset/dh_108907.pdf [accessed 2020-06-15]
16. Nursing and Midwifery Council. Standards for Pre-Registration Nursing Education. London, UK: Nursing and Midwifery Council; 2010 Sep 16. URL: https://www.nmc.org.uk/globalassets/sitedocuments/standards/nmc-standards-for-pre-registration-nursing-education.pdf [accessed 2020-06-15]
17. Vickers KS, Kircher KJ, Smith MD, Petersen LR, Rasmussen NH. Health behavior counseling in primary care: provider-reported rate and confidence. Fam Med 2007;39(10):730-735 [FREE Full text] [Medline: 17987416]
18. Nursing and Midwifery Council. Standards for Competence for Registered Nurses. London, UK: Nursing and Midwifery Council; 2010. URL: https://www.nmc.org.uk/globalassets/sitedocuments/standards/nmc-standards-for-competence-for-registered-nurses.pdf [accessed 2020-06-15]
19. General Medical Council. Outcomes for graduates. 2018 Jun. URL: https://www.gmc-uk.org/-/media/documents/outcomes-for-graduates-a4-6_pdf-78952372.pdf [accessed 2020-02-21]
20. Blake H, Patterson J. Paediatric nurses' attitudes towards the promotion of healthy eating. Br J Nurs 2015;24(2):108-112. [doi: 10.12968/bjon.2015.24.2.108] [Medline: 25615996]
21. Darch J, Baillie L, Gillison F. Nurses as role models in health promotion: a concept analysis. Br J Nurs 2017 Sep 28;26(17):982-988. [doi: 10.12968/bjon.2017.26.17.982] [Medline: 28956975]
22. Blake H, Harrison C. Health behaviours and attitudes towards being role models. Br J Nurs 2013;22(2):86-94. [doi: 10.12968/bjon.2013.22.2.86] [Medline: 23587891]
23. Lobelo F, de Quevedo IG. The evidence in support of physicians and health care providers as physical activity role models. Am J Lifestyle Med 2016 Jan;10(1):36-52 [FREE Full text] [doi: 10.1177/1559827613520120] [Medline: 26213523]
24. Oberg EB, Frank E. Physicians' health practices strongly influence patient health practices. J R Coll Physicians Edinb 2009 Dec;39(4):290-291 [FREE Full text] [doi: 10.4997/JRCPE.2009.422] [Medline: 21152462]
25. UK National Health Service. Every Mind Matters.: Public Health England; 2020. URL: https://www.nhs.uk/oneyou/every-mind-matters [accessed 2020-06-15]
26. Amstadter AB, Broman-Fulks J, Zinzow H, Ruggiero KJ, Cercone J. Internet-based interventions for traumatic stress-related mental health problems: a review and suggestion for future research. Clin Psychol Rev 2009 Jul;29(5):410-420 [FREE Full text] [doi: 10.1016/j.cpr.2009.04.001] [Medline: 19403215]
27. Morrell CJ, Sutcliffe P, Booth A, Stevens J, Scope A, Stevenson M, et al. A systematic review, evidence synthesis and meta-analysis of quantitative and qualitative studies evaluating the clinical effectiveness, the cost-effectiveness, safety and acceptability of interventions to prevent postnatal depression. Health Technol Assess 2016 May;20(37):1-414 [FREE Full text] [doi: 10.3310/hta20370] [Medline: 27184772]
28. Mohr DC, Burns MN, Schueller SM, Clarke G, Klinkman M. Behavioral intervention technologies: evidence review and recommendations for future research in mental health. Gen Hosp Psychiatry 2013 Aug;35(4):332-338 [FREE Full text] [doi: 10.1016/j.genhosppsych.2013.03.008] [Medline: 23664503]
29. Hilvert-Bruce Z, Rossouw PJ, Wong N, Sunderland M, Andrews G. Adherence as a determinant of effectiveness of internet cognitive behavioural therapy for anxiety and depressive disorders. Behav Res Ther 2012 Aug;50(7-8):463-468. [doi: 10.1016/j.brat.2012.04.001] [Medline: 22659155]
30. Donkin L, Christensen H, Naismith SL, Neal B, Hickie IB, Glozier N. A systematic review of the impact of adherence on the effectiveness of e-therapies. J Med Internet Res 2011;13(3):e52 [FREE Full text] [doi: 10.2196/jmir.1772] [Medline: 21821503]
31. Hayes SC. Acceptance and commitment therapy, relational frame theory, and the third wave of behavioral and cognitive therapies. Behav Ther 2016 Nov;47(6):869-885. [doi: 10.1016/j.beth.2016.11.006] [Medline: 27993338]
32. Hayes S, Barnes-Holmes D, Roche B. Relational Frame Theory: A Post-Skinnerian Account of Human Language and Cognition. New York, NY: Springer Science & Business Media; 2001.
33. Carlbring P, Hägglund M, Luthström A, Dahlin M, Kadowaki A, Vernmark K, et al. Internet-based behavioral activation and acceptance-based treatment for depression: a randomized controlled trial. J Affect Disord 2013 Jun;148(2-3):331-337. [doi: 10.1016/j.jad.2012.12.020] [Medline: 23357657]

34. Pots WTM, Fledderus M, Meulenbeek PAM, ten Klooster PM, Schreurs KMG, Bohlmeijer ET. Acceptance and commitment therapy as a web-based intervention for depressive symptoms: randomised controlled trial. Br J Psychiatry 2016 Jan;208(1):69-77. [doi: 10.1192/bjp.bp.114.146068] [Medline: 26250745]
35. Lappalainen P, Granlund A, Siltanen S, Ahonen S, Vitikainen M, Tolvanen A, et al. ACT Internet-based vs face-to-face? A randomized controlled trial of two ways to deliver acceptance and commitment therapy for depressive symptoms: an 18-month follow-up. Behav Res Ther 2014 Oct;61:43-54. [doi: 10.1016/j.brat.2014.07.006] [Medline: 25127179]
36. Lappalainen P, Langrial S, Oinas-Kukkonen H, Tolvanen A, Lappalainen R. Web-based acceptance and commitment therapy for depressive symptoms with minimal support: a randomized controlled trial. Behav Modif 2015 Aug 6. [doi: 10.1177/0145445515598142] [Medline: 26253644]
37. Jones HA, Heffner JL, Mercer L, Wyszynski CM, Vilardaga R, Bricker JB. Web-based acceptance and commitment therapy smoking cessation treatment for smokers with depressive symptoms. J Dual Diagn 2015;11(1):56-62 [FREE Full text] [doi: 10.1080/15504263.2014.992588] [Medline: 25671683]
38. Levin ME, Pistorello J, Seeley JR, Hayes SC. Feasibility of a prototype web-based acceptance and commitment therapy prevention program for college students. J Am Coll Health 2014;62(1):20-30 [FREE Full text] [doi: 10.1080/07448481.2013.843533] [Medline: 24313693]
39. Levin ME, Hayes SC, Pistorello J, Seeley JR. Web-based self-help for preventing mental health problems in universities: comparing acceptance and commitment training to mental health education. J Clin Psychol 2016 Mar;72(3):207-225. [doi: 10.1002/jclp.22254] [Medline: 26784010]
40. Trompetter HR, Bohlmeijer ET, Veehof MM, Schreurs KMG. Internet-based guided self-help intervention for chronic pain based on acceptance and commitment therapy: a randomized controlled trial. J Behav Med 2015 Feb;38(1):66-80. [doi: 10.1007/s10865-014-9579-0] [Medline: 24923259]
41. Brown M, Glendenning A, Hoon AE, John A. Effectiveness of web-delivered acceptance and commitment therapy in relation to mental health and well-being: a systematic review and meta-analysis. J Med Internet Res 2016 Aug 24;18(8):e221 [FREE Full text] [doi: 10.2196/jmir.6200] [Medline: 27558740]
42. Ruiz FJ. Acceptance and commitment therapy versus traditional cognitive behavioral therapy: a systematic review and meta-analysis of current empirical evidence. Int J Psychol Psychol Ther 2012;12(3):333-358 [FREE Full text]
43. Bohlmeijer ET, Fledderus M, Rokx TAJJ, Pieterse ME. Efficacy of an early intervention based on acceptance and commitment therapy for adults with depressive symptomatology: evaluation in a randomized controlled trial. Behav Res Ther 2011 Jan;49(1):62-67. [doi: 10.1016/j.brat.2010.10.003] [Medline: 21074752]
44. Fledderus M, Bohlmeijer ET, Pieterse ME, Schreurs KMG. Acceptance and commitment therapy as guided self-help for psychological distress and positive mental health: a randomized controlled trial. Psychol Med 2012 Mar;42(3):485-495. [doi: 10.1017/S0033291711001206] [Medline: 21740624]
45. Fledderus M, Bohlmeijer ET, Smit F, Westerhof GJ. Mental health promotion as a new goal in public mental health care: a randomized controlled trial of an intervention enhancing psychological flexibility. Am J Public Health 2010 Dec;100(12):2372. [doi: 10.2105/AJPH.2010.196196] [Medline: 20966360]
46. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. J Med Internet Res 2011;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]
47. Eldridge S, Kerry S, Torgerson DJ. Bias in identifying and recruiting participants in cluster randomised trials: what can be done? BMJ 2009 Oct 09;339:b4006. [doi: 10.1136/bmj.b4006] [Medline: 19819928]
48. Osrin D, Azad K, Fernandez A, Manandhar DS, Mwansambo CW, Tripathy P, et al. Ethical challenges in cluster randomized controlled trials: experiences from public health interventions in Africa and Asia. Bull World Health Organ 2009 Oct;87(10):772-779 [FREE Full text] [doi: 10.2471/blt.08.051060] [Medline: 19876544]
49. Peri K, Kerse N, Robinson E, Parsons M, Parsons J, Latham N. Does functionally based activity make a difference to health status and mobility? A randomised controlled trial in residential care facilities (The Promoting Independent Living Study; PILS). Age Ageing 2008 Jan;37(1):57-63. [doi: 10.1093/ageing/afm135] [Medline: 17965045]
50. Sedgwick P. Cluster randomized controlled trials: sample size calculations. BMJ 2013;346:f2839. [doi: 10.1136/bmj.f2839 NEW]
51. Rosenstock IM, Strecher VJ, Becker MH. Social learning theory and the health belief model. Health Educ Q 1988;15(2):175-183. [Medline: 3378902]
52. Godin G, Kok G. The theory of planned behavior: a review of its applications to health-related behaviors. Am J Health Promot 1996;11(2):87-98. [doi: 10.4278/0890-1171-11.2.87] [Medline: 10163601]
53. Coury J, Schneider JL, Rivelli JS, Petrik AF, Seibel E, D'Agostini B, et al. Applying the Plan-Do-Study-Act (PDSA) approach to a large pragmatic study involving safety net clinics. BMC Health Serv Res 2017 Jun 19;17(1):411 [FREE Full text] [doi: 10.1186/s12913-017-2364-3] [Medline: 28629348]
54. Leventhal H, Nenyamini Y, Brownlee S. Illness representations: theoretical foundations. In: Petrie KJ, Weinman JA, editors. Perceptions of Health and Illness. Amsterdam, Netherlands: Harwood Academic Publishers; 1997:1-18.
55. Spinuzzi C. The methodology of participatory design. Appl Theor 2005 May;52(2):163-174 [FREE Full text]

56. Kinzie MB, Cohn WF, Julian MF, Knaus WA. A user-centered model for web site design: needs assessment, user interface design, and rapid prototyping. J Am Med Inform Assoc 2002;9(4):320-330 [FREE Full text] [Medline: 12087113]
57. Rogers Y, Sharp H, Preece J. Interactive Design: Beyond Human Computer Interaction. Third edition. Hoboken, NJ: Wiley; Jun 26, 2011:9780470665763.
58. Lindsay S, Jackson D, Schofield G, Olivie P. Engaging older people using participatory design. 2012 Presented at: SIGCHI conference on human factors in computing systems; May 5-10, 2012; Austin, TX, USA. [doi: 10.1145/2207676.2208570]
59. Brown M, O'Neill N, van Woerden H, Eslambolchilar P, Jones M, John A. Gamification and adherence to web-based mental health interventions: a systematic review. JMIR Ment Health 2016;3(3):e39 [FREE Full text] [doi: 10.2196/mental.5710] [Medline: 27558893]
60. Deterding S, Dixon D, Khaled R, Nacke L. From game design elements to gamefulness: defining gamification. New York, NY: ACM; 2011 Presented at: 15th International Academic Mind Trek Conference: Envisioning Future Media Environments 2011; Sep 28-30, 2011; Tampere, Finland p. 9-15. [doi: 10.1145/2181037.2181040]
61. Tennant R, Hiller L, Fishwick R, Platt S, Joseph S, Weich S, et al. The Warwick-Edinburgh Mental Well-being Scale (WEMWBS): development and UK validation. Health Qual Life Outcomes 2007;5:63 [FREE Full text] [doi: 10.1186/1477-7525-5-63] [Medline: 18042300]
62. Maheswaran H, Weich S, Powell J, Stewart-Brown S. Evaluating the responsiveness of the Warwick Edinburgh Mental Well-Being Scale (WEMWBS): group and individual level analysis. Health Qual Life Outcomes 2012 Dec 27;10:156 [FREE Full text] [doi: 10.1186/1477-7525-10-156] [Medline: 23270465]
63. The Warwick-Edinburgh Mental Well-Being Scale (WEMWBS): development of clinical cut-off scores. URL: https://pdfs.semanticscholar.org/c32e/59006fbd769c7e595e23e5b20591de0ee018.pdf [accessed 2019-06-24]
64. Kroenke K, Spitzer RL, Williams JBW, Löwe B. An ultra-brief screening scale for anxiety and depression: the PHQ-4. Psychosomatics 2009;50(6):613-621. [doi: 10.1176/appi.psy.50.6.613] [Medline: 19996233]
65. Bond FW, Hayes SC, Baer RA, Carpenter KM, Guenole N, Orcutt HK, et al. Preliminary psychometric properties of the Acceptance and Action Questionnaire-II: a revised measure of psychological inflexibility and experiential avoidance. Behav Ther 2011 Dec;42(4):676-688. [doi: 10.1016/j.beth.2011.03.007] [Medline: 22035996]
66. Shari NI, Zainal NZ, Guan NC, Ahmad Sabki Z, Yahaya NA. Psychometric properties of the acceptance and action questionnaire (AAQ II) Malay version in cancer patients. PLoS One 2019;14(2):e0212788 [FREE Full text] [doi: 10.1371/journal.pone.0212788] [Medline: 30807594]
67. Thabane L, Ma J, Chu R, Cheng J, Ismaila A, Rios LP, et al. A tutorial on pilot studies: the what, why and how. BMC Med Res Methodol 2010;10:1 [FREE Full text] [doi: 10.1186/1471-2288-10-1] [Medline: 20053272]
68. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
69. Cosh H, Davies G, Francis I, Griffiths E, May L, Roberts C, et al. Tobacco and Health in Wales. Cardiff, UK: Public Health Wales Observatory, Welsh Government; 2012 Jun. URL: http://www2.nphs.wales.nhs.uk:8080/PubHObservatoryProjDocs.nsf/85c50756737f79ac80256f2700534ea3/509486bfd300fdef80257a29003c3c67/$FILE/Eng%20Smoking%20Report%20LowRes.pdf [accessed 2020-06-15]
70. NHS Digital. National Study of Health and Wellbeing. 2016. URL: https://digital.nhs.uk/data-and-information/areas-of-interest/public-health/national-study-of-health-and-wellbeing [accessed 2020-06-15]
71. Australian Institute of Health and Welfare. Mental health services in Australia. Canberra, Australia: Australian Institute of Health and Welfare; 2016. URL: https://www.aihw.gov.au/reports/mental-health-services/mental-health-services-in-australia [accessed 2020-06-15]
72. McManus S, Meltzer H, Brugha T, Bebbington P, Jenkins R. Adult Psychiatric Morbidity in England 2007: Results of a Household Survey. London, UK: NHS Information Centre for Health and Social Care; 2009.
73. McCrone P, Dhanasiri S, Patel A, Knapp M, Lawton-Smith S. Paying the Price: The Cost of Mental Health Care in England to 2026. London, UK: Kings Fund; 2008. URL: https://www.kingsfund.org.uk/sites/default/files/Paying-the-Price-the-cost-of-mental-health-care-England-2026-McCrone-Dhanasiri-Patel-Knapp-Lawton-Smith-Kings-Fund-May-2008_0.pdf [accessed 2020-06-15]
74. Kelders SM, Bohlmeijer ET, Van Gemert-Pijnen JE. Participants, usage, and use patterns of a web-based intervention for the prevention of depression within a randomized controlled trial. J Med Internet Res 2013;15(8):e172 [FREE Full text] [doi: 10.2196/jmir.2258] [Medline: 23963284]
75. Doherty G, Coyle D, Sharry J. Engagement with online mental health interventions: an exploratory clinical study of a treatment for depression. New York, NY: Association for Computing Machinery; 2012 Presented at: SIGCHI Conference on Human Factors in Computing Systems; May 2012; Austin, TX, USA p. 1421-1430. [doi: 10.1145/2207676.2208602]
76. Cook RF, Billings DW, Hersch RK, Back AS, Hendrickson A. A field test of a web-based workplace health promotion program to improve dietary practices, reduce stress, and increase physical activity: randomized controlled trial. J Med Internet Res 2007;9(2):e17 [FREE Full text] [doi: 10.2196/jmir.9.2.e17] [Medline: 17581811]
77. Cobb NK, Poirier J. Effectiveness of a multimodal online well-being intervention: a randomized controlled trial. Am J Prev Med 2014 Jan;46(1):41-48 [FREE Full text] [doi: 10.1016/j.amepre.2013.08.018] [Medline: 24355670]

Abbreviations

AAQ-II: Acceptance and Action Questionnaire—Revised
ACT: acceptance and commitment therapy
NHS: National Health Service
PHQ-4: 4-item Patient Health Questionnaire
WEMWBS: Warwick-Edinburgh Mental Well-Being Scale

Edited by J Torous, G Eysenbach; submitted 06.03.20; peer-reviewed by P Lappalainen, J Vermeir; comments to author 08.04.20; revised version received 29.04.20; accepted 26.05.20; published 07.08.20.

Please cite as:
Brown M, Hooper N, James P, Scott D, Bodger O, John A
A Web-Delivered Acceptance and Commitment Therapy Intervention With Email Reminders to Enhance Subjective Well-Being and Encourage Engagement With Lifestyle Behavior Change in Health Care Staff: Randomized Cluster Feasibility Stud
JMIR Form Res 2020;4(8):e18586
URL: https://formative.jmir.org/2020/8/e18586
doi:10.2196/18586
PMID:32763887

Original Paper

The Relationship Between Smartphone-Recorded Environmental Audio and Symptomatology of Anxiety and Depression: Exploratory Study

Daniel Di Matteo[1], MASc; Kathryn Fotinos[2], BSc (Hons); Sachinthya Lokuge[2], BSc (Hons); Julia Yu[2], MSc; Tia Sternat[2,3], MPsy; Martin A Katzman[2,3,4,5], MD, FRCPC; Jonathan Rose[1], PhD, FRSC

[1]The Centre for Automation of Medicine, The Edward S Rogers Sr Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON, Canada

[2]Stress Trauma Anxiety Rehabilitation Treatment Clinic for Mood and Anxiety Disorders, Toronto, ON, Canada

[3]Department of Psychology, Adler Graduate Professional School, Toronto, ON, Canada

[4]Department of Psychology, Lakehead University, Thunder Bay, ON, Canada

[5]The Northern Ontario School of Medicine, Thunder Bay, ON, Canada

Corresponding Author:
Daniel Di Matteo, MASc
The Centre for Automation of Medicine
The Edward S Rogers Sr Department of Electrical and Computer Engineering
University of Toronto
DL Pratt Building
6 King's College Road
Toronto, ON, M5S 3H5
Canada
Phone: 1 416 978 6992
Email: dandm@ece.utoronto.ca

Abstract

Background: *Objective* and *continuous* severity measures of anxiety and depression are highly valuable and would have many applications in psychiatry and psychology. A collective source of data for objective measures are the sensors in a person's smartphone, and a particularly rich source is the microphone that can be used to sample the audio environment. This may give broad insight into activity, sleep, and social interaction, which may be associated with quality of life and severity of anxiety and depression.

Objective: This study aimed to explore the properties of passively recorded environmental audio from a subject's smartphone to find potential correlates of symptom severity of social anxiety disorder, generalized anxiety disorder, depression, and general impairment.

Methods: An Android app was designed, together with a centralized server system, to collect periodic measurements of the volume of sounds in the environment and to detect the presence or absence of English-speaking voices. Subjects were recruited into a 2-week observational study during which the app was run on their personal smartphone to collect audio data. Subjects also completed self-report severity measures of social anxiety disorder, generalized anxiety disorder, depression, and functional impairment. Participants were 112 Canadian adults from a nonclinical population. High-level features were extracted from the environmental audio of 84 participants with sufficient data, and correlations were measured between the 4 audio features and the 4 self-report measures.

Results: The regularity in daily patterns of activity and inactivity inferred from the environmental audio volume was correlated with the severity of depression (r=−0.37; P<.001). A measure of sleep disturbance inferred from the environmental audio volume was also correlated with the severity of depression (r=0.23; P=.03). A proxy measure of social interaction based on the detection of speaking voices in the environmental audio was correlated with depression (r=−0.37; P<.001) and functional impairment (r=−0.29; P=.01). None of the 4 environmental audio-based features tested showed significant correlations with the measures of generalized anxiety or social anxiety.

Conclusions: In this study group, the environmental audio was shown to contain signals that were associated with the severity of depression and functional impairment. Associations with the severity of social anxiety disorder and generalized anxiety disorder

were much weaker in comparison and not statistically significant at the 5% significance level. This work also confirmed previous work showing that the presence of voices is associated with depression. Furthermore, this study suggests that sparsely sampled audio volume could provide potentially relevant insight into subjects' mental health.

(JMIR Form Res 2020;4(8):e18751) doi:10.2196/18751

KEYWORDS

depression; anxiety; mobile phone; ecological momentary assessment; mobile apps; mobile health; digital signal processing; acoustics; speech recognition software

Introduction

Background

Depression and anxiety disorders are some of the most prevalent mental health disorders [1], yet access to services and treatment for these disorders is lacking. It is common for many Ontario residents with mental health problems to wait for 6 months to 1 year for treatment [2]. Automation of a part of the mental health care process may help address this service gap in our health care system, as automated assessment could alleviate some of the workload that is currently being carried by health care workers.

The health care process can be modeled as beginning with assessment and measurement, followed by diagnosis, and finally treatment. Subsequent rounds of measurement or assessment occur with the final goal of achieving remission. This work focuses on the measurement and diagnosis components by working toward building an automated and objective severity measurement of anxiety, depression, and functional impairment associated with poor mental health.

Research in both psychiatry and clinical psychology traditionally involves assessments of subjects' state (eg, mood and behavior) in clinical or research settings where they are removed from their natural home and living environment. Often, these assessments were performed retrospectively, where the subjects were asked to recollect behaviors and feelings over several weeks in the past. Ecological momentary assessment (EMA) [3] is an alternative approach that endeavors to assess subjects' mood and behaviors in a naturalistic setting. It occurs in real time and has a higher frequency of measurement, possibly multiple times a day. These assessments are frequent and occur in subjects' natural setting(s), removing clinicians (and the potential for bias), from the measurement process. Furthermore, EMA is not limited to self-report data but can also use data collected from sensors (physiological sensors and smartphone sensors). Sensor-based data are especially interesting in an EMA context because they can be collected *passively* without any interaction from the study subject, essentially addressing concerns of self-report biases surrounding self-reported data [4,5]. Researchers now use the term *passive EMA* [6] to refer to EMA systems in which sensor-based data are collected without any interaction from the user. It is also sometimes referred to as *unobtrusive* EMA or *passive sensing*.

The feasibility of building passive EMA systems has been greatly improved by the smartphone revolution. Smartphones are ubiquitous and affordable consumer electronics, which are equipped with a wide range of sensors that can enable the type of sensing or monitoring necessary to perform passive EMA [7]. A brief survey of some of the existing work using passive EMA in the mental health space follows, including a more detailed review of studies that have used ecological audio data to predict mental health state.

Previous Work

A general methodology in these passive EMA or mobile sensing studies, and one that is used in this work, is to compute metrics, or *features*, from objective data sources that are designed to capture behaviors or traits of subjects that are known or suspected to be predictive of mental state. These features condense a large number of data points from a data source (eg, thousands of GPS coordinates measured over weeks of a study) into a single metric of behavior. This metric of behavior can then be tested for correlation with clinical measures of subjects' mental state. One such example of a GPS location–derived feature is the proportion of time a subject spends outside home. A low proportion might be indicative of avoidance behavior or low energy and, therefore, relevant to depression, for example.

A systematic review by Rohani et al [8] examined correlations between passively sensed smartphone data and symptoms of depression. *Homestay*, the proportion of time spent by the subject at home (computed from GPS data), and *screen active duration*, the proportion of time spent using the phone were 2 of the most strongly correlated features with depression. These features were reported as significant by numerous studies included in the review [8].

This general methodology of sampling objective data from subjects' smartphones (or other digital sensors) to infer health characteristics has been used in numerous studies, across many conditions. Although we will provide a focused review of relevant works that have used audio data to predict or measure mood and anxiety disorders, there is a wealth of research that has looked at using many different data sources to investigate, predict, or measure the severity of many characteristics of health and mental health disorders. Interested readers are directed to work that has investigated subjects' general mood and mental health [9-15], substance abuse [16,17], depression [18-24], bipolar disorder [25-29], anxiety disorders [30-32], and schizophrenia [33,34]. The most commonly used sources of passively collected smartphone data in these works include subjects' geolocation (ie, GPS data), screen activity and phone usage time, SMS and phone metadata, and physical activity and motion sensor data.

As all smartphones are equipped with microphones, they can be used to detect audio-based features of a subject's environment. Several works have investigated the recording

and analysis of speech audio from subjects' smartphones. There are different strategies to record audio, ranging from (1) actively prompting users to speak into a microphone, (2) passively recording subjects' phone calls, and (3) to passively recording environmental audio with no interaction from the user.

Using the active prompt-style methodology, Dickerson et al [35] conducted a study of depression in which subjects were asked twice daily to respond verbally to a prompt in free-form speech recorded by a microphone, yielding responses that were, on average, 1 to 2 min long. These audio recordings of prompted speech were then analyzed to produce 2 features: the fundamental frequency of subjects' speaking voices (F_0) and subjects' speech pause time. These 2 features were used to build a linear model for predicting the mood of the subjects. Mood was measured on a 1 to 10 scale (on the continuum of extremely depressed to extremely elevated mood), and the linear model was able to predict mood scores with a residual error of 0.092 (12 degrees of freedom, *P*=.011) [35]. Similarly, Guidi et al [36] investigated the fundamental frequency of speech (F_0) actively recorded from subjects via prompts in a study of bipolar disorder. This study was able to distinguish between individual bipolar subject's mood states. A total of 7 features extracted from speech audio were subjected to Kruskal-Wallis tests, and all features showed significant differences (at a 5% significance level) across angry, neutral, bored, and happy emotional states.

Using the passive phone call recording–style methodology, Faurholt-Jepsen et al [29] conducted a study of 28 outpatient subjects with bipolar disorder. Voice features produced from patients' phone calls were used to build 2 classification models that classified patients' states as manic or mixed versus euthymic (area under the curve=0.89) and depressive versus euthymic (area under the curve=0.78). Another study of bipolar subjects conducted by Grünerbl et al [25] used features of subjects' voices produced from the recordings of phone calls to predict mood state with 70% accuracy.

Finally, audio can be sampled in a much more passive and pervasive manner by using a smartphone's microphone to record environmental (ambient) audio. The StudentLife study by Wang et al [37] sampled ambient audio and used audio analysis techniques to detect the presence of human voices in the environment as a proxy measure of conversation frequency. They found that conversation frequency has a significant negative correlation with self-reported measures of depression severity [37]. The work by Abdullah et al [26] used the same approach in a study of 7 subjects with bipolar disorder. The conversation frequency feature was found to be weakly correlated with mood patterns (r=0.16; P=.06) as measured by the social rhythm metric. Ben-Zeev et al [10] used the amount of time proximal to human speech as a predictive feature in a study of general mental health. In a functional regression analysis, the amount of time spent proximal to human voices was found to be significantly associated with changes in a self-reported measure of depressive symptoms over the course of the study (P=.048).

Goal of This Study

This exploratory study seeks to discover potential correlates of anxiety and depression symptomatology from environmental audio acquired using passive smartphone sensing. Although previous research has studied how some features of the audio environment relate to depression and bipolar mood disorders, we will extend this to include anxiety disorders. In addition, one aspect of our study is the exploration of the sampled *average volume of the environment* over time, which has privacy-preserving attributes. We are not aware of any other study of mental health that makes use of this measurement. We hypothesize that the time series of the volume of subjects' environments reveals some qualities and characteristics of their daily activities, which are associated with symptoms of depression and anxiety, a hypothesis that we believe has not been investigated in the literature. In addition, we explore the effect of the presence of voices in the sampled audio.

Methods

Overview

Subjects from a nonclinical population were recruited for a 2-week observational study in which a custom app was installed on their personal Android phone. Self-report measures of anxiety, depression, and general quality of life and impairment were collected at the beginning and end of the study. Throughout the duration of the study, the smartphone app passively collected the average volume of environmental audio and the presence of voice activity (whether or not speech was detected in the environment at the time of recording). A set of features was designed and used to extract higher-level information from this set of data, and a statistical analysis was performed to determine if a significant relationship existed between subjects' self-reported anxiety, depression, and general impairment and these features. The study was approved by the University of Toronto Health Sciences Research Ethics Board (Protocol #36687).

Recruitment

Subjects were recruited from Prolific [38], a web-based platform for recruiting study participants. Prolific maintains an active pool of subjects who wish to engage in research activities and enables researchers to deploy web-based tasks to subjects with specified demographics. It is similar to other services, such as Amazon's Mechanical Turk [39], but has some properties that make it more attractive to academic researchers. These include ethical payment requirements and a comprehensive database of subjects' demographic data to enable prescreening.

The study inclusion criteria were as follows: subjects should (1) reside in Canada, (2) be fluent in English, (3) own an Android phone, (4) have completed at least 95% of their previous Prolific studies successfully, and (5) have previously participated in at least 20 Prolific studies. The final criterion was used to ensure that subjects were proficient at using the Prolific system and were generally technology literate. There were no exclusion criteria for the study. Subjects were paid Can $18.50 (US $14) for participating in the study.

Study Procedure

Members of the Prolific community who met the inclusion criteria could read a description of the study, which included an informed consent guide. Those who consented to the study were then directed to a webpage that acted as the study entry point. This website directed subjects to install the app from the Google Play app store and provided them with log-in credentials for using the study app. Once installed, the study app guided subjects through a short setup, where they were asked to provide the app with the necessary permissions to access their data, followed by a log-in. Immediately following setup and log-in, subjects were asked to complete a set of 4 self-report measures in digital form within the study app. At this point, following the completion of the self-report measures, the app began to periodically collect data in the background. No further actions or interactions with the study app were performed until the end of the study, exactly 14 days later, at the same time of day as the app installation/self-report work. At this time, subjects received a notification on their phone, informing them that the study had ended and requesting that they complete the same set of 4 self-report measures done at the beginning, again in the smartphone app. Following completion of this task, subjects were directed to uninstall the app from their phone and mark their study tasks as complete on the Prolific website. Subjects were then paid through Prolific's payment system.

Self-Report Measures

Subjects completed 4 self-report measures in digital form within the study app at the beginning and end of the 14-day study. A review by Belisario et al [40] found that self-administered survey scores do not differ when deployed by app versus other delivery modes. The 4 measures were the Liebowitz Social Anxiety Scale (LSAS) [41], the 7-item Generalized Anxiety Disorder Scale (GAD-7) [42], the 8-item Patient Health Questionnaire Scale (PHQ-8) [43], and the Sheehan Disability Scale (SDS) [44]. The LSAS is a 24-item scale used to assess the symptoms of social anxiety disorder by measuring respondents' fear and avoidance of various social situations [41]. The LSAS was originally developed as a clinician-administered instrument, yet the self-report version has been shown to have good psychometric properties [45]. The GAD-7 is a 7-item self-report scale used as a screener and severity measure of generalized anxiety disorder [42]. The PHQ-8 is an abbreviated form of the Patient Health Questionnaire 9-item depression scale [46], which omits the final item of the PHQ-9, a question that assesses suicidal ideation. The PHQ-8, similar to the PHQ-9, has been shown to be a valid diagnostic and severity measure for depressive disorders [43]. The PHQ-8 was chosen instead of the PHQ-9 because, owing to the anonymous and remote nature of the data collection in this study, the study investigators would be unable to properly intervene in the case of any evidence of risk for self-harm. Finally, the SDS is a 3-item self-report measure of general impairment, which has been shown to be a sensitive tool for measuring mental health–related functional impairment [44].

Both the GAD-7 and PHQ-8 instruments ask subjects to evaluate their symptoms over the past 2 weeks, whereas the LSAS and SDS ask subjects to evaluate their symptoms over the past week. Therefore, 2 weeks was the shortest duration possible to encompass the largest time window of assessment of the self-report measures, which is the rationale behind a 2-week study duration.

Smartphone Data Collection

An Android app was designed and created to collect all study data. This includes both the self-reported measures, described earlier, and the passively collected audio data—the volume of environmental audio and the presence or absence of speaking voices in the environment.

The study app records audio every 5 min, for a duration of 15 seconds, by turning on the microphone and recording the environment. This recording process occurs without any interaction from the user and with no notification to the user. Audio recordings are then securely transmitted from subjects' smartphones over the internet to a computer server where 2 further processing steps are performed. First, the average volume of each 15-second audio recording was calculated using the FFmpeg audio processing software framework [47]. Second, the presence of voices in the audio recording was detected using the Google Cloud Speech-to-Text software product [48]. This software generates text transcripts from audio recordings of speech; it was used to detect the presence of speech in audio recordings by simply noting whether each audio recording generated a transcript. Audio files containing silence, noise, or unintelligible speech do not successfully generate a text transcript, whereas recordings that contain intelligible speech produce a transcript.

The audio sampling period was chosen to be 5 min as a good trade-off between large amounts of data (with a shorter period) and the preservation of battery life of subjects' smartphones (with a longer period). Internal testing before the study showed a 5-min sampling period to be satisfactory for preserving battery life. Although a shorter period could yield more data, versions of the Android operating system since version 6 prevent this. Specifically, devices are prevented from performing background processing (such as this type of microphone sampling) while the device is sleeping more than once in a 9-min period [49]. Thus, for many devices, we are already at the limit of how frequently we can sample data in the background.

Data Preprocessing

Preprocessing of the volume time series was performed before feature extraction to account for missing data and to perform normalization. Periodic audio recordings were not reliably produced at a precise period of 5 min by the study app, so the volume time series were resampled to a period of 5 min, and missing samples were imputed using linear interpolation. After resampling and interpolation, volume samples were clipped at the ceiling and floor of 3 SDs from the mean of the volume time series to remove outliers (using subject mean and subject SD, not group). Finally, the volume time series were scaled linearly to ensure that all volume measurements were within the range of 0 to 1. No preprocessing of the voice presence time series was performed.

Feature Extraction

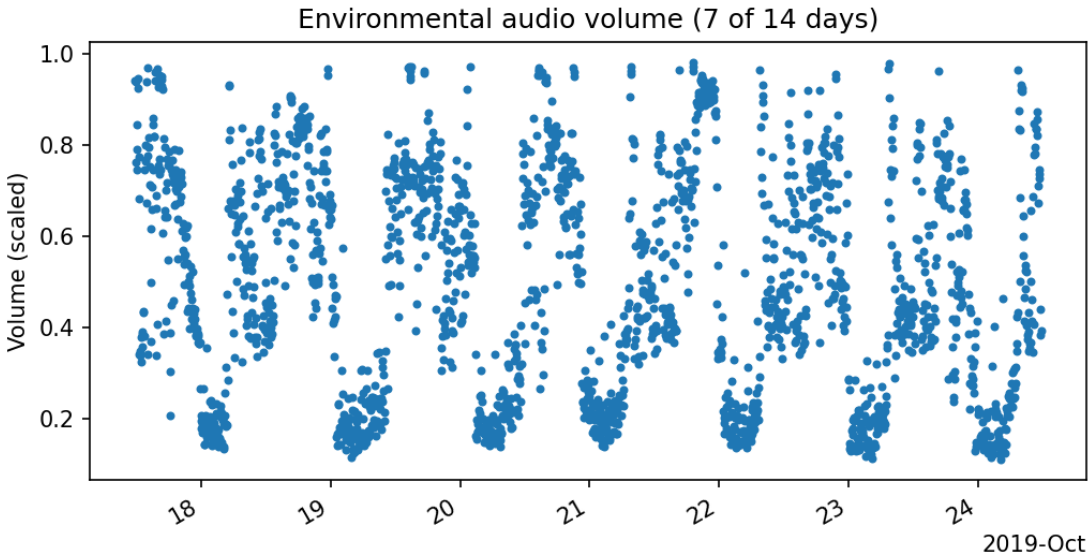
This subsection describes the methods used to compute the 4 correlates of anxiety and depression symptomatology derived from subjects' environmental audio recordings. These correlates, or *features*, were extracted from the volume and speech presence time series data to test for the association with symptoms of anxiety and depression as measured by the LSAS, GAD-7, PHQ-8, and SDS. In the sections below, we describe 3 features from the volume time series, called *daily similarity*, *sleep disturbance-all nights*, and *sleep disturbance-weeknights*. A fourth feature was extracted from the speech presence time series called the *speech presence ratio*.

Daily Similarity

The daily similarity feature was designed to infer the consistency of the subjects' sequence of daily activities. Visualizations of the volume time series clearly show distinct periods of activity (characterized by large spikes in volume) and inactivity (characterized by quieter volume with less variance). These periods coincide roughly with daytime and nighttime, respectively. Furthermore, these patterns are periodic and repeat daily. Figure 1 shows a visualization of 7 days of a subject's environmental audio volume data.

A link between regularity in daily activities, including sleep, and anxiety and depression is commonly described in the literature [50,51], and therefore, it follows that this feature may be associated with symptoms of anxiety and depression. To quantify the regularity of this pattern of daily activity, the *autocorrelation* function was computed for each subject's volume time series. This is a signal processing technique that computes the correlation between a signal and a time-delayed copy of itself [52]. The autocorrelation function of a signal is a time-dependent Pearson correlation coefficient of the signal and its copy for varying degrees of time lag between the two. As we are interested in quantifying the similarity between a subject's days, the value of the autocorrelation function at a time lag of 24 hours is most relevant. The *daily similarity* feature is, therefore, defined as the value of the autocorrelation function of the volume time series evaluated at a lag time of 24 hours.

Figure 1. Visualization of a subject's environmental audio volume data (7 of 14 days).



Sleep Disturbance

Majority of the periods of apparent inactivity that are visible in the volume time series appear to coincide with sleep. It was hypothesized that a proxy measure of sleep quality can be inferred by measuring how chaotic the volume of subjects' environments are during sleep times, replicating the link between sleep and mental health reported in the literature. For example, the prevalence of mood disorders has been shown to be much higher in populations with chronic sleep problems [53], and insomnia may be a state marker of anxiety disorders [54]. It follows, therefore, that a feature that infers the quality of subjects' sleep may be associated with symptoms of anxiety and depression.

To quantify sleep quality, the volume time series was examined with periods of quiet noted to be characterized by low variance in volume. Although the absolute value of the volume is also low at quiet times, the threshold for what can be considered quiet is greatly dependent on the specific microphone and phone placement; therefore, variance was considered a more appropriate measure of the noisiness of the environment. The *sleep disturbance* feature is defined as the SD of the volume time series between the hours of 12:00 AM and 06:00 AM, local time. This is a proxy measure of sleep quality. More specifically, it is a measure of the noisiness of a subject's environment during common hours of sleep. In addition, as people's patterns of sleep can vary between weeknights and weekends, we compute 2 versions of this feature: (1) using volume data between the hours of 12:00 AM and 06:00 AM on all nights of the week, and (2) using only volume data from weeknights (Monday through Friday). These choices are made without specific knowledge of the subjects' workday or workweek schedule.

Speech Presence Ratio

Previous studies of mental health using mobile sensing have computed proxy measures of social interaction as a feature predictive of depression severity [10,37]. To replicate these results and to ascertain if this also holds for measures of social

anxiety, generalized anxiety, and general impairment, a proxy measure of social interaction was computed. The *speech presence ratio* feature is computed from the speech presence time series data and is defined as the proportion of audio recordings containing the presence of speech to the total number of audio recordings.

Privacy

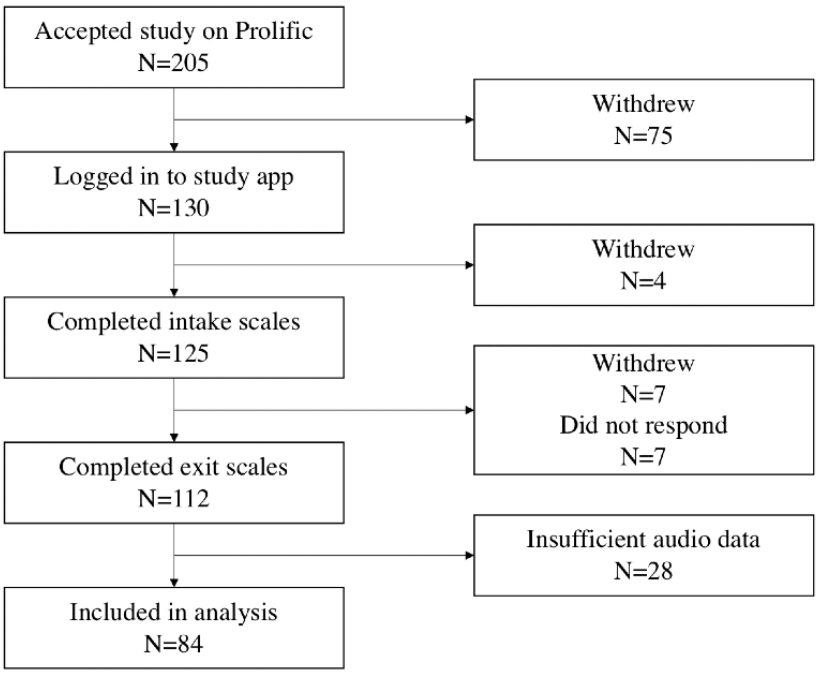
A number of privacy considerations drove the design of the study procedure, app, and data collection. Prolific, the platform from which subjects were recruited, anonymizes subjects. Subjects were provided with app log-in credentials, which were provided on demand to each subject as they enrolled in the study to avoid subjects using their name, email address, or some other potentially identifying information as their log-in name. Audio recordings were encrypted both at rest (on subjects' phones) and in transit to the server. Once processed on the server side, audio files were deleted. The speech transcripts generated to detect the presence of speech were processed in the following way: each transcript was broken into pairs of words (ie, bigrams) and then stored in random order for use in later studies. The stripping of ordering and time information from bigrams was done to prevent later re-creation of transcripts.

Results

Overview and Data Inclusion

From July 2019 to December 2019, 205 eligible Prolific members entered the study. Withdrawals were common, with 86 subjects choosing to withdraw at some point in the study (commentary on the high withdrawal rate is provided in the Limitations subsection of the Discussion section). Of the 119 subjects who did not withdraw, 112 completed both sets of self-report questionnaires. Finally, 84 of the 112 completed subjects yielded sufficient audio data for analysis based on the criterion that at least 50% of the expected number of audio recordings were made by the study app. Figure 2 provides an illustration of the study recruitment results. Although self-report measures were collected at both the beginning and end of the 2-week study, all results that followed used the values collected at the end of the study. The postobservation values of the self-report measures were used in the analysis because if such a smartphone-based assessment system were to exist, it would be more useful to predict trajectories or upcoming symptom severity, not severity 2 weeks before the start of data recording (as would be the case if correlations/associations with the prestudy scales were measured).

Figure 2. Flow chart of study recruitment.



Study Group Characteristics

The study sample had an average age of 30 years (SD 8.6) and 42% (35/84) of subjects were female. The mean and SD of the 4 self-reported measures are presented in Table 1. To understand whether the self-reported measures may be related to the age and sex of subjects, 2 statistics were computed. The Pearson correlation coefficient was computed to measure the correlation between self-reported measures and subject age, and an independent samples *t* test was performed to test whether the mean scale score differed between the male and female sexes. The LSAS, GAD-7, and SDS were negatively correlated with subject age, but there was no significant difference in the mean scale scores of the 2 sexes (at a 5% significance level). Table 1 also lists the results of these tests.

To further characterize the mental health of our study subjects, self-report measures were used to screen for social anxiety disorder, generalized anxiety disorder, and major depressive disorder. A cutoff of 60 was used with the LSAS scores to screen for social anxiety disorder (generalized subtype), as recommended by Mennin et al [55]. A cutoff of 10 was used with the GAD-7 scores to screen for generalized anxiety disorder. This cutoff was shown to optimize sensitivity (89%)

and specificity (82%) [42]. A cutoff of 10 was used with the PHQ-8 scores to screen for depression, as recommended by Kroenke et al [43]. Table 2 summarizes the results of these screenings on our study sample. The rates of social anxiety (32%), generalized anxiety (26%), and depression (37%) in our study sample were all compared with the rates in the general Canadian population. This will be elaborated in the Discussion section.

Table 1. Descriptive statistics for self-report measures of anxiety and depression (n=84).

Measures	Score, mean (SD)	Correlation with age		Difference between mean scores of the sexes	
		r	*P* value	*t* test (*df*)	*P* value
Liebowitz Social Anxiety Scale	53.7 (25.8)	−0.27	.01	−1.68 (82)	.10
Generalized Anxiety Disorder seven-item scale	6.6 (4.6)	−0.29	.01	−1.37 (82)	.17
Patient Health Questionnaire eight-item scale	8.5 (5.6)	−0.19	.09	−1.18 (82)	.24
Sheehan Disability Scale	10.8 (7.7)	−0.26	.02	−1.12 (82)	.27

Table 2. Results of screening the study sample for depression and anxiety disorders (n=84).

Disorders	Screening criteria	Positive screenings, n (%)
Social anxiety disorder	Liebowitz Social Anxiety Scale score ≥60	32 (38)
Generalized anxiety disorder	Generalized Anxiety Disorder seven-item scale score ≥10	22 (26)
Major depressive disorder	Patient Health Questionnaire eight-item scale score ≥10	31 (37)

Objective Audio Features

The objective audio features described in subsection *Feature Extraction* of the Methods section were computed using the environmental audio time series data of the 84 study subjects. Descriptive statistics for the 4 features are presented in Table 3. These descriptive statistics include the mean and SD of the features observed in the study sample, and additionally the minimum value observed, the maximum value observed, and the values of the first three quartiles. The values of the speech presence ratio appear to align reasonably with intuition, with subjects spending anywhere from 1% to 30% of their time in the presence of speech.

Table 3. Descriptive statistics for objective audio features (n=84).

Features	Mean (SD)	Minimum	Q1	Q2	Q3	Maximum
Daily similarity	0.80 (0.07)	0.45	0.77	0.83	0.85	0.90
Sleep disturbance—all nights	0.14 (0.06)	0.03	0.10	0.13	0.17	0.32
Sleep disturbance—weeknights	0.14 (0.06)	0.03	0.09	0.12	0.18	0.34
Speech presence ratio	0.15 (0.06)	0.01	0.11	0.16	0.20	0.30

Correlations Between Audio Features and Self-Report Measures

To test the association between the audio features and the self-reported measures of anxiety, depression, and functional impairment, the Pearson correlation coefficient was computed between each feature and scale. The daily similarity feature is negatively correlated with all 4 self-report measures, which supports the hypothesis that regularity in daily activity and circadian rhythm is associated with more positive mental health (ie, lower scale scores). This feature was most strongly correlated with depressive symptoms. The sleep disturbance feature, whether computed using all nighttime audio or only weeknight audio, was positively correlated with all 4 self-report measures, which is in line with the hypothesis that better sleep quality (ie, less sleep disturbance) is associated with positive mental health. The strength of the correlation is improved when only the weeknight audio is considered. Finally, the speech presence ratio feature was negatively correlated with all 4 self-report measures, where the correlation with depressive symptoms was the strongest for all observed correlations (r=−0.37; P<.001). Table 4 summarizes the results of the correlation analysis. No correction for multiple tests was performed, as we view this work as exploratory.

Table 4. Pearson correlation between objective audio features and self-reported measures of anxiety and depression (n=84).

Features	Liebowitz Social Anxiety Scale		Generalized Anxiety Disorder seven-item scale		Patient Health Questionnaire eight-item scale		Sheehan Disability Scale	
	r	*P* value	*r*	*P* value	*r*	*P* value	*r*	*P* value
Daily similarity	−0.20	.07	−0.19	.09	−0.37	<.001	−0.18	.10
Sleep disturbance—all nights	0.00	.99	0.07	.52	0.17	.13	0.15	.17
Sleep disturbance—weeknights	0.05	.65	0.12	.26	0.23	.03	0.18	.11
Speech presence ratio	−0.19	.08	−0.16	.14	−0.37	<.001	−0.29	.01

Discussion

Population

The self-report measures completed by the subjects revealed that this study's sample, despite being recruited from a healthy population, had a high prevalence of depression and anxiety. Data reported by the Government of Canada in 2006 estimate a 12-month prevalence of major depressive disorder at 4.8% [1], compared with a positive screening rate of 37% in our sample. The same 2006 report lists a 12-year prevalence of the combined class of anxiety disorders at 4.8%, a figure that is significantly lower than the positive screening rates for generalized anxiety disorder (26%) and social anxiety disorder (38%) observed in our sample. Given that the cutpoints used for screening were all shown to have high specificity, it seems unlikely that these high rates are solely a result of false-positive screenings. Instead, 2 possible explanations are proposed, which may be jointly responsible. First, the Canadian population of subjects on the Prolific recruitment platform may have elevated rates of mood and anxiety disorders with respect to the general Canadian population. It may be that people who choose to find work on the platform go there because of these conditions, which might prevent them from doing outside-the-home work. Second, that subject sampling was impacted by self-selection bias. In other words, Prolific participants who struggle with mental health to some degree may be more likely to have chosen to enroll and remain in a study that focuses on mental health.

Negative correlations were measured between age and self-reported measures of social anxiety (r=−0.27; P=.01), generalized anxiety (r=−0.29; P=.01), depression (r=−0.19; P=.09), and general impairment owing to poor mental health (r=−0.26; P=.02) in our sample of subjects. This indicates that younger subjects generally displayed worse mental health than older subjects. It is not clear if this observation is part of a more general trend in the greater population or if it is specific to the Prolific population. A review of studies examining the occurrence of anxiety, depression, and distress found some evidence that aging is associated with less susceptibility to anxiety and depression, but it is unclear if that was because of aging or cohort effects [56]. A study of mood disorders and suicide using data on American adolescents and adults from 2005 to 2017 observed an increase in mood disorders and suicidal ideation since the mid-2000s, and it is suggested that cultural trends in the use of digital media among younger individuals may be responsible for creating a cohort effect [57].

Correlations

The key finding of this work is the development and evaluation of a set of features, computed from subjects' environmental audio, as potential correlates of symptoms of anxiety, depression, and functional impairment. Correlation analysis of these features and self-reported measures, summarized in Table 4, shows that all 4 features are associated more strongly with symptoms of depression (measured by the PHQ-8) than with any other symptoms. The daily similarity and speech presence ratio were both significantly correlated with PHQ-8 scores at a 5% significance level. Speech presence ratio, but not daily similarity, was likewise significantly correlated with SDS scores at a 5% significance level. None of the 4 features were found to be significantly correlated with either LSAS or GAD-7 scores at a 5% significance level.

We note that while associations with the LSAS and GAD-7 are weak or nonexistent, the associations with the SDS are nearly as strong as those with the PHQ-8. This may suggest that the impairment that is being measured by the SDS may, in large part, be due to symptoms of depression. Indeed, we measured a stronger correlation between the SDS and the PHQ-8 scores of subjects (r=0.76; P<.001) than between the SDS and the LSAS (r=0.48; P<.001) or the GAD-7 (r=0.62; P<.001).

The fact that some features are associated with depressive symptomatology but none are associated with the symptomatology of generalized anxiety or social anxiety disorder is interesting, and we offer some speculation as to why this may be the case. First, we must observe that our features are very coarse—they measure sleep, activity, and speech, but with no specific context (they are measured on a gross scale). Depression is broadly debilitating on energy and activity, and if this impact is independent of a specific context, our features are appropriately designed to detect this impact. This is in contrast to anxiety disorders, which often have specific triggers that are context dependent. Individuals with anxiety can avoid contexts that trigger their anxiety and, therefore, present as if they do not suffer from the effects of anxiety as long as they continue to exhibit avoidance behaviors. Although it is unlikely that anxious individuals are able to avoid all triggers, especially those with generalized subtypes, avoidance behavior may be partially responsible for weakening the associations between inferred behavior and mental state. To passively measure the severity of anxiety disorders, it seems that any feature must capture avoidance behavior as a proxy for anxiety itself and

also have some measure of state anxiety to detect when a subject is in a specific context that acts as an anxiety trigger.

Comparison With Other Studies

To our knowledge, no other studies have inferred daily patterns of activity and inactivity solely from volume samples of ambient audio, which is captured by the daily similarity feature, so direct comparisons of the daily similarity feature with other known works are not possible. However, a feature called circadian movement, first proposed by Saeb et al [20], captures the degree to which the pattern of a subject's visits to different geolocations follows a 24-hour or circadian rhythm. The study by Saeb et al [20] measured a significant negative correlation between circadian movement and PHQ-9 scores (r=−0.63; P=.01; n=18). Our ambient volume-based measure of daily regularity in activity, called daily similarity, was also negatively correlated with PHQ-8 scores (r=−0.37; P<.001; n=84). This further supports the hypothesis that regularity in daily activities is associated with lower severity of depressive symptoms. It is interesting to consider that the same underlying signal may be present in both GPS data and ambient audio data. A study by Ware et al [58] also used the circadian movement feature in the prediction of depression. Although this feature was one of many used in a classifier that achieved an F1 score as high as 0.86, it is unclear to what degree this feature alone was associated with the severity of depressive symptoms.

The association between sleep and mental health has been investigated in a number of studies. Self-reported measures of sleep quality have been shown to be associated with state anxiety [59], anxiety disorders [60], and depression [61,62]. Sleep duration has been measured objectively in mobile sensing studies and shown to have a significant association with the severity of depressive symptoms [10,23,37]. although sleep duration and sleep disturbance features are different measures and as such cannot be directly compared, these studies support our results in demonstrating a general association between sleep and depression.

Finally, although the speech presence ratio feature does not appear in identical form in the literature, there are other studies that have used similar proxy measures of social interaction as correlates of depression severity. The pioneering study by Wang et al [37] inferred the number of conversations that subjects encountered throughout their days by performing an analysis of ambient audio and also found a significant negative correlation with PHQ-9 scores (r=−0.39; P=.02; n=48). This specific feature by Wang et al is more contextually aware than our speech presence ratio because not all audio that they found to contain speech was considered conversational in their system. Specifically, they ignored speech if it was detected during lecture or group meeting hours of their student subject population. Despite the fact that our proxy measure of social interaction is context unaware, we also show a negative correlation between the speech presence ratio and PHQ-8 scores (r=−0.37; P<.001; n=84). Time spent proximal to human speech was also found to be significantly associated with changes in PHQ-9 scores (P=.048) in a study by Ben-Zeev et al [10].

Limitations

A fundamental limitation of the study design is that as a cross-sectional study, it is not possible to make any claim regarding causation between the observed features and severity of anxiety, depression, or functional impairment. It cannot be determined, for example, if avoiding social contact (as inferred by the speech presence ratio feature) causes increased depression severity or if an increase in depression severity owing to some other factor causes individuals to retreat socially and engage in less speaking.

A high proportion of subject withdrawals can also be considered as a limitation of this study (86/205 subjects, 42%, chose to withdraw). The majority of withdrawals, 87% (75/86), occurred before a successful log-in to the study app (see Figure 2 for an illustration of the withdrawals throughout the study timeline). Having observed how most withdrawals occurred early in the study, we provide 2 possible hypotheses for the high number of withdrawals. The first is the relative difficulty in setting up the study app on a personal smartphone, which is more complex than the typical task asked of subjects recruited on the Prolific platform. Our setup procedure included turning off the smartphone's battery optimizations for the study app, which requires some facility with Android settings. The second is the possibility that subjects are unwilling to provide the app with the permissions necessary for data collection. The study app asks users to grant permissions before log-in, which might explain why 87% (75/86) of the withdrawals occurred without logging in. Our previous study explored clinical patients' willingness to consent to the collection of different forms of data collection for mental health purposes [54]. The number of withdrawals is in line with the data from the previous study.

A further limitation surrounding subject withdrawals is the possibility of sample bias. The 86 individuals who withdrew from the study may differ from those who remained in the study. We are unable to test this because Prolific removes researchers' access to the demographic data (age and sex) of participants who withdraw from studies. Nondemographic data (ie, digital data and self-report measures) may also differ between the groups, yet this is difficult to test because 92% (79/86) of withdrawals withdrew early enough in the study so as not to provide any digital data or self-report measures. Of the 112 individuals who remained in the study and completed all tasks, it is possible to test for differences between the group included in the analysis (ie, the group with sufficient audio data) and the group excluded from the analysis. These 2 groups did not differ significantly in age or on any of the 4 self-report measures, either at intake or exit, as tested by t tests at a 5% significance level (two-sided). These 2 groups also did not differ by sex (counts of men and women), as tested by a chi-square test (1 degree of freedom) for independence at a 5% significance level.

Some limitations also exist regarding the validity of the features. The speech presence ratio feature does not distinguish between recorded speech (eg, from a TV or radio) and human speech, it simply detects intelligible speech. This method does not distinguish between speakers, so in many cases, the subject themselves may not be the person speaking. The method also only detects English speech, so it will potentially miss speech

if, for example, a subject does not speak English at home. Finally, our technique for detecting speech using automatic speech recognition is more biased toward false negatives than false positives. If speech is detected by the system, it is highly likely that the speech is present, yet it is much more likely to miss speech in noisy environments or environments with multiple speakers speaking concurrently.

The sleep disturbance feature is affected by the subjects' specific mobile phone hardware, where different microphones with different automatic gain control functionality (which dynamically adjusts the volume and is not controllable by programmers) could produce different measurements in the same environment. To produce perfectly consistent features, one would be required to use a device such as a calibrated sound level meter, which measures volume as an absolute measure with no gain control.

Finally, it must be noted that the feasibility of completely passive mobile sensing with a high frequency of data sampling is becoming increasingly difficult on Android devices. Battery optimization features limit the rate at which apps can turn on in the background and sample their environment [63] because any app activity that *wakes up* a device not actively being used will drain device battery. Although Google does offer mechanisms for limiting battery optimizations to improve the frequency at which passive sensing can occur, many Android device vendors add third-party battery optimization software that can be difficult to disable systematically. This vendor-specific constraint on apps is a significant issue because of the fragmented nature of the Android ecosystem.

Conclusions

This work contributes toward the development of automated and objective severity measurements of anxiety, depression, and functional impairment associated with poor mental health. Focusing solely on environmental audio, which was passively sensed from subjects' smartphones, this work presents 2 new correlates of depressive symptoms and general impairment, which we refer to as the daily similarity and sleep disturbance features. Furthermore, this work supports previous findings by reproducing a measured association between time spent proximal to speech and severity of depression.

Acknowledgments

This research was funded in part by the Natural Science and Engineering Research Council Discovery Grant number RGPIN-2019-04395.

Conflicts of Interest

MAK has been a consultant or advisory board member for GlaxoSmithKline, Lundbeck, Eli Lilly, Boehringer Ingelheim, Organon, AstraZeneca, Janssen, Janssen-Ortho, Solvay, Bristol-Myers Squibb, Shire, Sunovion, Pfizer, Purdue, Merck, Astellas, Tilray, Bedrocan, Takeda, Eisai, and Otsuka. MAK conducted research for GlaxoSmithKline, Lundbeck, Eli Lilly, Organon, AstraZeneca, Jannsen-Ortho, Solvay, Genuine Health, Shire, Bristol-Myers Squibb, Takeda, Pfizer, Hoffman La Rosche, Biotics, Purdue, Astellas, Forest, and Lundbeck. MK has received honoraria from GlaxoSmithKline, Lundbeck, Eli Lilly, Boehringer Ingelheim, Organon, AstraZeneca, Janssen, Janssen-Ortho, Solvay, Bristol-Myers Squibb, Shire, Sunovion, Pfizer, Purdue, Merck, Astellas, Bedrocan, Tilray, Allergan, and Otsuka. MAK has received research grants from the Canadian Institutes of Health Research, Sick Kids Foundation, Centre for Addiction and Mental Health Foundation, Canadian Psychiatric Research Foundation, Canadian Foundation for Innovation, and the Lotte and John Hecht Memorial Foundation.

References

1. The Human Face of Mental Health and Mental Illness in Canada. Public Health Agency of Canada (PHAC). 2006. URL: https://www.phac-aspc.gc.ca/publicat/human-humain06/pdf/human_face_e.pdf [accessed 2020-02-18]
2. Annual Report. Office of the Auditor General of Ontario. 2016. URL: https://www.auditor.on.ca/en/content/annualreports/arreports/en16/2016AR_v1_en_web.pdf [accessed 2020-02-18]
3. Moskowitz DS, Young SN. Ecological momentary assessment: what it is and why it is a method of the future in clinical psychopharmacology. J Psychiatry Neurosci 2006 Jan;31(1):13-20 [FREE Full text] [Medline: 16496031]
4. van de Mortal TH. Faking it: social desirability response bias in self-report research. Aust J Adv Nurs 2008;25(4):40-48 [FREE Full text]
5. John OP, Robins RW. Accuracy and bias in self-perception: individual differences in self-enhancement and the role of narcissism. J Pers Soc Psychol 1994 Jan;66(1):206-219. [doi: 10.1037//0022-3514.66.1.206] [Medline: 8126650]
6. Ruwaard J, Kooistra L, Thong M. Ecological Momentary Assessment in Mental Health Research: A Practical Introduction With Examples in R. EMA Research Manual. 2018. URL: https://jruwaard.github.io/aph_ema_handbook/ [accessed 2020-02-18]
7. Vaid S, Harari G. Smartphones in personal informatics: a framework for self-tracking research with mobile sensing. In: Digital Phenotyping and Mobile Sensing: New Developments in Psychoinformatics. New York, USA: Springer; 2019.
8. Rohani DA, Faurholt-Jepsen M, Kessing LV, Bardram JE. Correlations between objective behavioral features collected from mobile and wearable devices and depressive mood symptoms in patients with affective disorders: systematic review. JMIR Mhealth Uhealth 2018 Aug 13;6(8):e165 [FREE Full text] [doi: 10.2196/mhealth.9691] [Medline: 30104184]

9. Sandstrom GM, Lathia N, Mascolo C, Rentfrow PJ. Putting mood in context: using smartphones to examine how people feel in different locations. J Res Personal 2017 Aug;69:96-101. [doi: 10.1016/j.jrp.2016.06.004]
10. Ben-Zeev D, Scherer EA, Wang R, Xie H, Campbell AT. Next-generation psychiatric assessment: using smartphone sensors to monitor behavior and mental health. Psychiatr Rehabil J 2015 Sep;38(3):218-226 [FREE Full text] [doi: 10.1037/prj0000130] [Medline: 25844912]
11. Asselbergs J, Ruwaard J, Ejdys M, Schrader N, Sijbrandij M, Riper H. Mobile phone-based unobtrusive ecological momentary assessment of day-to-day mood: an explorative study. J Med Internet Res 2016 Mar 29;18(3):e72 [FREE Full text] [doi: 10.2196/jmir.5505] [Medline: 27025287]
12. Pratap A, Atkins DC, Renn BN, Tanana MJ, Mooney SD, Anguera JA, et al. The accuracy of passive phone sensors in predicting daily mood. Depress Anxiety 2019 Jan;36(1):72-81. [doi: 10.1002/da.22822] [Medline: 30129691]
13. Baras K, Soares L, Paulo N, Barros R. 'Smartphine': Supporting Students' Well-Being According to Their Calendar and Mood. In: International Multidisciplinary Conference on Computer and Energy Science. 2016 Presented at: SpliTech'16; July 13-15, 2016; Split, Croatia. [doi: 10.1109/splitech.2016.7555919]
14. Mark G, Czerwinski M, Iqbal S, Johns P. Workplace Indicators of Mood: Behavioral and Cognitive Correlates of Mood Among Information Workers. In: Proceedings of the 6th International Conference on Digital Health Conference. 2016 Presented at: DH'16; April 11-13, 2016; Montreal, Canada. [doi: 10.1145/2896338.2896360]
15. Cho YM, Lim HJ, Jang H, Kim K, Choi JW, Shin C, et al. A cross-sectional study of the association between mobile phone use and symptoms of ill health. Environ Health Toxicol 2016;31:e2016022 [FREE Full text] [doi: 10.5620/eht.e2016022] [Medline: 27788568]
16. Lee H, Ahn H, Choi S, Choi W. The SAMS: smartphone addiction management system and verification. J Med Syst 2014 Jan;38(1):1. [doi: 10.1007/s10916-013-0001-1] [Medline: 24395031]
17. Naughton F, Hopewell S, Lathia N, Schalbroeck R, Brown C, Mascolo C, et al. A context-sensing mobile phone app (Q sense) for smoking cessation: a mixed-methods study. JMIR Mhealth Uhealth 2016 Sep 16;4(3):e106 [FREE Full text] [doi: 10.2196/mhealth.5787] [Medline: 27637405]
18. Burns MN, Begale M, Duffecy J, Gergle D, Karr CJ, Giangrande E, et al. Harnessing context sensing to develop a mobile intervention for depression. J Med Internet Res 2011 Aug 12;13(3):e55 [FREE Full text] [doi: 10.2196/jmir.1838] [Medline: 21840837]
19. Saeb S, Lattie EG, Schueller SM, Kording KP, Mohr DC. The relationship between mobile phone location sensor data and depressive symptom severity. PeerJ 2016;4:e2537 [FREE Full text] [doi: 10.7717/peerj.2537] [Medline: 28344895]
20. Saeb S, Zhang M, Karr CJ, Schueller SM, Corden ME, Kording KP, et al. Mobile phone sensor correlates of depressive symptom severity in daily-life behavior: an exploratory study. J Med Internet Res 2015 Jul 15;17(7):e175 [FREE Full text] [doi: 10.2196/jmir.4273] [Medline: 26180009]
21. Wahle F, Kowatsch T, Fleisch E, Rufer M, Weidt S. Mobile sensing and support for people with depression: a pilot trial in the wild. JMIR Mhealth Uhealth 2016 Sep 21;4(3):e111 [FREE Full text] [doi: 10.2196/mhealth.5960] [Medline: 27655245]
22. Farhan AA, Yue C, Morillo R, Ware S, Lu J, Bi J, et al. Behavior vs Introspection: Refining Prediction of Clinical Depression via Smartphone Sensing Data. In: Wireless Health. 2016 Presented at: WH'16; October 25-27, 2016; Bethesda, MD, USA p. 1-8. [doi: 10.1109/wh.2016.7764553]
23. Demasi O, Aguilera A, Recht B. Detecting Change in Depressive Symptoms From Daily Wellbeing Questions, Personality, and Activity. In: Wireless Health. 2016 Presented at: WH'16; October 25-27, 2016; Bethesda, MD p. 1-8. [doi: 10.1109/wh.2016.7764552]
24. Canzian L, Musolesi M. Trajectories of Depression: Unobtrusive Monitoring of Depressive States by Means of Smartphone Mobility Traces Analysis. In: Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing. 2015 Presented at: UbiComp'15; September 7-11, 2015; Osaka, Japan. [doi: 10.1145/2750858.2805845]
25. Grünerbl A, Muaremi A, Osmani V, Bahle G, Ohler S, Tröster G, et al. Smartphone-based recognition of states and state changes in bipolar disorder patients. IEEE J Biomed Health Inform 2015 Jan;19(1):140-148. [doi: 10.1109/JBHI.2014.2343154] [Medline: 25073181]
26. Abdullah S, Matthews M, Frank E, Doherty G, Gay G, Choudhury T. Automatic detection of social rhythms in bipolar disorder. J Am Med Inform Assoc 2016 May;23(3):538-543. [doi: 10.1093/jamia/ocv200] [Medline: 26977102]
27. 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]
28. Faurholt-Jepsen M, Vinberg M, Frost M, Debel S, Christensen EM, Bardram JE, et al. Behavioral activities collected through smartphones and the association with illness activity in bipolar disorder. Int J Methods Psychiatr Res 2016 Dec;25(4):309-323 [FREE Full text] [doi: 10.1002/mpr.1502] [Medline: 27038019]
29. Faurholt-Jepsen M, Busk J, Frost M, Vinberg M, Christensen EM, Winther O, et al. Voice analysis as an objective state marker in bipolar disorder. Transl Psychiatry 2016 Jul 19;6:e856 [FREE Full text] [doi: 10.1038/tp.2016.123] [Medline: 27434490]

30. Boukhechba M, Chow P, Fua K, Teachman BA, Barnes LE. Predicting social anxiety from global positioning system traces of college students: feasibility study. JMIR Ment Health 2018 Jul 4;5(3):e10101 [FREE Full text] [doi: 10.2196/10101] [Medline: 29973337]
31. Chow PI, Fua K, Huang Y, Bonelli W, Xiong H, Barnes LE, et al. Using mobile sensing to test clinical models of depression, social anxiety, state affect, and social isolation among college students. J Med Internet Res 2017 Mar 3;19(3):e62 [FREE Full text] [doi: 10.2196/jmir.6820] [Medline: 28258049]
32. Boukhechba M, Daros AR, Fua K, Chow PI, Teachman BA, Barnes LE. DemonicSalmon: monitoring mental health and social interactions of college students using smartphones. Smart Health 2018 Dec;9-10:192-203. [doi: 10.1016/j.smhl.2018.07.005]
33. Ben-Zeev D, Wang R, Abdullah S, Brian R, Scherer EA, Mistler LA, et al. Mobile behavioral sensing for outpatients and inpatients with schizophrenia. Psychiatr Serv 2016 May 1;67(5):558-561 [FREE Full text] [doi: 10.1176/appi.ps.201500130] [Medline: 26695497]
34. di Francesco S, Fraccaro P, van DV. Out-of-Home Activity Recognition from GPS Data in Schizophrenic Patients. In: 29th International Symposium on Computer-Based Medical Systems. 2016 Presented at: CBMS'16; June 20-24, 2016; Dublin, Ireland. [doi: 10.1109/cbms.2016.54]
35. Dickerson RF, Gorlin EI, Stankovic JA. Empath: A Continuous Remote Emotional Health Monitoring System for Depressive Illness. In: Proceedings of the 2nd Conference on Wireless Health. 2011 Presented at: WH'11; October 10-13, 2011; Atlanta, GA, USA. [doi: 10.1145/2077546.2077552]
36. Guidi A, Vanello N, Bertschy G, Gentili C, Landini L, Scilingo E. Automatic analysis of speech F0 contour for the characterization of mood changes in bipolar patients. Biomed Signal Process Control 2015 Mar;17:29-37. [doi: 10.1016/j.bspc.2014.10.011]
37. Wang R, Chen F, Chen Z. StudentLife: Assessing Mental Health, Academic Performance and Behavioral Trends of College Students Using Smartphones. In: Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing. 2014 Presented at: UbiComp'14; September 7-11, 2014; Osaka, Japan. [doi: 10.1145/2632048.2632054]
38. Prolific. URL: https://www.prolific.co/ [accessed 2019-07-01]
39. Mortensen K, Hughes TL. Comparing Amazon's mechanical Turk platform to conventional data collection methods in the health and medical research literature. J Gen Intern Med 2018 Apr;33(4):533-538 [FREE Full text] [doi: 10.1007/s11606-017-4246-0] [Medline: 29302882]
40. Belisario JS, Jamsek J, Huckvale K, O'Donoghue J, Morrison CP, Car J. Comparison of self-administered survey questionnaire responses collected using mobile apps versus other methods. Cochrane Database Syst Rev 2015 Jul 27(7):MR000042. [doi: 10.1002/14651858.MR000042.pub2] [Medline: 26212714]
41. Liebowitz MR. Social phobia. Mod Probl Pharmacopsychiatry 1987;22:141-173. [doi: 10.1159/000414022] [Medline: 2885745]
42. Spitzer RL, Kroenke K, Williams JB, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006 May 22;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
43. Kroenke K, Strine TW, Spitzer RL, Williams JB, Berry JT, Mokdad AH. The PHQ-8 as a measure of current depression in the general population. J Affect Disord 2009 Apr;114(1-3):163-173. [doi: 10.1016/j.jad.2008.06.026] [Medline: 18752852]
44. Leon AC, Olfson M, Portera L, Farber L, Sheehan DV. Assessing psychiatric impairment in primary care with the Sheehan disability Scale. Int J Psychiatry Med 1997;27(2):93-105. [doi: 10.2190/T8EM-C8YH-373N-1UWD] [Medline: 9565717]
45. Baker SL, Heinrichs N, Kim H, Hofmann SG. The liebowitz social anxiety scale as a self-report instrument: a preliminary psychometric analysis. Behav Res Ther 2002 Jun;40(6):701-715. [doi: 10.1016/s0005-7967(01)00060-2] [Medline: 12051488]
46. 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]
47. FFmpeg Developers. FFmpeg. 2019. URL: http://ffmpeg.org/ [accessed 2020-02-07]
48. Speech-To-Text. Google Cloud Platform. URL: https://cloud.google.com/speech-to-text/ [accessed 2020-02-07]
49. Optimize for Doze and App Standby. Android Developers. URL: https://developer.android.com/training/monitoring-device-state/doze-standby [accessed 2020-06-08]
50. Shear MK, Randall J, Monk TH, Ritenour A, Tu X, Frank E, et al. Social rhythm in anxiety disorder patients. Anxiety 1994;1(2):90-95. [doi: 10.1002/anxi.3070010208] [Medline: 9160553]
51. Szuba MP, Yager A, Guze BH, Allen EM, Baxter LR. Disruption of social circadian rhythms in major depression: a preliminary report. Psychiatry Res 1992 Jun;42(3):221-230. [doi: 10.1016/0165-1781(92)90114-i] [Medline: 1496054]
52. Gubner J. Probability and Random Processes for Electrical and Computer Engineers. Cambridge, UK: Cambridge University Press; 2006.
53. Benca RM, Okawa M, Uchiyama M, Ozaki S, Nakajima T, Shibui K, et al. Sleep and mood disorders. Sleep Med Rev 1997 Nov;1(1):45-56. [doi: 10.1016/s1087-0792(97)90005-8] [Medline: 15310523]
54. Neckelmann D, Mykletun A, Dahl AA. Chronic insomnia as a risk factor for developing anxiety and depression. Sleep 2007 Jul;30(7):873-880 [FREE Full text] [doi: 10.1093/sleep/30.7.873] [Medline: 17682658]

55. Mennin DS, Fresco DM, Heimberg RG, Schneier FR, Davies SO, Liebowitz MR. Screening for social anxiety disorder in the clinical setting: using the Liebowitz social anxiety scale. J Anxiety Disord 2002;16(6):661-673. [doi: 10.1016/s0887-6185(02)00134-2] [Medline: 12405524]
56. Jorm AF. Does old age reduce the risk of anxiety and depression? A review of epidemiological studies across the adult life span. Psychol Med 2000 Jan;30(1):11-22. [doi: 10.1017/s0033291799001452] [Medline: 10722172]
57. Twenge JM, Cooper AB, Joiner TE, Duffy ME, Binau SG. Age, period, and cohort trends in mood disorder indicators and suicide-related outcomes in a nationally representative dataset, 2005-2017. J Abnorm Psychol 2019 Apr;128(3):185-199. [doi: 10.1037/abn0000410] [Medline: 30869927]
58. Ware S, Yue C, Morillo R, Lu J, Shang C, Bi J, et al. Predicting depressive symptoms using smartphone data. Smart Health 2020 Mar;15:100093. [doi: 10.1016/j.smhl.2019.100093]
59. Atalay H. Comorbidity of insomnia detected by the Pittsburgh sleep quality index with anxiety, depression and personality disorders. Isr J Psychiatry Relat Sci 2011;48(1):54-59 [FREE Full text] [Medline: 21572244]
60. Ramsawh HJ, Stein MB, Belik S, Jacobi F, Sareen J. Relationship of anxiety disorders, sleep quality, and functional impairment in a community sample. J Psychiatr Res 2009 Jul;43(10):926-933. [doi: 10.1016/j.jpsychires.2009.01.009] [Medline: 19269650]
61. Hayashino Y, Yamazaki S, Takegami M, Nakayama T, Sokejima S, Fukuhara S. Association between number of comorbid conditions, depression, and sleep quality using the Pittsburgh sleep quality index: results from a population-based survey. Sleep Med 2010 Apr;11(4):366-371. [doi: 10.1016/j.sleep.2009.05.021] [Medline: 20219425]
62. Ağargün MY, Kara H, Solmaz M. Subjective sleep quality and suicidality in patients with major depression. J Psychiatr Res 1997;31(3):377-381. [doi: 10.1016/s0022-3956(96)00037-4] [Medline: 9306295]
63. Power Management. Android Developers. URL: https://developer.android.com/about/versions/pie/power [accessed 2020-02-07]

Abbreviations

EMA: ecological momentary assessment
GAD-7: 7-item Generalized Anxiety Disorder Scale
LSAS: Liebowitz Social Anxiety Scale
PHQ-8: Patient Health Questionnaire eight-item scale
SDS: Sheehan Disability Scale

Edited by M Focsa; submitted 21.03.20; peer-reviewed by P Chow, B Teachman; comments to author 20.05.20; revised version received 17.06.20; accepted 07.07.20; published 13.08.20.

Please cite as:
Di Matteo D, Fotinos K, Lokuge S, Yu J, Sternat T, Katzman MA, Rose J
The Relationship Between Smartphone-Recorded Environmental Audio and Symptomatology of Anxiety and Depression: Exploratory Study
JMIR Form Res 2020;4(8):e18751
URL: https://formative.jmir.org/2020/8/e18751
doi:10.2196/18751
PMID:32788153

Original Paper

Formative Evaluation of Consumer-Grade Activity Monitors Worn by Older Adults: Test-Retest Reliability and Criterion Validity of Step Counts

Stephanie A Maganja[1], BSc, MSc; David C Clarke[1], BSc, MSc, PhD; Scott A Lear[1,2,3], BSc, PhD; Dawn C Mackey[1,4], BSc, MSc, PhD

[1]Department of Biomedical Physiology and Kinesiology, Simon Fraser University, Burnaby, BC, Canada

[2]Faculty of Health Sciences, Simon Fraser University, Burnaby, BC, Canada

[3]Division of Cardiology, Providence Health Care, Vancouver, BC, Canada

[4]Centre for Hip Health and Mobility, University of British Columbia, Vancouver, BC, Canada

Corresponding Author:
Dawn C Mackey, BSc, MSc, PhD
Department of Biomedical Physiology and Kinesiology
Simon Fraser University
Shrum Science Centre Building K
8888 University Drive
Burnaby, BC, V5A 4Z2
Canada
Phone: 1 778 782 9330
Fax: 1 778 782 3040
Email: dmackey@sfu.ca

Abstract

Background: To assess whether commercial-grade activity monitors are appropriate for measuring step counts in older adults, it is essential to evaluate their measurement properties in this population.

Objective: This study aimed to evaluate test-retest reliability and criterion validity of step counting in older adults with self-reported intact and limited mobility from 6 commercial-grade activity monitors: Fitbit Charge, Fitbit One, Garmin vívofit 2, Jawbone UP2, Misfit Shine, and New-Lifestyles NL-1000.

Methods: For test-retest reliability, participants completed two 100-step overground walks at a usual pace while wearing all monitors. We tested the effects of the activity monitor and mobility status on the absolute difference in step count error (%) and computed the standard error of measurement (SEM) between repeat trials. To assess criterion validity, participants completed two 400-meter overground walks at a usual pace while wearing all monitors. The first walk was continuous; the second walk incorporated interruptions to mimic the conditions of daily walking. Criterion step counts were from the researcher tally count. We estimated the effects of the activity monitor, mobility status, and walk interruptions on step count error (%). We also generated Bland-Altman plots and conducted equivalence tests.

Results: A total of 36 individuals participated (n=20 intact mobility and n=16 limited mobility; 19/36, 53% female) with a mean age of 71.4 (SD 4.7) years and BMI of 29.4 (SD 5.9) kg/m^2. Considering test-retest reliability, there was an effect of the activity monitor (P<.001). The Fitbit One (1.0%, 95% CI 0.6% to 1.3%), the New-Lifestyles NL-1000 (2.6%, 95% CI 1.3% to 3.9%), and the Garmin vívofit 2 (6.0%, 95 CI 3.2% to 8.8%) had the smallest mean absolute differences in step count errors. The SEM values ranged from 1.0% (Fitbit One) to 23.5% (Jawbone UP2). Regarding criterion validity, all monitors undercounted the steps. Step count error was affected by the activity monitor (P<.001) and walk interruptions (P=.02). Three monitors had small mean step count errors: Misfit Shine (−1.3%, 95% CI −19.5% to 16.8%), Fitbit One (−2.1%, 95% CI −6.1% to 2.0%), and New-Lifestyles NL-1000 (−4.3%, 95 CI −18.9% to 10.3%). Mean step count error was larger during interrupted walking than continuous walking (−5.5% vs −3.6%; P=.02). Bland-Altman plots illustrated nonsystematic bias and small limits of agreement for Fitbit One and Jawbone UP2. Mean step count error lay within an equivalence bound of ±5% for Fitbit One (P<.001) and Misfit Shine (P=.001).

Conclusions: Test-retest reliability and criterion validity of step counting varied across 6 consumer-grade activity monitors worn by older adults with self-reported intact and limited mobility. Walk interruptions increased the step count error for all

monitors, whereas mobility status did not affect the step count error. The hip-worn Fitbit One was the only monitor with high test-retest reliability and criterion validity.

(JMIR Form Res 2020;4(8):e16537) doi:10.2196/16537

KEYWORDS

aged; gait; mobility limitation; exercise; movement; wearable electronic devices; mobile phone; reproducibility of results; bias; dimensional measurement accuracy

Introduction

Background and Rationale

In Canada, almost 90% of older adults (aged ≥65 years) do not meet the national physical activity recommendation of ≥150 min per week of moderate-to-vigorous aerobic physical activity [1]. Worldwide, physical inactivity is linked to an increased risk of type 2 diabetes, cardiovascular disease, colon cancer, osteoporosis, and postmenopausal breast cancer [2-7]. In addition, physically inactive older adults are at risk for falls, dependence in activities of daily living, and mobility limitation [8]. Mobility limitation affects approximately 30% of older adults in Canada and the United States and is linked to adverse health outcomes, including mobility disability and nursing home admission [9-11]. Older adults with a mobility limitation could especially benefit from physical activity interventions and corresponding physical activity monitoring [9-11].

Monitoring physical activity in older adult populations in both research and clinical settings is useful for several reasons: to detect longitudinal changes in physical activity levels [12], to determine the effects of interventions [13-19], to assess adherence to physical activity programs [14,18], to quantify daily physical activity patterns [20,21], and to motivate older adults to meet physical activity goals [22]. Consumer-grade activity monitors are a relatively affordable type of wearable technology that count steps in addition to quantifying other metrics of physical activity behavior. Older adults accept activity monitors, find them helpful for motivation, and often prefer them over simple pedometers [23].

To use a commercial-grade activity monitor to count the steps of older adults in research and clinical settings, the measured step counts must be reliable and valid [24]. If step counts exhibit poor test-retest reliability (eg, measurement errors vary from day to day), this limits the ability to detect changes in an individual's physical activity over time [25]. If step counts exhibit poor criterion validity (eg, systematic under or over counting of steps), this may lead to incorrect conclusions about the effectiveness of physical activity interventions or the effects of physical activity on health outcomes [24].

Prior Work

Substantial evidence indicates that step counts from consumer-grade activity monitors exhibit high interdevice reliability and criterion validity in healthy adults [26]. However, age-related changes in gait may affect the precision and accuracy of step counting [27]. To this end, emerging evidence from studies of older adults shows that the criterion validity of step counts from consumer-grade activity monitors is high during short-distance walks conducted in controlled laboratory settings at walking speeds >0.8 m per second [28]. However, consumer-grade activity monitors tend to overcount the steps of older adults during longer distance walking in free-living conditions [28-31] and undercount the steps when older adults walk with an assistive device, such as a walker [8,28,32,33].

Important gaps in evidence remain to be addressed. First, the test-retest reliability of step counts from consumer-grade activity monitors has not been evaluated in older adults [28]. Second, the influence of self-reported mobility limitation on the reliability and validity of activity monitor step counts in older adults has not been investigated. Finally, although aspects of the walking environment, including interruptions to continuous walking, have been suggested to influence the reliability and validity of step counting in adults [34-37], the effects of interruptions on walking have not been studied in older adults.

Study Aims

This study was motivated by our need to select a consumer-grade activity monitor for a randomized trial of a physical activity intervention for older adults, and the necessary data on the reliability and validity of step counts were not available. Thus, the purpose of this study was to evaluate the reliability and validity of step counts from consumer-grade activity monitors when worn by community-dwelling older adults during overground walking. The first aim was to determine how the *test-retest reliability* of step counting varied across 6 consumer-grade activity monitors and was affected by the presence of self-reported mobility limitations. The second aim was to determine how the *criterion validity* of step counting varied across 6 consumer-grade activity monitors and was affected by the presence of self-reported mobility limitations and walk interruptions.

Methods

Recruitment

Older adults were recruited through a variety of methods: study flyers posted around the community (eg, libraries, community and seniors' centers, and coffee shops); presentations by researchers and fitness instructors to groups of older adults (eg, at exercise classes); advertisements in local newspapers and recreation program guides; and email messages to previous research participants, fitness class attendees, and university alumni.

Individuals were eligible for inclusion, determined through telephone screening, if they were aged 65 years or older, community dwelling, and able to speak, read, and write English. We purposely recruited individuals with and without self-reported limited mobility. Individuals were classified as

having limited mobility if they self-reported difficulty walking one-quarter mile (2 to 3 blocks) outside on level ground or going up a flight of stairs (about 10 steps) without resting [10,38,39]; otherwise, they were classified as having intact mobility. Individuals were excluded if they reported an inability to walk 400 meters independently or scored below 26 (indicative of cognitive impairment) on the Montreal Cognitive Assessment [40,41]. If the Physical Activity Readiness Questionnaire for Everyone [42] indicated any medical contraindication to physical activity, the individual had to receive physician approval to participate in the study.

The study was approved by Simon Fraser University's Research Ethics Board and the University of British Columbia's Clinical Research Ethics Board. All participants provided verbal consent to telephone screening and written informed consent to participate in the study.

Descriptive Measures

Participant demographics including age, sex, racial background, level of education, and smoking history were obtained through a self-report questionnaire. Participants also self-rated their health compared with others of a similar age on a 5-point scale (excellent, good, fair, poor, or very poor). Height was measured with a portable stadiometer (seca GmbH & Co. model 217 1821009), and weight was measured with a digital scale (seca GmbH & Co. model 874 1321009). The BMI (kg/m^2) was then calculated. Lower extremity physical function was assessed using the Short Physical Performance Battery (SPPB) [10,11], which involved tests of standing balance, 6-meter gait speed, and chair stands to assess leg strength. The SPPB was scored out of 12, with a higher score indicating better function. Additional descriptive information was collected through self-report questionnaires, including physical activity, comorbidities (Functional Comorbidity Index) [43], and computer and cellphone use.

Outcome Measures

Activity Monitors

Six activity monitors were evaluated (Table 1). Three monitors were worn on the hip: Fitbit One, Misfit Shine, and New-Lifestyles NL-1000 Pedometer. The other 3 monitors, the Fitbit Charge, Garmin vívofit 2, and the Jawbone UP2, were worn on the wrist. The settings for each monitor were customized to the participant's height, weight, and age and were simultaneously placed on the nondominant side of their body according to the manufacturer's instructions. Wrist-worn monitors were randomized to their location on the arm (closest to the wrist, middle, or farthest from the wrist). Two of the hip-worn monitors were randomly assigned to 1 of 2 sites, either closer to the belly button or to the hip. The position of the New-Lifestyles NL-1000 hip-worn monitor was not randomized and was always placed halfway between the belly button and the hip, according to the manufacturer's recommendation. The randomization procedure was performed before testing.

Table 1. Description of activity monitors.

Monitors	Manufacturers	Body placement	Digital display	Step counting instruments
Fitbit Charge	Fitbit, San Francisco, California, United States	Wrist	Yes	Three-axis accelerometer
Fitbit One	Fitbit, San Francisco, California, United States	Hip	Yes	Three-axis accelerometer
Garmin vívofit 2	Garmin, Olathe, Kansas, United States	Wrist	Yes	Three-axis accelerometer
Jawbone UP2	JAWBONE, San Francisco, California, United States	Wrist	No	Three-axis accelerometer
Misfit Shine	Misfit, Burlingame, California, United States	Hip	No	Three-axis accelerometer and magnetometer
New-Lifestyles NL-1000	New-Lifestyles, Lee's Summit, Missouri, United States	Hip	Yes	Piezoelectric pedometer

Walking Trials

Participants completed 4 walking trials in a long hallway (Figure 1). Testing was conducted on weekends to avoid weekday foot traffic, and signs were displayed to minimize disruptions. For each walk, one researcher instructed the participant to start and stop walking, whereas another researcher timed the walk with a stopwatch and recorded the time to complete the walk. During the walks, the 2 researchers walked slightly behind the participants and counted their steps using tally counters. Tally counts were used as the criterion measure, which is common in activity monitor assessment [24,37,44]. When discrepancies occurred between steps counted by the 2 researchers, the median value was used and rounded up to the nearest whole number unless one researcher believed they miscounted the steps, in which case the other researcher's number was used. In total, 6 activity monitor step counts were recorded immediately before and after each walk.

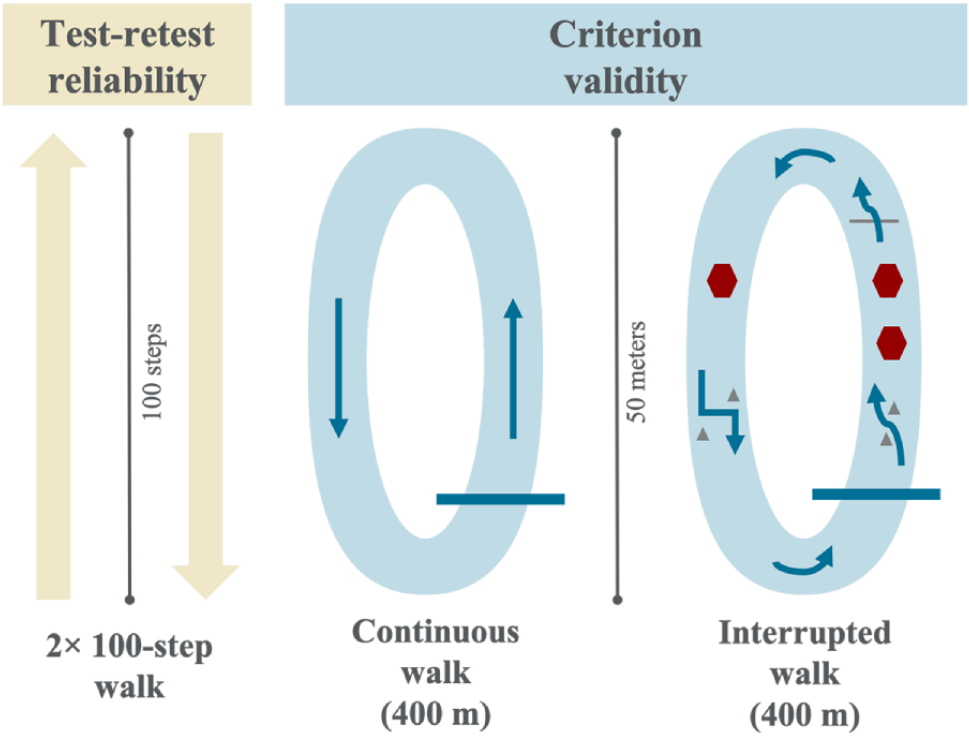
For all trials, participants were instructed to walk at their preferred walking speed, defined as a comfortable speed that they could maintain for the duration of the walk. To prevent fatigue, participants were provided with adequate rest time between the walks (5 to 15 min). The 4 walking trials were typically completed within 1 hour, within which approximately 15 min of walking was completed (approximately 1 min for each 100-step walk and approximately 5 to 7 min for each 400-meter walk).

The first 2 reliability walks (RW1 and RW2) required the participant to walk 100 steps (Figure 1). A researcher notified the participants when they had 5 steps left to walk and provided a verbal countdown to the end of the walk. If a participant did not walk exactly 100 steps on their first walk, the participant was instructed to walk the same number of steps for their second walk.

For the 400-meter continuous walk (CW), a 100-meter course was defined using pylons (Figure 1). Participants completed 4 laps of the course without stopping, beginning, and ending at the same point on the course.

Research suggests that the walking environment and interruptions can affect the validity and reliability of step counts from activity monitors [34-36], so a 400-meter interrupted walk (IW) was included to mimic the conditions of daily walking more closely than the CW (Figure 1). Participants walked the same 400 meters as in the CW, but 7 interruptions were incorporated into each lap using additional pylons and signs. These interruptions included an S-curve, 2 consecutive 5-second stops, object avoidance (stepping over a tree branch), a sharp turn to change the direction, one 5-second stop, 2 successive 90-degree angle turns, and an additional sharp turn. In completing 4 laps, participants encountered each interruption 4 times for 36 interruptions in total.

Figure 1. Walking trials completed in a level hallway. For criterion validity testing, participants walked 4 laps of the continuous and interrupted courses to reach 400 meters.



Measurements

Walking trial step counts for each activity monitor were calculated by subtracting the step count recorded at the beginning of the walk from the step count recorded at the end of the walk (eg, end of CW step count − beginning of CW step count = CW step count). To account for participants walking a different number of steps per trial, all step counts were converted to step count percent errors, which were calculated as follows:

Step count percent errors closer to zero are more desirable. Positive step count percent errors indicated that the activity monitor was overcounting steps relative to the tally count (criterion), whereas negative step count percent errors indicated undercounting relative to the tally count.

Statistical Analysis

Sample Size

We calculated that a total sample size of 34 participants (17 within each group of intact mobility and limited mobility) would provide 80% statistical power to detect an effect size of 5% step count error within each group with significance level alpha of .05, assuming that the SD in step count error was similar to what was observed in our pilot data (SD 3.3, n=5 young adults). We aimed to recruit 20 participants within each group to account for potential missing data.

Descriptive Analysis

Descriptive data for participant characteristics are presented as means and SDs for normally distributed continuous variables and medians and IQRs for skewed continuous variables. Judgments of normality were based on the visual inspection of frequency distributions. Categorical variables are reported as frequencies and percentages. To assess differences in descriptive characteristics between the groups with intact and limited mobility, 3 types of statistical tests were used depending on how the data were distributed: independent sample *t* tests for normally distributed continuous variables, Wilcoxon rank-sum test for skewed continuous variables, and chi-square tests for categorical variables. These tests were performed using JMP software (SAS Institute; version 13.1; 2016). Descriptive data for step count errors are presented as means and 95% CI. Statistical modeling was conducted using RStudio version 1.0.136. The family-wise significance level for statistical tests was set at an alpha of .05.

Test-Retest Reliability

As a measure of trial-to-trial consistency, the absolute difference between step count percent errors from RW1 and RW2 was calculated as follows:

A two-way analysis of variance (ANOVA) was used to assess the effects of the *activity monitor* and *mobility status* on the mean absolute difference in step count percent error between RW1 and RW2. A post hoc analysis of pairwise differences was conducted using the Tukey honest significant difference (HSD) test, where appropriate, which held the experiment-wise error rate constant at an alpha level of .05. To assess the normality of the step count percent error distributions, we visually inspected the quantiles of the distribution, histograms, and density plots and ran a Shapiro-Wilk normality test. Owing to suggestions of nonnormality, we also ran a nonparametric test, Kruskal-Wallis, which produced the same results and led to the same conclusions as the ANOVA. For ease of interpretation, we reported only the results of the ANOVA.

In addition, the standard error of measurement (SEM) was calculated as a descriptive measure of test-retest reliability. SEM was calculated as the SD of the differences between the step count percent errors of RW1 and RW2, divided by the square root of the number of walks, in accordance with Hopkins [25].

Criterion Validity

A three-way ANOVA was used to determine whether the *activity monitor*, *interruptions to walking*, and *mobility status* had effects on the mean step count percent error. A post hoc analysis of pairwise comparisons was conducted using the Tukey HSD test, where appropriate.

Bland-Altman plots [45,46] were produced to assess for systematic bias and limits of agreement in step counts for each activity monitor and for the CW and IW. The mean step count from the 2 measures was plotted on the x-axis (eg, [activity monitor step count]+[tally counter step count]/2), and the error between the 2 measures was plotted on the y-axis (eg, [activity monitor step count]−[tally counter step count]). Reference lines indicate the mean step count error, trend, and 95% limits of agreement (mean +1.96 SD and −1.96 SD).

In accordance with previous studies [33], equivalence testing was conducted to evaluate whether mean step count percent errors were equivalent to a zero step count percent error for each activity monitor and for both 400-meter walks. We defined the equivalence bound as −5.0% to +5.0% step count error, which we deemed to be clinically relevant. Two one-sided *t* tests were conducted to evaluate both sides of the equivalence interval. If there was sufficient evidence to reject both the null hypothesis of the upper threshold (mean error ≤5%) and the null hypothesis of the lower threshold (mean ≥−5%), then the mean step count error was interpreted as *practically* equivalent to a zero step count error.

Results

Participants

A total of 36 individuals participated in the study, including 20 with self-reported intact mobility (7 females) and 16 with self-reported limited mobility (12 females; Table 2). The mean age of the participants was 71.4 years (SD 4.7), and the mean BMI was 29.4 kg/m^2 (SD 5.9). For most characteristics, there were no significant differences between the groups with intact and limited mobility. However, the group with limited mobility had significantly slower gait speed than the group with intact mobility for the 6-meter (P<.001) and continuous 400-meter (P<.001) walks. In addition, the group with limited mobility had a greater number of comorbidities (P=.02).

Table 2. Participant characteristics.

Characteristics	Groups			*P* value[a]
	Overall (n=36)	Intact mobility (n=20)	Limited mobility (n=16)	
Female, n (%)	19 (53)	7 (35)	12 (75)	.02
Age (years), mean (SD)	71.4 (4.7)	73.1 (3.7)	71.6 (5.8)	.40
Weight (kg), mean (SD)	82.0 (16.8)	78.1 (17.4)	87.0 (15.1)	.11
BMI (kg/m^2), mean (SD)	29.4 (5.9)	27.7 (4.4)	31.5 (7.0)	.07
White, n (%)	25 (69)	13 (65)	12 (75)	.46
University education, n (%)	16 (44)	9 (45)	7 (44)	>.99
Montreal Cognitive Assessment (of 30), median (IQR)	27 (26-28)	27 (26-27)	28 (27-29)	.03[b]
Smoked previously, n (%)	17 (47)	8 (40)	9 (56)	.33[c]
Good or excellent self-rated health, n (%)	29 (81)	18 (90)	11 (69)	.20
Short Physical Performance Battery (of 12), median (IQR)	11.0 (10.0-11.0)	11.0 (10.0-12.0)	10.0 (9.8-11.0)	.06[b]
6-meter gait speed (m/s), mean (SD)	1.2 (0.2)	1.3 (0.2)	1.1 (0.2)	<.001
400-meter gait speed (m/s), mean (SD)	1.3 (0.2)	1.4 (0.1)	1.2 (0.2)	<.001
400-meter continuous walk step count, mean (SD)	600 (70)	562 (29)	648 (76)	<.001
400-meter interrupted walk step count, mean (SD)	656 (73)	617 (41)	703 (77)	<.001
Self-reported moderate-to-vigorous physical activity[d] (min/week), mean (SD)	343 (368)	343 (272)	343 (471)	.10
Self-reported walking (min/week), mean (SD)	208 (197)	240 (208)	167 (182)	.27
Number of comorbidities, median (IQR)	2.0 (0.0-3.0)	0.0 (0.0-2.0)	2.5 (1.0-4.0)	.02[b]
≥1 comorbidity, n (%)	19 (53)	8 (40)	11 (69)	.04
≥2 comorbidities, n (%)	12 (33)	4 (20)	8 (50)	.09[c]
Arthritis, n (%)	13 (36)	4 (20)	9 (56)	.04
Obesity, n (%)	10 (28)	3 (15)	7 (44)	.07
Visual impairments[e], n (%)	10 (28)	7 (35)	3 (19)	.46
Degenerative disc disease[f], n (%)	6 (17)	2 (10)	4 (25)	.37
Depression, n (%)	4 (11)	1 (5)	3 (19)	.30
Diabetes (type 1 or 2), n (%)	4 (11)	1 (5)	3 (19)	.30
Osteoporosis, n (%)	3 (8)	0 (0)	3 (19)	.08
Access to a computer with internet, n (%)	31 (86)	17 (85)	14 (88)	>.99
Access to cellphone or smartphone, n (%)	33 (92)	18 (90)	14 (88)	.57

[a]*P* values comparing intact mobility versus limited mobility, from a chi-square Fisher exact test for categorical variables and from an independent sample *t* test for continuous variables.

[b]From a Wilcoxon rank-sum test.

[c]From a chi-square Pearson test.

[d]Moderate-to-vigorous physical activity includes self-reported walking.

[e]For example, cataracts, glaucoma, and macular degeneration.

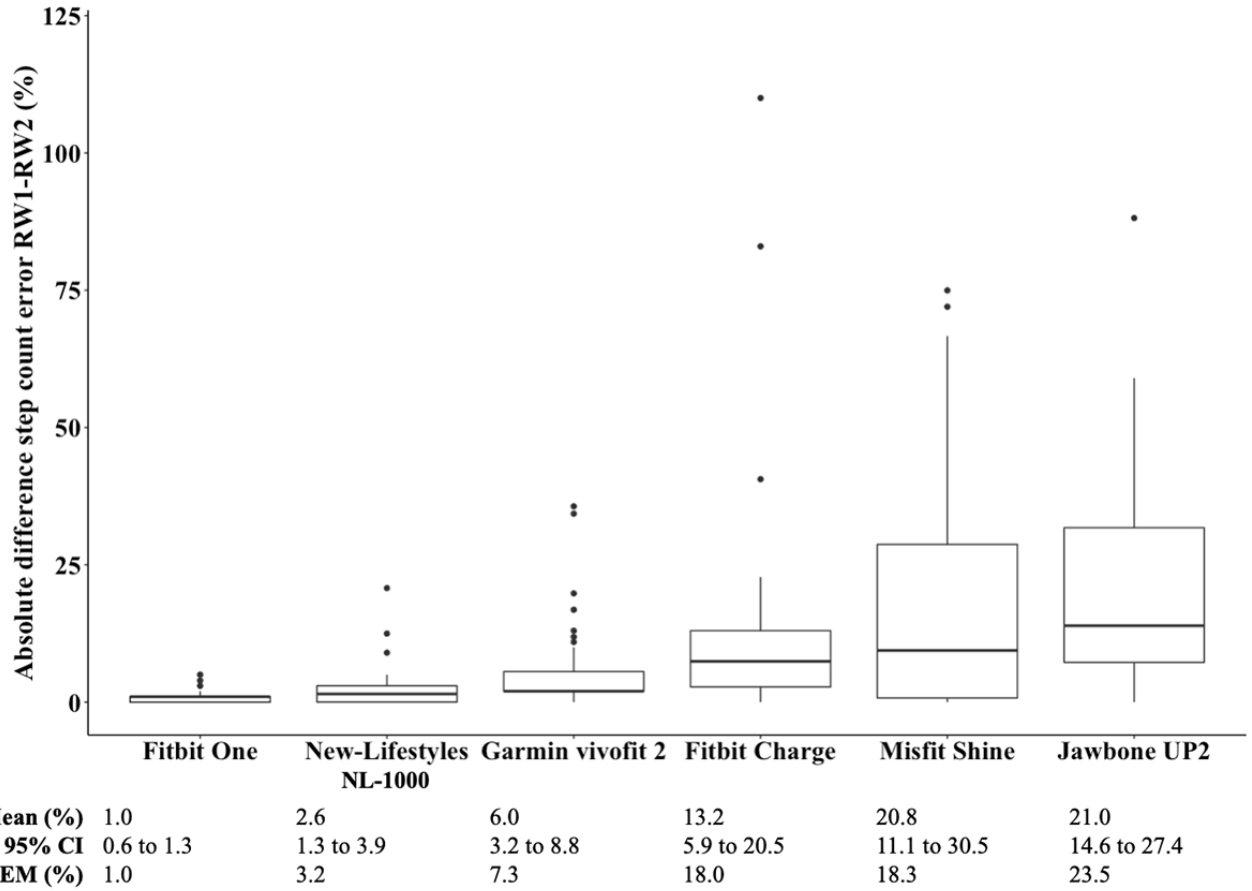
[f]For example, back disease, spinal stenosis, or severe chronic back pain.

Test-Retest Reliability

We found a significant main effect of the *activity monitor* on the absolute difference between the step count percent errors of RW1 and RW2 (*P*<.001), but we found no main effect of *mobility status* (*P*=.31) and no interaction between the *activity monitor* and *mobility status* (*P*=.29). We found the smallest mean absolute differences in step count percent errors for the Fitbit One (1.0%, 95% CI 0.6% to 1.3%), New-Lifestyles NL-1000 (2.6%, 95% CI 1.3% to 3.9%), and Garmin vívofit 2

(6.0%, 95% CI: 3.2% to 8.8%; Figure 2). Post hoc tests revealed that the Fitbit Charge (P=.02), Jawbone UP2 (P<.001), and Misfit Shine (P<.001) exhibited significantly higher mean absolute differences than the Fitbit One. In addition, the Jawbone UP2 (P<.001) and Misfit Shine (P<.001) had greater mean absolute differences than the New-Lifestyles NL-1000. Finally, the Jawbone UP2 (P=.002) and Misfit Shine (P=.004) had greater mean absolute differences than the Garmin vívofit 2. The SEM values ranged from 1.0% (Fitbit One) to 23.5% (Jawbone UP2; Figure 2).

Figure 2. Box plots illustrating median absolute difference in step count percent errors between the 100-step test-retest reliability walk 1 (RW1) and walk 2 (RW2) for the 6 activity monitors (n=36 for all monitors except Misfit Shine, n=28). Central rectangle spans the IQR, and the whiskers represent the inner fence (upper: Q3+1.5×IQR and lower: Q1−1.5×IQR). Fitbit Charge different than Fitbit One (P=.02); Misfit Shine different from Fitbit One (P<.001), New-Lifestyles NL-1000 (P<.001), and Garmin vívofit 2 (P=.004); Jawbone UP2 different from Fitbit One (P<.001), New-Lifestyles NL-1000 (P<.001), and Garmin vívofit 2 (P=.002). SEM: standard error of measurement.



	Fitbit One	New-Lifestyles NL-1000	Garmin vivofit 2	Fitbit Charge	Misfit Shine	Jawbone UP2
Mean (%)	1.0	2.6	6.0	13.2	20.8	21.0
95% CI	0.6 to 1.3	1.3 to 3.9	3.2 to 8.8	5.9 to 20.5	11.1 to 30.5	14.6 to 27.4
SEM (%)	1.0	3.2	7.3	18.0	18.3	23.5

Criterion Validity

The mean (SD) step count from the criterion tally counter on the 400-meter CW was 600 (SD 79) steps and on the 400-meter IW was 656 (SD 73) steps (Table 2). All monitors undercounted steps relative to the criterion (tally) counts (Figure 3), with the Misfit Shine exhibiting the lowest mean step count percent error (−1.3%). We found significant main effects of the *activity monitor* (P<.001) and *walk interruptions* (P=.02) on step count percent error, but no main effect of *mobility status* (P=.65). We observed no interactions between any of the factors. Regarding the main effect of the activity monitor, post hoc tests revealed that the Fitbit Charge (P<.001) and Garmin vívofit 2 (P=.02) exhibited significantly higher mean step count percent errors than the Misfit Shine. In addition, the Fitbit Charge exhibited a greater mean step count percent error than the Fitbit One (P<.001) and the New-Lifestyles NL-1000 (P=.03). Regarding the main effect of *interruptions*, the IW resulted in a greater mean step count percent error than the CW (mean difference 1.9%; P=.02).

Bland-Altman plots revealed nonsystematic bias across the range of observed step counts for the Fitbit One and Jawbone UP2 (Figure 4). Systematic bias and wide limits of agreement were observed for Misfit Shine, New-Lifestyles NL-1000, Garmin vívofit 2, and Fitbit Charge. In addition, Bland-Altman plots indicated systematic bias across the range of observed step counts and wide limits of agreement for both the CW and IW.

Equivalence tests indicated that the mean step count percent errors of 2 monitors lay within the −5% and +5% equivalence bound, the Fitbit One (P<.001) and Misfit Shine (P=.001); thus, step counts from these monitors were deemed equivalent to a zero step count percent error (Figure 5). The CW mean step count percent error was statistically equivalent to zero (P=.002), whereas the IW mean step count percent error lay outside the equivalence bounds (P=.28).

Figure 3. Box plots illustrating median step count percent errors for the 6 activity monitors (n=72 for all monitors except the Misfit Shine, n=67) and for the 2 different 400-meter walks (n=213 for continuous and n=214 interrupted) from 36 older adults. Central rectangle spans the IQR, and the whiskers represent the inner fence (upper: Q3+1.5×IQR and lower: Q1−1.5×IQR). Horizontal dotted lines represent zero step count percent error. Garmin vívofit 2 different from Misfit Shine (*P*=.02); Fitbit Charge different from Misfit Shine (*P*<.001), Fitbit One (*P*<.001), and New-Lifestyles NL-1000 (*P*=.03); interrupted different from continuous (*P*=.02).

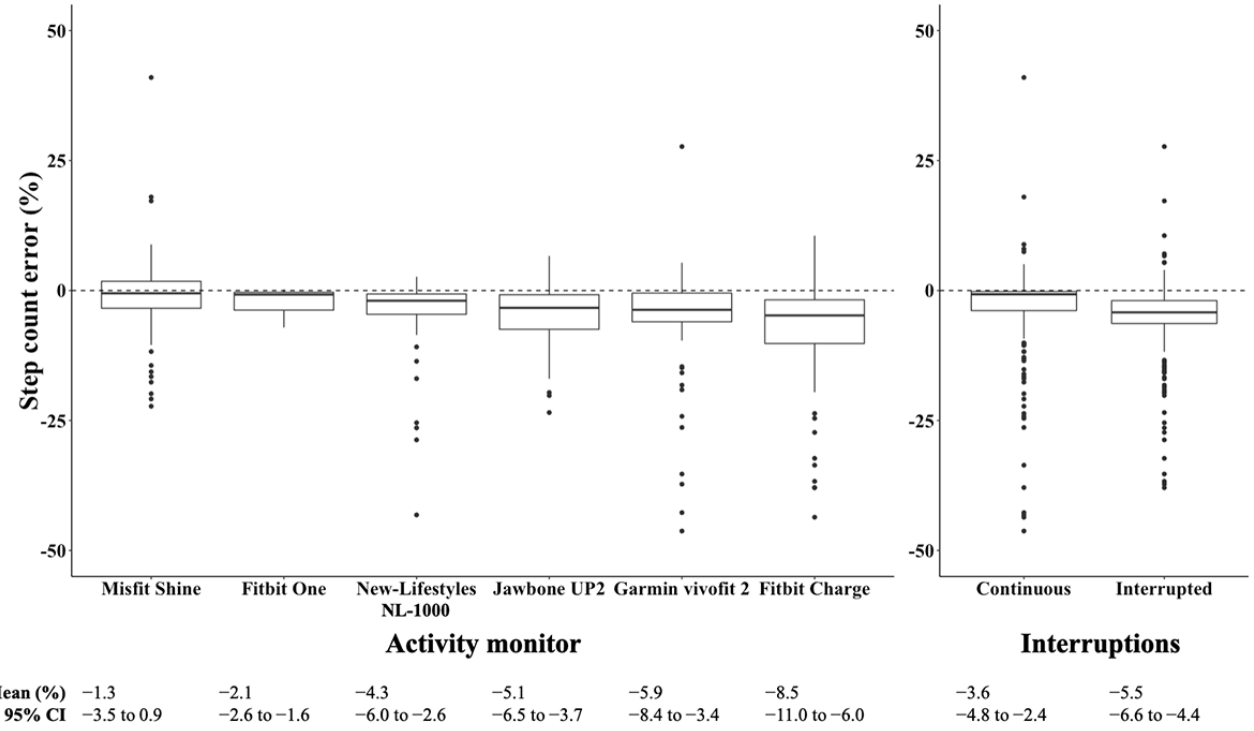


Figure 4. Bland-Altman plots for each activity monitor (n=72 for all monitors except the Misfit Shine, n=67) and for walk interruptions (n=213 for continuous and n=214 for interrupted) compared with the criterion tally counts from 36 older adults. The solid lines represent the mean step count error (horizontal) and line of best fit (trend line). Dotted lines represent the limits of agreement (mean±1.96 SD).

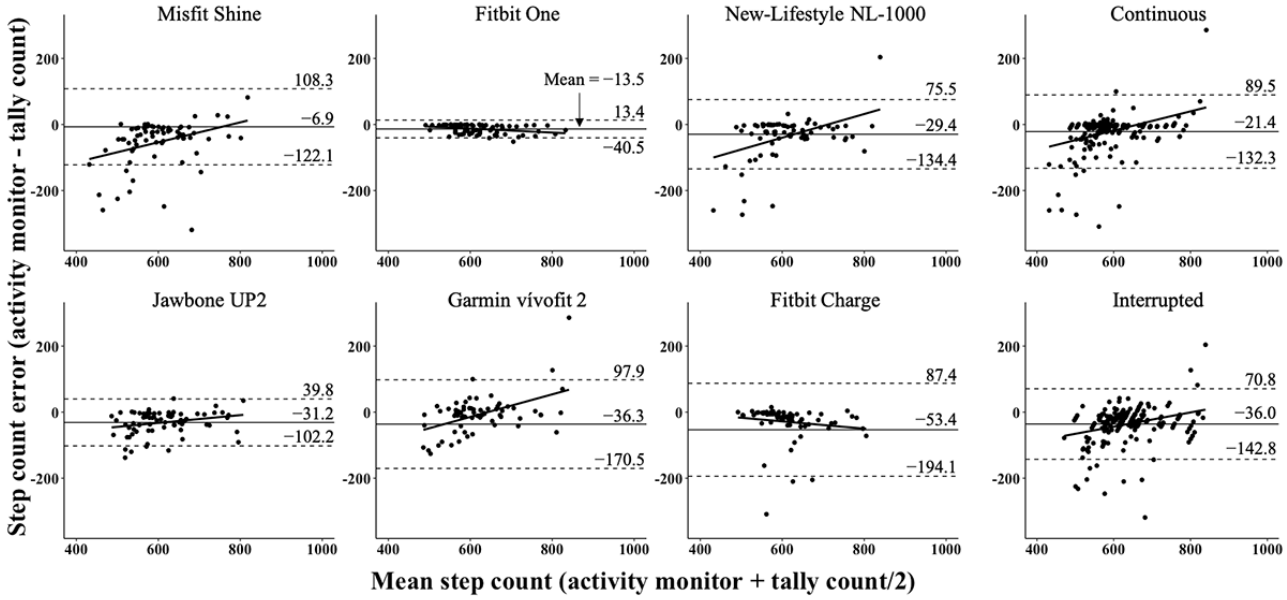
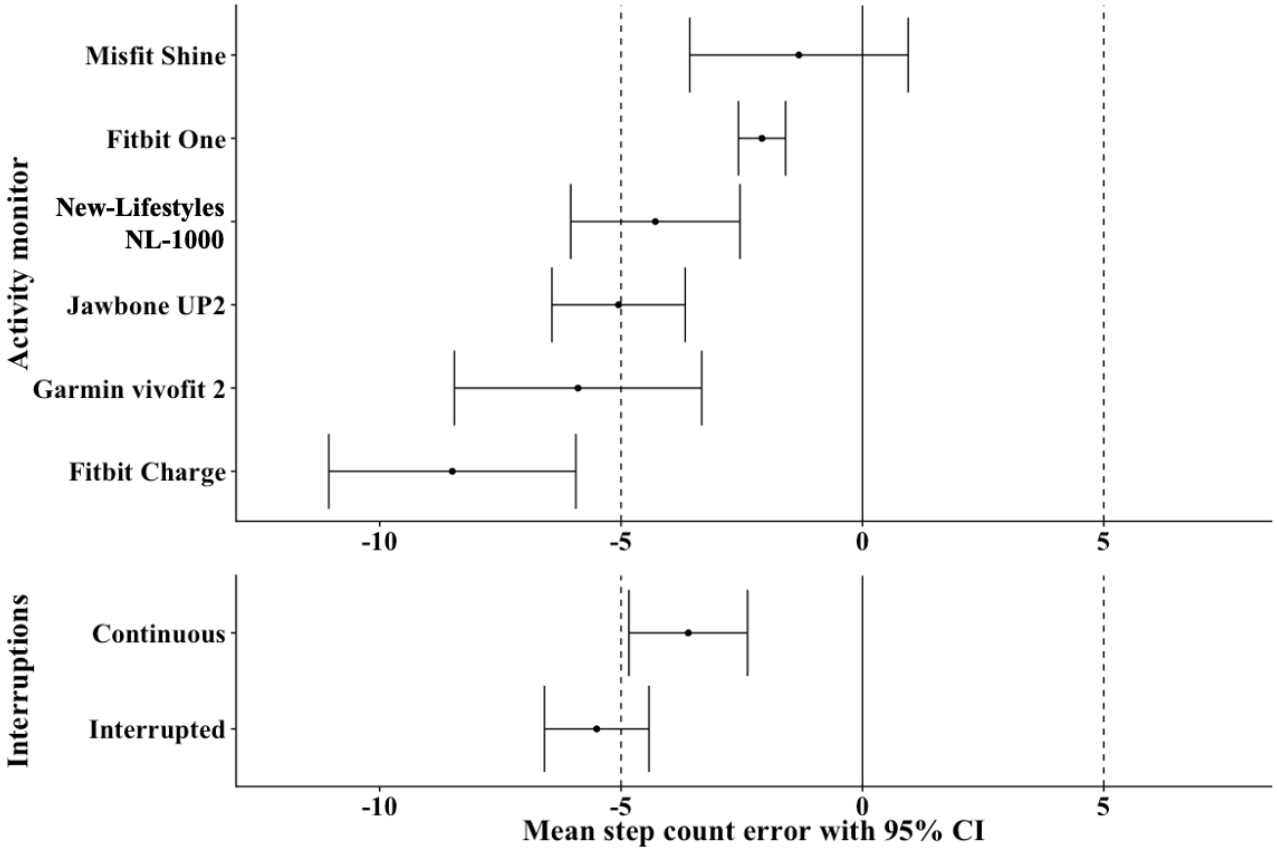


Figure 5. Equivalence testing plots for activity monitor (n=72 for all monitors except the Misfit Shine, n=67) and walk interruptions (n=213 for continuous and n=214 for interrupted). Mean step count errors (%) with 95% CI. Area between dotted vertical lines represents equivalence bounds (+/−5.0%).



Discussion

Principal Findings

Our study aimed to determine (1) how test-retest reliability of step counting by 6 consumer-grade activity monitors was affected by the presence of self-reported mobility limitations in community-dwelling older adults during overground walking and (2) how the criterion validity of step counting by these 6 activity monitors was affected by the presence of self-reported mobility limitations and walk interruptions in community-dwelling older adults during overground walking. We found that test-retest reliability varied across activity monitors (highest for the Fitbit One and lowest for the Jawbone UP2) but was unaffected by the self-reported mobility status. The monitors featured varying degrees of criterion validity, with the Fitbit One exhibiting the highest and the Fitbit Charge, the lowest. Criterion validities were negatively impacted by walk interruptions but were unaffected by self-reported mobility status. The hip-worn Fitbit One was the only monitor that exhibited both high test-retest reliability and criterion validity.

Comparison With Prior Work

To the best of our knowledge, our study is the first to report on the test-retest reliability of consumer-grade activity monitors in a community-dwelling older adult population [28]. We found that test-retest reliability of step counting, measured by mean absolute percent difference in step count error between repeated 100-step walks, varied across activity monitors. Specifically, only 2 monitors had small mean absolute percent differences in the step count error of less than 5.0%: the Fitbit One and the New-Lifestyles NL-1000. Three monitors (Fitbit Charge, Jawbone UP2, and Misfit Shine) were significantly less reliable than either or both the Fitbit One and New-Lifestyles NL-1000. Finally, the SEM of the Fitbit One was small, within −2.5% and +2.5%, which translates into a between-trial difference of −4.9% and +4.9% step count error in 95 of 100 instances (95% likely range of −4.9% to 4.9%). All other monitors had SEM values indicative of poor reproducibility. Therefore, we conclude that only the Fitbit One had sufficiently high test-retest reliability.

We found that the criterion validity of step counting was affected by both the activity monitor and walk interruptions during 400-meter walks, with no interaction observed between these 2 factors. Fitbit One was the only monitor with high criterion validity. This interpretation is based on the Fitbit One's small mean step count percent error (less than −5.0% or +5.0%), lack of systematic bias, and small limits of agreement, and it was deemed equivalent to a zero step count percent error (equivalence bound of −5.0% to +5.0%). Three of the other monitors (Misfit Shine, New-Lifestyles NL-1000, and Jawbone UP2) exhibited moderate correspondence to the criterion, whereas both the Garmin vívofit 2 and Fitbit Charge had poor correspondence with the criterion. Our results for criterion validity are consistent with previous research by Floegel et al [33] who found that the Fitbit One had the lowest mean step count percent error and outperformed other monitors (StepWatch, Omron HJ-112, Fitbit Flex, and Jawbone UP) when

compared with direct observation during a 100-meter walk involving both older adults with and without mobility impairments.

For all activity monitors that we tested, walking with interruptions resulted in greater mean step count percent errors than walking continuously. In addition, the mean step count percent error for interrupted walking was not equivalent to zero (equivalence bound of −5.0% to +5.0%), whereas it was equivalent to zero for continuous walking. We observed a systematic bias in step count errors for both walking conditions; specifically, step count errors increased in proportion to the number of steps taken, and the limits of agreement were wide. Previous studies have not tested interrupted walking in older adults in controlled settings, as we did. However, in previous studies that investigated activity monitors during free-living walking conditions in older adults [29-31,47,48], 5 of 8 consumer-grade hip- and wrist-worn activity monitors were found to overcount steps relative to criterion measures [29-31,47,48]. These results are inconsistent with our finding that consumer-grade activity monitors undercounted steps during continuous and interrupted walking. A possible reason for this discrepancy is that during free-living conditions, movements other than stepping (eg, movements during eating or conversation) may reach accelerations that exceed the monitor algorithm thresholds, causing steps to be erroneously recorded [49]. In support of this notion, Tudor-Locke et al [50] compared the hip- and wrist-worn ActiGraph accelerometers during controlled treadmill walking and in free-living conditions. During treadmill testing, they found that the wrist-worn monitor detected fewer steps than the hip-worn monitor; however, during free-living conditions the wrist-worn monitor counted more steps than the hip-worn monitor.

Regarding self-reported mobility, we found that test-retest reliability and criterion validity of step counting were unaffected by the presence of a self-reported mobility limitation, suggesting that older adults with a self-reported mobility limitation can expect similar performance from the activity monitors tested in this study as older adults with self-reported intact mobility. Consistent with our results, Floegel et al [33] reported that mean step count errors for most monitors they tested were similar and small for older adults with or without walking impairment who did not walk with a cane or walker (StepWatch −4.42% vs −3.45%, Fitbit One −2.59% vs −1.71%, Omron −4.48% vs −3.15%, Fitbit Flex −26.94% vs −16.31%, and Jawbone UP −2.86% vs −8.43%). In contrast to our results, Lauritzen et al [8] reported that mobility limitations decreased activity monitor validity when comparing a small group of walker-dependent older adults in nursing homes to healthy older adults [8]. In that study, lower gait assessment scores were significantly correlated with larger absolute percent errors, whereas longer walk times and larger step counts were significantly correlated with larger absolute percent errors. Our study population differed because participants did not use walking aids.

Previous literature indicates that slow gait speed significantly affects the criterion validity of activity monitors [32]. Simpson et al [32] reported that the Fitbit One, when worn on the hip, recorded zero steps when participants walked at speeds between 0.3 m/s and 0.5 m/s, and it had a mean percent error smaller than 10% only when walking speed was 0.8 m/s and 0.9 m/s [32]. Our participants walked, on average, at 1.2 m/s (intact mobility 1.3 m/s and limited mobility 1.1 m/s); thus, speed should not have negatively impacted activity monitor performance in our study. If participants with a self-reported mobility limitation had very slow gait speed or more severe asymmetries in their gait, we may have detected differences in test-retest reliability and criterion validity of step counts from the monitors based on the self-reported mobility status. Future studies of older adults with slower gait speeds (eg, 0.4 m/s to 0.8 m/s) are still needed to understand the performance of consumer-grade activity monitors in the growing population of older adults living with frailty and more severe mobility limitation than this study population.

Limitations and Strengths

This study had certain limitations. First, the results have limited generalizability with respect to the activity monitors. We tested a single monitor of each activity monitor model with a relatively small sample size. Thus, the results obtained from this study may not be applicable to all versions of the activity monitor model tested or other monitors produced by the same brand. A poorly calibrated monitor (in relation to the average monitor) would have led us to underestimate monitor validity, whereas a better-than-average calibrated monitor would have led us to overestimate monitor validity. Ideally, multiple versions of each monitor would have been tested, and the difference between the monitors was assessed. We had to limit the number and distance of walks performed with our older adult study population to manage participant burden and prevent fatigue, so it was not feasible to conduct additional testing. However, we believe that interdevice variation would likely have been minimal based on a systematic review of consumer-grade activity monitors that reported high interdevice reliability for step counts from 4 studies testing 3 Fitbit models (Classic, One, and Ultra; intraclass correlation coefficients ranged from 0.76 to 1.00) [26]. Second, as the reliability of consumer-grade activity monitors had not been previously evaluated in older adults, we chose to begin by assessing test-retest reliability on short, 100-step CWs. Future studies are needed to examine the effects of walking interruptions and longer distances on test-retest reliability. Our results suggest that reliability under these conditions would likely be better for a hip-worn Fitbit monitor, such as the Fitbit One, than for other monitors. Third, we did not consider the contributions of sex, walking speed, participant height, or stride length on test-retest reliability or criterion validity. In addition, we did not investigate how common daily tasks, other than walking, might affect activity monitor step counts. It will be important for future studies to evaluate the reliability and validity of step counting by consumer-grade activity monitors during a wider range of daily movements than was tested in this study. Further, future studies should seek to determine sources of error during activities of daily living, which often result in overcounting during free-living assessment of consumer-grade activity monitors.

This study also has several strengths. First, all walking tests were performed during overground walking, which represented natural walking conditions more closely than treadmill walking. Treadmill walking has been used frequently in previous studies

to evaluate the measurement properties of activity monitors because it enables monitors to be tested at controlled walking speeds. However, older adults who are unfamiliar with treadmill walking exhibit increased heart rate and oxygen consumption while walking on a treadmill compared with overground walking [51]. Moreover, treadmills impose greater symmetry in gait than may be observed naturally, which could, in turn, influence the measurement of reliability and validity. Second, this study tested 6 different activity monitors, and, to our knowledge, 4 of the 6 (Fitbit Charge, Garmin vívofit 2, Jawbone UP2, and New-Lifestyles NL-1000) have not been previously tested in older adults. Unfortunately, because of fast product cycles for consumer-grade monitors, only the Misfit Shine and the New-Lifestyles NL-1000 are currently available for purchase; Fitbit advanced from the One to Inspire and from the Charge to Charge 4, Garmin replaced the vívofit 2 with vívofit 4, and Jawbone went out of business. Nevertheless, the strength of this study is that it presents systematic methods that other researchers can adopt or modify to evaluate the performance of current versions of consumer-grade activity monitors before their use in trials, observational studies, or surveillance systems. Finally, we studied older adults with self-reported mobility limitations, which is important because they are a relevant population for physical activity interventions and surveillance and comprise a sizable proportion of the older adult population.

Conclusions

The results of this study contribute to the growing literature on consumer-grade activity monitors in the older adult population. This study provides information about the test-retest reliability and criterion validity of step counting by several consumer-grade activity monitors in older adults with either self-reported intact or limited mobility. The results of this study may assist in the selection of an activity monitor for future studies designed to detect changes in physical activity levels, assess adherence to physical activity programs, quantify daily physical activity patterns (in conjunction with self-report questionnaires), or motivate physical activity behavior via goal setting in older adult study populations.

We found variations in step count test-retest reliability and criterion validity across 6 consumer-grade activity monitors when worn by a sample of older adults with self-reported intact and limited mobility. Walk interruptions increased the step count error for all monitors but did not affect any monitor to a greater extent than the others. The presence of self-reported mobility limitations did not affect the step count error. Only one monitor exhibited both high test-retest reliability and criterion validity, the hip-worn Fitbit One, and it is recommended for use in groups of older adults with self-reported intact and limited mobility.

Acknowledgments

The study received funding from the Natural Sciences and Engineering Research Council of Canada. SM was supported by a Graduate Fellowship from Simon Fraser University. SL holds the Pfizer/Heart and Stroke Foundation Chair in Cardiovascular Prevention Research at St. Paul’s Hospital. DM was supported by the Michael Smith Foundation for Health Research Scholar Award. The sponsors of the study had no role in the study design, data collection, data analysis, data interpretation, or writing of the report. The corresponding author had full access to all the data in the study and had the final responsibility for the decision to submit for publication. The authors would like to thank the study participants and research assistants for their time, effort, and enthusiasm.

Authors' Contributions

SM and DM conceived the study. SM, DC, and DM designed the study protocol. SM and DM coordinated and oversaw data collection. SM, DC, and DM conducted the statistical analysis. SM, DC, SL, and DM interpreted the data. SM created the tables and figures and wrote the first draft of the manuscript. DC, SL, and DM revised the manuscript critically for intellectual content. DC, SL, and DM provided scientific oversight and direction throughout the study. All authors approved the final version to be submitted for publication.

Conflicts of Interest

None declared.

References

1. Directly Measured Physical Activity of Adults, 2012 and 2013. Statistics Canada. URL: https://www150.statcan.gc.ca/n1/pub/82-625-x/2015001/article/14135-eng.htm [accessed 2019-03-01] [WebCite Cache ID 76YW64uI9]
2. Pedersen BK. The diseasome of physical inactivity--and the role of myokines in muscle--fat cross talk. J Physiol 2009 Dec 1;587(Pt 23):5559-5568 [FREE Full text] [doi: 10.1113/jphysiol.2009.179515] [Medline: 19752112]
3. Warburton DE, Charlesworth S, Ivey A, Nettlefold L, Bredin SS. A systematic review of the evidence for Canada's physical activity guidelines for adults. Int J Behav Nutr Phys Act 2010 May 11;7:39 [FREE Full text] [doi: 10.1186/1479-5868-7-39] [Medline: 20459783]
4. Warburton DE, Katzmarzyk PT, Rhodes RE, Shephard RJ. Evidence-informed physical activity guidelines for Canadian adults. Can J Public Health 2007;98(Suppl 2):S16-S68. [Medline: 18213940]

5. Katzmarzyk PT, Janssen I. The economic costs associated with physical inactivity and obesity in Canada: an update. Can J Appl Physiol 2004 Feb;29(1):90-115. [doi: 10.1139/h04-008] [Medline: 15001807]
6. Canadian Physical Activity Guidelines. For Older Adults - 65 Years & Older. Canadian Society For Exercise Physiology. URL: https://csepguidelines.ca/wp-content/uploads/2018/03/CSEP_PAGuidelines_older-adults_en.pdf [accessed 2019-02-28] [WebCite Cache ID 76YeK0w1b]
7. Tremblay MS, Warburton DE, Janssen I, Paterson DH, Latimer AE, Rhodes RE, et al. New Canadian physical activity guidelines. Appl Physiol Nutr Metab 2011 Feb;36(1):36-46; 47. [doi: 10.1139/H11-009] [Medline: 21326376]
8. Lauritzen J, Muñoz A, Sevillano JL, Civit A. The usefulness of activity trackers in elderly with reduced mobility: a case study. Stud Health Technol Inform 2013;192:759-762. [Medline: 23920659]
9. Disability in Canada: Initial Findings From the Canadian Survey on Disability. Statistics Canada. URL: https://www150.statcan.gc.ca/n1/pub/89-654-x/89-654-x2013002-eng.htm [accessed 2019-03-01] [WebCite Cache ID 76YX7mqI7]
10. Guralnik JM, Ferrucci L, Simonsick EM, Salive ME, Wallace RB. Lower-extremity function in persons over the age of 70 years as a predictor of subsequent disability. N Engl J Med 1995 Mar 2;332(9):556-561. [doi: 10.1056/NEJM199503023320902] [Medline: 7838189]
11. Guralnik JM, Simonsick EM, Ferrucci L, Glynn RJ, Berkman LF, Blazer DG, et al. A short physical performance battery assessing lower extremity function: association with self-reported disability and prediction of mortality and nursing home admission. J Gerontol 1994 Mar;49(2):M85-M94. [doi: 10.1093/geronj/49.2.m85] [Medline: 8126356]
12. Boisvert-Vigneault K, Payette H, Audet M, Gaudreau P, Bélanger M, Dionne IJ. Relationships between physical activity across lifetime and health outcomes in older adults: results from the NuAge cohort. Prev Med 2016 Oct;91:37-42. [doi: 10.1016/j.ypmed.2016.07.018] [Medline: 27471024]
13. Cyarto EV, Moorhead GE, Brown WJ. Updating the evidence relating to physical activity intervention studies in older people. J Sci Med Sport 2004 Apr;7(1 Suppl):30-38. [doi: 10.1016/s1440-2440(04)80275-5] [Medline: 15214599]
14. King AC, Rejeski WJ, Buchner DM. Physical activity interventions targeting older adults. A critical review and recommendations. Am J Prev Med 1998 Nov;15(4):316-333. [doi: 10.1016/s0749-3797(98)00085-3] [Medline: 9838975]
15. Conn VS, Valentine JC, Cooper HM. Interventions to increase physical activity among aging adults: a meta-analysis. Ann Behav Med 2002;24(3):190-200. [doi: 10.1207/S15324796ABM2403_04] [Medline: 12173676]
16. Conn VS, Minor MA, Burks KJ, Rantz MJ, Pomeroy SH. Integrative review of physical activity intervention research with aging adults. J Am Geriatr Soc 2003 Aug;51(8):1159-1168. [doi: 10.1046/j.1532-5415.2003.51365.x] [Medline: 12890083]
17. Müller AM, Khoo S. Non-face-to-face physical activity interventions in older adults: a systematic review. Int J Behav Nutr Phys Act 2014 Mar 10;11(1):35 [FREE Full text] [doi: 10.1186/1479-5868-11-35] [Medline: 24612748]
18. Ashworth NL, Chad KE, Harrison EL, Reeder BA, Marshall SC. Home versus center based physical activity programs in older adults. Cochrane Database Syst Rev 2005 Jan 25(1):CD004017 [FREE Full text] [doi: 10.1002/14651858.CD004017.pub2] [Medline: 15674925]
19. van der Bij AK, Laurant MG, Wensing M. Effectiveness of physical activity interventions for older adults: a review. Am J Prev Med 2002 Feb;22(2):120-133. [doi: 10.1016/s0749-3797(01)00413-5] [Medline: 11818183]
20. Colley RC, Garriguet D, Janssen I, Craig CL, Clarke J, Tremblay MS. Physical activity of Canadian adults: accelerometer results from the 2007 to 2009 Canadian health measures survey. Health Rep 2011 Mar;22(1):7-14 [FREE Full text] [Medline: 21510585]
21. Loprinzi PD. Light-intensity physical activity and all-cause mortality. Am J Health Promot 2017 Jul;31(4):340-342. [doi: 10.4278/ajhp.150515-ARB-882] [Medline: 26730555]
22. Tudor-Locke C, Lutes L. Why do pedometers work?: a reflection upon the factors related to successfully increasing physical activity. Sports Med 2009;39(12):981-993. [doi: 10.2165/11319600-000000000-00000] [Medline: 19902981]
23. Mercer K, Giangregorio L, Schneider E, Chilana P, Li M, Grindrod K. Acceptance of commercially available wearable activity trackers among adults aged over 50 and with chronic illness: a mixed-methods evaluation. JMIR Mhealth Uhealth 2016 Jan 27;4(1):e7 [FREE Full text] [doi: 10.2196/mhealth.4225] [Medline: 26818775]
24. Falck RS, McDonald SM, Beets MW, Brazendale K, Liu-Ambrose T. Measurement of physical activity in older adult interventions: a systematic review. Br J Sports Med 2016 Apr;50(8):464-470. [doi: 10.1136/bjsports-2014-094413] [Medline: 26276362]
25. Hopkins WG. Measures of reliability in sports medicine and science. Sports Med 2000 Jul;30(1):1-15. [doi: 10.2165/00007256-200030010-00001] [Medline: 10907753]
26. Evenson KR, Goto MM, Furberg RD. Systematic review of the validity and reliability of consumer-wearable activity trackers. Int J Behav Nutr Phys Act 2015 Dec 18;12:159 [FREE Full text] [doi: 10.1186/s12966-015-0314-1] [Medline: 26684758]
27. Ko S, Hausdorff JM, Ferrucci L. Age-associated differences in the gait pattern changes of older adults during fast-speed and fatigue conditions: results from the Baltimore longitudinal study of ageing. Age Ageing 2010 Nov;39(6):688-694 [FREE Full text] [doi: 10.1093/ageing/afq113] [Medline: 20833863]
28. Straiton N, Alharbi M, Bauman A, Neubeck L, Gullick J, Bhindi R, et al. The validity and reliability of consumer-grade activity trackers in older, community-dwelling adults: a systematic review. Maturitas 2018 Jun;112:85-93. [doi: 10.1016/j.maturitas.2018.03.016] [Medline: 29704922]

29. Alharbi M, Bauman A, Neubeck L, Gallagher R. Validation of Fitbit-Flex as a measure of free-living physical activity in a community-based phase III cardiac rehabilitation population. Eur J Prev Cardiol 2016 Sep;23(14):1476-1485. [doi: 10.1177/2047487316634883] [Medline: 26907794]
30. Boeselt T, Spielmanns M, Nell C, Storre JH, Windisch W, Magerhans L, et al. Validity and usability of physical activity monitoring in patients with chronic obstructive pulmonary disease (COPD). PLoS One 2016;11(6):e0157229 [FREE Full text] [doi: 10.1371/journal.pone.0157229] [Medline: 27305105]
31. Farina N, Lowry RG. The validity of consumer-level activity monitors in healthy older adults in free-living conditions. J Aging Phys Act 2018 Jan 1;26(1):128-135. [doi: 10.1123/japa.2016-0344] [Medline: 28595019]
32. Simpson LA, Eng JJ, Klassen TD, Lim SB, Louie DR, Parappilly B, et al. Capturing step counts at slow walking speeds in older adults: comparison of ankle and waist placement of measuring device. J Rehabil Med 2015 Oct 5;47(9):830-835 [FREE Full text] [doi: 10.2340/16501977-1993] [Medline: 26181670]
33. Floegel TA, Florez-Pregonero A, Hekler EB, Buman MP. Validation of consumer-based hip and wrist activity monitors in older adults with varied ambulatory abilities. J Gerontol A Biol Sci Med Sci 2017 Feb;72(2):229-236 [FREE Full text] [doi: 10.1093/gerona/glw098] [Medline: 27257217]
34. Storm FA, Heller BW, Mazzà C. Step detection and activity recognition accuracy of seven physical activity monitors. PLoS One 2015;10(3):e0118723 [FREE Full text] [doi: 10.1371/journal.pone.0118723] [Medline: 25789630]
35. Takacs J, Pollock CL, Guenther JR, Bahar M, Napier C, Hunt MA. Validation of the Fitbit One activity monitor device during treadmill walking. J Sci Med Sport 2014 Sep;17(5):496-500. [doi: 10.1016/j.jsams.2013.10.241] [Medline: 24268570]
36. Fulk GD, Combs SA, Danks KA, Nirider CD, Raja B, Reisman DS. Accuracy of 2 activity monitors in detecting steps in people with stroke and traumatic brain injury. Phys Ther 2014 Feb;94(2):222-229. [doi: 10.2522/ptj.20120525] [Medline: 24052577]
37. Nelson MB, Kaminsky LA, Dickin DC, Montoye AH. Validity of consumer-based physical activity monitors for specific activity types. Med Sci Sports Exerc 2016 Aug;48(8):1619-1628. [doi: 10.1249/MSS.0000000000000933] [Medline: 27015387]
38. Simonsick EM, Newman AB, Visser M, Goodpaster B, Kritchevsky SB, Rubin S, Health, Aging and Body Composition Study. Mobility limitation in self-described well-functioning older adults: importance of endurance walk testing. J Gerontol A Biol Sci Med Sci 2008 Aug;63(8):841-847 [FREE Full text] [doi: 10.1093/gerona/63.8.841] [Medline: 18772472]
39. Wolinsky FD, Miller DK, Andresen EM, Malmstrom TK, Miller JP. Further evidence for the importance of subclinical functional limitation and subclinical disability assessment in gerontology and geriatrics. J Gerontol B Psychol Sci Soc Sci 2005 May;60(3):S146-S151. [doi: 10.1093/geronb/60.3.s146] [Medline: 15860791]
40. MoCA Montreal - Cognitive Assessment. URL: https://www.mocatest.org/splash/ [accessed 2019-03-01] [WebCite Cache ID 76YXQoReg]
41. Nasreddine ZS, Phillips NA, Bédirian V, Charbonneau S, Whitehead V, Collin I, et al. The Montreal cognitive assessment, MoCA: a brief screening tool for mild cognitive impairment. J Am Geriatr Soc 2005 Apr;53(4):695-699. [doi: 10.1111/j.1532-5415.2005.53221.x] [Medline: 15817019]
42. PAR-Q+ The Physical Activity Readiness Questionnaire for Everyone. URL: http://eparmedx.com/wp-content/uploads/2013/03/PARQPlus2019ImageVersion2.pdf [accessed 2019-03-05] [WebCite Cache ID 76eeYe4Fv]
43. Groll DL, To T, Bombardier C, Wright JG. The development of a comorbidity index with physical function as the outcome. J Clin Epidemiol 2005 Jun;58(6):595-602. [doi: 10.1016/j.jclinepi.2004.10.018] [Medline: 15878473]
44. Case MA, Burwick HA, Volpp KG, Patel MS. Accuracy of smartphone applications and wearable devices for tracking physical activity data. J Am Med Assoc 2015 Feb 10;313(6):625-626. [doi: 10.1001/jama.2014.17841] [Medline: 25668268]
45. Bland JM, Altman DG. Measuring agreement in method comparison studies. Stat Methods Med Res 1999 Jun;8(2):135-160. [doi: 10.1177/096228029900800204] [Medline: 10501650]
46. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986 Feb 8;1(8476):307-310. [Medline: 2868172]
47. Paul SS, Tiedemann A, Hassett LM, Ramsay E, Kirkham C, Chagpar S, et al. Validity of the activity tracker for measuring steps in community-dwelling older adults. BMJ Open Sport Exerc Med 2015;1(1):e000013 [FREE Full text] [doi: 10.1136/bmjsem-2015-000013] [Medline: 27900119]
48. Thorup CB, Andreasen JJ, Sørensen EE, Grønkjær M, Dinesen BI, Hansen J. Accuracy of a step counter during treadmill and daily life walking by healthy adults and patients with cardiac disease. BMJ Open 2017 Mar 31;7(3):e011742 [FREE Full text] [doi: 10.1136/bmjopen-2016-011742] [Medline: 28363918]
49. How Does My FitBit Device Calculate My Daily Activity? Fitbit Help. URL: http://help.fitbit.com/articles/en_US/Help_article/1141/?q=accelerometer&l=en_US&fs=Search&pn=1 [accessed 2019-10-07] [WebCite Cache ID 76eerrlXI]
50. Tudor-Locke C, Barreira TV, Schuna JM. Comparison of step outputs for waist and wrist accelerometer attachment sites. Med Sci Sports Exerc 2015 Apr;47(4):839-842. [doi: 10.1249/MSS.0000000000000476] [Medline: 25121517]
51. Parvataneni K, Ploeg L, Olney SJ, Brouwer B. Kinematic, kinetic and metabolic parameters of treadmill versus overground walking in healthy older adults. Clin Biomech (Bristol, Avon) 2009 Jan;24(1):95-100. [doi: 10.1016/j.clinbiomech.2008.07.002] [Medline: 18976839]

Abbreviations

ANOVA: analysis of variance
CW: continuous walk
HSD: honest significant difference
IW: interrupted walk
RW: reliability walk
SEM: standard error of measurement
SPPB: Short Physical Performance Battery

Edited by G Eysenbach, K Fortuna; submitted 07.10.19; peer-reviewed by B Heller, S Kara; comments to author 23.04.20; revised version received 08.05.20; accepted 03.06.20; published 18.08.20.

Please cite as:
Maganja SA, Clarke DC, Lear SA, Mackey DC
Formative Evaluation of Consumer-Grade Activity Monitors Worn by Older Adults: Test-Retest Reliability and Criterion Validity of Step Counts
JMIR Form Res 2020;4(8):e16537
URL: http://formative.jmir.org/2020/8/e16537/
doi:10.2196/16537
PMID:32651956

Original Paper

A Mobile Health Intervention for Adolescents Exposed to Secondhand Smoke: Pilot Feasibility and Efficacy Study

Natalie Nardone[1], PhD; Jeremy Giberson[1], BS; Judith J Prochaska[2], PhD, MPH; Shonul Jain[3], MD; Neal L Benowitz[1,4,5], MD

[1]Clinical Pharmacology Research Program, Division of Cardiology, Department of Medicine, University of California, San Francisco, CA, United States

[2]Stanford Prevention Research Center, Department of Medicine, Stanford University, Stanford, CA, United States

[3]Department of Pediatrics, Zuckerberg San Francisco General Hospital, University of California, San Francisco, CA, United States

[4]Departments of Medicine, University of California, San Francisco, CA, United States

[5]Department of Bioengineering and Therapeutic Sciences, University of California, San Francisco, CA, United States

Corresponding Author:
Neal L Benowitz, MD
Clinical Pharmacology Research Program, Division of Cardiology
Department of Medicine
University of California
Building 30, Room 3316
1001 Potrero Ave
San Francisco, CA, 94110
United States
Phone: 1 628 206 8324
Email: neal.benowitz@ucsf.edu

Abstract

Background: Secondhand smoke (SHS) exposure in children and adolescents has adverse health effects. For adolescents of lower socioeconomic status (SES), exposure is widespread, evidenced in the measurement of urinary cotinine, a major metabolite of nicotine. Direct intervention with exposed children has been proposed as a novel method, yet there is minimal evidence of its efficacy. Combining this approach with a mobile health (mHealth) intervention may be more time and cost-effective and feasible for adolescent populations.

Objective: In this pilot study, we assessed the feasibility and preliminary evidence of efficacy of a 30-day text message–based mHealth intervention targeted at reducing SHS exposure in adolescent populations of low SES.

Methods: For the study, 14 nonsmoking and nonvaping participants between the ages of 12-21 years exposed to SHS were enrolled. The intervention consisted of a daily text message sent to the participants over the course of a month. Text message types included facts and information about SHS, behavioral methods for SHS avoidance, or true-or-false questions. Participants were asked to respond to each message within 24 hours as confirmation of receipt. Feasibility outcomes included completion of the 30-day intervention, receiving and responding to text messages, and feedback on the messages. Efficacy outcomes included a reduction in urinary cotinine, accuracy of true-or-false responses, and participants' perceptions of effectiveness.

Results: Of the 14 participants that were enrolled, 13 completed the intervention. Though not required, all participants had their own cell phones with unlimited text messaging plans. Of the total number of text messages sent to the 13 completers, 91% (372/407) of them received on-time responses. Participant feedback was generally positive, with most requesting more informational and true-or-false questions. In terms of efficacy, 54% (6/11) of participants reduced their cotinine levels (however, change for the group overall was not statistically significant (P=.33) and 45% (5/11) of participants increased their cotinine levels. Of the total number of true-or-false questions sent across all completers, 77% (56/73) were answered correctly. Participants' ratings of message effectiveness averaged 85 on a scale of 100.

Conclusions: In this pilot study, the intervention was feasible as the majority of participants had access to a cell phone, completed the study, and engaged by responding to the messages. The efficacy of the study requires further replication, as only half of the participants reduced their cotinine levels. However, participants answered the majority of true-or-false questions accurately and reported that the messages were helpful.

(JMIR Form Res 2020;4(8):e18583) doi:10.2196/18583

KEYWORDS

secondhand smoke; adolescents; cotinine; mHealth; intervention

Introduction

Secondhand smoke (SHS) exposure is associated with adverse health effects in children and adolescents, including respiratory disease and asthma [1-4]. In a study of adolescents of low socioeconomic status (SES) in San Francisco, California, 76% were found to have recent light or heavy SHS exposure based on biochemical screening of the major nicotine metabolite, cotinine, with ranges of 0.05-30 ng/mL [5]. In a follow-up study, these adolescents reported common exposure to SHS in public areas; however, reported SHS in homes and cars significantly predicted biochemical exposure [6]. Interventions targeted at reducing SHS may improve adolescent health.

Past interventions to reduce SHS exposure in children and adolescents have demonstrated mixed results. Many have focused on intervening on parental or caregiver smoking status, with a recent meta-analytic review demonstrating that some studies were effective at reducing SHS exposure while many others were not [7]. In addition, the focus on caregivers may not translate to patient care, as pediatricians are often reluctant to intervene on parental smoking [8].

Intervening directly with exposed children without the inclusion of cessation counseling for parents is a novel approach [9,10]. One proposed study for decreasing SHS exposure describes an educational intervention developed from behavioral change theory, consisting of three 40-minute educational sessions on the adverse effects of tobacco, international laws regarding tobacco, and methods to prevent tobacco exposure; weekly take-home brochures; and 3 reinforcement lessons [9]. A school-based study conducted in Bangladesh with children aged 10-12 years evaluated a similar method, consisting of two 45-minute educational sessions over a 2-day period followed by four 15-minute refresher sessions over the following month; saliva cotinine at the 2-month follow-up showed a significant difference between the intervention (0.53 ng/mL) and control groups (2.02 ng/mL) [10].

Delivery of SHS prevention educational interventions for youth via mobile health (mHealth) technologies may be more time and cost-effective and may be particularly acceptable to an adolescent population. MHealth interventions, such as sending daily text messages, have been developed utilizing behavioral [11] and social cognitive theories [12], and they have been effective for increasing physical activity among adolescents with attention deficit hyperactivity disorder (ADHD) [13], for promoting weight loss in adolescents with obesity [14], and for type 1 diabetes management in children and adolescents, with a focus on healthy eating and activity [11].

While mHealth interventions have been efficacious for a variety of indications [9,10,13], a downside to these methods may be low levels of engagement. Studies have provided participants with cell phones or prepaid phone cards to facilitate their participation [12,15]. Others have reported that technical problems led to a loss of data and affected the participants' levels of motivation to continue in the study [15-18]. Past studies have assessed feasibility as participant satisfaction and rates of participant retention [15,16]. In determining the efficacy of mHealth interventions, outcome measures are varied and include biomarker reductions, self-reported efficacy, and changes in disease or condition knowledge [11,13,14]. To date, mHealth interventions have not been developed to assist adolescents in reducing SHS exposure.

In this pilot study, we examined the feasibility and efficacy of a novel mHealth intervention to reduce measured SHS exposure in a small sample of urban adolescents of low SES. Feasibility was measured by intervention completion, receipt of text messages, participant responses to messages, and participant feedback on the quality and delivery of the messages. Efficacy was measured by reduction in urinary cotinine, SHS knowledge, and participant perception of the effectiveness of the messages for changing their behavior to reduce SHS exposures.

Methods

Participants

Participants of the ages of 12-21 years were recruited from July 2017 to May 2018. Recruitment occurred through (1) invitation from prior research studies [5,6], if the individual's cotinine level was within the heavy SHS range (urinary cotinine 0.25-30 ng/mL) during their prior study involvement; (2) social media ads on Facebook and Instagram; and (3) flyers posted in the Children's Health Center (CHC) at the Zuckerberg San Francisco General Hospital. We included participants over the age of 18 years, as the CHC services adolescent patients up to the age of 21 years. A screening measure assessed tobacco use and SHS exposure [6]. Self-reported active use of cigarettes, electronic cigarettes, or other tobacco products was an exclusion.

The University of California, San Francisco (UCSF), Institutional Review Board approved the research. Informed consent was obtained from the youths and from parents of those under the age of 18 years. Parents were asked if their adolescents had their own cell phones with unlimited text messaging services. The study team was prepared to provide a study phone on loan and gift cards for messaging fees, if applicable.

Measures

At baseline and the 30-day follow-up, the adolescents provided a urine sample for cotinine testing, and they self-reported their exposure to SHS across several environments for the past 7 days and the past 24 hours [6]. At the follow-up visit, through an online questionnaire, the adolescents were asked about their experiences with the intervention, with questions like "Have the text messages made you more active in reducing your contact with secondhand smoke?" They were also asked to rate how helpful the messages were (on a scale of 0-100), how they felt about receiving them, if they shared the messages with others,

what they liked most about the study, how the study could be improved, and why they may not have responded to the messages. Adolescents' age, sex, and race were recorded, and nicotine and tobacco product ever use were assessed. All questionnaires were created for this study using standard items where available.

Text Message Intervention

The intervention consisted of one text message sent per day via Outlook's SMS text message service, a free feature available as part of the Outlook package (version 16.0; Microsoft Corp). Messages were cued in Outlook to be sent from the research center's email address to the participants' phones at a time of their choosing. Participants were asked to respond to each message within 24 hours as confirmation of receipt. Text message categories included facts and information about SHS (eg, "Breathing secondhand smoke for a short period of time can hurt your body"), behavioral methods to avoid SHS exposure (eg, "Avoid SHS today by walking away from someone who is smoking"), and "bonus" true-or-false questions (eg, "True or False: SHS is annoying, but it's not really a health concern"). All messages included emojis, which were pretested to ensure compatibility with differing phone models. A total of 35 text messages were created to ensure that there were enough for additional days if participants were not able to return for their follow-up visit on day 30. A 30-day duration was selected as it was similar to prior interventions designed for an adolescent population [9,10]; in addition, 30 days is long enough to see a change in steady-state cotinine levels [19].

The primary sources for the information presented in the messages were websites from government agencies (http://cdc.gov/, http://smokefree.gov, http://cancer.gov) and the 2006 Surgeon General Report [1].

Analytical Chemistry

Baseline and follow-up urine samples were analyzed for cotinine by liquid chromatography–tandem mass spectrometry [20]. Cotinine, the main proximate metabolite of nicotine, has a half-life of about 16 hours [19] and is a biomarker of ongoing or recent exposure (past 5-6 days). The limit of quantitation (LOQ) for cotinine was 0.05 ng/mL. Cotinine cutpoints were 0.05-30 ng/mL for light or heavy SHS and >30 ng/mL for active smoking or vaping [5].

Statistical Analysis

We utilized descriptive statistics to summarize participant characteristics and rates of response to text messages. To assess changes in cotinine levels pre- and postintervention, we utilized a Wilcoxon signed rank test, a nonparametric equivalent to a paired sample *t* test. Biomarkers falling below the limit of quantitation (BLQ) were replaced with the LOQ divided by the square root of 2.

Results

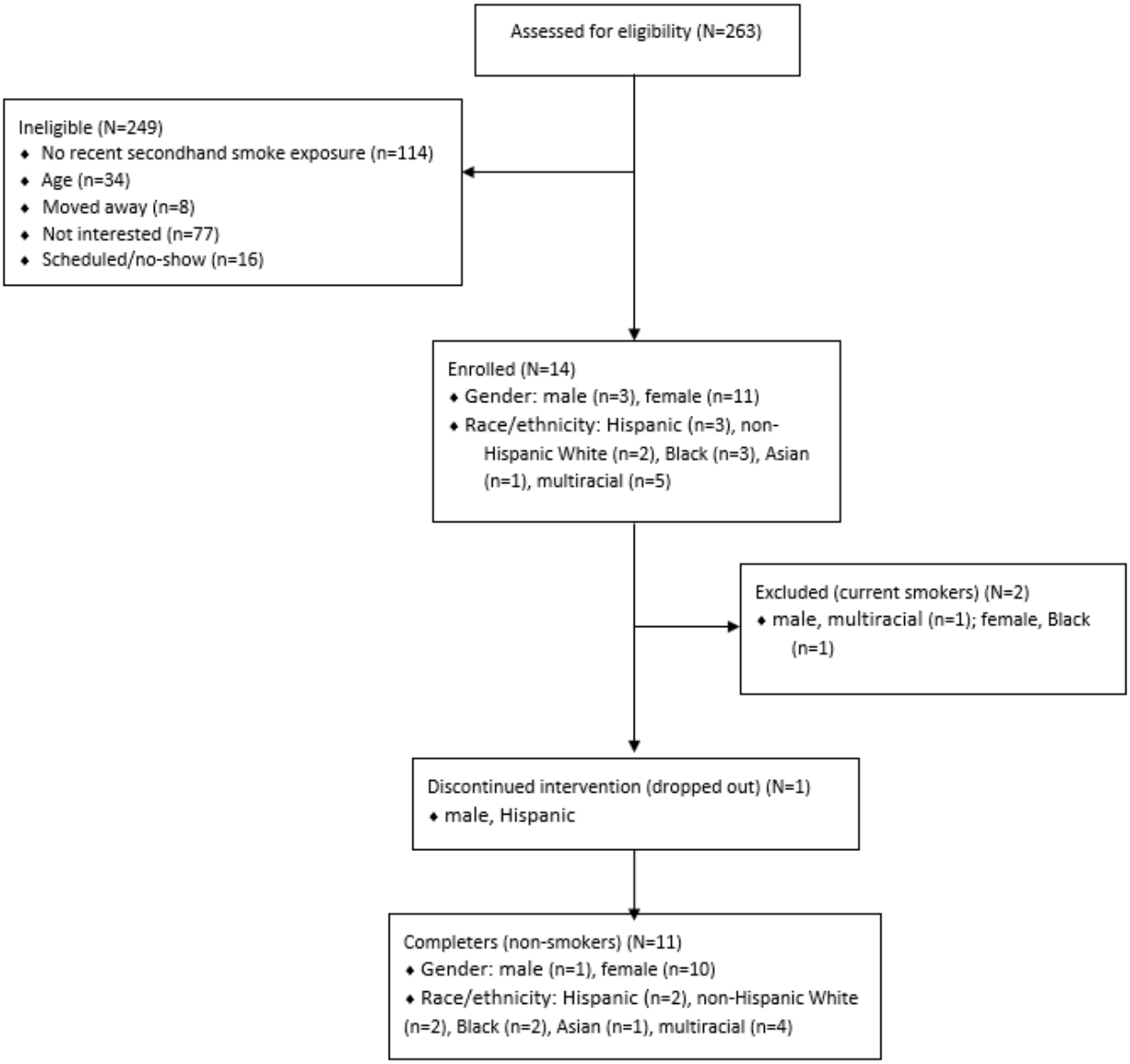
Demographics and SHS Exposure at Baseline

A total of 263 participants were screened, and 14 consented and were enrolled in the study. The primary reason for ineligibility was not meeting the criteria for recent SHS exposure (43%, 114/263); other individuals were ineligible for age (13%, 34/263), no longer lived in the area (3%, 8/263), or were not interested in participating (29%, 77/263). Some were eligible and scheduled for a baseline visit but did not arrive (6%, 16/263).

Of the 14 enrolled participants, 3 (21%) were male adolescents and 11 (79%) were female adolescents; when asked to self-report race/ethnicity, 3 (21%) were Hispanic, 2 (15%) were non-Hispanic White, 3 (21%) were Black, 1 (7%) was Asian, and 5 (36%) were multiracial. Participants' ages ranged from 12 to 20 years, with a mean of 17 (SD 2.39) years. All participants reported SHS exposure within the last 7 days, with the most highly reported environments being public areas (30%), residences (30%), and cars (21%).

Nicotine and other tobacco product use was reported as ever use of cigarettes (50%, 7/14), blunts (64%, 9/14), electronic cigarettes (43%, 6/14), pipes (36%, 5/14), cigars or cigarillos (21%, 3/14), or hookah (36%, 5/14). Figure 1 shows the consolidated standards of reporting trials (CONSORT) flow diagram.

Figure 1. Consolidated standards of reporting trials (CONSORT) flow diagram for participants recruited (N=263) and enrolled (n=14).



Intervention Feasibility

Of the 14 participants enrolled, 13 (93%) completed the intervention, as 1 participant dropped out midstudy. The remaining analyses focus on the 13 adolescents who completed the intervention (completers).

Due to variability in the scheduling of the follow-up visits and occasional missed messages due to Outlook shutdowns during routine software updates, the adolescents received a range of 19 to 34 text messages (mean 31.3, SD 7.71). On average, 91% (SD 12.61; range 55%-100%) of the messages received a reply from the participants within 24 hours of sending.

Participant feedback indicated that participants generally liked receiving the messages and would have wanted to see more, especially from the informational and true-or-false categories (Table 1). The majority (12/13) shared the messages with friends and family. One participant reported having technical difficulties.

Table 1. Participant feedback from the intervention completers (n=13).

Question & Response Options	Participant Response Selections
How have the text messages made you more active in reducing your contact with secondhand smoke? n (%)	
I have avoided people smoking around me.	11 (85)
I noticed more people smoking around me.	8 (62)
I asked friends not to smoke around me.	4 (31)
I stopped smoking myself.	1 (8)
I wasn't active in reducing my contact.	1 (8)
On a scale of 0-100 (0=not helpful, 100=very helpful), how would you rate the text messages in helping you remember the dangers of being in contact with secondhand smoke? Mean (SD)	85.8 (23)
On a scale from 0-100 (0=too many to read, 100=I would have liked more), how did you feel about getting the text messages? Mean (SD)	85.2 (20.2)
Did you talk to anyone else about the information in the text messages? n (%)	
I shared with both family and friends.	5 (38)
I shared with family only.	4 (31)
I share with friends only.	3 (23)
I didn't share any messages.	1 (8)
What did you like most about this study? n (%)	
Getting gift cards	12 (92)
Receiving text messages	9 (69)
What suggestions do you have for us to make this study better? n (%)	
Nothing	8 (62)
More true-or-false questions	2 (15)
More information on side effects of secondhand smoke	1 (8)
More text messages	1 (8)
Make sure there are no technical issues	1 (8)
Please let us know why you may have not responded to our messages; n (%)	
I didn't have my cell phone with me.	5 (38)
My phone service changed.	3 (23)
No particular reason.	3 (23)
I was too busy.	2 (15)

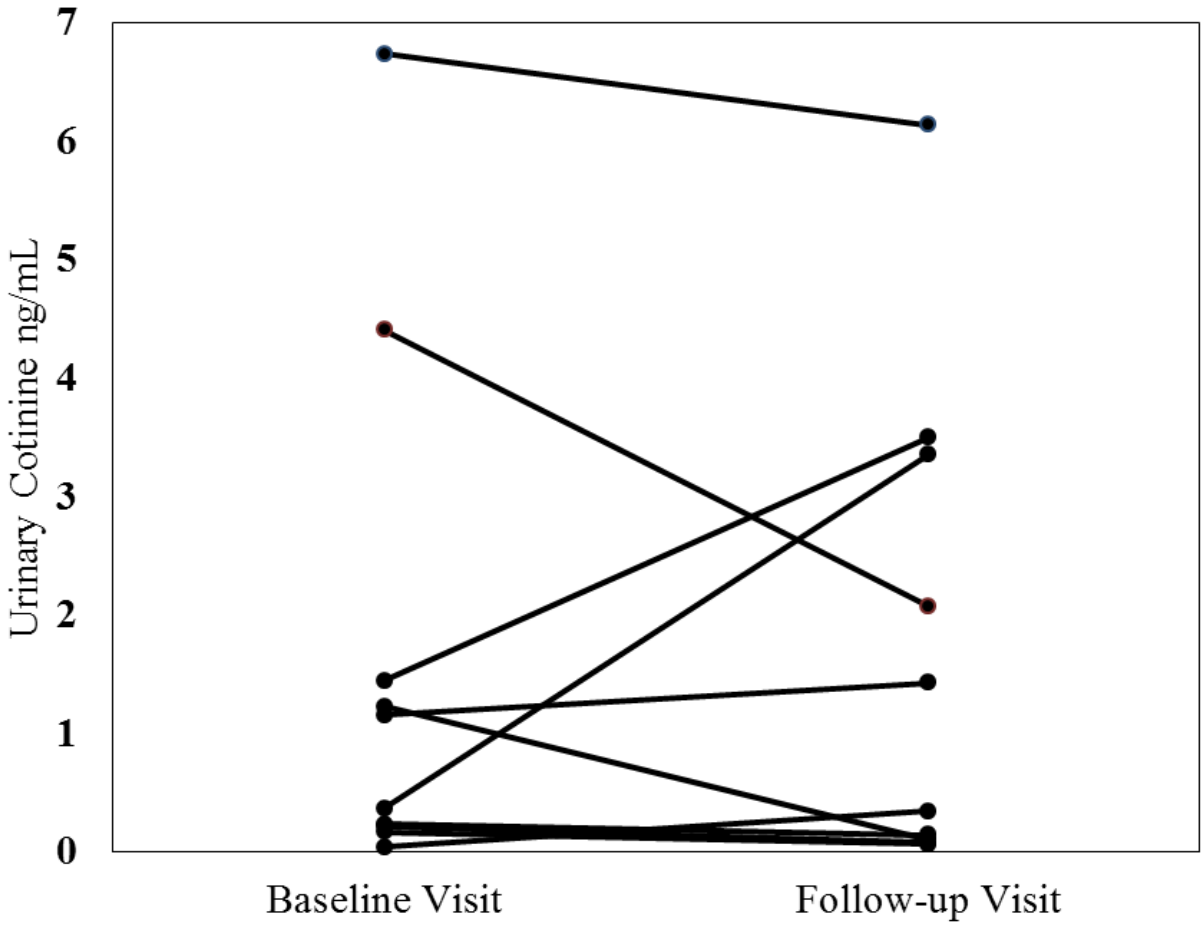
Intervention Efficacy

At baseline, despite denying recent tobacco use on the screener, 15% (2/13) of the participants had cotinine levels indicative of active smoking or vaping (283.1 ng/mL and 31.0 ng/mL). Further, 46% (6/13) of the participants had cotinine levels indicative of high SHS exposure (0.25-30 ng/mL), 31% (4/13) had cotinine levels indicative of low SHS exposure (range 0.05-0.25 ng/mL), and 8% (1/13) had BLQ cotinine levels. All further cotinine results reflect only the 11 completers of the intervention who were nonsmokers at baseline.

The geometric mean cotinine levels for the 11 completers were 0.48 ng/mL at baseline and 0.83 ng/mL at follow-up, including 1 participant who was nearing active smoking levels (cotinine=29.42 ng/mL) and 1 participant who was BLQ (cotinine=0.04 ng/mL). With the BLQ participant and the near-active smoking levels participant excluded, the geometric mean cotinine levels were 0.67 ng/mL at baseline and 0.09 ng/mL at follow-up.

Of the 11 completers, 5 had an increase in cotinine and 6 reduced their cotinine levels; 4 participants reduced their levels by >50%. The Wilcoxon signed rank test indicated no significant difference for the entire sample (P=.66). The 2 participants excluded as active smokers at baseline had cotinine levels consistent with active smoking at follow-up (473.8 and 43.8 ng/mL). Cotinine levels pre- and postintervention are shown in Figure 2, excluding the participant with near-active smoking levels and the participant who was BLQ at follow-up.

Figure 2. Urinary cotinine levels pre- and postintervention for nonsmoking intervention completers (n=9).



As a reflection of the gains made by participants in SHS knowledge, the accuracy of the true-or-false responses averaged 77% (SD 28; range 40% - 100%) across participants. Almost all the participants (12/13, 92%) perceived the messages to be helpful in reducing their SHS exposure. Avoiding people smoking was the most commonly endorsed strategy participants used to limit their SHS exposure. One participant said the messages did not encourage them to be active in reducing their SHS (Table 1).

Discussion

Principal Findings

In this pilot study, we sought to determine the feasibility and preliminary effectiveness of a 30-day mHealth intervention for adolescents to avoid SHS exposure. This was a novel approach to help adolescents limit their exposure to SHS, as we sought to intervene directly with the nonsmoking, SHS-exposed youth rather than their caregivers.

Our results support that mHealth interventions are feasible in an adolescent sample. Most participants completed the study, and all had technology accessible to them to receive the text messages. Nearly all participants remained in the study for the 30-day period. There was a high rate of response to the messages, and if participants did not respond, it was mostly due to not having their phones at the time the message was sent. Participants reported that they would have liked to see more messages, especially in the information and true-or-false categories.

The efficacy findings are equivocal. Approximately half of the sample reduced their cotinine levels while the other half increased. With many participants starting in the low range of SHS exposure, a floor effect may have occurred in that the lower limit did not allow for a significant change to be exhibited. The increased biochemical exposure may reflect that the intervention was not successful in impacting specific types of exposures, or the timing of exposures in conjunction with the baseline and follow-up urine collections could have varied. Participants scored well on the true-or-false messages, indicating they were reading and processing the information. The majority of participants rated the messages as effective in changing their behavior to avoid SHS; however, these behavioral changes did not result in a significant cotinine reduction for the sample overall.

Limitations

The limitations of this pilot study included a small and mostly female sample, a generally low level of SHS exposure, and recent tobacco use in 2 of the participants at baseline despite screening procedures. In addition, our study was conducted in a single geographic setting with progressive clean air laws but unavoidable SHS exposure in the urban environment. Although all participants reported exposure to SHS in the past 7 days, 1 (8%) participant did not meet the cotinine criteria for biochemical exposure.

Although the results of this study support the feasibility of conducting an mHealth intervention in adolescents, we cannot generalize the results to a broader population due to the small sample size. Moreover, we had only one follow-up visit at the end of the 30-day period, which did not allow us to evaluate the long-term effects of this intervention.

Due to time constraints on performing the assays, we were unable to analyze baseline cotinine levels prior to entrance in the study. Our study experienced periodic technical difficulties with occasional Outlook updates delaying or eliminating the deployment of text messages. We did not ask about nicotine or tobacco product use at the follow-up visit, so we cannot determine if some increases in cotinine were due to SHS exposure or to one's personal product use. Additionally, a few of our participants were over the age of 18 and may reflect more of a young adult population, with different levels of access to nicotine and tobacco products. At the time this study was conducted, the minimum tobacco sales age in California was 21 years, and all study participants were aged 20 years and younger.

Suggestions for Future Research

Researchers utilizing mHealth interventions in adolescent populations should consider providing interactive and informational messages. In addition, they should consider utilizing a follow-up period longer than 30 days to evaluate if there are long-term effects of the intervention. If applying this intervention strategy to limit SHS exposure, researchers should utilize an expired carbon monoxide monitor to exclude active tobacco users, or a dipstick cotinine test for active electronic cigarette users during the screening process. Researchers should ensure a reliable method for text message transmissions. Technical issues occurred utilizing Outlook; other text messaging services such as Twilio may be a more reliable deployment method.

Conclusions

In this pilot mHealth intervention, we found text messages to be a feasible method of encouraging avoidance of SHS exposure among adolescents. However, the sample did not demonstrate an overall significant reduction in biochemical exposure.

Acknowledgments

This work was supported by the Bland Lane Center of Excellence on Secondhand Smoke at the University of California, San Francisco, funded by the Flight Attendants Medical Research Institute. We thank Daisy Mendoza-Zepeda, Laura Carrillo, and the nurses in the Children's Health Center at Zuckerberg San Francisco General Hospital for clinical research assistance, and Trisha Mao, Lita Ramos, and Lawrence Chan for performing analytical chemistry.

Authors' Contributions

NB and NN designed the study, supervised the implementation of the study, and coordinated the research activities. NN and JG performed the analyses. NN and JG drafted the manuscript and incorporated feedback from coauthors. JJP and SJ provided intellectual input to the study design, and edited and provided feedback on versions of the manuscript. SJ assisted in the implementation and conduction of the study.

Conflicts of Interest

NB and JJP are consultants to pharmaceutical companies that market or are developing medications to aid smoking cessation, and have served as paid expert witnesses in litigation against tobacco companies. All other authors have no conflicts to declare.

References

1. Office on Smoking and Health (US). The Health Consequences of Involuntary Exposure to Tobacco Smoke: A Report of the Surgeon General. Atlanta (GA): Centers for Disease Control and Prevention (US); 2006.
2. Centers for Disease Control and Prevention (US), National Center for Chronic Disease Prevention and Health Promotion (US), Office on Smoking and Health (US). How Tobacco Smoke Causes Disease: The Biology and Behavioral Basis for Smoking-Attributable Disease: A Report of the Surgeon General. Atlanta (GA): Centers for Disease Control and Prevention (US); 2010.
3. Cuthbertson L, Britton J. Passive smoking and children's health. Clin Med 2010 Apr 01;10(2):113-114. [doi: 10.7861/clinmedicine.10-2-113]
4. Lang JE, Dozor AJ, Holbrook JT, Mougey E, Krishnan S, Sweeten S, et al. Biologic Mechanisms of Environmental Tobacco Smoke in Children with Poorly Controlled Asthma: Results from a Multicenter Clinical Trial. The Journal of Allergy and Clinical Immunology: In Practice 2013 Mar;1(2):172-180. [doi: 10.1016/j.jaip.2012.11.006]
5. Benowitz N, Jain S, Dempsey D, Nardone N, Helen G, Jacob P. Urine Cotinine Screening Detects Nearly Ubiquitous Tobacco Smoke Exposure in Urban Adolescents. Nicotine Tob Res 2017;19(9):1048-1054. [doi: 10.1093/ntr/ntw390]
6. Nardone N, Jain S, Addo N, St.Helen G, Jacob P, Benowitz NL. Sources and Biomarkers of Secondhand Tobacco Smoke Exposure in Urban Adolescents. Academic Pediatrics 2020 May;20(4):493-500. [doi: 10.1016/j.acap.2019.12.006]
7. Daly JB, Mackenzie LJ, Freund M, Wolfenden L, Roseby R, Wiggers JH. Interventions by Health Care Professionals Who Provide Routine Child Health Care to Reduce Tobacco Smoke Exposure in Children. JAMA Pediatr 2016 Feb 01;170(2):138-147. [doi: 10.1001/jamapediatrics.2015.3342]

8. Winickoff JP, Tanski SE, McMillen RC, Ross KM, Lipstein EA, Hipple BJ, et al. Acceptability of Testing Children for Tobacco-Smoke Exposure: A National Parent Survey. Pediatrics 2011;127(4):628-634. [doi: 10.1542/peds.2010-2462]
9. Rao A, B U, Mithra P, M N, Shenoy R, Rungta N. Effectiveness of a School-Based 'Tobacco Free' Intervention on Adolescents' Knowledge and Exposure to Second Hand Tobacco Smoke - A Multiphase Study. Asian Pac J Cancer Prev 2019 Dec 01;20(12):3533-3537. [doi: 10.31557/apjcp.2019.20.12.3533]
10. Siddiqi K, Huque R, Kanaan M, Ahmed F, Ferdous T, Shah S, et al. Children Learning About Secondhand Smoke (CLASS II): A Pilot Cluster Randomized Controlled Trial. Nicotine Tob Res 2019;21(5):670-677. [doi: 10.1093/ntr/nty090]
11. Herbert L, Owen V, Pascarella L, Streisand R. Text message interventions for children and adolescents with type 1 diabetes: a systematic review. Diabetes Technol Ther 2013 May;15(5):362-370. [doi: 10.1089/dia.2012.0291] [Medline: 23550554]
12. Franklin VL, Waller A, Pagliari C, Greene SA. A randomized controlled trial of Sweet Talk, a text-messaging system to support young people with diabetes. Diabet Med 2006 Dec;23(12):1332-1338. [doi: 10.1111/j.1464-5491.2006.01989.x] [Medline: 17116184]
13. Schoenfelder E, Moreno M, Wilner M, Whitlock KB, Mendoza JA. Piloting a mobile health intervention to increase physical activity for adolescents with ADHD. Preventive Medicine Reports 2017 Jun;6:210-213. [doi: 10.1016/j.pmedr.2017.03.003]
14. Vidmar AP, Pretlow R, Borzutzky C, Wee CP, Fox DS, Fink C, et al. An addiction model-based mobile health weight loss intervention in adolescents with obesity. Pediatric Obesity 2018 Aug 16;14(2):e12464. [doi: 10.1111/ijpo.12464]
15. Han Y, Faulkner MS, Fritz H, Fadoju D, Muir A, Abowd GD, et al. A Pilot Randomized Trial of Text-Messaging for Symptom Awareness and Diabetes Knowledge in Adolescents With Type 1 Diabetes. Journal of Pediatric Nursing 2015 Nov;30(6):850-861. [doi: 10.1016/j.pedn.2015.02.002]
16. Deacon AJ, Edirippulige S. Using mobile technology to motivate adolescents with type 1 diabetes mellitus: A systematic review of recent literature. J Telemed Telecare 2015 Dec;21(8):431-438. [doi: 10.1177/1357633X15605223] [Medline: 26377124]
17. Hood KK, Hilliard M, Piatt G, Ievers-Landis CE. Effective strategies for encouraging behavior change in people with diabetes. Diabetes Management 2015 Nov;5(6):499-510. [doi: 10.2217/dmt.15.43]
18. Konrath S, Falk E, Fuhrel-Forbis A, Liu M, Swain J, Tolman R, et al. Can Text Messages Increase Empathy and Prosocial Behavior? The Development and Initial Validation of Text to Connect. PLoS ONE 2015 Sep 10;10(9):e0137585. [doi: 10.1371/journal.pone.0137585]
19. Hukkanen J, Jacob P, Benowitz NL. Metabolism and Disposition Kinetics of Nicotine. Pharmacol Rev 2005 Feb 24;57(1):79-115. [doi: 10.1124/pr.57.1.3]
20. Jacob 3rd P, Yu L, Duan M, Ramos L, Yturralde O, Benowitz NL. Determination of the nicotine metabolites cotinine and trans-3′-hydroxycotinine in biologic fluids of smokers and non-smokers using liquid chromatography–tandem mass spectrometry: Biomarkers for tobacco smoke exposure and for phenotyping cytochrome P450 2A6 activity. Journal of Chromatography B 2011 Feb 01;879(3-4):267-276. [doi: 10.1016/j.jchromb.2010.12.012]

Abbreviations

ADHD: attention deficit hyperactivity disorder
BLQ: below the limit of quantitation
CHC: Children's Health Center
LOQ: limit of quantitation
mHealth: mobile health
SES: socioeconomic status
SHS: secondhand smoke

Edited by G Eysenbach; submitted 05.03.20; peer-reviewed by S Hugh-Jones, J Alvarez Pitti; comments to author 27.04.20; revised version received 22.06.20; accepted 14.07.20; published 19.08.20.

Please cite as:
Nardone N, Giberson J, Prochaska JJ, Jain S, Benowitz NL
A Mobile Health Intervention for Adolescents Exposed to Secondhand Smoke: Pilot Feasibility and Efficacy Study
JMIR Form Res 2020;4(8):e18583
URL: http://formative.jmir.org/2020/8/e18583/
doi:10.2196/18583
PMID:32812888

Viewpoint

A Digital Health Intervention for Weight Management for Latino Families Living in Rural Communities: Perspectives and Lessons Learned During Development

Zenong Yin[1], PhD; Vanessa L Errisuriz[2], PhD; Martin Evans[2], PhD; Devasena Inupakutika[3], PhD; Sahak Kaghyan[3], PhD; Shiyu Li[1], MS; Laura Esparza[2], MS, MCHES; David Akopian[3], PhD; Deborah Parra-Medina[2], PhD, MPH

[1]Department of Kinesiology, Health, and Nutrition, The University of Texas at San Antonio, San Antonio, TX, United States

[2]Latino Research Institute, The University of Texas at Austin, Austin, TX, United States

[3]Department of Electrical and Computer Engineering, The University of Texas at San Antonio, San Antonio, TX, United States

Corresponding Author:
Deborah Parra-Medina, PhD, MPH
Latino Research Institute
The University of Texas at Austin
210 W 24th St
Mailcode F9200
Austin, TX, 78712
United States
Phone: 1 512 471 4557
Email: parramedina@austin.utexas.edu

Abstract

Rural residents face numerous challenges in accessing quality health care for management of chronic diseases (eg, obesity, diabetes), including scarcity of health care services and insufficient public transport. Digital health interventions, which include modalities such as internet, smartphones, and monitoring sensors, may help increase rural residents' access to health care. While digital health interventions have become an increasingly popular intervention strategy to address obesity, research examining the use of technological tools for obesity management among rural Latino populations is limited. In this paper, we share our experience developing a culturally tailored, interactive health intervention using digital technologies for a family-oriented, weight management program in a rural, primarily Latino community. We describe the formative research that guided the development of the intervention, discuss the process of developing the intervention technologies including issues of privacy and data security, examine the results of a pilot study, and share lessons learned. Our experience can help others design user-centered digital health interventions to engage underserved populations in the uptake of healthy lifestyle and disease management skills.

(JMIR Form Res 2020;4(8):e20679) doi:10.2196/20679

KEYWORDS

mhealth; digital intervention; Latino families; rural population; weight; self-management; diet; lifestyle; chronic disease

Introduction

The ownership of cellular phones, smartphones, and tablet computers has reached historic heights across all US population groups in recent years [1]. In 2018, at least 75% of Blacks/African Americans, Latinos, and non-Latino Whites and 65% of rural residents owned a smartphone. Advances in information technology (IT), along with the diminishing digital divide among population groups, have increased the use of digital technologies in health promotion and disease management [2]. As a result, individuals from low-income and racial/minority backgrounds, who traditionally had less access to health information, are now able to more easily access health information via Internet with mobile devices [3,4]. Diffusion of digital technologies has also popularized the use of digital health interventions to reach rural and resource-poor communities with limited access to traditional tools and programs for health promotion and disease prevention and management [5,6].

In the realm of digital health, Internet- and mobile-based IT platforms (eg, SMS text messages, websites, social media, wearable technology) have been used to store, manage, and transmit data using hardware, software, and networks for delivering health care services and health promotion programs

[7]. Mobile health (mHealth) technology is an important subset of digital health using mobile phones, tablets, apps, and other wireless technologies. Compared with traditional face-to-face health care delivery, digital health interventions have demonstrated great promise in enhancing program reach and participant engagement, and achieving similar effectiveness in health-related outcomes. [7,8]. However, mHealth interventions often limit the number of behavior change techniques (BCTs) and rarely offer problem solving, social support, and didactic education [9,10]. Engagement of historically underserved populations in digital health interventions is influenced by a variety of factors, such as the extent to which program content is tailored and formatted, the alignment between devices used for program delivery and population needs [6], and the appropriateness of program context [11]. Additional research indicates that personalization of content, contact with program users [12], and preference of digital devices to receive content are also important indicators of program engagement [13]. For example, Latino users are more likely to stop using health-related mobile apps due to access issues (eg, rely primarily on mobile phones versus tablets or computers), lack of interest in the content, and high cost [14,15]. The development of digital health interventions must address these barriers to increase participation among underserved populations.

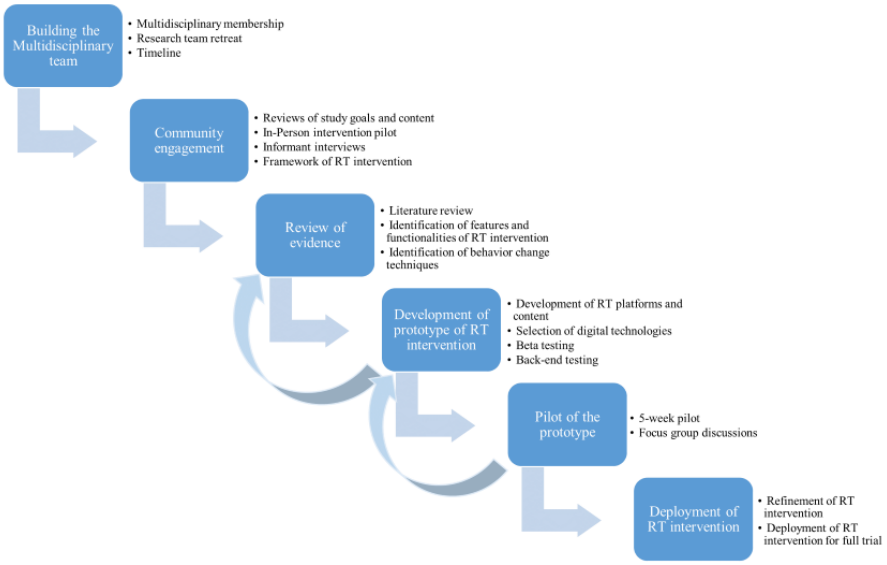
Healthy Frio is a translation study of the Y-Living Program (Y-Living), an evidence-based, family-focused intervention designed for urban Latino families (Trial Registration: ClinicalTrials.gov NCT03186885) [16]. The translation of Y-Living is a two-phased study guided by the principles of community-based participatory research [17]. The first phase of the study engaged community partners in formative research to adapt Y-Living for rural Latino families. Formative research included (1) a pilot of an in-person (IP) group program at a community center, and (2) a pilot of a home-based program remotely delivered by digital technologies. While we designed both forms of program delivery to address unique social, cultural, and environmental factors facing rural Latino families, the latter takes advantage of digital technologies to increase program availability, accessibility, and engagement. More specifically, the objectives of using remote delivery were to (1) increase access to evidence-based health education content and resources; (2) address learning needs with content design; (3) address the need for individualized support, and (4) support behavior change with wearable technologies grounded in behavior change theory. The second phase of the study is a 12-month randomized controlled trial (RCT) that tests the comparative effectiveness of the Y-Living IP and remote technology (RT) interventions on weight loss and energy balance behaviors among overweight and obese rural Latino adults relative to a control group. Following baseline assessments and randomization, intervention participants receive a 12-week IP or RT program, whereas control participants receive 1 brief health education session and educational materials. Study outcomes will be assessed again immediately following the intervention at 3 months, 6 months after intervention, and 12-month follow-up.

In this paper, we describe the process undertaken by a multidisciplinary research team to develop the RT intervention for the Healthy Frio study. Key elements of the process include the formative research that guided the development of the RT intervention, the selection and integration of digital technologies [18], and a pilot study. Our experience and lessons learned can help inform decisions other researchers may face when designing user-centered digital health interventions to engage underserved populations in the uptake of healthy lifestyle and disease management skills.

Overview of the Development Process

The development of the RT intervention took place over 16 months, from the formation of the research team to the deployment of the RT intervention for the full RCT. The process included several steps: (1) formulating the research team, (2) engaging community partners and residents to adapt the Y-Living for rural Latino families, (3) reviewing the evidence on use of digital technologies for behavior change, (4) developing a prototype of the RT intervention in parallel with IP intervention, and (5) conducting a pilot test of the prototype. Figure 1 shows the development process of Healthy Frio RT intervention.

Figure 1. Healthy Frio remote technology (RT) intervention development process.



Formulation of RT Development Team

A multidisciplinary team including researchers with expertise in computer engineering, database and instructional design, and public and behavioral health, and primary care practitioners (community/family medicine and nursing) developed the RT intervention. The instructional design specialist was designated as the coordinator. The technology modules and components integrated as a part of the RT intervention required development of necessary links, incorporation of application program interfaces (APIs), service interfaces, dashboards, and testing of interoperability, compatibility, and robustness. Thus, digital health RT interventions demand the expertise of a professional team of computer engineers to develop and manage seamless functioning and efficient intervention delivery in collaboration with behavioral health researchers who understand behavior change processes. A team science approach was critical in the communication and collaboration between computer engineers and behavioral health researchers and the development and management of the multidisciplinary team [19]. Ultimately, this approach helps in scaling and disseminating effective interventions. During the first year of development, the RT team met biweekly via conference call to discuss and develop the intervention and then met monthly after the first year. Additionally, the team met IP for 2 full-day retreats over the 16-month development period.

Community Engagement

Y-Living is a 12-week behavioral modification program grounded in the Social Cognitive Theory of behavior change to engage the whole family in lifestyle changes [20]. During this program, participants develop knowledge and skills to be physically active and eat healthy, set goals for changes, self-monitor their physical activity and diet, and create a supportive environment at home. Program components of Y-Living include semiweekly health education lessons; semiweekly group exercise sessions; self-monitoring of weight, physical activity, and diet; family wellness consultations; and special events. Trained Young Men's Christian Association (YMCA) staff and volunteers deliver the program activities in YMCA facilities. The Y-Living program was designed for urban Latino families. To translate Y-Living into an efficacious program for rural Latino families, the Healthy Frio intervention team, including the YMCA wellness coordinator, conducted an extensive review of Y-Living program activities and sought feedback from community partners regarding feasibility of program activities for a rural environment. The Healthy Frio intervention team reviewed and updated health education lessons to reflect current, evidence-based recommendations for weight-loss, physical activity, and diet that addressed the needs of rural populations. Some lessons were omitted and replaced with new topics that more closely aligned with behavior change strategies identified in the Social Cognitive Theory and the Coventry, Aberdeen, and London-Refined taxonomy [21,22], such as goal-setting, self-monitoring, feedback, rewards, social support, coaching, problem solving, and action planning. A 5-week pilot with 5 parent–child dyads was used to test the feasibility of the adapted IP intervention in the community. Table 1 displays the Healthy Frio intervention components for both the IP and RT interventions and demonstrates program alignment. The intervention team also conducted key informant interviews with community stakeholders to gather information on access and availability to physical activity and healthy eating in the community, availability of digital technologies and supportive resources in the community, and level of health technology literacy. The intervention team reviewed program content and delivery approaches prior to the deployment of the RT intervention in the pilot study and the full trial.

Table 1. Description of and alignment between in-person and remote technology intervention.

Mediators	Strategies	In-person intervention	Remote technology intervention
• Self-regulation	• Initial wellness report	• Provided at first wellness consultation. • Includes height, weight, BMI, average PA[a] minutes, and caloric intake recorded at baseline assessments. • A YMCA[b] wellness coach discusses with each participant and helps him/her set goals to increase PA and lower calorie intake, and set a target weight loss goal (5%) for 3-month assessment.	• Provided at first wellness consultation. • Includes height, weight, BMI, average PA minutes, and caloric intake recorded at baseline assessments • Health educator discusses with each participant and helps him/her set goals to increase PA and lower calorie intake, and set a target weight loss goal (5%) for 3-month assessment.
• Self-efficacy • Perceived barriers • Social support	• Orientation session	• Participants receive program materials (eg, program binder, self-monitoring logs) and information about the program schedule and goals.	• Participant is provided, and given instructions on how to use, all equipment (eg, tablet, Fitbit), apps, timeline, program goals, and expectations.
• Home environment • Self-efficacy • Self-regulation	• Healthy living lessons	• 12 weeks of lessons at community center: 2 lessons per week.	• 12 weeks of online lessons delivered via tablet computers: 2 lessons per week.
• Self-efficacy • Perceived barriers	• Cooking demonstrations	• Brief cooking demonstration by YMCA staff following health education lessons.	• Video recordings of cooking demonstrations conducted by YMCA staff.[c]
• Home environment • Self-efficacy • Self-regulation • Perceived barriers • Social support	• Physical activity sessions	• Immediately following Y Living lessons, a 1-hour PA session led by YMCA wellness coach. Examples of activities include circuit training, Zumba, and games from CATCH[d] curriculum. Children and adults are usually not separated.	• PA sessions (for both children and adults) are included in online lessons using videos curated from YouTube, as well as 5 videos produced by Healthy Frio and YMCA wellness coaches. • Separate videos are provided that are oriented toward adults and children.
• Self-efficacy • Self-regulation • Perceived barriers	• Self-monitoring	• Paper and pencil log for PA and food intake. • Weigh-ins during individual consultation.	• Use of Fitbit, scale, apps, etc.
• Self-efficacy • Self-regulation	• SMS text messaging	• 2 per week. • Healthy Lifestyle tips for increasing PA and healthy eating.	• Lesson reminders, health challenges, same healthy lifestyle tips as in-person, polling question regarding health challenges.
• Social support • Perceived barriers	• Grocery store tour	• Participants taken to a local grocery store where they are led on a tour and given tips for healthy shopping (eg, comparing labels, prices, and suggested routes through grocery store [perimeter first]).	• Virtual grocery store tour lesson using video segments for each section of grocery store.
• Home environment • Social support	• Wellness consultations	• Small group discussions and coaching sessions.	• Group video calls via Google Hangouts[e] or Skype (weeks 2, 6, and 10).
• Home environment • Self-efficacy • Social support	• One-on-one meetings	• 3 individual consultations between YMCA staff and participants to review progress report (weeks 2, 5, and 8).	• Biweekly phone calls with each family, and a midprogram progress report sent via email in week 7.

[a]PA: physical activity.

[b]YMCA: Young Men's Christian Association.

[c]Cooking demos performed at Y-Living Program in San Antonio (not part of Healthy Frio Trial).

[d]CATCH: Coordinated Approach to Child Health.

[e]Google Hangouts used for Cohort 1. Skype will be used for remaining cohorts.

State-of-the-Art Review and Integration of Research Literature

As the third step of the development process, the RT team conducted a review of the literature on the application of digital technologies in health promotion and disease management from the perspectives of health promotion practices and computer engineering. Furthermore, the team conducted a state-of-the-art review on modern health data collection and sharing approaches in the context of an emerging gig economy [23]. The goal was to understand the progress, best practices, and challenges in digital health intervention and provide guidance in designing a theory-driven RT intervention.

Application of digital health technologies has evolved from a single IT platform that focused on offering access to web-based intervention information to multiple platforms that extended exposure to the intervention in the context of the targeted behaviors. Research indicates that exposure is extended by (1) minimizing time-consuming data input and output activities, (2) automating delivery of task reminders and reinforced messages, (3) providing frequent and real-time feedback to intervention participants and staff, (4) offering the ability to tailor the intervention based on baseline information as well as ongoing changes of the behavior targets and the environment, and (5) leveraging the interactive capabilities of mHealth devices to incorporate BCTs [10,24]. By contrast, lack of structured social interactions and social support was a noted weakness of current digital technologies [23,24]. For example, studies have demonstrated that incorporating IP contact enhanced program compliance and participant engagement in mHealth interventions [25,26]. Early mHealth interventions have also been criticized for not using behavioral theories to inform the program design [7].

The RT team identified key features of successful digital health interventions such as two-way communication between the participants and intervention, social support from intervention staff and others, and some forms of competition [7,27,28]. For example, a systematic review found that successful interactive health IT systems for self-management in older, chronically ill, and underserved adults had a *completed feedback loop*. This feedback loop monitored the patient's health status, provided the progress toward patient's goals based on the monitoring, facilitated the adjustment of goals and management plans, and offered tailored advice and recommendations [29]. Using multiple IT platforms can increase intervention exposure and support for participants [27,28]. The RT team also identified important barriers influencing uptake of targeted health behaviors, including the program's ability to address community interests and cultural context, perceived effectiveness of the intervention, difficulty accessing and navigating the systems (eg, interface design), reliability of the systems (eg, technical issues), ability to deliver the program using technologies familiar to participants, the extent to which intervention activities fit into daily routines, and timely and frequent contact with intervention staff [7,29,30].

Finally, contemporary health behavior theories and models have been developed primarily based on face-to-face interactions for the management of behavior changes and do not address the time-intensive and interactive nature and asynchronous and dynamic process of digital health interventions [31]. While the design of the content of digital health interventions has benefited from these theories and models, the design of delivery and feedback processes remains poorly guided. This weakness speaks about the lack of cooperation between computer system engineers and behavioral researchers [22,31] and the need to establish a model to guide the incorporation of BCTs into the content and functionality of digital health interventions [32].

Based on the results of the review, the RT team decided to adopt a user-centered system design [33,34] utilizing an iterative, participatory process to address learning goals, communication styles, and community context. This system design required participation from instructional designers, computer engineers, behavioral scientists, and end users (interventionist and participants). Applying the principles of a user-centered approach has contributed to increases in acceptance and usability of digital technology-based interventions [35,36]. From a computer engineering perspective, the RT team concluded that a blended approach, combining digital technologies with limited human high-touch, may be most efficacious to advance the field of digital health interventions while awaiting the next generation of artificial intelligence-based solutions to address weaknesses associated with inability to provide personalized social support [23,37].

The Development of Prototype for Delivery of the RT Intervention

Delivery Platform for RT Intervention

The design of the Healthy Frio RT intervention was guided by a blended approach to create a program that is "both useful and easy to use, ie, at really serving the needs of the user," following the principles of the persuasive design system (PDS) [38,39]. The RT team reached a consensus early on in the development process to implement the 4 features of PDS (primary task support, dialogue support, system credibility, and social support) by blending numerous IT platforms with the facilitation of a live lifestyle coach to take advantage of digital technologies and the desire of human touch in supporting behavior changes [37,40]. Table 2 presents the operationalization of PDS techniques in alignment with the Social Cognitive Theory in Healthy Frio RT intervention. There are compromises between the rigorousness of the program content and the limitations of technology due to costs and scalability [38]. Figure 2 depicts the conceptual framework of RT intervention development that integrates PDS technologies to address moderators of self-efficacy to influence targeted health behaviors and outcomes, based on the Social Cognitive Theory [20]. The

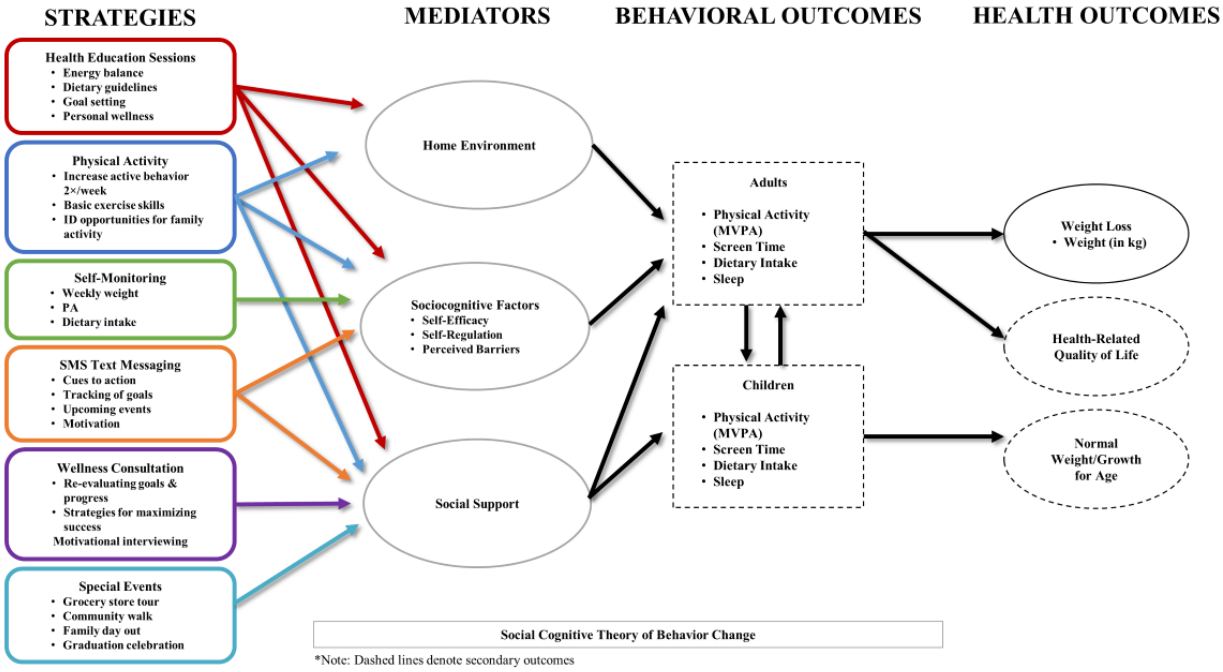
combination of digital technologies and live coaching affords the most efficient support to engage the study participants with BCTs [22] and close the feedback loop while providing constant exposure to the intervention content [29]. The Healthy Frio intervention team reviewed the content and delivery approaches of the RT intervention prior to its deployment for the pilot study and the full trial.

Table 2. Implementation of features and techniques for designing the content and functionality of a persuasive design system.

PDS[a] features	Operationalization of PDS techniques in Healthy Frio digital health intervention
Primary task support: Supporting and facilitating completion of the user's main tasks	Guided goal-setting; self-monitoring of weight, physical activity, and diet; using simulation and rehearsal in online interactive health education; online health challenges to target specific behavior
Dialogue support: Providing support, guidance, and feedback to users by verbal or other form	Suggestion and reminders via SMS text messaging; praise and reinforcement via health app (Fitbit app), and live online coaching
System credibility support: Presenting program content as trustworthy and authoritative to users	Embedding spirituality in online health education; using respected and authoritative sources; connecting program content to local community; affiliating the program with trusted local entities
Social support: Motivating and engaging users by leveraging social influence	Using parent–child dyad; providing physical activity trackers and weight tracking accounts to dyads; weekly calls by lifestyle coach

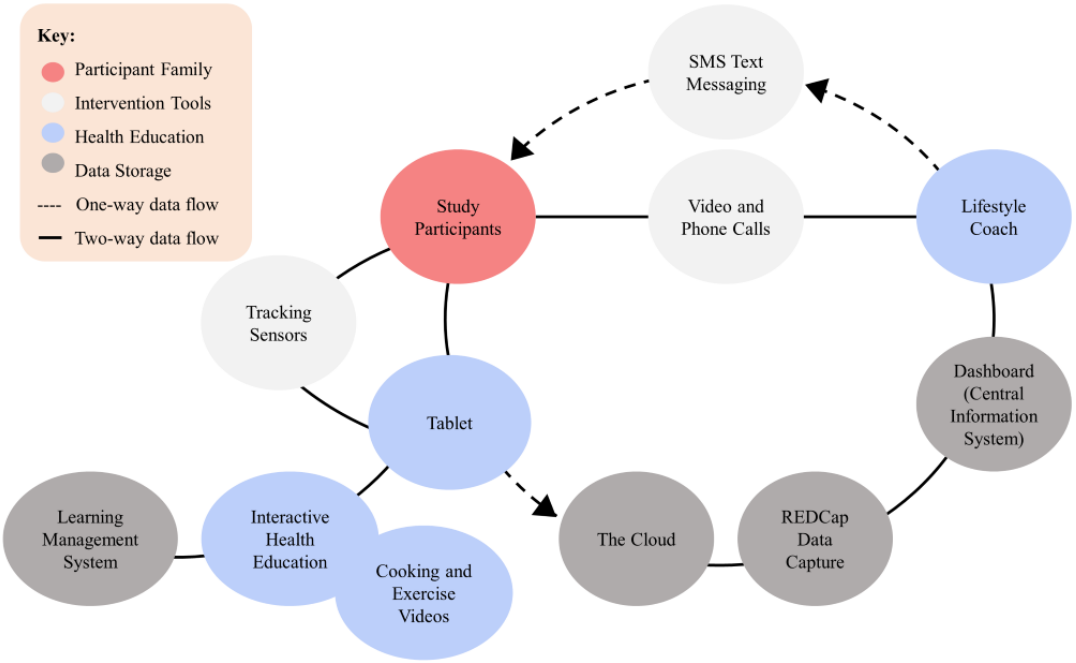
[a]PDS: persuasive design system.

Figure 2. Healthy Frio remote technology intervention conceptual model. MVPA: moderate to vigorous physical activity; PA: physical activity.



The implementation of the RT intervention is accomplished by a multiple-system integrated platform that is designed to address 3 elements of behavior change for persuasive design in human–computer interactions: (1) motivating participants with goal setting, positive reinforcement, and social support; (2) increasing participant's ability to adopt new behaviors by learning behavior change skills; and (3) providing cues to initiate and maintain new behavior via SMS text message reminders and feedback from digital technologies [38]. The platform included interactive, online health education lessons delivered via tablet; tracking sensors (activity tracker and Bluetooth weight scale) and apps to self-monitor activity, diet, and weight; social support and consultations in the form of phone calls, video chats, and motivation and support from automated SMS text messages; cooking and exercise videos; and a dashboard to allow the lifestyle coach to monitor and support participants (Figure 3). Embedding tools (ie, the Cloud, REDCap [Research Electronic Data Capture] data capture, the learning management system) into the platform to gather information on participant access and progress toward completing intervention activities allowed the lifestyle coach to assess participant engagement and success (eg, achievement of weight and behavior goals) [11,41].

Figure 3. Healthy Frio remote technology intervention delivery platform.



Creation of Interactive Health Education Content in RT Intervention

Once Y-Living health education lessons were updated for the rural context, the RT team developed the interactive version of the lessons. The design and delivery of the lesson content were tailored to address the learning needs and the cultural and community characteristics of low-income Latino families [6]. We used avatars resembling study participants (eg, Latino adults) to present content, modified activities to reduce numeracy and literacy levels, and used narration and graphics to enhance content comprehension and retention. We also selected activities to reflect community context, limited the amount of information on each screen to reduce overload, and only included information from trusted resources within the lesson [42,43]. Finally, we used gamification to increase motivation and engagement.

E-learning authoring software Articulate Rise [44] and Articulate Storyline [45] were used to create the health education content. Articulate Rise was relatively easy to use and was well-suited for rapid deployment; however, we learned that there were limited options for customizing the lesson content to be interactive and address intervention needs revealed in the RT pilot study. As a result, the RT team decided to use Articulate Storyline to enhance lesson customization and participants' interactive learning experience. The RT team converted PowerPoint slides that were created for the IP intervention as a basis for lesson content to develop each of the 24 RT lessons (2 lessons/week). Lesson titles remained the same, as did the lesson topics. RT intervention lessons followed the same structure as those for the IP intervention whereby each began with a motivational or spiritual message (ie, "Higher Thought"), followed by a recap of the previous lesson, the objectives for the current lesson, a pretest, lesson content, a posttest, and finally, physical activity videos. After completing a lesson, the participant is prompted to evaluate the lesson content and provide feedback on what they "liked" or "didn't like," and comments that could be used for improving the lessons for future versions.

The RT team developed interactive activities within each lesson to increase participant's knowledge and self-efficacy to change physical activity and diet behaviors. Examples of interactive skill-building activities include creation of SMART (ie, specific, measurable, adjustable, realistic, time-based) goals for physical activity and dietary intake, practice reading *nutrition facts* food labels to assist in selection of healthier choices when grocery shopping, and identifying healthier choices when ordering food at a restaurant. All lessons were narrated to increase participants' engagement with the lessons. By narrating the lesson content, the amount of text that the participants had to read from the screen was greatly reduced, and the information could be provided in more depth. The intervention staff wrote and recorded narration scripts for each lesson.

To provide the RT intervention participants an experience similar to that of the IP participants, the RT team video recorded physical activity sessions and cooking demonstrations conducted by Y-Living program staff from the YMCA. These videotaped activities were similar to the ones used with participants in the Healthy Frio IP program. Videos included Zumba, salsa dancing, boot camp style-physical activity routines, 22 cardiovascular fitness routines, and demonstrations for healthy ways to cook proteins vegetables, fats, and grains. The RT version included additional curated videos that demonstrated various physical activity routines available on YouTube. The criteria for selecting the videos were that they were primarily suitable for beginners, did not require any equipment, included a variety of physical activities for both strength training and cardiorespiratory fitness (ie, yoga, stretching, body weight exercises), and were fun. Each

lesson included 2 physical activity videos: 1 for adult participants, and 1 that was more suitable for children. Narrated instructions introduced each video.

Selection and Testing of Digital Technologies for Delivery of Healthy Frio RT Intervention

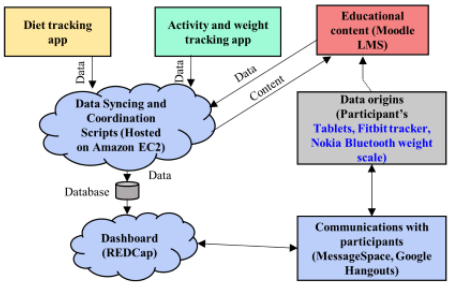
The RT team underwent a tedious process of testing, developing, and integrating various hardware and software tools that would be supportive and cost-efficient for delivery of the RT intervention. Major considerations included (1) ease of use with user-friendly interface and minimum data entry burden, (2) product reliability (hardware durability and signal strength), and (3) ability to collect information from the tools via web-based API [7,29]. Table 3 lists the hardware and software utilized for delivering the intervention with descriptions of functions. With the exception of the smartphone, all of the equipment is provided to participants for a 12-week intervention period. Following the completion of the intervention, participants return the tablet computer and the weight scale, but allowed to keep their Fitbits and receive a digital weight scale.

Table 3. Technology components and functions used by remote technology intervention.

Technology component	Function
Hardware	
Verizon Ellipsis 10 tablet computer with mobile Wi-Fi hotspot and data plan	Intervention content delivery device to access the interactive health lessons; gathers and uploads information from monitoring devices
Fitbit Flex 2 activity tracker	Personal sensor to track physical activity and sleep
Nokia Body digital scale	Bluetooth scale to track weight
Smartphones (participants' personal)	Device used to view SMS text messages from the study
Software	
MessageSpace	Automated SMS text messaging system to send motivational texts and reminders to participants
Moodle Learning Management System	Delivery of intervention content; tracks participant viewing of program lessons
Fitbit and Nokia Health Mate apps	Apps to help participants self-monitor physical activity, food intake, water consumption, sleep, and weight. Weight syncing through Nokia Health Mate app.
Google Hangouts/Skype	Video conferencing apps to conduct group counseling calls with health educator
MX Player	Video player to view physical activity videos
Electronic games	Games related to physical activity and nutrition for child participants

Figure 4 shows the system architecture constituting the selected digital technologies and their integration. Such architecture enables successful data collection and communication with participants during the RT intervention. The green components integrated in the system interface with Google Hangouts [46] or Skype [47] are for lifestyle coaching while collecting and processing data and communicating with the participants. These advanced messaging tools, when connected to REDCap [48], provide a communication channel to interact with participants. Data synchronization is typically done using a Bluetooth wireless connection with a smartphone or computer app. The participants can set their specific goals as reference marks and challenge themselves to achieve higher goals. Many of these apps allow tracker owners to connect with peer groups or friends, and some allow sharing data with other users. We chose not to use this function because the RT intervention is not group based. Some tools also have an option to connect to other third-party, variable-specific (eg, calorie counting or meal tracking app) tracking apps (eg, Fitbit [49]) that allow for tracking dietary intake. There exists significant value in observing diet trends over time and understanding the correlation between calorie expenditure and activity levels. During our evaluation study for activity trackers, we observed that Fitbit trackers offer the convenience of tracking users' calorie intake, sleep duration, steps, miles, and activity minutes for the day.

Figure 4. Healthy Frio remote technology intervention system architecture. EC2, Elastic Cloud Compute; LMS, learning management system; REDCap, Research Electronic Data Capture.



To provide access to educational content during the intervention, we used Moodle as the learning management system to deliver program content, manage health education lessons, and track progress [50]. In other words, the lessons (content) created on Articulate Storyline were uploaded to Moodle for participants to access. Completed content is hosted on Amazon Elastic Cloud Compute (EC2) server [51] and is accessible to participants via a Moodle app on the tablets. Data from the aforementioned services are collected and retrieved in real time using corresponding manufacturer's APIs (Fitbit, Moodle SQL database on EC2, and Nokia [52]) and reside in the database hosted on Amazon EC2 server. Data synchronization and coordination scripts take care of this process. Data are further postprocessed and are accessible to the lifestyle coach and intervention team on the Dashboard supported by REDCap. Alternatively, one can use Microsoft Azure [53] or Google Cloud [54] as the centralized server for incoming data from the components involved in the integrated digital health system. Table 4 presents lessons learned and design implications to address various design goals that are deemed necessary to engage participants and strengthen the human touch in the intervention.

Table 4. Summary of goals, design implications, and lessons learned in the design of remote technology intervention.

Design goal	Lesson learned	Design implication
• To investigate whether usage of the tablet by parent and child affects user engagement and frequency of regular synchronizing of hardware devices with respective apps on the tablet.	• Parents and children felt a burden with regard to synchronizing their activity, entering dietary data and weight, due to repeated log-in and logout from their respective accounts on the apps. This extra step led to very few, if any, syncs of their data.	• Parents and children should be able to synchronize their data without the hassle of logging in and out. We created clones of the apps that are used by both parents and children, and their devices are configured accordingly to their cloned version and accounts. This improved data synchronizing and we are able to get frequent updates on their syncs.
• To check the ease of access to educational content and health knowledge via Moodle, enhance access to information, get proper feedback for improving the content.	• Navigating the Moodle app is challenging as there are many steps to take before landing on the course page to access lessons. This dissuaded participants from accessing weekly lessons. • Due to poor network connectivity, downloading of lessons from the app is inhibited or is slow, further dissuading participants from accessing the lessons.	• The app's configuration settings on the Amazon EC2[a] server as well as on the tablet should be changed. With the configuration changes, participants are able to directly see the list of lessons upon opening the Moodle app. • Providing offline access to content (ie, via SD[b] cards, download of complete 12 weeks of content from Moodle app onto the tablet) should resolve the network-related issue affecting access to lessons. This is essential when internet access is poor.
• To investigate whether regular one-on-one communication can fill the gaps due to missing data collected remotely.	• Due to diverse technical capabilities of participants in the intervention, they tend to forget some steps in constantly synchronizing the devices, accessing the content, and entering nutrition details.	• SMS text messages as reminders, polling questions, surveys, and one-on-one video calls with health educator help remind participants about various aspects of the intervention, which in turn helps collect a more robust set of data.
• To observe active logging of water consumption and tracking calories by entering dietary details and calories consumed.	• Logging food and consumed calories requires manually entering or scanning the food item. Participants mostly either skip entering it or do it less often than required.	• Regular SMS text messages as reminders and one-on-one video calls with Health Educator motivates and helps participants log food intake, keep track of their daily consumption, and stick to weight goals.
• To evaluate tools and methods to enhance tablet usage with ease of navigation for regular synchronizing of data as well as for constant motivation and participant engagement.	• Direct access to apps with little to no information to enter manually, easy methods of synchronizing their devices to the apps, automatic notifications from Fitbit (including device lights and vibrations), Health Mate, and Moodle apps seem to motivate participants to monitor their healthy living.	• Tools for the intervention should be as integrated as closely as possible with the participants' daily schedule. User-centered design will support end-user engagement, improve user's experience, and encourage attentiveness toward the intervention.

[a]EC2: Elastic Cloud Compute.

[b]SD: secure digital.

Limitations of Using Commercial Products in Research: Access, Security, and Privacy

The potential for digital tools such as mobile apps and wearable sensors to enhance behavioral health is substantial. However, the use of tablets, wearable trackers, and other technologies for digital health interventions introduces additional complexity for participant's data security and privacy [55]. One should carefully evaluate security issues, conforming to rules and regulations in the settings where these interventions are implemented [23]. Participants signed a consent form, which

informs them about the heterogeneous RT environment and the type of data collected.

All the constituent apps in the Healthy Frio RT intervention have different security aspects. The current RT intervention delivery leverages Fitbit, Nokia Health Mate, and Moodle apps, which rely on remote servers for storing and processing the participant's data, and adds to security challenges. Hence, the computer engineers and developers need to consistently secure servers, the transmission of data, and the involved software through their proper configuration. To address these issues, we performed a thorough evaluation of the complete RT intervention delivery infrastructure by observing the data and sensitive information flowing through multiple apps, modules that we developed leveraging corresponding software development kits, libraries, and APIs. This helps to anticipate potential security issues in order to take preventative measures. The third-party libraries and code have been assessed for any security vulnerabilities or other reported problems. Other safety practices include observing the differences between different APIs that enabled us to configure the apps properly for the security-related features and handle the permissions for data collection. This research helped us in adapting the code base accordingly to avoid compromising the security of participants' information. We also generated participants' credentials for the involved apps securely and ensured the passwords are not stored in the remote server in plaintext. We avoided storing their passwords and instead let them authorize our system to collect their activity, weight, and educational content access data prior to the start of the intervention. Finally, we used Amazon EC2 to deploy and maintain our remote server that communicates with the apps. While Amazon provides a secure environment including privacy protection, we took appropriate measures to configure the server for Fitbit, Nokia Health Mate, and Moodle course data-retrieval systems.

Pilot Study

The prototype of the Healthy Frio RT intervention was tested with 5 parent–child dyads in a 3-week pilot, using content from the first 3 weeks of the intervention (ie, 6 lessons). The purposes were to (1) test the performance, management, and maintenance of digital equipment—especially with regard to connectivity and syncing of Fitbit trackers and weight scales; (2) obtain feedback about the design of the lessons; and (3) obtain feedback on the experience of engaging in a remotely delivered lifestyle modification program with digital technologies. Dyads were recruited from a convenience sample of eligible participants from the study population. The adult participants were all females aged 36 to 47. The ages of the children ranged from 7 to 14 years; 4 of the 5 families self-identified as Latino. Some of the participants had previously participated in an early IP intervention pilot. Therefore, they were able to provide feedback on the delivery of the RT intervention in comparison to the IP format.

The week prior to the start of the pilot study, participants came to an orientation session in which they were given an overview of the RT intervention, received their equipment as well as an instructional manual. The orientation session included a description of the 3-week pilot intervention, demonstrations of the use of equipment (eg, syncing Fitbit trackers and weight scales to tablets), introduction to apps installed on the tablet that would be used for self-monitoring, and a tutorial on accessing health education content on the tablet. RT pilot participants engaged in 3 focus group discussions in a community center (1 per week). Each family received a US $60 gift card for each focus group they attended. A focus group guide was developed prior to the pilot study that included questions related to the orientation session, lesson content and format, weekly SMS text messages received, use of equipment, and participants' overall experience. Participants' responses were transcribed in field notes which were then analyzed using content analysis [56]. RT team members decided to focus on manifesting content and develop categories using an inductive approach (ie, deriving categories directly from the text data). One member of the RT team was responsible for reviewing field notes, immersing herself in the data by reading through the notes several times. She then noted categories emerging from the data in an Excel file and shared them with other members of the RT team for verification. Several categories emerged from the data including hands-on training, user-friendliness (equipment), user-friendliness (apps), lesson content and aesthetics, and family engagement. These categories formed the basis for lessons learned from our pilot study.

Hands-On Training

Participants stated that the orientation session needed more time and many wanted more one-on-one time with staff members to guide them to use the equipment. Participants wanted to spend more time going through the apps on the tablet and more demonstrations on how to log-in/log out of the apps for syncing the equipment—especially because 2 people were using the apps (parent–child dyad). Many participants suggested creating a hotline number that they could call for help troubleshooting any problems experienced.

User-Friendliness (Equipment)

A majority of the participants stated that the tablet was not very user-friendly. For instance, many found the tablet bulky and difficult to carry around. Participants indicated that they preferred a smaller tablet that they could carry around with them (in their purse). However, participants also stated that they really liked using the tablet to view the exercise videos that were provided during the second part of each online lesson. Most participants stated that they wore the Fitbit all the time (except when charging) and like using it to keep track of their steps and sleep.

User-Friendliness (Apps)

Participants expressed that there were "too many applications or things to do." In particular, participants did not like having to log-in/log out of the apps, for instance, to sync their child's data in addition to their own in the Fitbit app. Participants also had a hard time navigating within the Moodle app to open and view the interactive lessons. Participants also had difficulty downloading the lessons to view. This issue was likely related to their cellular data connection speed.

Lesson Content and Aesthetics

Overall, the participants stated that they liked the lesson content, including the interface design, colors, and layout. Several participants requested SMS text message reminders to view the lessons because it was "easy to put off."

Family Engagement

Though participants were encouraged to view the lessons with their child, most participants viewed the lessons during the evening or on the weekends and did not include their child. However, they stated that they shared the relevant information with their child later.

Lessons Learned

Based on the findings from our pilot study, we learned that all of the equipment should be presynced and apps ready to be used prior to the orientation session. For the pilot study, when the participants were given the equipment during the orientation session, staff had to help them sync their Fitbit activity trackers and scales to the tablet for both the parent and child. Because the trackers and scales connect with the tablet using Bluetooth, this proved to be challenging due to multiple devices being in close proximity to one another. It was especially difficult for less technologically savvy participants. Additionally, both the parent and the child were using the same Fitbit app on each tablet, requiring them to login and logout each time they wanted to sync their equipment. This issue became a problem that caused their data to sometimes get mixed up, especially with the weight data from the scales. "Too many devices in one space using Bluetooth to sync Fitbits and scales to tablets" as one participant reported. We solved this issue by creating clones of the Fitbit app so that two copies of the app could be installed on the tablet: 1 for the parent and 1 for the child.

Many participants expressed that they had issues downloading the lessons to view. This issue was likely related to their cellular data connection speed. Because the participants lived in rural, south Texas, their cellular connectivity was less than optimal. Having a weaker cellular signal resulted in significantly long times for each lesson to download to the tablet, and caused the videos that were embedded into each of the lessons to often not be able to play. To solve these issues, we embedded all of the videos within the lessons rather than stream the lessons from their original source (primarily YouTube). This required having to download all of the videos from YouTube, and then importing them into the lessons. To address the issue of the lengthy time for downloading the lessons, which was exacerbated once videos were embedded due to larger file size, all of the lessons were predownloaded to the tablet prior to the equipment being given to the participants. It should be noted that making these changes resulted in the amount of data to be stored on the tablet hard drive exceeding the available space (16 GB). Therefore, we installed a micro SD (secure digital) data card (32 GB) into each tablet.

By addressing issues participants had syncing their equipment to the tablet and viewing online lessons, we were able to address many of the challenges that occurred during the pilot. For instance, we were able to provide more time during the orientation session so participants have more hands-on training in using the equipment and apps and we enhanced the user-friendliness of the tablet, Fitbits, and apps.

Conclusion

We designed a digital technology–based intervention to deliver a lifestyle modification program commonly delivered IP, by a lifestyle coach. The combination of digital technologies and live coaching affords the most efficient support to engage the study participants and close *the feedback loop* while providing constant exposure to the intervention content [29]. Digital health interventions are promising, but sometimes demonstrate mixed effectiveness [6,57], often due to low participant engagement and high study attrition [27,57] and a lack of integration of BCTs into interventions [32]. To address these concerns, we designed the Healthy Frio RT intervention by considering both the behavioral and psychological theories as well as systems engineering models [31,58]. The latter was critical in formulating technologies for efficient delivery and dynamic feedback in the RT intervention. We also incorporated key features identified in successful digital health interventions (eg, live lifestyle coaching) to enhance the efficacy of the RT intervention [28]. The RT development team was able to overcome many challenges through close collaboration of a multidisciplinary team consisting of behavioral health experts, computer scientists, an instruction specialist, and primary care providers. Finally, we embedded tools to collect information on participant adherence and engagement as part of the RT intervention. We hope that sharing our development process and lessons learned will help other researchers understand the factors influencing behavior changes in digital health interventions and guide the development of the next generation of digital health interventions.

Acknowledgments

This work was supported by the National Institute of Nursing Research of the National Institutes of Health under Award Number R01NR016269. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Authors' Contributions

All authors wrote, reviewed, and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Pew Research Center Internet & Technology Mobile Fact Sheet. Pew Research Center. 2018. URL: http://www.pewinternet.org/fact-sheet/mobile/ [accessed 2020-08-09]
2. Pérez-Stable EJ, Jean-Francois B, Aklin CF. Leveraging Advances in Technology to Promote Health Equity. Med Care 2019 Jun 8;57 Suppl 6 Suppl 2(32-33):S101-S103. [doi: 10.1097/MLR.0000000000001112] [Medline: 31095045]
3. Fox S, Duggan M. Pew Internet (2015-01-30). Health Online. 2013. URL: http://www.pewinternet.org/2013/01/15/health-online-2013/ [accessed 2020-08-09]
4. Vangeepuram N, Mayer V, Fei K, Hanlen-Rosado E, Andrade C, Wright S, et al. Smartphone ownership and perspectives on health apps among a vulnerable population in East Harlem, New York. Mhealth 2018;4:31. [doi: 10.21037/mhealth.2018.07.02] [Medline: 30221166]
5. Mallow JA, Theeke LA, Barnes ER, Whetsel T, Mallow BK. Using mHealth Tools to Improve Rural Diabetes Care Guided by the Chronic Care Model. Online J Rural Nurs Health Care 2014;14(1):43-65 [FREE Full text] [doi: 10.14574/ojrnhc.v14i1.276] [Medline: 26029005]
6. Anderson-Lewis C, Darville G, Mercado RE, Howell S, Di Maggio S. mHealth Technology Use and Implications in Historically Underserved and Minority Populations in the United States: Systematic Literature Review. JMIR Mhealth Uhealth 2018 Jun 18;6(6):e128 [FREE Full text] [doi: 10.2196/mhealth.8383] [Medline: 29914860]
7. Sawesi S, Rashrash M, Phalakornkule K, Carpenter JS, Jones JF. The Impact of Information Technology on Patient Engagement and Health Behavior Change: A Systematic Review of the Literature. JMIR Med Inform 2016 Jan 21;4(1):e1 [FREE Full text] [doi: 10.2196/medinform.4514] [Medline: 26795082]
8. World Health Organization. WHO Guideline: Recommendations on Digital Interventions for Health System Strengthening. Geneva, Switzerland: World Health Organization; 2019.
9. Pagoto S, Bennett GG. How behavioral science can advance digital health. Transl Behav Med 2013 Sep;3(3):271-276 [FREE Full text] [doi: 10.1007/s13142-013-0234-z] [Medline: 24073178]
10. Pagoto S, Schneider K, Jojic M, DeBiasse M, Mann D. Evidence-based strategies in weight-loss mobile apps. Am J Prev Med 2013 Nov;45(5):576-582. [doi: 10.1016/j.amepre.2013.04.025] [Medline: 24139770]
11. Perski O, Blandford A, West R, Michie S. Conceptualising engagement with digital behaviour change interventions: a systematic review using principles from critical interpretive synthesis. Transl Behav Med 2016 Dec 13. [doi: 10.1007/s13142-016-0453-1] [Medline: 27966189]
12. Lentferink AJ, Oldenhuis HK, de GM, Polstra L, Velthuijsen H, van GJE. Key Components in eHealth Interventions Combining Self-Tracking and Persuasive eCoaching to Promote a Healthier Lifestyle: A Scoping Review. J Med Internet Res 2017 Aug 01;19(8):e277 [FREE Full text] [doi: 10.2196/jmir.7288] [Medline: 28765103]
13. Granger D, Vandelanotte C, Duncan MJ, Alley S, Schoeppe S, Short C, et al. Is preference for mHealth intervention delivery platform associated with delivery platform familiarity? BMC Public Health 2016 Jul 22;16:619 [FREE Full text] [doi: 10.1186/s12889-016-3316-2] [Medline: 27450240]
14. Krebs P, Duncan DT. Health App Use Among US Mobile Phone Owners: A National Survey. JMIR Mhealth Uhealth 2015;3(4):e101 [FREE Full text] [doi: 10.2196/mhealth.4924] [Medline: 26537656]
15. Chang E, Blondon K, Lyles CR, Jordan L, Ralston JD. Racial/ethnic variation in devices used to access patient portals. Am J Manag Care 2018 Jan 01;24(1):e1-e8 [FREE Full text] [Medline: 29350513]
16. Parra-Medina D, Liang Y, Yin Z, Esparza L, Lopez L. Weight Outcomes of Latino Adults and Children Participating in the Y Living Program, a Family-Focused Lifestyle Intervention, San Antonio, 2012-2013. Prev Chronic Dis 2015 Dec 10;12:E219 [FREE Full text] [doi: 10.5888/pcd12.150219] [Medline: 26652219]
17. Strong LL, Israel BA, Schulz AJ, Reyes A, Rowe Z, Weir SS, et al. Piloting interventions within a community-based participatory research framework: lessons learned from the healthy environments partnership. Prog Community Health Partnersh 2009;3(4):327-334 [FREE Full text] [doi: 10.1353/cpr.0.0094] [Medline: 20097994]
18. 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:. Health Educ Behav 2006 Feb;33(1):25-39 [FREE Full text] [doi: 10.1177/1090198105282412] [Medline: 16397157]
19. Stokols D, Hall KL, Taylor BK, Moser RP. The science of team science: overview of the field and introduction to the supplement. Am J Prev Med 2008 Aug;35(2 Suppl):S77-S89. [doi: 10.1016/j.amepre.2008.05.002] [Medline: 18619407]
20. Bandura A. Health promotion by social cognitive means. Health Educ Behav 2004 Apr;31(2):143-164. [doi: 10.1177/1090198104263660] [Medline: 15090118]
21. Hagger M, Keatley D, Chan D. CALO-RE taxonomy of behavior change techniques. In: Eklund RC, Tenenbaum GT, editors. Encyclopedia of sport and exercise psychology. Thousand Oaks: Sage Publications, Inc; 2014:99-104.

22. Sullivan AN, Lachman ME. Behavior Change with Fitness Technology in Sedentary Adults: A Review of the Evidence for Increasing Physical Activity. Front Public Health 2016 Jan 11;4:289 [FREE Full text] [doi: 10.3389/fpubh.2016.00289] [Medline: 28123997]
23. Morales J, Inupakutika D, Kaghyan S, Akopian D, Yin Z, Evans M, et al. Technology-based health promotion: Current state and perspectives in emerging gig economy. Biocybern Biomed Eng 2019 Jul;39(3):825-842 [FREE Full text] [doi: 10.1016/j.bbe.2019.07.006] [Medline: 32313347]
24. 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]
25. Rodearmel SJ, Wyatt HR, Barry MJ, Dong F, Pan D, Israel RG, et al. A family-based approach to preventing excessive weight gain. Obesity (Silver Spring) 2006 Aug;14(8):1392-1401 [FREE Full text] [doi: 10.1038/oby.2006.158] [Medline: 16988082]
26. Catenacci V, Barrett C, Odgen L, Browning R, Schaefer CA, Hill J, et al. Changes in physical activity and sedentary behavior in a randomized trial of an internet-based versus workbook-based family intervention study. J Phys Act Health 2014 Feb;11(2):348-358 [FREE Full text] [doi: 10.1123/jpah.2012-0043] [Medline: 23364318]
27. Morrison LG. Theory-based strategies for enhancing the impact and usage of digital health behaviour change interventions: A review. Digit Health 2015;1:2055207615595335. [doi: 10.1177/2055207615595335] [Medline: 29942544]
28. Morrison LG, Yardley L, Powell J, Michie S. What design features are used in effective e-health interventions? A review using techniques from Critical Interpretive Synthesis. Telemed J E Health 2012 Mar;18(2):137-144. [doi: 10.1089/tmj.2011.0062] [Medline: 22381060]
29. Jimison H, Gorman P, Woods S, Nygren P, Walker M, Norris S, et al. Barriers and drivers of health information technology use for the elderly, chronically ill, and underserved. Evid Rep Technol Assess (Full Rep) 2008 Nov(175):1-1422. [Medline: 19408968]
30. Zhang X, Hailu B, Tabor DC, Gold R, Sayre MH, Sim I, et al. Role of Health Information Technology in Addressing Health Disparities: Patient, Clinician, and System Perspectives. Med Care 2019 Jun 8;57 Suppl 6 Suppl 2(32-33):S115-S120 [FREE Full text] [doi: 10.1097/MLR.0000000000001092] [Medline: 31095049]
31. 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]
32. Fulton EA, Kwah KL, Wild S, Brown KE. Lost in Translation: Transforming Behaviour Change Techniques into Engaging Digital Content and Design for the StopApp. Healthcare (Basel) 2018 Jul 06;6(3) [FREE Full text] [doi: 10.3390/healthcare6030075] [Medline: 29986396]
33. Baek EO, Cagiltay K, Boling ETF. User-centered design and development. In: Spector JMM, Merrienboer J, Driscoll M, editors. Handbook of Research on Educational Communications and Technology (3rd ed.). New York, NY: Lawrence Earlbaum; 2008:659-670.
34. Abras C, Maloney-Krichmar DJP, Preece J. User-centered design. In: Bainbridge W, editor. Encyclopedia of Human-Computer Interaction. Thousand Oaks, CA: Sage; 2004:445-456.
35. Schnall R, Rojas M, Bakken S, Brown W, Carballo-Dieguez A, Carry M, et al. A user-centered model for designing consumer mobile health (mHealth) applications (apps). J Biomed Inform 2016 Apr;60:243-251 [FREE Full text] [doi: 10.1016/j.jbi.2016.02.002] [Medline: 26903153]
36. Griffin L, Lee D, Jaisle A, Carek P, George T, Laber E, et al. Creating an mHealth App for Colorectal Cancer Screening: User-Centered Design Approach. JMIR Hum Factors 2019 May 08;6(2):e12700 [FREE Full text] [doi: 10.2196/12700] [Medline: 31066688]
37. Fairburn CG, Patel V. The impact of digital technology on psychological treatments and their dissemination. Behav Res Ther 2017 Jan;88:19-25 [FREE Full text] [doi: 10.1016/j.brat.2016.08.012] [Medline: 28110672]
38. Oinas-Kukkonen H, Harjumaa M. Persuasive Systems Design: Key Issues, Process Model, and System Features. CAIS 2009;24. [doi: 10.17705/1cais.02428]
39. Matthews J, Win KT, Oinas-Kukkonen H, Freeman M. Persuasive Technology in Mobile Applications Promoting Physical Activity: a Systematic Review. J Med Syst 2016 Mar 9;40(3):72. [doi: 10.1007/s10916-015-0425-x] [Medline: 26748792]
40. Pellegrini CA, Verba SD, Otto AD, Helsel DL, Davis KK, Jakicic JM. The comparison of a technology-based system and an in-person behavioral weight loss intervention. Obesity (Silver Spring) 2012 Feb 10;20(2):356-363 [FREE Full text] [doi: 10.1038/oby.2011.13] [Medline: 21311506]
41. Sieverink F, Kelders SM, van GJE. Clarifying the Concept of Adherence to eHealth Technology: Systematic Review on When Usage Becomes Adherence. J Med Internet Res 2017 Dec 06;19(12):e402 [FREE Full text] [doi: 10.2196/jmir.8578] [Medline: 29212630]
42. Chesser A, Burke A, Reyes J, Rohrberg T. Navigating the digital divide: A systematic review of eHealth literacy in underserved populations in the United States. Inform Health Soc Care 2016 Feb 24;41(1):1-19. [doi: 10.3109/17538157.2014.948171] [Medline: 25710808]

43. Mackert M, Mabry-Flynn A, Champlin S, Donovan EE, Pounders K. Health Literacy and Health Information Technology Adoption: The Potential for a New Digital Divide. J Med Internet Res 2016 Oct 04;18(10):e264 [FREE Full text] [doi: 10.2196/jmir.6349] [Medline: 27702738]
44. Articulate. Articulate Rise [Software]. URL: https://articulate.com/360/rise [accessed 2020-08-15]
45. Articulate. Articulate Storyline [Software]. URL: https://articulate.com/360/storyline [accessed 2020-08-15]
46. Google. Google Hangouts [Mobile Application Software]. 2018. URL: https://play.google.com/store/apps/details?id=com.google.android.talk&hl=en_US [accessed 2020-08-09]
47. Microsoft. Skype [Mobile Application Software]. 2017. URL: https://skype.com/en/
48. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
49. Fitbit. Fitbit (Version 2.52) [Mobile Application Software]. 2016. URL: https://www.fitbit.com/my/flex2 [accessed 2020-08-09]
50. Moodle. Moodle LMS [Mobile Application Software]. 2017. URL: https://moodle.com
51. Amazon Web Services Inc. Amazon Elastic Compute Cloud [Computer Program]. 2017. URL: https://aws.amazon.com/ec2/ [accessed 2020-08-09]
52. Withings. Withings Health Mate: a Fitness, Activity and Health Tracker App [Mobile Application Software]. 2018. URL: https://www.withings.com/us/en/health-mate [accessed 2020-08-09]
53. Microsoft. Microsoft Azure [Computer Software]. 2017. URL: [accessed 2020-08-09]
54. Google Inc. Google Cloud [Mobile Application Software]. 2017. URL: https://cloud.google.com/ [accessed 2020-09-08]
55. Hughes PP, Goldstein MM. Behavioral health care and technology using science-based innovations to transform practice. In: Privacy, Security, and Regulatory Considerations as Related to Behavioral Health Information Technology. Oxford, UK: Oxford University Press; 2016:224-238.
56. Vaismoradi M, Turunen H, Bondas T. Content analysis and thematic analysis: Implications for conducting a qualitative descriptive study. Nurs Health Sci 2013 Sep 11;15(3):398-405. [doi: 10.1111/nhs.12048] [Medline: 23480423]
57. Marcolino MS, Oliveira JAQ, D'Agostino M, Ribeiro AL, Alkmim MBM, Novillo-Ortiz D. The Impact of mHealth Interventions: Systematic Review of Systematic Reviews. JMIR Mhealth Uhealth 2018 Jan 17;6(1):e23 [FREE Full text] [doi: 10.2196/mhealth.8873] [Medline: 29343463]
58. 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;16(6):e146 [FREE Full text] [doi: 10.2196/jmir.3077] [Medline: 24905070]

Abbreviations

API: application programming interface
BCT: behavior change techniques
EC2: Elastic Cloud Compute
IP: in-person
IT: information technology
mHealth: mobile health technology
PA: physical activity
PDS: persuasive design system
REDCap: Research Electronic Data Capture
RCT: randomized controlled trial
RT: remote technology
SD: secure digital
SQL: structured query language

Edited by G Eysenbach; submitted 25.05.20; peer-reviewed by G Darville, C Carrion, J Alvarez Pitti, S Huat; comments to author 17.06.20; revised version received 10.07.20; accepted 26.07.20; published 20.08.20.

Please cite as:
Yin Z, Errisuriz VL, Evans M, Inupakutika D, Kaghyan S, Li S, Esparza L, Akopian D, Parra-Medina D
A Digital Health Intervention for Weight Management for Latino Families Living in Rural Communities: Perspectives and Lessons Learned During Development
JMIR Form Res 2020;4(8):e20679
URL: http://formative.jmir.org/2020/8/e20679/
doi:10.2196/20679
PMID:32726748

Original Paper

Primary Care Peer-Supported Internet-Mediated Psychological Treatment for Adults With Anxiety Disorders: Mixed Methods Study

Linnea Nissling[1,2], Psy M; Claudia Fahlke[1], Prof; Josefine L Lilja[1,2], PhD; Ingmarie Skoglund[2,3], PhD; Sandra Weineland[1,2], Docent

[1]Department of Psychology, University of Gothenburg, Gothenburg, Sweden

[2]R&D Primary Health Care, Västra Götaland, Sweden

[3]General Medicine, School of Public Health and Community Medicine, Institute of Medicine, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden

Corresponding Author:
Linnea Nissling, Psy M
Department of Psychology
University of Gothenburg
Haraldsgatan 1
Gothenburg, 413 14
Sweden
Phone: 46 725 029 216
Email: linnea.nissling@vgregion.se

Abstract

Background: The effect of internet-delivered cognitive behavioral therapy (iCBT) on anxiety in adults is well-known. However, patient dropouts and poor adherence to treatment are common. Feelings of belonging and empowerment from the treatment might be key to the completion of iCBT. Peer support workers are people with a personal experience of mental health problems, trained to provide professional support to people who require mental health care.

Objective: This study aims to assess patient experiences; the feasibility, safety, and acceptability; and preliminary effectiveness on anxiety and depression, empowerment, and adherence to treatment in an 8-week peer-supported iCBT program for patients with anxiety disorders treated in primary care.

Methods: This was a single-arm mixed methods feasibility study. Participants were patients referred to a central unit for iCBT in primary care. Quantitative data were collected pre-, post-, and 3 months postintervention. Qualitative data were collected through semistructured interviews.

Results: A total of 9 participants completed the quantitative outcome assessment. Statistically significant improvements were observed in perceived empowerment at a 3-month follow-up, and significant decreases in anxiety, depression, and psychological distress at the end of the treatment were maintained at a 3-month follow-up. In total, 8 of the 9 patients showed improvement in the severity of their symptoms of anxiety. Adherence to treatment was good among the participants. No serious adverse events were reported. Eight participants were enrolled in the qualitative analysis. The qualitative results showed 3 main themes: (1) real contact in an online world, (2) empowering experiences, and (3) being behind the wheel. Qualitative results largely emphasized the personal relationship and supported the acceptability of adding peer support to iCBT.

Conclusions: Peer support in digital treatment seems to be a safe and acceptable intervention. The preliminary results suggest the effectiveness of peer support on patient empowerment, anxiety, depression, psychological distress, and adherence to treatment. The results indicate the need for future studies to evaluate the effect of adding peer support to iCBT in larger randomized controlled trials.

(JMIR Form Res 2020;4(8):e19226) doi:10.2196/19226

KEYWORDS

iCBT; cognitive behavioral therapy; internet-based intervention; anxiety disorders; primary care; treatment adherence and compliance; peer support

Introduction

Background

Anxiety disorders are mental health problems commonly encountered in primary care that may lead to impaired functioning in daily life. Data from the Swedish National Public Health Survey show that in 2018, 39% of adult respondents reported symptoms of anxiety [1]. Anxiety disorders are mainly treated with antidepressants in combination with various psychological treatments; however, the lack of access to such treatments is a general problem in Sweden's routine health care [2].

Cognitive behavioral therapy (CBT) is a well-documented and effective method for various anxiety disorders [3] and is often recommended as a first-hand treatment in Sweden [2]. Therapist-guided internet-delivered cognitive behavioral therapy (iCBT) may be an effective medium for disseminating psychological treatment by providing greater availability and reaching out to patients who would otherwise not seek care due to fear of stigmatization [4]. Studies show that iCBT is effective for treating a number of psychiatric disorders, including several anxiety disorders, depression, substance abuse, and bipolar disorder, among others [5-8]. A growing number of studies also show that the effects of such treatment persist at long-term follow-ups [8-11] and that it may be as efficient as face-to-face therapy [7,12]. Moreover, people who have completed iCBT generally show a positive attitude toward the treatment and its effectiveness [13]. However, more research is needed on the efficacy of iCBT in clinical settings, as only a limited number of studies have been conducted to date [12].

Despite the positive results shown for iCBT, patient dropouts and poor adherence to treatment are common [14-16]. To increase patient compliance and completion of iCBT, additional interventions might be needed to help patients cope with continued symptoms during treatment. The need for additional interventions is also supported by findings that therapist-guided internet treatments have better outcomes [6,7] and more patient adherence [6] than unguided interventions. Feelings of belonging and empowerment gained from the treatment might be key to patients' completion of iCBT. Peer networks might be effective in increasing adherence to treatment.

The mental health sector has recently seen an increase in the use of peer support workers (PSWs) [17,18]. A peer support worker is a person with lived experience of mental health problems and rehabilitation who is employed in the mental health sector and becomes professionally active in recovery and support services for patients with mental health problems [19-21]. The PSW publicly identifies as a person who has received or is receiving mental health services [20] and has recovered enough from their mental illness to manage that illness and live a fulfilling life [19]. Peer support interventions focus on strengthening patients' resources, rather than focusing simply on symptom reduction [22]. As peer supporters share their own experiences of the path from mental illness to recovery, they can function as role models and give patients hope [20,22]. This sharing of experiential knowledge might empower patients to become active agents in their own self-care [20,23]. Peer support services might also be effective in providing social support, which prevents isolation and acts as a buffer against stressors [20,24]. Through the helper-therapy principle, it has also been proposed that helping others might benefit PSWs themselves by strengthening their self-esteem, sense of empowerment, and confidence in their overall capability and ability to manage their illness [20].

The effectiveness of peer support services is promising; however, the evidence base is insufficient, and the results are mixed [19,25]. A recent systematic review [26] showed that the inclusion of peer support interventions in general clinical care is as effective as usual care conditions on traditional clinical outcomes, such as symptom severity and hospitalization rates, and that peer interventions have a positive effect on measures of hope, empowerment, and quality of life. Mahlke et al [22] found that the addition of peer support to standard treatment for patients with severe mental illness was related to higher scores of self-efficacy than standard treatment alone.

Recently, peer support services have been tested and offered in digital settings [21]. Some studies have investigated digital peer support in online interventions aimed at patients with bipolar disorders [27,28]. The participants in the supported groups were guided in the program by people with personal experiences of bipolar disorder. In both studies, adherence to treatment was significantly higher among patients receiving peer support [27,28]. A recent systematic review found preliminary evidence for the effectiveness of digital peer support on patients' functioning, symptom reduction, and program utilization [21].

However, to date, few researchers have evaluated digital peer support in iCBT interventions. The results of a study in elderly adults with depression indicated that peer-supported CBT-informed intervention programs are acceptable and equivalent to individually delivered internet interventions. Less time from the therapist was needed in the group with peer support [29].

Objectives

To our knowledge, no study has investigated peer support in primary care iCBT for adults with anxiety disorders. Therefore, the aim of this explorative study was to assess patients' experiences, the feasibility, safety, and acceptability, and preliminary effectiveness of an 8-week primary care, peer-supported iCBT program for patients with anxiety disorders. The hypotheses were that contact with a peer supporter would enhance participants' feelings of empowerment and increase their adherence to treatment.

Methods

Design

This was a single-arm, pre-post mixed methods intervention study to test the feasibility of an 8-week iCBT program for patients with anxiety disorders treated in primary care. This mixed methods study had a convergent design [30]. The intent of a convergent mixed methods design is to collect quantitative and qualitative data to allow the different methods to complement each other in terms of strengths and weaknesses and provide a fuller understanding of the research problem [30].

One motivation for using a mixed methods design in this study was to triangulate the research question using different methods to investigate whether the results of the different methods aligned with each other. The other motivation was to augment the quantitative data with qualitative data of participants' experiences, as few prior studies have investigated using peer support in iCBT treatment, and thus to reveal new information that could be useful in future work. Outcome assessments were conducted at baseline, postintervention, and at the 3-month follow-up. Semistructured interviews were conducted postintervention to investigate the participants' experiences of treatment. Quantitative and qualitative data were analyzed separately and integrated and interpreted during the discussion of the results. Multimedia Appendix 1 shows the flowchart of the study.

Participants

The study was approved by the Regional Ethics Committee of Gothenburg (Dnr: 845-18). Participants were recruited from patients referred to a central unit for primary care iCBT in the Västra Götaland region, Sweden. Participants were adults aged 18 years or older with an anxiety disorder diagnosed according to ICD-10 [31]. The inclusion criteria were having reached the age of 18 years, having access to a computer with an internet connection, being able to speak and understand Swedish, and meeting the diagnostic criteria for an anxiety disorder (social anxiety, generalized anxiety disorder [GAD], panic disorder, obsessive compulsive disorder, or unspecified anxiety disorder). The exclusion criteria were having started pharmacological treatment for mental health problems or made major changes in the medication in the past few months, having serious suicidal ideation or suicide plans, having severe or complex comorbidity, or needing other care or receiving ongoing psychological treatment during the treatment period.

Procedure

All participants were interviewed before and at the end of treatment by a psychologist using the structured diagnostic interview instrument, PRIME-MD [32]. Before starting treatment, all participants had a physical visit to the health care center and a somatic examination by a general practitioner (GP). The assessment interview was conducted by telephone upon referral from the health care center. The psychologist conducted an in-depth interview with the PRIME-MD and conducted a clinical assessment. After the interview, all patients who met the inclusion criteria were asked to participate in the study. They were asked verbally, and written information was sent by post to participants who wished to participate. The recruitment period for the study was 4 weeks in the spring of 2019. Of a total of 41 patients booked for an assessment interview, 21 met the inclusion criteria and were asked to participate. Of these 21 respondents, 15 agreed to participate. One of the initial 15 participants chose to withdraw from treatment before it began, and one chose to discontinue participation in the study after reading the information letter, but continued the usual iCBT. Two participants were excluded from data analysis because they were never active in the treatment program despite several reminders. Two participants failed to send in their written consent to participate in the study and were thus excluded from the data analysis. The results are thus based on information from 9 participants who agreed to the study and took part in the iCBT with additional peer support. Table 1 shows demographic variables of the participants.

Table 1. Demographic variables of patients (n=9) participating in the study.

Characteristics	Frequency, n (%)
Age (years)	
18-25	4 (44)
25-35	4 (44)
35-45	1 (12)
>45	N/A[a]
Gender	
Male	4 (44)
Female	5 (56)
Employment status	
Employed	6 (67)
Student	1 (11)
Sick leave	2 (22)
Previously received psychological treatment (lifetime)	
Yes	5 (56)
No	4 (44)
Other psychological treatment during time for iCBT[b]	
Yes (counseling)	1 (11)
No	8 (89)
Current medication for mental health problems during time for iCBT	
Yes (2 begun just before treatment)	3 (33)
No	6 (67)

[a]N/A: not applicable.

[b]iCBT: internet-delivered cognitive behavioral therapy.

Quantitative data were collected through assessment forms sent to participants through a digital link sent to their email. The participants filled in assessments of the primary outcomes at the start of the treatment, at the middle of treatment, at the end of treatment, and 3 months after completing the treatment. The participants also completed a weekly assessment of symptoms of general psychological distress in connection with each new module in the treatment program. In addition, the participants regularly responded to a question regarding suicidal ideation as part of the weekly assessment of psychological distress. At the end of treatment, a follow-up assessment was made by the therapist in charge of the treatment to determine whether the participant still fulfilled the diagnosis received at the assessment interview before starting treatment. After completion of the treatment, qualitative data were collected. A trained research assistant, with no relationship to the participants, interviewed the participants about their experiences and attitudes toward treatment. The interviews were conducted by telephone 2 to 5 weeks after completion of treatment through a semistructured interview guided with open-ended questions such as “How did you experience the contact with a peer supporter?” The participants’ descriptions led the interviewer to add follow-up and in-depth questions such as “Can you tell me more?” The interviewer assured data reliability by repeatedly asking participants if they understood and by summing up what the participant had said. The interviews lasted 30 to 45 min. All interviews were audio-recorded and transcribed verbatim. The research assistant transcribed the interviews. Of the 9 participants who completed the treatment, only one declined to participate in the interview, and one responded to the interview questions in writing. The qualitative result is thus based on a total of 8 participants. In addition to quantitative measures and participant interviews, text messages sent from the peer supporters to the participants were collected and analyzed.

Intervention

Two peer supporters were recruited through a Swedish patient organization, the National Cooperation for Mental Health in Gothenburg (NSPHiG). This organization has an established program for Swedish peer support education and a national platform of guidelines and frameworks. The organization has experience implementing peer support in psychiatric care and participated in recruiting and educating PSWs in this study. The 2 PSWs recruited for the study had experience working at an inpatient psychiatric clinic and were temporarily employed in primary care during the study. They worked 16 hours per week and supported 5 to 7 participants each. They had weekly scheduled meetings with a supervising psychologist, who is also the first author of this study. Supervision included discussing the treatment content, the participants’ answers on

the questionnaires, and reflections on the written messages from the participants.

Peer supporters and participants interacted in the treatment program via asynchronous secure messages. Peer supporters were able to provide support and feedback on exercises in the iCBT program. The participants received a follow-up telephone call from the peer supporter in the middle of the treatment. In addition, the interactions in the text messages between peer supporters and participants were tracked by the supervising psychologist, who had joint access with the peer supporters to the treatment program in the digital system. The purpose of this was to ensure the safety of treatment. The psychologist, however, to stay true to how peer support operates in nondigital interventions [33] and to ensure that the content solely reflected the peer support intervention, did not try to influence the content of the peer supporters' messages.

Participants also had limited contact with 2 licensed psychologists in the iCBT. Both psychologists worked at the participating clinic and were experienced in working with iCBT. All participants received a follow-up call from the psychologist after the completion of treatment. The psychologist also made telephone calls on the request of participants or if they thought it necessary (eg, if the psychologist noticed high scores on suicidal ideations or the participant remained inactive in the program for longer than 2 weeks—a routine intervention at the participating clinic).

In all cases, peer supporters had the most contact with the participants. The contact between the participants and the psychologists was limited to messages informing the participants that new modules had been activated in the treatment program. More detailed feedback on exercises and written messages was given only upon request from the participants. The peer supporters stated that they spent between 5 and 6 hours per week on the intervention. These hours included supporting the participants through the program as well as getting familiar with the treatment program and the content of each module. The psychologists spent roughly 10 to 15 min per participant per week (ie, 120 min per patient during the total treatment period).

The iCBT program used in this study was developed by Livanda-Internetkliniken AB for people affected by anxiety problems such as panic disorder, social phobia, and generalized anxiety. The program is based on both CBT methods and acceptance and commitment therapy (ACT) interventions and is a transdiagnostic program aimed at treating people with mild-to-moderate anxiety problems [34]. The program is designed as a course that includes education about symptoms common to anxiety disorders and as a training on different tools that have been shown to have a positive effect on such disorders. The program includes 13 different tools, and the treatment consists of 8 modules meant to be completed within 8 weeks. The tools presented are based on ACT principles such as exposure, acceptance, valued action, mindfulness, and defusion. The program contains psychoeducational text sections and video clips, assessments, and home assignments.

Measurements

Anxiety was measured using the GAD 7-item scale (GAD-7) instrument. The instrument has 7 items measured on thresholds for mild, medium, and severe anxiety [35]. Symptoms of psychological distress were measured using Clinical Outcomes in Routine Evaluation 10 (CORE-10) [36]. The instrument consists of 10 items and has clinical cutoff scores for general psychological distress. Symptoms of depression were measured on the 9-item Montgomery-Åsberg Depression Rating Scale–Self report (MADRS-S) [37] with cutoff points for symptom severity. Empowerment was measured on the Empowerment Scale [38], which consists of 28 items. Acceptability was measured by using 4 questions, "To what degree have you experienced the treatment as helpful?" "How meaningful did you perceive the contact with the peer supporter?" "Would you recommend iCBT to someone else?" and "Would you recommend iCBT with peer support to someone else?" All questions were scored on a scale of 1 to 5. Higher values indicate greater acceptability.

Data Analysis

Quantitative Data Analysis

The quantitative data analysis for repeated measures was performed using the Friedman analysis of variance (ANOVA) [39,40], which is a nonparametric correspondence to a one-way ANOVA with repeated measures. Post hoc analyses were performed using the Wilcoxon signed rank test for related samples. A Bonferroni-adjusted significance level was calculated to minimize the risk of type 1 error since multiple post hoc comparisons were made. The analysis was performed based on the intention-to-treat approach. The last observation carried forward was used to deal with missing data.

Qualitative Data Analysis

The qualitative data analysis was performed according to a thematic analysis [41]. The method was chosen because it provides the researcher with a flexible framework for finding themes and patterns in data [41]. The method is not bound to any theoretical foundation and can thus be used to analyze data both deductively and inductively. As this was an explorative study about the participants' attitudes and experiences of a new treatment intervention, an inductive bottom-up approach was used to capture experiences and opinions as unconditionally as possible. The analysis was performed using a realist approach with the aim of identifying the manifest content of the participants' attitudes and experiences. The first author's pre-existing understanding of the topic is that of a working clinical psychologist and scientist as well as the project leader of the study. This might have influenced the understanding of the content of the interviews, but it also provided a deeper understanding of the intervention as a whole, which might have made the interpretation of the participants' experiences deeper and more realistic. To ensure that the interpretation of the material was as close to reality as possible, the third and last authors were also involved in the interpretation process described below. In addition, the last author also read the original interview material to ensure a fit between the generated themes and the content of the interviews.

The data analysis was guided by the 6 steps described by Braun and Clarke [41]. The interviews were first read several times by the first author to become familiarized with the entire data set and formulate ideas about the initial codes and themes. All materials related to the research question were then coded by the first author. The software program, NVivo 12, was used to facilitate the coding process. After the initial coding, the codes were manually gathered by the first author into a thematic map, and preliminary themes and subthemes were identified. The themes were then reviewed by the first and last authors and revised when necessary. The first thematic map consisted of several themes and subthemes that were later refined. Examples of themes in this phase were experiences of peer support, effects of treatment, the treatment medium, the treatment program, and external factors. After this stage, the first author reviewed the codes and corresponding themes and subthemes with the third author, who had not been involved in the study design, recruitment, treatment, or initial analysis nor to that point had been familiar with the content of the interviews. After discussion, the initial themes were condensed into 3 main themes with corresponding subthemes. The codes were rechecked and reorganized by the first author according to the new themes. This process was guided by the dual criteria for judging internally homogenous and externally heterogeneous categories of [41]. Finally, the first author and last author read the original transcriptions of the interview material and reviewed the codes and corresponding themes and subthemes to ensure a fit between the interview content and the formulated themes.

Peer Supporters' Text Messages

The text messages sent from the peer supporters to the participants were also analyzed using the same procedure as described above for the interviews. The first step of the analysis was performed by the first, fourth, and last authors and was compiled into themes by the first author. The fit between themes and content was rechecked by the last author.

Results

Quantitative Results

Quantitative results are based on information from a total of 9 participants who went through the iCBT with additional support from peer support. Five participants stated that they spent 0-2 hours per week on the treatment and 4 participants spent 2-4 hours per week. No participant dropped out after starting treatment. Some participants did not complete all of the modules, but they continued to have written contact with their peer supporter through messages in the program and thus remained active in the treatment. Two participants completed only one module, and 67% (6/9) completed more than half of the modules. Table 2 shows the number of modules completed for all participants.

Participants showed levels of anxiety and depression above the thresholds before treatment. Table 3 shows descriptive statistics for the different measurement points for GAD-7, MADRS-S, CORE-10, and the Empowerment Scale.

Table 2. Number of modules completed (n=9).

Participants, n (%)	Modules completed, n
2 (22)	1
1 (11)	4
3 (33)	5
2 (22)	6
1 (11)	8

Table 3. Descriptive statistics for measurement points on GAD-7, MADRS-S, CORE-10, and the Empowerment Scale.

Measurements	Median	Mean (SD)	Variance	Minimum scores	Maximum scores
GAD-7[a]					
Pre	15	12.9 (4.3)	18.6	5	18
Mid	9	9.8 (4.8)	23.2	4	18
Post	5	6.8 (4.6)	20.9	2	13
3-month follow-up	5	6.4 (4.6)	20.8	1	13
MADRS-S[b]					
Pre	27	23.0 (6.4)	41.0	11	29
Middle	23	19.9 (7.7)	59.9	5	28
Post	15	15.8 (9)	80.9	4	28
3-month follow-up	12	14.3 (10.4)	108.5	1	30
CORE-10[c]					
Pre	22	22.9(5.6)	31.6	16	31
Post	15	15.3(5.7)	32.5	8	25
Empowerment Scale					
Pre	71	73.2 (9.7)	94.4	59	91
Middle	72	73.8 (11.8)	138.7	59	95
Post	80	77.6 (9.3)	86.0	62	90
3-month follow-up	85	83.2 (12)	143.7	62	100

[a]GAD-7: generalized anxiety disorder 7-item scale.

[b]MADRS-S: Montgomery-Åsberg Depression Rating Scale–Self report.

[c]CORE-10: Clinical Outcomes in Routine Evaluation 10.

Anxiety

The results from Friedman test for GAD-7 showed a statistically significant difference between measurement points, X^2_2=11.6; *P*=.003. A post hoc analysis with a Wilcoxon signed rank test for related samples was performed with a Bonferroni correction applied, resulting in a significance level set at *P*<.017. Results from the post hoc analysis showed a statistically significant reduction in anxiety symptoms from pretest to 3-month follow-up (*Z*=−2.552; *P*=.01; *r*=0.60) as well as from pretest to posttest (*Z*=−2.668; *P*=.01; *r*=0.63). There was no statistical difference in anxiety symptoms between postmeasurement and 3-month follow-up (*Z*=−0.170; *P*=.87; *r*=0.04).

Depression

For MADRS-S, the results from Friedman test showed a statistically significant difference between measurement points (X^2_2=9.9; *P*=.01). Results from the post hoc analysis with the Wilcoxon signed rank test, with the Bonferroni correction applied (*P*<.017), showed a statistically significant reduction in depressive symptoms from pretest to the 3-month follow-up measurement (*Z*=−2.429; *P*=.02; *r*=0.57) and from pretest to posttest (*Z*=−2.521; *P*=.01; *r*=0.59). There was no statistical difference in depressive symptoms between postmeasurement and 3-month follow-up for MADRS-S (*Z*=−0.491; *P*=.62; *r*=0.12).

Psychological Distress

Results from the Wilcoxon signed rank test for CORE-10 showed a statistically significant reduction in symptoms of psychological distress from pretest to posttest (*Z*=−2.524; *N*-ties=8; *P*=.01; *r*=0.59). The 3-month follow-up test for CORE-10 was not assessed since the participants filled in this measurement in connection with new modules of the treatment program and, by the 3-month follow-up, no longer had access to the treatment program.

Empowerment

The results from Friedman test for the Empowerment Scale showed a statistically significant difference between measurement points (X^2_2=10.1; *P*=.01). Results from the post hoc analysis with a Wilcoxon signed rank test with the Bonferroni correction (*P*<.017) showed a statistically significant increase in experienced empowerment from pretest to 3-month follow-up measurement (*Z*=−2.547; *P*=.01; *r*=0.60). There was no statistically significant difference between pre- and postmeasurement with the Bonferroni correction applied (*Z*=−2.082; *P*=.04; *r*=0.49) or between postmeasurement and 3-month follow-up with the Bonferroni correction applied (*Z*=−2.371; *P*=.02; *r*=0.56).

Clinically Significant Improvement

Clinically significant improvement was determined by comparing the scores on GAD-7 and MADRS-S against the

thresholds on the respective scales for mild, moderate, and severe anxiety and depression at the start of treatment and at the end of treatment. Of the 9 patients, 8 were considered "improved" on the severity of their symptoms of anxiety; 4 were considered "improved" on symptoms of depression based on the thresholds for mild, moderate, and severe depression on MADRS-S; 5 were considered "unchanged" for symptoms of depression, although 1 did not meet the cutoff threshold for depression at startup or at the 3-month follow-up. Table 4 shows the changes in the cutoff scores for the GAD-7 and MADRS-S.

Table 4. Clinically significant improvement.

Participant	Symptom severity for anxiety based on cutoff scores on GAD-7[a]	Improvement GAD-7	Symptom severity for depression based on cutoff scores on MADRS-S[b]	Improvement MADRS-S
1	From severe to moderate	Improved	From moderate to moderate	Unchanged
2	From mild to absent	Improved	From absent to absent	Unchanged
3	From moderate to mild	Improved	From moderate to mild	Improved
4	From severe to absent	Improved	From mild to absent	Improved
5	From moderate to mild	Improved	From mild to mild	Unchanged
6	From severe to mild	Improved	From mild to absent	Improved
7	From severe to moderate	Improved	From moderate to moderate	Unchanged
8	From mild to moderate	Unchanged (worsening)	From moderate to moderate	Unchanged
9	From severe to mild	Improved	From moderate to absent	Improved

[a]GAD-7: generalized anxiety disorder 7-item scale.

[b]MADRS-S: Montgomery-Åsberg Depression Rating Scale–Self report.

Based on the diagnostic interview with Prime-MD conducted at the start and the end of the treatment, 5 participants no longer met the criteria for their main anxiety diagnosis after treatment; 4 were assessed as still meeting the criteria for their main anxiety diagnosis, one of whom was referred back to the health care center for further treatment. With that one exception, none of the participants were considered in need of further treatment for anxiety. Table 5 shows the participants' diagnoses at the start and end of the treatment.

At the 3-month follow-up assessment, 4 participants stated that they had sought continued care for their mental health; 3 stated that they had started some form of counseling, and 1 had started pharmacological treatment for mental illness. Of the 4 participants who sought continued care after termination of iCBT, 1 was classified as unchanged, while 3 were classified as improved according to the results above.

Table 5. Diagnosis at start and end of treatment assessed by the diagnostic interview, PRIME-MD.

Participant	Diagnosis at start of the treatment	Meets criteria for diagnosis after treatment	Referred for further treatment
1	• Anxiety, unspecified • Depressive episode, unspecified	• Yes • No	N/A[a]
2	• Panic disorder • Anxiety, unspecified • Specific phobia (emetophobia)	• No • No • Yes	N/A
3	• Panic disorder • Generalized anxiety disorder • Social phobia • Depressive episode, unspecified • Obsessive compulsive disorder • Eating disorder, unspecified	• No • Yes • Yes • Yes • No • Not assessed	N/A
4	• Social phobia • Generalized anxiety disorder	• No • Yes	N/A
5	• Anxiety, unspecified • Depressive episode, unspecified	• No • No	N/A
6	• Generalized anxiety disorder • Social phobia	• Yes • Yes	N/A
7	• Generalized anxiety disorder • Social phobia • Depressive episode, mild	• Yes • Yes • Yes	Referred to health care center
8	• Generalized anxiety disorder • Recurring depressive episode, moderate	• Yes • Yes	N/A
9	• Anxiety, unspecified	• No	N/A

[a]N/A: not applicable.

Safety and Acceptability

No serious adverse events were reported during the treatment period or in the interviews with the participants. The participants (n=9) scored a median value of 4 (on a scale of 1-5) on the question "To what degree have you experienced the treatment as helpful?" The participants also scored a median value of 3 (on a scale of 1-5) on the question "How meaningful did you perceive the contact with the peer supporter?"

For the question "Would you recommend iCBT to someone else?" 7 of the 9 participants stated that they would recommend it, 1 would not, and 1 did not know or had no opinion. For the question "Would you recommend iCBT with peer support to someone else?" 7 of the 9 participants answered affirmative and 2 stated that they did not know or had no opinion.

Qualitative Results

Three main themes with associated subthemes were generated from the qualitative results: real contact in an online world (subthemes: "Support and encouragement" and "Sharing experiences provides personal contact"); empowering experiences (subthemes: "Changes in psychological well-being," "An ongoing task," and "Acquired psychological strategies"); and being behind the wheel (subthemes: "Flexibility and responsibility" and "Barriers to treatment").

The quotes in the text have been translated into English.

Real Contact in the Online World

This theme relates to the participants' experiences of their contact with the peer supporter and their thoughts on what had been helpful or not about having the peer support in treatment.

Support and Encouragement

The participants generally described having had a good experience with the peer supporter. Several described their perception of peer support as very good and said that contact with a peer supporter could help someone to see another side to anxiety and to imagine the possibility of feeling better. Several participants said that it felt good when the peer supporter checked in on them and showed that they were there by emailing every week, asking them questions about how things had been going. They said it was useful to know there was someone there who they could turn to. Many thought it was nice to be able to write about anything they felt like and that the treatment program felt more real when there were real people to write to:

> *I absolutely believe that it is a very good idea and that you know that you are not alone and that it is possible, there is a second side to the problem as well. You can crawl out of this, so that's what I think but it was, it was a good experience on the whole.* [Participant #8]

> *I thought it was very, it was nice that the person, the peer supporter noticed if I had written that—I haven't had anyone to talk to […] and then he wrote directly saying that you can write to me when it is needed, and then I wrote to him, so I thought it was very nice, to just be able to write to someone like that.* [Participant #3]

Several participants were able to point to situations where contact with peer support was extra helpful; however, some stated that they had not had much contact with their peer supporter. A couple of participants described wanting more verbal contact with the peer supporter and thought that this had facilitated their connection with each other. In line with this, some participants described how it had been extra helpful when they were able to speak with their peer supporter on the phone. One participant asked for a physical meeting with the peer supporter at the beginning of treatment and thought this had facilitated their connection. One other participant, however, felt that digital contact reduced the pressure of social settings and made it easier to open up. Another described having wanted to know more about the peer supporter's background and concrete experience. In line with this, another participant described how the contact was made more difficult because of uncertainty about the peer supporters. Yet another participant felt that the treatment worked equally well without peer support.

Sharing Experiences Provide Personal Contact

Many of the participants felt that it was positive to share experiences with someone who had been in a similar situation. Some described situations during treatment when it had been helpful when the peer supporter shared how they had handled similar situations in the past:

> *The peer supporter could write that I've also felt like that sometimes, and I usually think this way and this, that if you are afraid to say something that will sound wrong, then try and do it and see what happens.* [Participant #3]

Several of the participants said that it felt like a more personal contact to talk to someone with similar experiences than to health care professionals (therapists, psychologists, or doctors). The peer supporter was seen as a fellow human being and someone who could understand their problems differently than a therapist because of their personal experiences:

> *It was pretty nice to talk to someone. She had gone through everything, and it wasn't like someone or a therapist was in charge, but maybe a fellow human being, who knows where you are at.* [Participant #1]

Empowering Experiences

This theme relates to the participants' experiences of being strengthened in themselves and how the treatment contributed to their feeling able to handle their anxiety in a different way.

Changes in Psychological Well-Being

The majority of the participants described feeling less anxious in general after the treatment and said that situations that used to provoke anxiety did not do so anymore. Many participants also described changes in cognition as they thought of their emotions as "just feelings" and thus felt that they could handle their anxiety differently than before starting treatment:

> *Well that, in fact, it might not be so dangerous. It's just a feeling. It can be very difficult, and to try to tell yourself that it's actually just a feeling, you probably won't not die anyway.* [Participant #1]

Several participants described how the treatment had helped them do things they had previously avoided because of their anxiety. They described pushing themselves to talk in situations that made them feel uncomfortable, daring to make mistakes, or to state their opinion more clearly. Some described how going through treatment had created a positive feedback loop and that doing things they had previously avoided made them feel stronger about themselves. Some also described how exposing themselves to situations they had previously avoided helped them to realize that those things were not as dangerous as they had thought and to realize that if they did not try, things would not get better:

> *…when I had anxiety, I often put things off. I mean I didn't want to meet people, but now I force myself to just meet people, because it is not getting better from me not doing it.* [Participant #4]

> *The treatment has helped me to be able to do things that I previously was anxious about. So it has strengthened me, it has strengthened other parts of my person and my inner well-being, which has made me less uncomfortable in those situations that made me uncomfortable in the past.* [Participant #9]

A couple of the participants felt that through the treatment they had realized that they were not the only ones to deal with anxiety, and this realization made them feel less alone. They also felt that they could think of their unhelpful thoughts more as symptoms of anxiety than reflections of real conditions, and could therefore feel less odd.

Several participants described feeling they had taken hold of themselves by deciding to go through treatment, and a couple described having clearer thoughts about themselves as the people who had to deal with their problems.

An Ongoing Task

Most participants described feeling that they would continue working on themselves even after treatment. Some pointed to different interventions in the treatment program that they were going to remind themselves to continue using. Others stated that they would continue to do exercises from the program to become better at taking care of themselves:

> *[There were] some practical exercises to do, and to bring those with you, to keep doing those exercises when needed. Or maybe not even when needed, but on a regular basis.* [Participant #9]

Acquired Psychological Strategies

Many participants described the program as having given them helpful tools and new knowledge about thoughts, feelings, and physical reactions to anxiety. Some participants emphasized having a different perspective on their own thoughts and having learned to question them. Other strategies that patients acquired

were being able to put their thoughts and feelings into words and to set goals for themselves. The participants mentioned helpful interventions such as breathing exercises or postponing rumination to a set time.

Being Behind the Wheel

This theme concerns the participants' perceptions of the form of the treatment and the perceived advantages and disadvantages of mediating the treatment digitally.

Flexibility and Responsibility

Many participants had a positive attitude toward the treatment program, and several thought the treatment had helped them. None of the participants felt that the treatment had resulted in a negative change for them. However, a few participants felt that the treatment was not suitable for them.

Several participants emphasized the increased flexibility of the digital treatment and appreciated being able to access the treatment when it suited them and to reflect on the content at their own pace:

> *It was actually the whole concept of having someone all the time that you can have contact with and at the same time be able to read up on and do things yourself, but at the same time be able to write to someone if you feel that you need it.* [Participant #3]

In contrast, a few participants described how the flexibility of the treatment could also be a disadvantage, as they postponed working with the treatment, had less time to sit with the program than they had intended, or felt too tired to engage with the program after a day of working. In line with this, some participants reasoned that the digital medium for treatment placed greater demands on their own responsibility and self-awareness. They reasoned that this might place a greater demand on patients to have a functional everyday life and that the treatment might be more suitable for patients with less severe mental illness:

> *My first CBT treatment in group, for example, did not work because I was too bad and what was required there was way beyond what I could handle, so I would say that maybe it is something to try when things start to stabilize so that you are able to take that responsibility yourself.* [Participant #8]

Barriers to Treatment

The participants also described some difficulties related to the digital form of treatment. Some participants felt that there was too much information in the treatment program to read and listen to. Several participants also mentioned that the pace of the treatment was too quick, and they wished for more time to go through the program. A few participants described having difficulty engaging in the treatment because it came at an inappropriate time in life, they were not prepared for how much the treatment would require of them, and they found it hard to take charge of doing things for themselves. Some participants said that they had wished for more verbal contact with the therapist or the peer supporter and thought this might have facilitated their engagement with the treatment program:

> *It was nothing that suited me because a lot of what I needed to do was a little more, I got stuck in trying to understand it myself and doing it, and it was that, that was the problem from the beginning, so it's like, I never got started so to speak.* [Participant #7]

Peer Supporters' Text Messages

Interviews with peer supporters showed that they were satisfied with their role in the treatment program. The peer supporters generally felt that the pace of their support could have been quicker and thought they could have supported 8 to 12 patients per day rather than 5 to 7.

The qualitative analysis of the peer supporters' text messages to the participants (n=81) resulted in 3 themes: reinforcement of resources, being present for the patient, and being personal.

Significant for the theme "Reinforcement of resources" was how the peer supporters reinforced the participants' work with the treatment by paying attention to and encouraging the participants to work with the treatment program. They reinforced participants' engagement in positive behaviors or behaviors related to their valued direction or goals (such as self-exposure to anxiety or using strategies from the treatment program) and encouraged them to express and share their opinions about the treatment's form, content, and specific exercises. The peer supporters also invited dialogue with participants through questions such as "What do you think about the treatment program?" or "How do you feel you are doing with the treatment?" The theme "Being present for the patient" represents the various ways in which peer supporters showed the participants that they were to support them. The peer supporters validated the participants' difficulties, expressed empathy, and encouraged them to get in touch with them if they had any questions or difficulties. The peer supporters asked about the participants' lives, but also created connection with the participants by using everyday expressions such as "Happy weekend" or "I wish you a happy Easter." The peer supporters also encouraged the participants to continue working with the treatment program through written reminders and by prompting upcoming exercises. The theme "Being personal" refers not only to how the peer supporters shared their own thoughts and opinions about the treatment program and various exercises that they thought were helpful for them, but also to how they shared their own experiences of dealing with and handling similar difficulties in various situations. The peer supporters also made use of self-disclosure by telling the participants about situations they found difficult and emphasizing self-acceptance and self-compassion when confronting difficulties in life.

Discussion

Principal Findings

To our knowledge, this is the first study to embed peer supporters in primary care iCBT treatment for anxiety. We used a naturalistic single-arm mixed methods approach to investigate the feasibility, safety, experiences, acceptability, and preliminary effectiveness of a peer-supported 8-week iCBT program on anxiety, depression, psychological distress, empowerment, and adherence to treatment. The results show that such a program

is feasible and that participants appreciated the peer supporters' help. The study was conducted in a clinical setting and thus supports the feasibility of adding peer support to iCBT treatment in real-world settings. A common criticism of efficacy studies (ie, studies conducted in research settings with rigorous design methods and with strict selection of research participants) is that research conditions do not accurately simulate the real world and thus have high internal validity but limited external validity [42]. Our sample was largely representative of the patient profile in primary care, which includes a range of socioeconomic and clinical backgrounds. This study also provides evidence of the general feasibility and effectiveness of iCBT in clinical (specifically primary care) settings, as only a few studies on iCBT have been conducted in clinical settings [12]. The quantitative results further support previous research showing that iCBT is effective both for mild-to-moderate problems [5-8] and for more severe mental illness [43]. The participants in this study had symptoms of anxiety and depression above the cutoff thresholds, indicating that iCBT treatment might be effective for more severe psychological distress.

The quantitative and qualitative results of this study support the presumptive positive effect of peer support. The quantitative results showed a significant reduction in anxiety and depression after treatment, which was maintained at the 3-month follow-up assessment. There was also a significant reduction in general psychological distress at the end of the treatment. Statistical data were also supported by clinical measures and qualitative data. The participants felt that the treatment contributed to their feeling of being able to handle their anxiety differently. Participants also declared that they felt their lives had expanded and, that after treatment, they could do things they had previously avoided.

The combined results show an increase in the participants' sense of empowerment connected both to their contact with the peer supporter and the actual tools in the iCBT program. Several participants described how contact with a peer supporter helped them see that it was possible to feel better. The treatment program was seen as helpful because it provided tools to overcome psychological barriers associated with anxiety, and overcoming these barriers allowed the participants to feel more empowered. The quantitative results also showed a significant increase in the participants' sense of empowerment between the start of treatment and the 3-month follow-up after completion. The quantitative and qualitative results in this study showing increased empowerment are in line with previous research showing that peer support interventions are related to increased measures of self-efficacy, hope, and empowerment [22,26]. These observations reflect several key learning points in peer-supported interventions, such as focusing on strengthening the patient's resources. In addition, when peer supporters share experiences of their own paths from mental illness to recovery, they can function as role models and thus provide hope for patients.

Prior research shows that guided iCBT is generally more effective than unguided [6,7]; however, little is known about what contributes to effective guidance in internet-mediated treatment. The internet provides enormous possibilities for disseminating evidence-based psychological treatment; however, it provides little personalized contact and might seem more effortful to patients. This possible disadvantage was evident in this study with participants saying that although internet-mediated treatment provides flexibility, it demands more responsibility from patients than face-to-face treatment. This is the only study that we know of that has qualitatively analyzed participants' experiences of peer support in iCBT, but several prior qualitative studies have analyzed participants' experiences of therapist-supported iCBT. The results of this study regarding patient perceptions of the treatment medium are in line with previous qualitative studies. The flexibility of iCBT is often perceived both positively, by contributing to patients' experiences as the primary agents of their own change and their ability to choose when and how to receive the treatment [44-46], and negatively, by placing more responsibility on patients, which can be experienced as demanding and might contribute to difficulties in engaging with the treatment [47,48]. Many studies also emphasize support from the therapist as an important factor in treatment outcomes [44-46,49].

Based on this study's results, one factor that can contribute to effective guidance in psychological internet treatment might be more personal and self-disclosing messages, which become even more important in a digital context than in normal face-to-face treatment, where alliance-forming factors such as body language, tone, and implicit validation strategies are lost. The analysis of the text messages sent from peer supporters to participants in the treatment program in this study shows how peer supporters made use of self-disclosure, shared their own experiences of dealing with difficulties in life, and shared their personal reflections on the content and tools they thought were helpful in the program. The usefulness of this was strengthened by the qualitative results, in which several participants emphasized the personal connection they felt when they were able to share their difficulties with someone who had similar experiences.

Adherence to treatment in this study was above that commonly seen in other studies on iCBT, with a recent systematic review and meta-analysis showing an average dropout rate of 15.7% in guided iCBT treatment for psychiatric and somatic conditions [7]. In this study, despite its small sample size, no participant dropped out of treatment and 67% (6/9) completed more than half of the treatment modules. These results are in line with previous studies on digital peer specialists in internet treatment [21,27,28], supporting the hypothesis that peer support can enhance treatment engagement. Moreover, in this study, less time was required by the therapist than usual in iCBT at the participating unit since the study was designed to limit support from the clinicians in favor of peer support. These results support the feasibility of this study. This is also in line with a previous study on peer-supported iCBT interventions, which showed that less support from therapists was needed in the peer-supported group [29]. A recent meta-analysis on iCBT for anxiety and depression disorder reported that therapists reporting spending from 18 to 352 min per patient ranged over the treatment period, indicating large general between-therapist differences in time allotted to iCBT [50]. In this study, clinicians spent roughly 120 min per patient over the entire treatment

period. Less time from the therapist can potentially increase the scalability of iCBT and increase access to care for patients.

The qualitative results from this study also pointed to some improvements that could be made in the design of the study. Some participants stated that their contact with the peer supporter was not fully utilized. Stated reasons for difficulties in engaging with the peer supporter included a lack of knowledge about the peer supporter's background and training, and a treatment period too short to create a real connection with each other. Some participants and peer supporters (in discussions during supervision) suggested that more verbal contact could be useful for creating a sense of connection, especially early in treatment. Prior research suggests that to be effective, it is important that peer support services reflect cultural diversity [20]. Research on self-help groups shows that when participants perceive the other people in the group to be similar to themselves, they are more likely to continue attending the group [20]. For administrative reasons, this study had only 2 PSWs, which limited their cultural diversity and might have influenced the perceived helpfulness of the peer supporters for some participants.

Future studies could contribute further by exploring how interactions between participants and peer supporters can be enhanced.

Limitations

This study was an uncontrolled feasibility study, so it is not possible to rule out factors other than the treatment affecting the outcomes. We also do not know whether similar results would have been seen in iCBT treatment without peer support for this specific group of participants. The study also has a limited sample size, which limits both the conclusions that can be drawn from the quantitative assessments and the transferability of the qualitative data. Data saturation is defined as the point when additional data no longer add new information to the analysis, and it is used to determine when the recruitment of new interview participants can be terminated. The themes formulated in this study thus cannot claim transferability but can guide the development of future peer-supported services embedded within iCBT treatments. Despite the limited sample size in this study, a clear strength is the combination of several different analytic methods, including statistical analysis, clinical measures, and qualitative analysis, all of which support the conclusion that peer support is a feasible addition to iCBT.

In light of its limitations, this study should be replicated in larger sample sizes, different populations, different contexts, and randomized controlled trials.

Conclusions

This study supports the feasibility of adding peer support to iCBT for adults with anxiety disorders. The results suggest preliminary support for the effectiveness of peer support on patient empowerment, reduction in anxiety, depression, and psychological distress, and adherence to treatment. Qualitative results also suggest that clinicians may be more effective by allowing themselves, such as PSWs, to be more personal and self-disclosing in their messages in the treatment program. Peer support might therefore contribute to more effective guidance in internet-based psychological treatment and might counter the loss of alliance-forming factors such as body language, tone, and implicit validation present in physical encounters. Future studies should validate the findings of this study with larger sample sizes and randomized controlled trials.

Acknowledgments

The authors would like to offer their special thanks to the participants for taking part in the study and to the peer supporters for their admirable work with supporting the participants. The authors extend a warm thank you to the patient organization, NSPHiG, with chairman Sonny Wåhlstedt and coordinator Filippa Gagnér Jenneteg, for their collaboration and for making it possible to carry out this study. They also extend thanks to Agneta Petersson, Anders Almingefelt, and Ann-Sofie Larsson from the participating iCBT unit, who carried out diagnostic assessment interviews, iCBT treatments, and interviews with the participants, and to Anna Larsson as team coordinator. The authors would also like to thank Lars Ström and Livanda-Internetkliniken AB, developer of the iCBT program used in the study; the management group for Närhälsan, primary care, region of Södra Älvsborg for providing the necessary resources; and R&D Primary Health Care, Västra Götaland region for financial sponsoring, and Kunskapsstöd för psykisk hälsa (KPH) for contributing valuable opinions on the study design.

The study was sponsored by R&D Primary Health Care, Västra Götaland region. R&D Primary Health Care had no role in the study design, data collection or analysis, writing of the report, or decision to submit the paper for publication.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Flowchart of the study.
[DOCX File , 33 KB - formative_v4i8e19226_app1.docx]

References

1. Ängslan, Oro Eller Angest [Unease, Worry or Anxiety]. Folkhälsomyndigheten — Myndigheten För Folkhälsofrågor (Public Health Agency of Sweden). 2020. URL: https://www.folkhalsomyndigheten.se/folkhalsorapportering-statistik/tolkad-rapportering/folkhalsans-utveckling/resultat/halsa/angslan-oro-eller-angest/ [accessed 2020-08-05]
2. National Board of Health. Nationella Riktlinjer För Vård Vid Depression Och Ångestsyndrom: Stöd För Styrning Och Ledning [National Guidelines for Care in Depression and Anxiety Disorders: Support for Governance and Management]. Stockholm, Sweden: National Board of Health; 2017.
3. Otte C. Cognitive behavioral therapy in anxiety disorders: current state of the evidence. Dialogues Clin Neurosci 2011;13(4):413-421 [FREE Full text] [Medline: 22275847]
4. O'Dea B, Calear AL, Perry Y. Is e-health the answer to gaps in adolescent mental health service provision? Curr Opin Psychiatry 2015 Jul;28(4):336-342. [doi: 10.1097/YCO.0000000000000170] [Medline: 26001926]
5. Kumar V, Sattar Y, Bseiso A, Khan S, Rutkofsky IH. The effectiveness of internet-based cognitive behavioral therapy in treatment of psychiatric disorders. Cureus 2017 Aug 29;9(8):e1626 [FREE Full text] [doi: 10.7759/cureus.1626] [Medline: 29098136]
6. Andersson G, Titov N. Advantages and limitations of internet-based interventions for common mental disorders. World Psychiatry 2014 Feb;13(1):4-11 [FREE Full text] [doi: 10.1002/wps.20083] [Medline: 24497236]
7. Carlbring P, Andersson G, Cuijpers P, Riper H, Hedman-Lagerlöf E. Internet-based vs face-to-face cognitive behavior therapy for psychiatric and somatic disorders: an updated systematic review and meta-analysis. Cogn Behav Ther 2018 Jan;47(1):1-18. [doi: 10.1080/16506073.2017.1401115] [Medline: 29215315]
8. Andersson G. Internet interventions: past, present and future. Internet Interv 2018 Jun;12:181-188 [FREE Full text] [doi: 10.1016/j.invent.2018.03.008] [Medline: 30135782]
9. Carlbring P, Nordgren LB, Furmark T, Andersson G. Long-term outcome of internet-delivered cognitive-behavioural therapy for social phobia: a 30-month follow-up. Behav Res Ther 2009 Oct;47(10):848-850. [doi: 10.1016/j.brat.2009.06.012] [Medline: 19631312]
10. Paxling B, Almlöv J, Dahlin M, Carlbring P, Breitholtz E, Eriksson T, et al. Guided internet-delivered cognitive behavior therapy for generalized anxiety disorder: a randomized controlled trial. Cogn Behav Ther 2011;40(3):159-173. [doi: 10.1080/16506073.2011.576699] [Medline: 21770848]
11. El Alaoui S, Hedman E, Kaldo V, Hesser H, Kraepelien M, Andersson E, et al. Effectiveness of internet-based cognitive-behavior therapy for social anxiety disorder in clinical psychiatry. J Consult Clin Psychol 2015 Oct;83(5):902-914. [doi: 10.1037/a0039198] [Medline: 26009780]
12. Olthuis JV, Watt MC, Bailey K, Hayden JA, Stewart SH. Therapist-supported Internet cognitive behavioural therapy for anxiety disorders in adults. Cochrane Database Syst Rev 2015 Mar 5(3):CD011565. [doi: 10.1002/14651858.CD011565] [Medline: 25742186]
13. Titov N, Andrews G, Johnston L, Schwencke G, Choi I. Shyness programme: longer term benefits, cost-effectiveness, and acceptability. Aust N Z J Psychiatry 2009 Jan;43(1):36-44. [doi: 10.1080/00048670802534424] [Medline: 19085526]
14. Eriksson MC, Kivi M, Hange D, Petersson E, Ariai N, Häggblad P, et al. Long-term effects of Internet-delivered cognitive behavioral therapy for depression in primary care - the PRIM-NET controlled trial. Scand J Prim Health Care 2017 Jun;35(2):126-136 [FREE Full text] [doi: 10.1080/02813432.2017.1333299] [Medline: 28585868]
15. van Ballegooijen W, Cuijpers P, van Straten A, Karyotaki E, Andersson G, Smit JH, et al. Adherence to Internet-based and face-to-face cognitive behavioural therapy for depression: a meta-analysis. PLoS One 2014;9(7):e100674 [FREE Full text] [doi: 10.1371/journal.pone.0100674] [Medline: 25029507]
16. 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]
17. Farkas M, Boevink W. Peer delivered services in mental health care in 2018: infancy or adolescence? World Psychiatry 2018 Jun;17(2):222-224 [FREE Full text] [doi: 10.1002/wps.20530] [Medline: 29856537]
18. Ibrahim N, Thompson D, Nixdorf R, Kalha J, Mpango R, Moran G, et al. A systematic review of influences on implementation of peer support work for adults with mental health problems. Soc Psychiatry Psychiatr Epidemiol 2020 Mar;55(3):285-293. [doi: 10.1007/s00127-019-01739-1] [Medline: 31177310]
19. Chinman M, George P, Dougherty RH, Daniels AS, Ghose SS, Swift A, et al. Peer support services for individuals with serious mental illnesses: assessing the evidence. Psychiatr Serv 2014 Apr 1;65(4):429-441. [doi: 10.1176/appi.ps.201300244] [Medline: 24549400]
20. Solomon P. Peer support/peer provided services underlying processes, benefits, and critical ingredients. Psychiatr Rehabil J 2004;27(4):392-401. [doi: 10.2975/27.2004.392.401] [Medline: 15222150]
21. Fortuna KL, Naslund JA, LaCroix JM, Bianco CL, Brooks JM, Zisman-Ilani Y, et al. Digital peer support mental health interventions for people with a lived experience of a serious mental illness: systematic review. JMIR Ment Health 2020 Apr 3;7(4):e16460 [FREE Full text] [doi: 10.2196/16460] [Medline: 32243256]
22. Mahlke CI, Priebe S, Heumann K, Daubmann A, Wegscheider K, Bock T. Effectiveness of one-to-one peer support for patients with severe mental illness - a randomised controlled trial. Eur Psychiatry 2017 May;42:103-110. [doi: 10.1016/j.eurpsy.2016.12.007] [Medline: 28364685]

23. Rogers ES, Teague GB, Lichenstein C, Campbell J, Lyass A, Chen R, et al. Effects of participation in consumer-operated service programs on both personal and organizationally mediated empowerment: results of multisite study. J Rehabil Res Dev 2007;44(6):785-799 [FREE Full text] [doi: 10.1682/jrrd.2006.10.0125] [Medline: 18075937]
24. Dennis C. Peer support within a health care context: a concept analysis. Int J Nurs Stud 2003 Mar;40(3):321-332. [doi: 10.1016/s0020-7489(02)00092-5] [Medline: 12605954]
25. Pitt V, Lowe D, Hill S, Prictor M, Hetrick SE, Ryan R, et al. Consumer-providers of care for adult clients of statutory mental health services. Cochrane Database Syst Rev 2013 Mar 28(3):CD004807. [doi: 10.1002/14651858.CD004807.pub2] [Medline: 23543537]
26. Bellamy C, Schmutte T, Davidson L. An update on the growing evidence base for peer support. Mental Health Soc Incl 2017 Jun 12;21(3):161-167. [doi: 10.1108/mhsi-03-2017-0014]
27. Proudfoot J, Parker G, Manicavasagar V, Hadzi-Pavlovic D, Whitton A, Nicholas J, et al. Effects of adjunctive peer support on perceptions of illness control and understanding in an online psychoeducation program for bipolar disorder: a randomised controlled trial. J Affect Disord 2012 Dec 15;142(1-3):98-105. [doi: 10.1016/j.jad.2012.04.007] [Medline: 22858215]
28. Simon GE, Ludman EJ, Goodale LC, Dykstra DM, Stone E, Cutsogeorge D, et al. An online recovery plan program: can peer coaching increase participation? Psychiatr Serv 2011 Jun;62(6):666-669 [FREE Full text] [doi: 10.1176/ps.62.6.pss6206_0666] [Medline: 21632737]
29. Tomasino KN, Lattie EG, Ho J, Palac HL, Kaiser SM, Mohr DC. Harnessing peer support in an online intervention for older adults with depression. Am J Geriatr Psychiatry 2017 Oct;25(10):1109-1119 [FREE Full text] [doi: 10.1016/j.jagp.2017.04.015] [Medline: 28571785]
30. Creswell JW, Plano Clark VL. Designing and Conducting Mixed Methods Research. Third Edition. Thousand Oaks, CA: Sage Publications; 2018.
31. World Health Organization. International Statistical Classification of Diseases and Related Health Problems. Tenth Edition. Geneva, Switzerland: World Health Organization; 2015.
32. Tamburrino MB, Lynch DJ, Nagel RW, Smith MK. Primary care evaluation of mental disorders (PRIME-MD) screening for minor depressive disorder in primary care. Prim Care Companion J Clin Psychiatry 2009;11(6):339-343 [FREE Full text] [doi: 10.4088/PCC.08.m00711] [Medline: 20098526]
33. Repper J, Carter T. A review of the literature on peer support in mental health services. J Ment Health 2011 Aug;20(4):392-411. [doi: 10.3109/09638237.2011.583947] [Medline: 21770786]
34. Oromendia P, Orrego J, Bonillo A, Molinuevo B. Internet-based self-help treatment for panic disorder: a randomized controlled trial comparing mandatory versus optional complementary psychological support. Cogn Behav Ther 2016 Jun;45(4):270-286. [doi: 10.1080/16506073.2016.1163615] [Medline: 27007256]
35. Spitzer RL, Kroenke K, Williams JB, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006 May 22;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
36. Barkham M, Bewick B, Mullin T, Gilbody S, Connell J, Cahill J, et al. The CORE-10: a short measure of psychological distress for routine use in the psychological therapies. Couns Psychother Res 2013 Mar;13(1):3-13. [doi: 10.1080/14733145.2012.729069]
37. Montgomery SA, Asberg M. A new depression scale designed to be sensitive to change. Br J Psychiatry 1979 Apr;134:382-389. [doi: 10.1192/bjp.134.4.382] [Medline: 444788]
38. Rogers ES, Chamberlin J, Ellison ML, Crean T. A consumer-constructed scale to measure empowerment among users of mental health services. Psychiatr Serv 1997 Aug;48(8):1042-1047. [doi: 10.1176/ps.48.8.1042] [Medline: 9255837]
39. Friedman M. The use of ranks to avoid the assumption of normality implicit in the analysis of variance. J Am Stat Assoc 1937 Dec;32(200):675-701. [doi: 10.1080/01621459.1937.10503522]
40. Field A. Discovering statistics using IBM SPSS statistics. London, UK: Sage Publications; 2018.
41. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
42. Möller HJ. Effectiveness studies: advantages and disadvantages. Dialogues Clin Neurosci 2011;13(2):199-207 [FREE Full text] [Medline: 21842617]
43. Andrews G, Williams AD. Up-scaling clinician assisted internet cognitive behavioural therapy (iCBT) for depression: a model for dissemination into primary care. Clin Psychol Rev 2015 Nov;41:40-48. [doi: 10.1016/j.cpr.2014.05.006] [Medline: 25043445]
44. Lenhard F, Vigerland S, Engberg H, Hallberg A, Thermaenius H, Serlachius E. 'On my own, but not alone' - adolescents' experiences of internet-delivered cognitive behavior therapy for obsessive-compulsive disorder. PLoS One 2016;11(10):e0164311 [FREE Full text] [doi: 10.1371/journal.pone.0164311] [Medline: 27711249]
45. Holst A, Nejati S, Björkelund C, Eriksson MC, Hange D, Kivi M, et al. Patients' experiences of a computerised self-help program for treating depression - a qualitative study of Internet mediated cognitive behavioural therapy in primary care. Scand J Prim Health Care 2017 Mar;35(1):46-53 [FREE Full text] [doi: 10.1080/02813432.2017.1288813] [Medline: 28277055]

46. Lillevoll KR, Wilhelmsen M, Kolstrup N, Høifødt RS, Waterloo K, Eisemann M, et al. Patients' experiences of helpfulness in guided internet-based treatment for depression: qualitative study of integrated therapeutic dimensions. J Med Internet Res 2013 Jun 20;15(6):e126 [FREE Full text] [doi: 10.2196/jmir.2531] [Medline: 23786763]
47. Johansson O, Michel T, Andersson G, Paxling B. Experiences of non-adherence to internet-delivered cognitive behavior therapy: a qualitative study. Internet Interv 2015 May;2(2):137-142. [doi: 10.1016/j.invent.2015.02.006]
48. Halmetoja CO, Malmquist A, Carlbring P, Andersson G. Experiences of internet-delivered cognitive behavior therapy for social anxiety disorder four years later: a qualitative study. Internet Interv 2014 Jul;1(3):158-163 [FREE Full text] [doi: 10.1016/j.invent.2014.08.001]
49. Asplund RP, Jäderlind A, Björk IH, Ljótsson B, Carlbring P, Andersson G. Experiences of internet-delivered and work-focused cognitive behavioral therapy for stress: a qualitative study. Internet Interv 2019 Dec;18:100282 [FREE Full text] [doi: 10.1016/j.invent.2019.100282] [Medline: 31737490]
50. Andrews G, Basu A, Cuijpers P, Craske MG, McEvoy P, English CL, et al. Computer therapy for the anxiety and depression disorders is effective, acceptable and practical health care: an updated meta-analysis. J Anxiety Disord 2018 Apr;55:70-78 [FREE Full text] [doi: 10.1016/j.janxdis.2018.01.001] [Medline: 29422409]

Abbreviations

ACT: acceptance and commitment therapy
ANOVA: analysis of variance
CBT: cognitive behavioral therapy
CORE-10: Clinical Outcomes in Routine Evaluation 10
GAD: generalized anxiety disorder
iCBT: internet-delivered cognitive behavioral therapy
MADRS-S: Montgomery-Åsberg Depression Rating Scale–Self report
NSPHiG: National Cooperation for Mental Health in Gothenburg
PSW: peer support worker

Edited by J Torous, G Eysenbach; submitted 09.04.20; peer-reviewed by K Fortuna, K Mathiasen; comments to author 12.05.20; revised version received 02.07.20; accepted 02.07.20; published 20.08.20.

Please cite as:
Nissling L, Fahlke C, Lilja JL, Skoglund I, Weineland S
Primary Care Peer-Supported Internet-Mediated Psychological Treatment for Adults With Anxiety Disorders: Mixed Methods Study
JMIR Form Res 2020;4(8):e19226
URL: http://formative.jmir.org/2020/8/e19226/
doi:10.2196/19226
PMID:32815819

Original Paper

Understanding the Experience of Cancer Pain From the Perspective of Patients and Family Caregivers to Inform Design of an In-Home Smart Health System: Multimethod Approach

Virginia LeBaron[1], PhD, APRN, FAANP; Rachel Bennett[1], MSN, RN; Ridwan Alam[2], MS; Leslie Blackhall[3], MD, MTS; Kate Gordon[4], MSN, RN; James Hayes[2], BS; Nutta Homdee[2], BS; Randy Jones[1], PhD, RN, FAAN; Yudel Martinez[2]; Emmanuel Ogunjirin[2]; Tanya Thomas[1], MSN, APRN; John Lach[5], PhD

[1]University of Virginia School of Nursing, Charlottesville, VA, United States

[2]University of Virginia School of Engineering & Applied Science, Charlottesville, VA, United States

[3]University of Virginia School of Medicine, Charlottesville, VA, United States

[4]Virginia Commonwealth University Health, Richmond, VA, United States

[5]The George Washington University School of Engineering & Applied Science, Washington, DC, United States

Corresponding Author:
Virginia LeBaron, PhD, APRN, FAANP
University of Virginia School of Nursing
225 Jeannette Lancaster Way
McLeod Hall
Charlottesville, VA, 22908
United States
Phone: 1 434 243 9291
Email: vtl6k@virginia.edu

Abstract

Background: Inadequately managed pain is a serious problem for patients with cancer and those who care for them. Smart health systems can help with remote symptom monitoring and management, but they must be designed with meaningful end-user input.

Objective: This study aims to understand the experience of managing cancer pain at home from the perspective of both patients and family caregivers to inform design of the Behavioral and Environmental Sensing and Intervention for Cancer (BESI-C) smart health system.

Methods: This was a descriptive pilot study using a multimethod approach. Dyads of patients with cancer and difficult pain and their primary family caregivers were recruited from an outpatient oncology clinic. The participant interviews consisted of (1) open-ended questions to explore the overall experience of cancer pain at home, (2) ranking of variables on a Likert-type scale (0, no impact; 5, most impact) that may influence cancer pain at home, and (3) feedback regarding BESI-C system prototypes. Qualitative data were analyzed using a descriptive approach to identity patterns and key themes. Quantitative data were analyzed using SPSS; basic descriptive statistics and independent sample *t* tests were run.

Results: Our sample (n=22; 10 patient-caregiver dyads and 2 patients) uniformly described the experience of managing cancer pain at home as stressful and difficult. Key themes included (1) unpredictability of pain episodes; (2) impact of pain on daily life, especially the negative impact on sleep, activity, and social interactions; and (3) concerns regarding medications. Overall, taking pain medication was rated as the category with the highest impact on a patient's pain (=4.79), followed by the categories of wellness (=3.60; sleep quality and quantity, physical activity, mood and oral intake) and interaction (=2.69; busyness of home, social or interpersonal interactions, physical closeness or proximity to others, and emotional closeness and connection to others). The category related to environmental factors (temperature, humidity, noise, and light) was rated with the lowest overall impact (=2.51). Patients and family caregivers expressed receptivity to the concept of BESI-C and reported a preference for using a wearable sensor (smart watch) to capture data related to the abrupt onset of difficult cancer pain.

Conclusions: Smart health systems to support cancer pain management should (1) account for the experience of both the patient and the caregiver, (2) prioritize passive monitoring of physiological and environmental variables to reduce burden, and (3) include

functionality that can monitor and track medication intake and efficacy; wellness variables, such as sleep quality and quantity, physical activity, mood, and oral intake; and levels of social interaction and engagement. Systems must consider privacy and data sharing concerns and incorporate feasible strategies to capture and characterize rapid-onset symptoms.

(JMIR Form Res 2020;4(8):e20836) doi:10.2196/20836

KEYWORDS

cancer; pain; sensors; smart health; caregiver; home based; palliative care; opioids; smart watch

Introduction

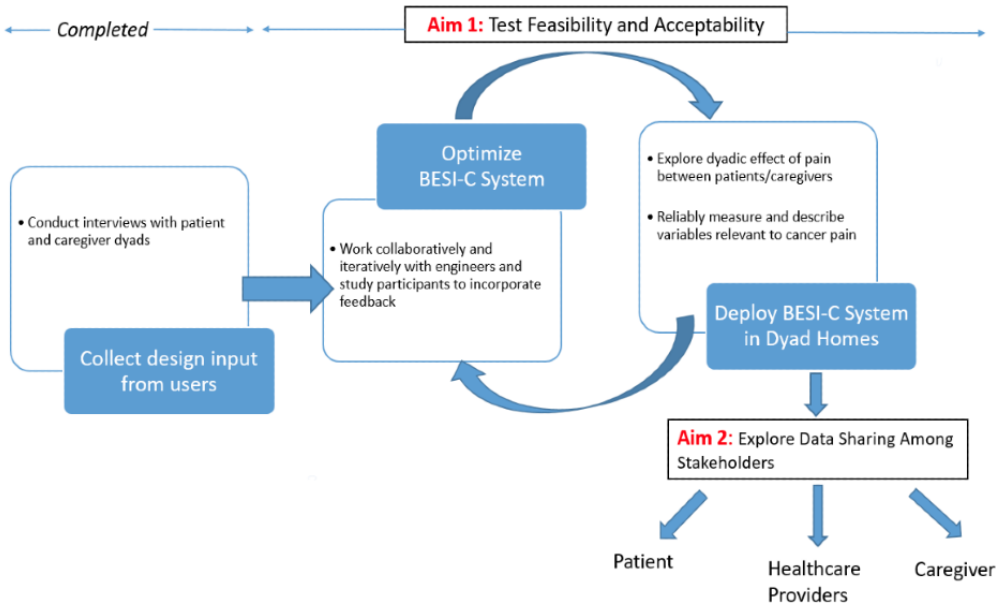
Inadequately managed pain continues to be a serious problem for patients with cancer and those who help care for them. An estimated 40% to 90% of patients with cancer experience pain across the illness continuum [1-3], negatively affecting sleep, adherence to treatment, mood, and overall quality of life [2,4]. Even patients with cancer enrolled in home hospice programs, which are uniquely designed to provide comprehensive support at the end of life, risk experiencing poorly managed symptoms [5-7]. One study found that >50% of hospice patients experience moderate to severe pain in the last week of life [8]. The majority of cancer symptom management occurs in the home setting, where family caregivers commonly play a key role in supporting patients. However, family caregivers are often required to make decisions about symptom management with limited information and support, which can significantly increase emotional distress [4,9,10]. In fact, working to control difficult pain is consistently rated as one of the most stressful tasks performed by family caregivers [11-15].

Ensuring quality home-based symptom management support is especially relevant for patients with advanced disease who may wish to forego aggressive curative treatments, avoid trips to the emergency department and hospitalizations, and focus on comfort care at home. For example, pain that escalates without adequate, prompt treatment can cause significant patient and caregiver distress as well as unplanned health care utilization/emergency department visits, which may not be compatible with patient goals at the end of life [16-19]. Recent studies have estimated that between 25% and 55% of emergency department visits for patients with advanced cancer are avoidable [16,17,20], and uncontrolled pain at home is a major reason that patients disenroll from hospice programs [21-23]. As health care adapts to the challenges and realities of COVID-19, home-based monitoring strategies are likely to become even more essential for seriously ill and immunocompromised patients who will be at higher risk for adverse outcomes if they must present to acute care settings for symptom management.

Although the literature richly describes the experience and consequences of poorly managed cancer pain within the home setting [4,5,24,25], gaps exist in understanding real-time, dynamic contextual factors that may worsen or mitigate the experience of cancer pain from the perspectives of patients *and* family caregivers [4,15,26-30]. Smart health (eg, wireless/mobile technology and user interfaces) is increasingly being utilized to improve remote symptom monitoring and management [31,32], but it is not always designed with meaningful end-user input [33] and may not be appropriate or feasible for the unique needs of patients and caregivers coping with the stressors of advanced, late-stage illness, limiting its ultimate utility and effectiveness [34,35]. Relatedly, ever-evolving technological capabilities can capture a large range of data, but it is not always clear which variables, especially environmental, are most essential and how they should be prioritized [30].

This research represents a multiphase effort to design and pilot test an in-home smart health system, known as the Behavioral and Environmental Sensing and Intervention for Cancer (BESI-C) system, with a palliative care oncology population to support patients and family caregivers in monitoring and managing cancer pain. The overall research protocol is described in detail elsewhere [36], but, briefly, BESI-C includes a package of environmental and wearable sensors and user interfaces deployed in patient homes to gather physiological, behavioral, contextual, and environmental data regarding pain events from the perspective of both patients and family caregivers. The ultimate goal of BESI-C is to successfully predict pain episodes and deliver real-time tailored interventions to both patients and caregivers as well as share relevant data with stakeholders. This manuscript presents results from phase I of the project (Figure 1), which aimed, from the perspective of both patients and family caregivers, to (1) explore the general experience and challenges of managing cancer pain in the home setting, (2) evaluate the role of specific variables that may influence cancer pain in the home setting, and (3) gather end-user input to inform BESI-C system design.

Figure 1. Overall study design. BESI-C: Behavioral and Environmental Sensing and Intervention for Cancer.



Methods

Study Design

This was a descriptive pilot study using a multimethod approach.

Setting

Patients and family caregivers were recruited from an academic palliative care outpatient oncology clinic from April to July 2018.

Sample

Our goal was to recruit patients and family caregivers managing difficult cancer-related pain in the home setting. Therefore, we used a purposive sampling technique [37], and patient inclusion criteria included (1) diagnosis of locally advanced or metastatic malignancy, (2) currently taking prescribed opioid medications (eg, morphine-type medications) for cancer-related pain, (3) a score of 6 on the National Institutes of Health Patient-Reported Outcomes Measurement Information System (NIH PROMIS) Cancer Pain Interference measures (a composite score assessed at each palliative care clinic visit to identify patients experiencing difficult pain) [38,39], and (4) a primary informal (unpaid; family, defined broadly) caregiver who helped manage their care and symptoms at home. Both patients and caregivers were aged >18 years, English speaking, and did not have cognitive or visual deficits that would preclude the ability to participate in the study. Palliative care clinicians helped screen and verify the clinical eligibility of possible study participants.

Data Collection Procedures

Before data collection, approval was granted by the University of Virginia Health Sciences Institutional Review Board. Both patients and caregivers provided informed consent. A study guide was created, informed by the literature and the research study aims (Multimedia Appendix 1). In addition to basic demographic questions, the study guide consisted of 3 parts.

Part 1

Part 1 consisted of open-ended questions regarding general challenges and concerns in managing cancer pain at home. Patients and caregivers were asked (1) Have you/the patient experienced cancer pain at home in the past week or so? If so, can you describe the experience from your perspective?; (2) What has been the most difficult part of managing pain at home?; and (3) What would help make managing the pain at home easier?

Part 2

Part 2 consisted of a list of variables that may influence cancer pain in the home setting that participants were asked to rank regarding impact. The list of variables was created based on their known relationship with cancer pain (such as the connection between sleep and pain) [40,41], current technological capabilities of the parent BESI system [42-44], and our hypotheses that certain environmental variables (eg, light and noise) that have received scant attention in the literature [26] can influence cancer pain. Overall, 14 variables were included in the final list and grouped into 4 categories: medication, wellness, interaction, and environmental. The primary objective of the variable list was to help inform the design of the BESI-C system (ie, which sensors to include) and validate our data collection plan.

Participants were asked to rate, on a Likert-type scale of 0 (no impact) to 5 (significant impact), the degree to which they thought each variable may influence the patient's experience of cancer-related pain. For example, patients were asked how much they felt their mood or the temperature of the room impacted their pain on a scale from 0 to 5. Caregivers were presented with the same list of variables and asked to quantify how much each factor influenced the patient's pain, *from their perspective as the caregiver.* Patients and caregivers were instructed that the study team was interested in their individual opinion and perspective and that it was fine if their answers differed from those of their partner. If a participant felt the *correct* answer was between 2 discrete values on the scale, they

could indicate a half-way point, for example, 3.5. If participants remained unsure and did not feel they could quantify the impact of a particular variable, the item was skipped. (Note: participants were not asked about the direction of the variable [eg, does light make your pain worse?], but instead if they felt the variable impacted their pain, either positively or negatively [eg, how much of an impact, negative or positive, does light have on your pain?]). After the set list of variables was reviewed, participants were asked if there were any additional factors they felt influenced cancer pain at home that they were not asked about (eg, "Are there other factors we did not ask you about that you think influence the experience of cancer pain at home? What did we miss?")

Part 3

Part 3 consisted of structured questions regarding the desired features of the BESI-C system. Participants were shown physical prototypes or pictures of the proposed components of the BESI-C system, including environmental room sensors, wearable sensors (smart watches), and a laptop base station used for remote system monitoring and local data processing and storage (Figures 2-5). Patients and caregivers were then asked about their general impressions, concerns, and suggestions regarding each system component. A key objective of this part of the interview was to ascertain how willing participants would be to interact with specific components of the system, for example, how often they would be willing to answer ecological momentary assessments (EMAs) [45] (brief survey questions) on a smart watch or if they had concerns about wearing a smart watch, in general. We were particularly interested in answering these specific design questions, as our goal was to create an unobtrusive smart health monitoring system that was acceptable, user friendly, and did not increase burden in an already highly stressed and vulnerable patient population.

Figure 2. Behavioral and Environmental Sensing and Intervention for Cancer (BESI-C) initial environmental sensor.

Figure 3. Behavioral and Environmental Sensing and Intervention for Cancer (BESI-C) updated environmental sensor based on user design input.

Figure 4. Behavioral and Environmental Sensing and Intervention for Cancer (BESI-C) base station laptop.

Figure 5. Behavioral and Environmental Sensing and Intervention for Cancer (BESI-C) smart watch with custom app.

Dyads were interviewed together or separately, generally in the outpatient palliative care clinic, based on their preference and logistic considerations. (Note: we recognize that this difference may have influenced participant responses, which is discussed in more detail below.) For convenience, caregivers who were unable to accompany the patient to the clinic were given the option to be interviewed over the phone (to complete part 3 over the phone, caregivers were provided with detailed verbal descriptions by the interviewer and/or pictures of the BESI-C components for visual reference). Interviews were audio-recorded with permission; detailed notes and responses were also recorded using pen and paper during all interviews by the research team member. Interviews lasted approximately 30 min, and dyads received a US $10 gift card as compensation for their time.

Data Analysis Procedures

Qualitative Data

Interviews were transcribed verbatim and verified, and all identifiers were removed. All transcripts were read in entirety before analysis to understand the data set holistically. Open-ended responses (Parts 1 and 3 of the interview) were exported into Microsoft Word and organized by response to each corresponding interview question, by patient and by caregiver (eg, all patient responses to question 1 were grouped together, and all caregiver responses to question 1 were grouped

together). A deductive, descriptive qualitative approach was then used to analyze interview responses across the entire data set from the perspective of patients and family caregivers. In keeping with our study aims, codes (or descriptive labels) were applied to portions of text that discussed the general experience of managing cancer pain at home and system design feedback to help identify patterns and key themes. For example, if a participant discussed fears related to pharmacological management of pain, this was coded with the straightforward label *medication concerns*. Themes were identified by considering both frequency of codes (how often a similar message was conveyed) and intensity of response (the strength of an articulated opinion, either negative or positive). Our goal with the analysis of open-ended responses was not to conduct qualitative analysis with a high level of abstraction, but instead, consistent with a descriptive approach, to stay close to our data and more concretely understand participant responses to each interview question [46].

Quantitative Data

Quantitative responses (part 2 of the interview) were entered into SPSS (v25.0), and basic descriptive statistics were run, including frequency counts and percentages for demographic data and means calculated (overall, patient, and caregiver) for individual and category (medication, wellness, interaction, and environmental) pain variable impact scores. Independent sample *t* tests were performed across all individual and category pain variables to assess for statistically significant differences (α set at .05) between patient and caregiver mean scores.

Results

Interviews

A total of 22 individuals were interviewed (22/22, 100%), including 10 patient-caregiver dyads and 2 individual patients, whose caregivers did not accompany them to the original clinic visit and were unable to be contacted after 3 attempts. A total of 5 dyads were interviewed together (dyads 6, 7, 9, 10, and 12), and 5 dyads were interviewed separately (dyads 1, 2, 3, 5, and 11). Of the 5 dyads interviewed separately, 2 caregivers were interviewed over the phone owing to logistic constraints. All other interviews were conducted face-to-face. Results are presented below by demographics and then by section of the interview guide (Parts 1, 2, and 3) for clarity.

Demographic

Overall, almost half of the total sample (10/22, 46%) was aged between 50 and 59 years, with an equal number of females and males (11/22, 50%). The participants were primarily White (20/22, 91%) and non-Hispanic/Latino (21/22, 96%). The majority of patients (11/12, 92%) had a primary residence classified by the Centers for Medicare and Medicaid Services as rural [47]. Caregivers were predominantly female (7/10, 70%), lived full time with the patient (9/10, 90%), and were the significant other or spouse (5/10, 50%). The average patient-reported NIH PROMIS pain interference score was 7.16 (0, lowest interference to 10, highest interference), and half of the patients (6/12, 50%) self-reported their performance status as symptomatic, but ambulatory and able to complete their basic needs independently (Eastern Cooperative Oncology Group [48] score of 1). The most common malignancy was lung cancer (4/12, 33%), and 50% (6/12) of patients received their diagnosis <1 year ago. Table 1 presents the demographic data for the overall sample, patients, and caregivers. Table 2 presents cancer-related details for the patient sample.

Table 1. Demographic characteristics of the patient and caregiver sample.

Demographic variable	Total (N=22), n (%)	Patients (n=12), n (%)	Caregivers (n=10), n (%)
Age range (years)			
18-29	2 (9.1)	0 (0)	2 (20.0)
30-39	2 (9.1)	1 (8.3)	1 (10.0)
40-49	4 (18.2)	2 (16.7)	2 (20.0)
50-59	10 (45.5)	7 (58.3)	3 (30.0)
60-69	3 (13.6)	1 (8.3)	2 (20.0)
>70	1 (4.5)	1 (8.3)	0 (0.0)
Gender			
Female	11 (50)	4 (33.3)	7 (70.0)
Male	11 (50)	8 (66.7)	3 (30.0)
Race			
Black/African American	1 (4.5)	1 (8.3)	0 (0.0)
White	20 (90.9)	11 (91.7)	9 (90.0)
Missing (not asked)	1 (4.5)	0 (0)	1 (10.0)
Ethnicity			
Latino/Hispanic	0 (0)	0 (0)	0 (0.0)
Non-Latino/Hispanic	21 (95.5)	12 (100)	9 (90.0)
Missing (not asked)	1 (4.5)	0 (0)	1 (10.0)
Rural[a]	N/A[b]	11 (91.7)	N/A
Highest education level			
Less than high school	5 (22.7)	3 (25)	2 (20.0)
High school graduate	8 (36.4)	4 (33.3)	4 (40.0)
Some college	2 (9.1)	1 (8.3)	1 (10.0)
2-year degree	2 (9.1)	2 (16.7)	0 (0.0)
4-year degree	4 (18.2)	2 (16.7)	2 (20.0)
Professional/graduate degree	1 (4.5)	0 (0)	1 (10.0)
Doctorate	0 (0)	0 (0)	0 (0.0)
Current employment			
Full time	4 (18.2)	0 (0)	4 (40.0)
Part time	1 (4.5)	1 (8.3)	0 (0.0)
Retired	5 (22.7)	2 (16.7)	3 (30.0)
Unemployed	12 (54.5)	9 (75)	3 (30.0)
Caregiver lives with patient			
Yes, full time	N/A	N/A	9 (90.0)
Yes, part time	N/A	N/A	1 (10.0)
Caregiver relationship with patient			
Significant other/spouse	N/A	N/A	5 (50.0)
Sibling	N/A	N/A	1 (10.0)
Parent	N/A	N/A	1 (10.0)
Child	N/A	N/A	2 (20.0)
Other (daughter-in-law)	N/A	N/A	1 (4.5)

[a]Rural determined by the Centers for Medicare and Medicaid Services based on patient's address of primary residence.
[b]Not applicable.

Table 2. Patient sample cancer characteristics (N=12).

Patient cancer variable	Total, n (%)
Primary cancer diagnosis	
Breast	1 (8)
Gastrointestinal (other)	1 (8)
Gastrointestinal (pancreatic)	1 (8)
Gynecological	1 (8)
Head and neck	2 (17)
Hematological[a]	1 (8)
Lung	4 (33)
Prostate	1 (8)
Time since diagnosis (years)	
<1	6 (50)
1-5	4 (33)
5-10	1 (8)
>10	1 (8)
Patient self-reported ECOG[b] score	
0, normal activity	0 (0)
1, symptomatic and ambulatory	6 (50)
2, ambulatory 50%, some help needed	2 (17)
3, ambulatory <50%, nursing care needed	3 (25)
4, no self-care, bedridden	0 (0)
Not available	1 (8)
NIH PROMIS[c] pain interference score[d] (n)	7.16 (12)

[a]Multiple myeloma.
[b]ECOG: Eastern Cooperative Oncology Group; standard patient performance scale.
[c]NIH PROMIS: National Institutes of Health Patient-Reported Outcomes Measurement Information System.
[d]Patient self-reported NIH PROMIS pain interference composite score, scored for clinical use on a scale of 0 (least) to 10 (most).

Part 1: Understanding the Experience of Managing Cancer Pain at Home

Patients and caregivers uniformly described the experience of managing cancer pain at home as stressful and difficult. Key themes included (1) unpredictability and perceived inevitability of pain episodes; (2) impact of pain on daily life, especially the negative impact on sleep, activity, and social interactions; and (3) concerns regarding medications. All 3 themes overlapped as they did not occur in isolation. For example, the unpredictability of pain episodes could be worse at night when pain medications did not seem to be as effective, thus affecting sleep. When asked what could make managing pain at home easier, one caregiver simply stated, "when he heals, and this goes away," (CG1). Others suggested ideas such as more rapid-acting interventions to relieve pain; reduced back-and-forth travel for medical appointments; a more holistic, nonpharmacological approach to managing pain; and better ways to track and record medication use. Textbox 1 summarizes and presents exemplar quotes related to part 1 of the interview.

Textbox 1. Experience of managing cancer pain at home from the perspective of patients (Pt) and caregivers (CG).

What is the experience of managing cancer pain at home? What is most difficult?

Theme 1: Unpredictability and perceived inevitability of pain

- "That's one thing about the cancer pain, is that you never know what you're going to experience." Pt 5
- "I do about all I can do. I don't see it being any easier. It just stays, you know, it's going to be what it's going to be. It's not going to get any better or any worse." Pt 7
- "No, I don't think you can manage the pain." Pt 9
- "It hits me so bad sometimes it brings tears to my eyes…When I'm in really, really bad pain it gets me down.
- I get depressed and it's like, 'God, is this ever gonna quit?'" Pt 11
- "Well, I know it hurts. Some days it looks worse than others." CG 2
- "The experience to me, he gives me a headache," CG 1

Theme 2: Impact of pain on daily life

- "Well, you know, I'm miserable... I pretty much became a hermit since this happened. You know, I try to stay away from everybody, so I don't have to talk very much. I stay in the bedroom, you know, and watch TV most the time so I don't have to talk to people." Pt 1
- "The intensity is worse in the evening at night...and also when I wake up in the morning. [The pain pills] are less effective at night...Sleep [is the most difficult part]. It's most frustrating when it [pain] has kept me awake or wakes me up." Pt 3
- "I'm an active person and with the pain I could barely get my shirt off." Pt 4
- "Miserable. Miserable, I don't do nothing. I can't." Pt 9
- "The bottom line is just be secluded when I am in pain...When I'm really, really in pain, if I'm alone it seems to soothe it…nothing there to irritate me to make it worse." Pt 11
- "He was in real agony for a couple weeks, so bad he couldn't sleep." CG3
- "I find it a little difficult...like she appears to be in pain, definitely lethargic and I think between the pain and feeling tired that definitely affects her mental health...so it's just kind of all blurred together." CG 5
- "It's really pulled her down. You know, we went from being outside every day and doing things to, you know, pretty much watch watching her lay on the couch." CG 9

Theme 3: Concerns regarding medications

- *Fear of running out of medication or becoming addicted*
  - "I'm concerned that somebody will say you can't have it anymore when I still need it. I know they're addictive...but for me they're necessary for the pain." Pt 3
  - I haven't looked it [my pain medication] up on-line but I kind of worry about how dangerous it is for the rest of my body." Pt 4
  - "The only thing that changes my pain is my medication. Especially if I got it, I use it right, it makes a big difference. But when I run out, well, I got problems." Pt 8
  - "That's my biggest fear is getting addicted." Pt 9
  - "I don't really have enough medications, I guess...I take them just as they are prescribed to me...it's frustrating after a while. Either I won't be able to sleep, do I want to be in pain or do I want to conserve the medications and if I'm gonna have enough or God forbid I lose some or whatever...I ask for some [pain medication], 'No.' I ask again. 'No.' I don't even ask them [health care providers] anymore. I'm tired of it. I feel like a little child asking for a piece of f*...ing candy. It's frustrating as hell." Pt 10
  - "...he could be in a whole lot less pain, but that's regulation...What makes it difficult is that he knows that if he took another half a tablet he would be in less pain but if he takes that half a tablet extra than he's gonna have to be in more pain later [because there are not enough tablets]." CG 10
- *Coping with side effects*
  - "They've had me on so many different medications and you can take 30 pills a day and still not get the relief you need…and that's hard on your body. You're dealing with all the different side effects..." Pt 5
  - "...a lot of times I think I'd almost prefer to be in pain sometimes and live a little bit of life than sleep my life away." Pt 10
  - "Approaching things from a more holistic point [would help]...it seems that it's just very much medication based and then side effects and then you treat those side effects with medications and then those side effects with other medications..." CG 5
  - "I don't think any of it is doing its job. I mean, it did at first, but I think that her body's just gotten so used to it and it's not doing what it was doing." CG 9

- *Keeping up with, and keeping track of, medications*
  - "Managing my pain, being on so many different medicines for pain, trying to make sure I take them all...it's time consuming." Pt 1
  - "Most difficult? Taking my medicine. Sometimes I'll take more than I should if I'm really in a lot of pain, and I know I'm not supposed to but it's hard not to." Pt 2
  - "It [oxycodone] eases it a heck of a lot...I can sit and relax once it kicks in but once it wears off I'm going right back in the same state again. The pain comes back...Then I say when it gets to the point where I just take one pill to kill the pain, fine, but taking two, you know, it could be a problem." Pt 11
  - "I don't always ask when I need it [pain medicine]." Pt 6
  - "I can see when you need it [pain medicine] but I don't just automatically give it to you...you've got to ask for it too…[it would be easier] if [my husband] would ask for [pain medicine] instead of me saying, 'do you need your pain medicine?'" CG 6

What would make managing pain at home easier?

- "Having something that would provide instantaneous relief because sometimes it just seems like it takes a long time for anything to take effect." Pt 3
- "Not having to keep going back and forth to the doctor so many times. It hurts her riding in the car." CG 2
- "Some way to track when you've actually taken something because he's writing it down, but when he was really dopey, he either didn't remember to write it down or he couldn't read what he wrote because he was so doped up. Something that would, I don't know what kind of technology there would be, but something that scans the pill bottle or something and says 'Ah, you've taken another one, so great, good for you'." CG 3

The unpredictability of pain manifested in both the timing of the pain, which could occur abruptly and severely, and the lack of clarity regarding the origin of the pain. Patients often had multiple potential sources of pain, such as rheumatoid arthritis and cancer; this made it difficult for participants to sort out which pain was related to cancer and which was not, and then how to best and most appropriately intervene. The impact of pain on daily life was particularly noted in the areas of sleep, activity, and social engagement. Patients and caregivers reported a vicious cycle related to pain and sleep: pain intensity could flare up at night, causing insomnia, which resulted in worsening of pain and social withdrawal. Some participants, primarily patients, expressed a fatalistic attitude that cancer pain is inevitable and inherently unmanageable, regardless of what they tried to do to alleviate or mitigate the pain.

A strong theme in the interviews was related to challenges regarding pharmacological management of pain. Both patients and caregivers gave specific examples detailing the significant labor—both logistic and emotional—involved in managing cancer pain medications. Logistically, managing pain involved time and discomfort of multiple trips to the clinic for medical appointments; vigilance to coordinate, monitor, and remember complex and ever-changing pain medication regimens; *keeping ahead* of the pain by remembering to take pain medications before the previous dose wears off; and coping with, and balancing, side effects such as the perceived tradeoff between having pain better controlled but becoming too drowsy. Emotionally, participants discussed frustration and deep fears about running out of prescription opioid pain medications, being unable to obtain needed refills, or becoming *addicted*. These fears often resulted in the rationing of tablets, further exacerbating pain and distress. Medications were viewed by both patients and caregivers as an essential, but imperfect, tool that offered temporary relief from the pain, but that came with a (metaphorically) high price tag.

Part 2: Variables that Influence Cancer Pain at Home

Table 3 compares the mean impact scores (0, no impact; 5, highest impact) for factors that may influence a patient's pain at home by *category* (medication, wellness, interaction, and environmental). Table 4 presents the ranking of the *individual* variables from 1 (highest scored factor) to 14 (lowest scored factor) by comparing mean impact scores by the overall sample, patient, and caregiver. No statistically significant differences between patient and caregiver mean scores were detected across all variables.

Overall, and for patients and caregivers, *taking pain medication* was rated as the category with the highest impact on a patient's pain (=4.79), followed by the categories *wellness* (=3.60; sleep quality/quantity, physical activity, and mood and oral intake) and *interaction*, (=2.69; busyness of home, social/interpersonal interactions, physical closeness/proximity to others, and emotional closeness/connection to others). The category related to *environmental* factors (temperature, humidity, noise, and light/brightness) was rated with the lowest overall impact (=2.51).

Regarding individual variables within each category, in the wellness category, the individual variables of sleep quality, sleep quantity, and physical activity were rated as having the most impact on a patient's pain by the overall sample (=4.28, =3.98, and =3.90, respectively), by patients (=3.91, 1=3.96, and =3.91, respectively), and by caregivers (=4.72, =4.00, and =3.89, respectively). In the interaction category, the variable busyness of home was rated with the highest impact score by the overall sample (=3.21) and by caregivers (=3.28). Patients rated *social and interpersonal interactions* as having the highest impact on their pain (=3.50). In the

environmental category, temperature was the highest rated variable by the overall sample (=3.30) and by patients (=3.63); caregivers rated humidity as the highest impact environmental variable (=3.22).

The rank order of individual variables (Table 4) revealed that pain medication, sleep quality/quantity, physical activity, and mood occupied the top 5 spots for the overall sample and for both patients and caregivers. Other variables related to environmental and contextual factors were ranked more diversely. Interestingly, emotional closeness/connection to others was the lowest ranked variable by patients, whereas *noise* was the lowest ranked for caregivers (but #9 for patients).

When asked if there were other variables that influenced cancer pain not included on our list that we should measure with a home monitoring system, only 2 participants identified additional variables. One caregiver (CG 1) stated that the amount the patient talks influenced his pain (patient had a diagnosis of head and neck cancer) and felt this was an important variable to assess. One patient (Pt 4) added "good support group of people to help" as a broader interpretation of our questions regarding the impact of emotional connection and social interaction. A total of 3 patients responded to this question by reiterating the importance of pain medication as the most important variable that influenced their cancer pain.

Table 3. Comparison of mean impact scores of factors that influence a patient's cancer pain at home, by category and individual variable, rated from 0 (no impact) to 5 (highest impact).

Category and individual variables[a]	Overall (N=22), n (%)	Patients (n=12), n (%)	Caregivers (n=10), n (%)
	(n[b])	(n)	(n)
Medication, category mean	4.79	4.79	4.78
Taking pain medication	4.79 (21)	4.79 (12)	4.78 (9)
Wellness, category mean	3.60	3.45	3.80
Sleep quality (how well)	4.28 (20)	3.91 (11)	4.72 (9)
Sleep quantity (how much)	3.98 (21)	3.96 (12)	4.00 (9)
Physical activity	3.90 (20)	3.91 (11)	3.89 (9)
Mood	3.45 (20)	3.55 (11)	3.33 (9)
Oral intake (eating/drinking)	2.43 (20)	1.91 (11)	3.06 (9)
Interaction, category mean	2.69	2.82	2.52
Busyness of home	3.21 (19)	3.15 (10)	3.28 (9)
Social/interpersonal interactions	2.97 (16)	3.50 (9)	2.29 (7)
Physical closeness/proximity to others	2.38 (21)	2.73 (11)	2.00 (10)
Emotional closeness/connection to others	2.20 (20)	1.90 (10)	2.50 (10)
Environmental, category mean	2.51	2.48	2.50
Temperature	3.30 (22)	3.63 (12)	2.90 (10)
Humidity	2.61 (18)	2.00 (9)	3.22 (9)
Noise	2.07 (21)	2.21 (12)	1.89 (9)
Light/brightness	2.05 (20)	2.08 (12)	2.00 (8)

[a]Instructions provided to participants during the interview: Please think back over the past few weeks or months. Patient: for each item, on a scale of 0-5 (0=not at all, 5=a great deal), how much do you think it makes your pain better or worse? Caregiver: for each item, on a scale of 0-5 (0=not at all, 5=a great deal), how much do you think it makes the patient's pain better or worse?

[b]Where "n" is not equal to the total sample, participant either was unsure/could not answer or the item was not asked (social/interpersonal interaction factor question was added after dyad 3).

Table 4. Rank order of individual variable impact means (0=no impact; 5=highest impact) on patient's pain.

Rank	Overall		Patient		Caregiver	
	Variable	Mean	Variable	Mean	Variable	Mean
1	Pain medication	4.79	Pain medication	4.79	Pain medication	4.78
2	Sleep quality	4.28	Sleep quantity	3.96	Sleep quality	4.72
3	Sleep quantity	3.98	Sleep quality (tie); physical activity (tie)	3.91	Sleep quantity	4.00
4	Physical activity	3.90	Temperature	3.63	Physical activity	3.89
5	Mood	3.45	Mood	3.55	Mood	3.33
6	Temperature	3.30	Social/interpersonal interactions	3.50	Busyness of home	3.28
7	Busyness of home	3.21	Busyness of home	3.15	Humidity	3.22
8	Social/interpersonal interactions	2.97	Physical closeness/proximity to others	2.73	Oral intake	3.06
9	Humidity	2.61	Noise	2.21	Temperature	2.90
10	Oral intake	2.43	Light/brightness	2.08	Emotional closeness/connection to others	2.50
11	Physical closeness/proximity to others	2.38	Humidity	2.00	Social/interpersonal interactions	2.29
12	Emotional closeness/connection to others	2.20	Oral intake	1.91	Physical closeness/proximity to others (tie); light/brightness (tie)	2.00
13	Noise	2.07	Emotional closeness/connection to others	1.90	Noise	1.89
14	Light/brightness	2.05	N/A[a]	N/A	N/A	N/A

[a]N/A: not applicable.

Part 3: Feedback Regarding the BESI-C System Components

The results presented in Textbox 2 focus on the 2 primary components of the BESI-C system: environmental and wearable (smart watch) sensors as well as general system impressions, suggestions, and concerns. (Participants expressed minimal or no concerns about the laptop base station, which we are currently removing from the system architecture and replacing with a cloud-based service for a simpler and less-intrusive system deployment and to facilitate more efficient data management.) Overall, patients and family caregivers expressed interest and receptivity to the concept of BESI-C, validated the importance of monitoring cancer pain at home, were eager for innovative ways in which to do so, and provided constructive feedback regarding the system components. However, there was the acknowledgment that providing feedback would involve actual use and pilot testing of the system. There was also the acknowledgment that a system such as BESI-C could be particularly helpful in assisting caregivers to *tune in* to variables that may influence a patient's pain, but that may not be readily obvious. Participants expressed a strong preference for technology that is unobtrusive, simple, convenient, durable, and aesthetically pleasing and that involves minimal interference with daily activities, such as sleep. For both environmental and wearable sensors, participants expressed concerns regarding privacy and a desire for multifunctionality (eg, could environmental sensors measure variables beyond those focused on cancer pain, such as general air quality in the home, or other symptom management issues, such as sleep apnea, and could wearables measure additional factors such as blood pressure or blood glucose levels).

Textbox 2. Feedback from patients and caregivers regarding the Behavioral and Environmental Sensing and Intervention for Cancer system components.

General impressions and interest

- "I think it's exciting that somebody's coming up with this. I really do...the information that you can get from the sensors and the watch." Pt 5
- "I think what you showed me is a good idea." Pt 8
- "I do too [think it would be cool] because anything to try to help stabilize the pain." Pt 9
- "Anything you all can come up with to help, I'd approve of anything. Yes, I would, because cancer is bad. It's very painful... It's just terrible." CG 2
- "No big deal, we like stuff like that...it's useful." Pt 7
- "Those are things that even if it's a question that you just ask, those are things that we don't pay attention to every day. You know, the stress level in a room, issues like that. We're not really tuned in to that." CG 9

Environmental Sesnors

- *Importance of household buy in*
  - "Wouldn't bother me, but may bother my Dad...if it transmits information somewhere else and it monitors stuff in his house he probably wouldn't like it." Pt 1
  - "That would be good to me...It might not bother me, but it's going to bother her [caregiver]." Pt 8
  - "With me, if this is only gonna be in my room, fine. It wouldn't bother me one bit. But as far as the one in the kitchen, the living room, her [daughter in law's] bedroom, I don't think she's gonna agree with that...You just got to figure out if we've got enough outlets to put these things in that won't interfere with her cooking, appliances, stuff like that... what I'm saying is run it by her and see what she thinks." Pt 11
- *Desire for multifunctionality*
  - "It reminds me of a little robot...that's wonderful, so I'm glad if it will work. Could you put a smoke detector in there?" Pt 4
  - "Well, with his condition, I mean it is a good idea for something like that, not just the cancer, with his sleep apnea and all that stuff and like his asthma and stuff. I mean, that would be a good idea [to monitor too]." CG 1
  - "I think that's interesting. I would like to know exactly, you know, what I'm breathing, and you know, the air and stuff in the house." Pt 2
  - "Is there a way to monitor diet or when someone's eating or not eating? I just know with my mom sometimes when she's feeling a lot of pain, she could go an entire day without eating." CG 5
- *Privacy considerations and data sharing*
  - "Cool. I'd want to know–okay, so it measures all of that–then what does it do with it? Does it spit it out at the doctor's office? Can you get it through an app? Can you look at what it's doing?" CG 3
  - "As long as it ain't watching us." CG 9
  - "My major concern would be the privacy." Pt 9
  - "I would be quite concerned if it's recording what I'm saying." Pt 10

Smart watch

- *Desire for simplicity and comfort*
  - "Should be super simple like the old people's cellphone, the Jitterbug. I get up in the middle of the night and I don't have my glasses on...so it's got to be really self-explanatory. Like you look at it, and you go, 'Red is bad, green is good'." CG 3
  - "I don't like jewelry on me and stuff on my wrist...working on cars and stuff, a watch gets in the way." Pt 1
  - "I'd be concerned about how comfortable it is, how easy it would be to put on, about finding it if I took it off 'cause I tend to lose things like watches." Pt 3
  - "I see it possibly interfering just with work maybe. Just because of the work that I do [manual labor]. But other than that, I mean I think on her it can be beneficial." CG 5
  - "I wouldn't want it to take up my life, but I would be willing to try it." Pt 4
- *Privacy considerations and interfacing with the technology*
  - "People our age, it's stereotypical, but it's way too small. I cannot imagine trying to answer a question on that." CG3
  - "I mean she [the patient] could wear it, but I don't know if I could, to be honest with you. When it comes to this high tech stuff I don't know nothing about it. I wouldn't mind wearing it, if I could learn how to work it. You know, I would love to do it." CG 2
  -

"As long as it's not picking up on my personal stuff, I wouldn't have a problem with it...I just don't want my personal life put out there." Pt 9

- *General impressions of smart watch*
  - "I think it's a good idea myself. I think that would make a big different in monitoring some things." CG 6
  - [Could be tricky to remember to mark pain events]: "Would it give like an alert or something or some noise? I think once I get used to it I think it would be great [and I would remember]" Pt 2
  - [Willing to answer questions multiple times a day]: "whenever I needed to." CG 11
  - "It would be an annoyance to you because you'd have to answer it all the time...It does look nice, though. It looks nice and sleek." CG 12
  - "Convenient...doesn't take any space up." Pt 11
  - "I think it would be beneficial, but it would be a pain." Pt 12

Smart watch versus tablet to answer ecological momentary assessments (EMAs)

- "The watch because it's with me all the time, even when I'm not at home." Pt 2
- "Prefer iPad...it just looks easier to use. I assume it's bigger...I would find the watch the most inconvenient, because I'd always have to have it on or keep track of it or whatever. Something I could just put in one place and forget about would be better for me." Pt 3
- "iPad may not be as accurate because you're not going to remember everything after the fact." Pt 4
- "I'd rather do the watch...it would be easier for me...this is attached." Pt 5
- "I think I'd like the tablet more but that's me." Pt 6
- "It [the watch] would be a lot easier for me...I'm no electronic expert, you know what I'm saying? I don't deal with computers...[willing to interact with watch multiple times a day] anytime they [EMAs] popped up, as long as I'm not asleep." Pt 11
- "[I'd prefer the watch] because I can't stand an iPad. I had one at work and I just could not." Pt 12
- "I think having a watch, having it all together in one unit would just be more streamlined so you don't have to keep up with multiple devices." CG 5
- "Probably for a lot of people it would be easier on the tablet, but my only thought to that is...would somebody actually go for the tablet and answer it? You know, if it's on the watch you would do it automatically cause it's right there." CG 6
- "Either one is fine with me...in the summertime I really don't want a tan line...so I would want to take a watch off if I had to wear it." CG 11

System suggestions/concerns

- "You could set intervals at different times to remind you to take different types of medicine at different times; [that] would be about the best thing I know." Pt 1
- "For older people, as simple as possible. When you get to be a lot older you really don't want to have to fuss with a lot of things that aren't central to your condition...if you're in pain this kind of stuff's going to go out the window, so I should say as simple as possible, absolutely as simple as possible." CG 3
- "We just have really horrible internet...our internet is just off of a hotspot from my cell service. That's the only internet we have." CG 5
- "The only concern I would have is, is it going to be like making noise and stuff like that?" CG 11
- "Those are things that even if it's a question that you just ask, those are things that we don't pay attention to every day. You know, the stress level in a room, issues like that. We're not really tuned in to that." CG 9
- "I've got to see what all we gonna have set up in there. I've gotta see how comfortable I am with this by seeing how it all works, you know, and then I can give you my opinion, based on everything." Pt 11

Dyadic effect of technology

- "I think most people it wouldn't bother. For me it's different. I don't like jewelry." [Pt 6–then after hearing wife's receptiveness to the watch said:] "I think it would [make a difference]. I would make myself wear it if I had to." Pt 6
- "Yeah, she's [caregiver] taking it kind of hard..." Pt 8, [when explained how the watch would help monitor CG too.]
- "As long as she [patient] doesn't mind, I wouldn't mind." CG 9
- [Patient acknowledging importance of monitoring caregiver experience]: "I think he [caregiver] kind of puts on a show of handling it better than he does." Pt 5

Specific to the environmental sensors, patients discussed the importance of having everyone in the home consent to sensor placement. Primary privacy concerns related to possible audio and video data collection as well as data sharing (eg, where are

the data going and who has access to them). Participants also discussed practical issues related to internet connectivity and power outages (particularly in rural areas); the number of electrical outlets in the home needed to plug in environmental sensors; how much space the sensors would occupy; and placing sensors so that they are discrete, out of the way of small children, and any cords are safely secured to prevent trip hazards or falls.

Regarding wearable sensors, participants were intrigued by the smart watch platform but wondered if (1) it would be difficult or complicated to answer questions on the watch owing to the relatively small touch screen, (2) they would have trouble using the technology, and (3) EMAs would become annoying. Some participants expressed concerns about having multiple watches and having to charge them and the potential for them being misplaced. Similar to environmental sensors, concerns were expressed regarding privacy and data sharing (eg, what exactly is being collected and where are the data going and when), and some participants whose jobs required manual labor were concerned that the watch may interfere with their work or described themselves as individuals who just did not like wearing watches or *jewelry*. Participants emphasized the importance of a simple, clear user interface, and all but 1 participant reported a willingness to answer EMAs on a wearable device. A total of 85% (17/20; 2 participants not asked) of the participants reported that they would prefer to mark and characterize pain events on a smart watch compared with a tablet, as the watch would be *attached* to them and they would be less likely to forget important details (Table 5). Some participants felt that answering EMAs on a wearable device could be annoying and they may forget to do it, but others were willing to answer as many EMAs on the smart watch as needed.

One interesting finding involved divergent perceptions between patients and caregivers regarding aspects of the BESI-C system. When patients and caregivers were interviewed together, caregivers often helped encourage an initially skeptical or reluctant patient to try the technology or reassured them about practical aspects, such as where environmental sensors could be placed in the home or that they actually had enough electrical outlets. When interviewed separately, patients commonly expressed concerns about the technology they *thought* caregivers would have, but which the caregiver did not actually express when interviewed independently. In addition, conversations related to the technology often revealed additional dyadic dynamics, beyond what was expressed regarding the impact of general symptoms. For example, when patients were asked about specific aspects of the technology, it often prompted comments acknowledging the difficulty of their illness on their caregiver and their concern for the impact it has had on them. Patients expressed that the BESI-C system would be helpful in providing an objective picture of how their caregiver was actually coping and what support their caregiver needed.

Table 5. Participant preferences and concerns regarding a wearable device to answer ecological momentary assessments.

Participant	Willing to answer EMAs[a] on a wearable device, in general	Preference for tablet/smartphone versus smart watch for EMAs	Specific comments/concerns
Pt[b] 1	May be	No preference	Tablet would have to have durable case; watch needs to be unobtrusive
CG[c] 1	Yes	No preference	—[d]
Pt 2	Yes	Watch	Watch would be "awesome"
CG 2	Yes	No preference	Concern about ability to manage technology; would need to be easy
Pt 3	No	Tablet	Concerns about watch: comfort, loss, and potential burden
CG 3	Yes	Tablet	Worried about display size/visibility of watch and ease of button use on watch; worried about loss of watch
Pt 4	Yes	Watch	Concern about watch bulkiness versus size display
Pt 5	Yes	Watch	—
CG 5	Yes	Watch	Concern about wearing at work
Pt 6	Yes	Tablet	—
CG 6	Yes	Watch	—
Pt 7	Yes	Not asked	"Watch is high-tech, I like it"
CG 7	Yes	Not asked	—
Pt 8	Yes	Watch	Thinks CG would prefer tablet
Pt 9	Yes	Watch	Concerned about privacy; concerned about ability to be "outdoorsy"
CG 9	Yes	Watch	—
Pt 10	Yes	Watch	"I don't want to answer to anybody"
CG 10	Yes	Watch	—
Pt 11	Yes	Watch	Concern about sleep interruption
CG 11	Yes	No preference	—
Pt 12	Yes	Watch	Privacy concerns
CG 12	Yes	Watch	Privacy concerns

[a]EMA: ecological momentary assessment.

[b]Pt: Patient.

[c]CG: Caregiver.

[d]No additional comments provided.

Discussion

Principal Findings

This research contributes to a more complete understanding of the experience of cancer pain in the home context and adds an important dimension of considering the caregiver's perspective. It also fills an important gap in the evidence-based design of smart health monitoring and intervention systems to provide symptom support for patients and caregivers coping with advanced cancer, a population with often significant, and unmet, symptom management needs [35,49,50]. Our overall sample, although equally split between males and females, had a disproportionate number of female caregivers and, although not racially or ethnically diverse, represents a geographically underserved sample, as the majority of patients were from rural areas of Central Virginia (consistent with the general demographics of the cancer center recruitment study site). We recruited a final sample size of 22 individuals, which is consistent with the pilot, qualitative, and early stage smart health design work [33,51-60] and the aims of our study to explore proof of concept of the BESI-C system with end users and better understand the experience of cancer pain at home.

Experience of Cancer Pain at Home: How This Can Inform System Design

Our key qualitative themes that highlight the unpredictability and perceived inevitability of pain, the negative impact of pain on foundational aspects of daily life (especially sleep, activity, and social engagement), and the challenges of managing and monitoring complex medication regimens are not surprising and validate a large body of knowledge regarding the difficulty of managing advanced cancer pain at home [4,9,10,15,23,24,61,62]. What is particularly noteworthy regarding concerns about medication management is the emphasis participants placed, especially patients, on fears and

concerns regarding access to opioid pain medication, a mainstay therapy in the management of serious cancer pain. It is especially important to consider this finding in the context of the *opioid epidemic* [63], where increased scrutiny and stigma attached to opioid therapy, even for legitimate purposes, have created unintended and increased obstacles and barriers to pain relief [64,65]. This finding underscores the significant role of pharmacological strategies in the management of cancer pain at home and the importance of designing home-based monitoring systems equipped with capabilities to support patients and caregivers in tracking, monitoring, and using mediations safely and effectively, especially prescription opioids. Consistent with other literature [9,66-70], our interviews revealed that keeping track of changing and complex medication regimens is time consuming and stressful; home-based monitoring systems that can assist with this aspect of care are essential to optimally support patients and caregivers and should be thoughtfully designed to not contribute or exacerbate the stigma and fears associated with opioid therapy needed to treat legitimate cancer pain.

Another important finding from our interviews is the need for systems that capture the impact of unpredictable cancer pain that can escalate severely and without warning. This type of pain is commonly referred to as *breakthrough* pain and is notoriously difficult to manage [19,71,72]. Although patients and caregivers did not specifically use the term *breakthrough pain* during interviews, they described significant distress associated with abrupt pain of intense severity. Finding ways to effectively capture and characterize this type of rapid-onset pain requires systems that are simple, portable, and extremely quick and easy to use, and is a primary rationale for our interest in using wearable sensors to collect these data (vs smartphones or tablets). Breakthrough pain can be an *out of control experience* for patients and caregivers, resulting in feelings of hopelessness [19,71]. We suggest that a monitoring system designed to assist patients and caregivers in tracking, recording, characterizing, and, ultimately, treating breakthrough cancer pain episodes is vital to help restore a sense of control and empowerment over their situation.

In addition, although pharmacological management of cancer pain is critical to assess, participants also expressed a desire for a more holistic approach to managing cancer pain, especially given concerns regarding the multiple side effects of medications. We accounted for this finding by adding EMAs to our smart watch app that specifically asked about nonpharmacological approaches patients and caregivers use to manage cancer pain.

Variables That Influence Cancer Pain: How This Can Inform System Design

A better understanding of the impact of variables that influence cancer pain in the home setting can facilitate the design of tailored systems equipped to measure and assess the most salient variables. Although we did not detect statistically significant differences between patient and caregiver mean impact scores (most likely due to our small sample size and the possible influence of participants being interviewed together), our results make an important contribution and extend existing work [30,73] as they (1) focus on perceived influences on cancer pain from *both* the patient and caregiver perspective; (2) consider a holistic set of environmental *and* contextual variables; (3) suggest important data collection features to include in remote symptom monitoring systems; and (4) provide initial insights into the impact of critical variables that may influence cancer pain, which can be built upon for future inquiry. The lack of a significant difference may suggest that patients and caregivers are largely *in sync* about what impacts pain, which could be helpful and productive. For example, both patients and caregivers ranked variables related to taking pain medication, sleep, activity, and mood among the top 5, and they both ranked *taking pain medication* as the most impactful variable influencing cancer pain. These findings corroborate our qualitative findings and again underscore the importance of designing monitoring systems that account for ways to support patients and caregivers in tracking and managing pain medications.

It is not surprising that sleep quality and quantity were rated, among all the wellness variables, as having the highest impact on pain as the relationship between sleep and pain has been well established [74]. Likewise, our results also confirm the known connection between mood and pain [75]. Patients and caregivers also rated physical activity as having a significant impact on patients' pain, justifying the importance of including activity-monitoring features, such as pedometers and accelerometers. It is noteworthy that *sleep quality* is rated as the most impactful variable by caregivers, second only after taking pain medication. This underscores the need to design monitoring systems and interventions that do not further worsen or interrupt sleep, such as with low-battery reminders or bothersome sensor lights. Another particularly interesting finding within the category of wellness variables is the fact that caregivers rated the impact of oral intake (#8) higher than patients, who rated it as the second lowest rated variable (#12). For health care providers, this likely resonates as family caregivers often have strong opinions and beliefs/concerns about the nutritional intake of their loved ones, specifically how the lack of adequate oral intake may worsen distress and pain [76].

Interpreting the impact of interaction and environmental variables is more complex as there is less congruence between patient and caregiver scores. Regarding interaction variables, it is interesting that patients rated *emotional closeness to others/connection to others* as having the least impact on their pain, after noise, light, humidity, and oral intake. One interpretation of this finding is that patients do not value emotional connection or closeness or see little association in its role to their experience of pain, which seems unlikely. An alternative explanation could be that because we interviewed dyads (and most commonly spouses), emotional closeness/connection was assumed by the participants and therefore not considered to be a significant variable. If we had interviewed single patients with cancer, who may experience more notable fluctuations in available emotional support, this variable may have been rated differently. Another hypothesis, consistent with the principles of Maslow's hierarchy of needs, is that patients experiencing significant pain will naturally focus first on the physical aspects of their well-being (such as sleep),

with less priority given to higher level needs, such as emotional connection. However, patients and caregivers did rank *social/interpersonal interactions* and *busyness of home* higher on the list, suggesting that engagement within the home is important and can influence pain. These findings suggest that capturing the degree of social engagement within the home is important and should be incorporated into home monitoring systems with features that can track location and person-to-person interactions.

As a category, environmental variables were ranked as having the least influence on pain. However, as individual variables, patients ranked *temperature* as having the fourth highest impact, but all other environmental variables (noise, light, and humidity) were ranked in the bottom 5. For caregivers, all environmental variables were ranked in the lower half. One possibility is that environmental variables have little influence on a patient's pain. Instead, we argue that environmental variables *are* likely to be important factors (just ask anyone whose pain increases during rainy, humid weather), but that patients and caregivers may simply not be fully aware of the role these variables play as they are rarely (if ever) systematically monitored, tracked, or reported, making their impact less obvious and more difficult to quantify. As monitoring environmental variables can be performed passively, requiring minimal participant burden, we suggest it is important to collect these data so that we can more clearly understand the potential relationships between environmental factors and pain episodes.

Finally, it is reassuring that participants did not readily identify additional variables to measure with the BESI-C system. In other words, patients and caregivers validated our list of proposed variables and felt that we included a comprehensive list.

Feedback About the BESI-C Prototype Components: How This Can Inform System Design

Showing prototypes of the BESI-C system and discussing them with participants proved to be very effective and helpful in informing system design and refining the BESI-C system. For example, based on feedback from participants, we elected to use a smart watch to collect EMA data (vs a smartphone or table) and iterated our environmental sensors to make them smaller, sleeker, and more discrete. Another benefit of discussing the technology with participants is that it confirmed important dyad dynamics that reinforced our initial hypotheses about system design. For example, the ultimate goal of BESI-C is to improve communication between patients and caregivers, particularly around pain management. This was reinforced by comments from dyad 6, where the patient reported that he did not always ask for pain medication, even when he needed it, and the caregiver expressed that she was not always sure when the patient needed pain medication. Data from BESI-C could improve these types of interactions by providing helpful data in real time to patients and caregivers.

We paid particular attention to data privacy based on participant feedback. The system includes no cameras, and the microphone outputs are locally processed to extract relevant audio features (eg, loudness level and noise fluctuation) so that no interpretable audio data are stored or transmitted. Participants also had questions about data sharing and expressed interest in seeing their own data and having it be available to health care providers. Our initial pilot work does not fully explore this important question (although patients and caregivers will see selected extracted features of their data), but future work will examine how to optimally generate data visualizations and how to best share these data with relevant stakeholders.

Patients and caregivers expressed less concerns regarding passive monitoring and seemed to prefer elements of the system that could be left alone and just *do their thing*. This is important when considering the system design for this patient population and suggests that passive data collection, with environmental sensors or physiological monitoring with wearables, may be more acceptable and feasible and that active data collection with EMAs should be extremely judicious to reduce user burden. Simplicity and ease of use were critically (and not surprisingly) important to participants, and we designed our smart watch user interface to be extremely easy to use, intuitive, and to work well on a watch touchscreen [36]. Through all portions of the interviews, both patients and caregivers reinforced the importance of medication tracking and monitoring. On the basis of this feedback, the BESI-C EMAs include simple questions about medication use as well as reasons pain medication may not be taken even if patients are in pain (eg, concerns about running out of tablets).

Limitations

The primary limitation of this study is the sample size, which reduced generalizability and precluded our ability to detect statistical significance in our analysis of variable impact scores. However, our sample size is consistent with the scope of a pilot study related to early stage smart health design [33,53,54] and the aims of qualitative research [51,52,56] and provides an important rural perspective. Our initial intent was to interview all dyads together, but this proved difficult/impossible due to logistical constraints. To avoid increasing the participant burden, a critical consideration for this patient population, some dyads were interviewed together and some separately. Although participants were instructed that we were interested in their own individual opinion and perspective, it is likely that for dyads interviewed together, hearing their partner's responses may have influenced their answers. Finally, we do not know the direction of the impact of variables as we did not ask participants whether certain factors made pain better or worse, only whether they felt the variable had an important impact, either negative or positive. Therefore, for example, we cannot say that a high score of *physical activity* means physical activity improves pain or makes it worse—only that the variable of physical activity is perceived to have a significant influence on the patient's pain.

Future Directions

This study provides important foundational data that can inform future research, particularly related to understanding variable influence on the pain experience and how this can inform remote monitoring system design. Conducting a similar study in part 2 with a larger sample size of dyads would be helpful to detect statistically significant differences between how caregivers and patients rank variables that may influence pain. Relatedly, it would be interesting to explore whether, and how, concordance

between variable ratings among patients and caregivers changes relevant clinical outcomes. For example, are dyads with higher variable congruence (eg, more agreement regarding the impact of certain variables that may influence pain) more likely to experience lower levels of pain and overall distress? This could be evaluated by asking patients and caregivers to independently score a list of variables, deploy a monitoring system such as the BESI-C, and then compare the reported pain and distress levels with predeployment variable congruence levels. Such results would help further inform the design of smart health monitoring systems and personalized interventions. In addition, more recent concerns regarding home-based care in the context of COVID-19 are prompting adjustments in the BESI-C system design to facilitate contactless deployments.

Conclusions

Strategies to monitor and treat cancer pain outside the acute care setting are critical as most cancer symptom management occurs at home, often causing significant stress for both patients and family caregivers. Home-based smart health monitoring systems designed to support cancer pain management should account for the experience of both the patient and the caregiver; prioritize passive monitoring of physiological and environmental variables to reduce burden; and include functionality that can monitor and track medication intake and efficacy, wellness variables (such as sleep quality/quantity, physical activity, mood, and oral intake), and levels of social interaction and engagement. In addition, systems must consider concerns regarding privacy and data sharing and incorporate feasible strategies to capture and characterize rapid-onset symptoms.

Acknowledgments

The authors acknowledge Kara Fitzgibbon, Center for Survey Research, University of Virginia. Funding for this research was provided by The University of Virginia Engineering in Medicine Seed Pilot Program (February 1, 2018 to June 30, 2019).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Study guide.
[DOCX File , 17 KB - formative_v4i8e20836_app1.docx]

References

1. van den Beuken-van Everdingen MH, de Rijke JM, Kessels AG, Schouten HC, van Kleef M, Patijn J. Prevalence of pain in patients with cancer: a systematic review of the past 40 years. Ann Oncol 2007 Sep;18(9):1437-1449 [FREE Full text] [doi: 10.1093/annonc/mdm056] [Medline: 17355955]
2. Goodwin PJ, Bruera E, Stockler M. Pain in patients with cancer. J Clin Oncol 2014 Jun 1;32(16):1637-1639. [doi: 10.1200/JCO.2014.55.3818] [Medline: 24799489]
3. Deandrea S, Montanari M, Moja L, Apolone G. Prevalence of undertreatment in cancer pain. A review of published literature. Ann Oncol 2008 Dec;19(12):1985-1991 [FREE Full text] [doi: 10.1093/annonc/mdn419] [Medline: 18632721]
4. Smyth JA, Dempster M, Warwick I, Wilkinson P, McCorry NK. A systematic review of the patient- and carer-related factors affecting the experience of pain for advanced cancer patients cared for at home. J Pain Symptom Manage 2018 Feb;55(2):496-507. [doi: 10.1016/j.jpainsymman.2017.08.012] [Medline: 28843458]
5. Teunissen SC, Wesker W, Kruitwagen C, de Haes HC, Voest EE, de Graeff A. Symptom prevalence in patients with incurable cancer: a systematic review. J Pain Symptom Manage 2007 Jul;34(1):94-104 [FREE Full text] [doi: 10.1016/j.jpainsymman.2006.10.015] [Medline: 17509812]
6. Kutner JS, Kassner CT, Nowels DE. Symptom burden at the end of life: hospice providers' perceptions. J Pain Symptom Manage 2001 Jun;21(6):473-480 [FREE Full text] [doi: 10.1016/s0885-3924(01)00281-0] [Medline: 11397605]
7. Phongtankuel V, Teresi JA, Eimicke JP, Kong JX, Adelman RD, Prigerson HG, et al. Identifying the prevalence and correlates of caregiver-reported symptoms in home hospice patients at the end of life. J Palliat Med 2020 May;23(5):635-640. [doi: 10.1089/jpm.2019.0324] [Medline: 31873053]
8. de la Cruz M, Noguera A, San Miguel-Arregui MT, Williams J, Chisholm G, Bruera E. Delirium, agitation, and symptom distress within the final seven days of life among cancer patients receiving hospice care. Palliat Support Care 2015 Apr;13(2):211-216. [doi: 10.1017/S1478951513001144] [Medline: 24556057]
9. Mehta A, Cohen SR, Ezer H, Carnevale FA, Ducharme F. Striving to respond to palliative care patients' pain at home: a puzzle for family caregivers. Oncol Nurs Forum 2011 Jan;38(1):E37-E45. [doi: 10.1188/11.ONF.E37-E45] [Medline: 21186150]
10. Mehta A, Chan LS, Cohen SR. Flying blind: sources of distress for family caregivers of palliative cancer patients managing pain at home. J Psychosoc Oncol 2014;32(1):94-111. [doi: 10.1080/07347332.2013.856057] [Medline: 24428253]
11. Martín JM, Olano-Lizarraga M, Saracíbar-Razquin M. The experience of family caregivers caring for a terminal patient at home: a research review. Int J Nurs Stud 2016 Dec;64:1-12. [doi: 10.1016/j.ijnurstu.2016.09.010] [Medline: 27657662]

12. Chi N, Demiris G. Family caregivers' pain management in end-of-life care: a systematic review. Am J Hosp Palliat Care 2017 Jun;34(5):470-485. [doi: 10.1177/1049909116637359] [Medline: 26975303]
13. Kelley M, Demiris G, Nguyen H, Oliver DP, Wittenberg-Lyles E. Informal hospice caregiver pain management concerns: a qualitative study. Palliat Med 2013 Jul;27(7):673-682 [FREE Full text] [doi: 10.1177/0269216313483660] [Medline: 23612959]
14. Oliver DP, Wittenberg-Lyles E, Washington K, Kruse RL, Albright DL, Baldwin PK, et al. Hospice caregivers' experiences with pain management: 'I'm not a doctor, and I don't know if I helped her go faster or slower'. J Pain Symptom Manage 2013 Dec;46(6):846-858 [FREE Full text] [doi: 10.1016/j.jpainsymman.2013.02.011] [Medline: 23731855]
15. Ferrell BR. Family caregiving and cancer pain management. Anesth Analg 2019 Nov;129(5):1408-1413. [doi: 10.1213/ANE.0000000000003937] [Medline: 30531216]
16. Green E, Ward S, Brierley W, Riley B, Sattar H, Harris T. 'They shouldn't be coming to the ed, should they?': a descriptive service evaluation of why patients with palliative care needs present to the emergency department. Am J Hosp Palliat Care 2017 Dec;34(10):984-990. [doi: 10.1177/1049909116676774] [Medline: 27903774]
17. Wallace EM, Cooney MC, Walsh J, Conroy M, Twomey F. Why do palliative care patients present to the emergency department? Avoidable or unavoidable? Am J Hosp Palliat Care 2013 May;30(3):253-256. [doi: 10.1177/1049909112447285] [Medline: 22628898]
18. Oh TK, Jo YH, Choi JW. Associated factors and costs of avoidable visits to the emergency department among cancer patients: 1-year experience in a tertiary care hospital in South Korea. Support Care Cancer 2018 Nov;26(11):3671-3679. [doi: 10.1007/s00520-018-4195-0] [Medline: 29740693]
19. Deandrea S, Corli O, Consonni D, Villani W, Greco MT, Apolone G. Prevalence of breakthrough cancer pain: a systematic review and a pooled analysis of published literature. J Pain Symptom Manage 2014 Jan;47(1):57-76 [FREE Full text] [doi: 10.1016/j.jpainsymman.2013.02.015] [Medline: 23796584]
20. Delgado-Guay MO, Kim YJ, Shin SH, Chisholm G, Williams J, Allo J, et al. Avoidable and unavoidable visits to the emergency department among patients with advanced cancer receiving outpatient palliative care. J Pain Symptom Manage 2015 Mar;49(3):497-504 [FREE Full text] [doi: 10.1016/j.jpainsymman.2014.07.007] [Medline: 25131891]
21. Phongtankuel V, Paustian S, Reid MC, Finley A, Martin A, Delfs J, et al. Events leading to hospital-related disenrollment of home hospice patients: a study of primary caregivers' perspectives. J Palliat Med 2017 Mar;20(3):260-265 [FREE Full text] [doi: 10.1089/jpm.2015.0550] [Medline: 27893951]
22. Phongtankuel V, Scherban BA, Reid MC, Finley A, Martin A, Dennis J, et al. Why do home hospice patients return to the hospital? A study of hospice provider perspectives. J Palliat Med 2016 Jan;19(1):51-56 [FREE Full text] [doi: 10.1089/jpm.2015.0178] [Medline: 26702519]
23. Kehl KA, Kowalkowski JA. A systematic review of the prevalence of signs of impending death and symptoms in the last 2 weeks of life. Am J Hosp Palliat Care 2013 Sep;30(6):601-616. [doi: 10.1177/1049909112468222] [Medline: 23236090]
24. Barbera L, Seow H, Howell D, Sutradhar R, Earle C, Liu Y, et al. Symptom burden and performance status in a population-based cohort of ambulatory cancer patients. Cancer 2010 Dec 15;116(24):5767-5776 [FREE Full text] [doi: 10.1002/cncr.25681] [Medline: 21136579]
25. Chi N, Demiris G, Pike KC, Washington K, Oliver DP. Pain management concerns from the hospice family caregivers' perspective. Am J Hosp Palliat Care 2018 Apr;35(4):601-611 [FREE Full text] [doi: 10.1177/1049909117729477] [Medline: 28875732]
26. Zadeh RS, Eshelman P, Setla J, Kennedy L, Hon E, Basara A. Environmental design for end-of-life care: an integrative review on improving the quality of life and managing symptoms for patients in institutional settings. J Pain Symptom Manage 2018 Mar;55(3):1018-1034 [FREE Full text] [doi: 10.1016/j.jpainsymman.2017.09.011] [Medline: 28935129]
27. Littleton-Kearney MT, Grady PA. The science of caregiving bringing voices together: summary of national institute of nursing research's 2017 summit. Nurs Outlook 2018;66(2):157-159. [doi: 10.1016/j.outlook.2018.02.002] [Medline: 29576245]
28. von Ah D, Brown CG, Brown SJ, Bryant AL, Davies M, Dodd M, et al. Research agenda of the oncology nursing society: 2019-2022. Oncol Nurs Forum 2019 Nov 1;46(6):654-669. [doi: 10.1188/19.ONF.654-669] [Medline: 31626621]
29. McGuire DB, Grant M, Park J. Palliative care and end of life: the caregiver. Nurs Outlook 2012;60(6):351-6.e20. [doi: 10.1016/j.outlook.2012.08.003] [Medline: 23141194]
30. Rodríguez I, Herskovic V, Gerea C, Fuentes C, Rossel PO, Marques M, et al. Understanding monitoring technologies for adults with pain: systematic literature review. J Med Internet Res 2017 Oct 27;19(10):e364 [FREE Full text] [doi: 10.2196/jmir.7279] [Medline: 29079550]
31. Silva EH, Lawler S, Langbecker D. The effectiveness of mHealth for self-management in improving pain, psychological distress, fatigue, and sleep in cancer survivors: a systematic review. J Cancer Surviv 2019 Feb;13(1):97-107. [doi: 10.1007/s11764-018-0730-8] [Medline: 30635865]
32. Allsop MJ, Taylor S, Mulvey MR, Bennett MI, Bewick BM. Information and communication technology for managing pain in palliative care: a review of the literature. BMJ Support Palliat Care 2015 Dec;5(5):481-489. [doi: 10.1136/bmjspcare-2013-000625] [Medline: 24644214]

33. Signorelli GR, Lehocki F, Fernández MM, O'Neill G, O'Connor D, Brennan L, et al. A research roadmap: connected health as an enabler of cancer patient support. J Med Internet Res 2019 Oct 29;21(10):e14360 [FREE Full text] [doi: 10.2196/14360] [Medline: 31663861]
34. Dickman Portz J, Ford K, Bekelman DB, Boxer RS, Kutner JS, Czaja S, et al. 'We're taking something so human and trying to digitize': provider recommendations for mhealth in palliative care. J Palliat Med 2020 Feb;23(2):240-247. [doi: 10.1089/jpm.2019.0216] [Medline: 31526220]
35. Theile G, Klaas V, Tröster G, Guckenberger M. Mhealth technologies for palliative care patients at the interface of in-patient to outpatient care: protocol of feasibility study aiming to early predict deterioration of patient's health status. JMIR Res Protoc 2017 Aug 16;6(8):e142 [FREE Full text] [doi: 10.2196/resprot.7676] [Medline: 28814378]
36. LeBaron V, Hayes J, Gordon K, Alam R, Homdee N, Martinez Y, et al. Leveraging smart health technology to empower patients and family caregivers in managing cancer pain: protocol for a feasibility study. JMIR Res Protoc 2019 Dec 9;8(12):e16178 [FREE Full text] [doi: 10.2196/16178] [Medline: 31815679]
37. Etikan I. Comparison of convenience sampling and purposive sampling. Am J Thery Appl Stat 2016 Jun;5(1):1-11. [doi: 10.11648/j.ajtas.20160501.11]
38. Overview. NIH Common Fund. URL: https://commonfund.nih.gov/promis/overview [accessed 2020-08-03]
39. Blackhall LJ, Read P, Stukenborg G, Dillon P, Barclay J, Romano A, et al. CARE track for advanced cancer: impact and timing of an outpatient palliative care clinic. J Palliat Med 2016 Jan;19(1):57-63 [FREE Full text] [doi: 10.1089/jpm.2015.0272] [Medline: 26624851]
40. Cheng KK, Yeung RM. Impact of mood disturbance, sleep disturbance, fatigue and pain among patients receiving cancer therapy. Eur J Cancer Care (Engl) 2013 Jan;22(1):70-78. [doi: 10.1111/j.1365-2354.2012.01372.x] [Medline: 22805171]
41. Dodd MJ, Miaskowski C, Lee KA. Occurrence of symptom clusters. J Natl Cancer Inst Monogr 2004(32):76-78. [doi: 10.1093/jncimonographs/lgh008] [Medline: 15263044]
42. Alam R, Dugan J, Homdee N, Gandhi N, Ghaemmaghami B, Meda H, et al. BESI: Reliable and Heterogeneous Sensing and Intervention for In-Home Health Applications. In: International Conference on Connected Health: Applications, Systems and Engineering Technologies. 2017 Presented at: CHASE'17; July 17-19, 2017; Philadelphia, PA, USA URL: https://ieeexplore.ieee.org/document/8010628
43. Bankole A, Anderson MS, Homdee N, Alam R, Lofton A, Fyffe N, et al. BESI: behavioral and environmental sensing and intervention for dementia caregiver empowerment-phases 1 and 2. Am J Alzheimers Dis Other Demen 2020;35:1533317520906686. [doi: 10.1177/1533317520906686] [Medline: 32162529]
44. Homdee N, Smith-Jackson T, Bankole A, Anderson MS, Lach J, Alam R, et al. Agitation monitoring and prevention system for dementia caregiver empowerment. Computer 2019 Nov;52(11):30-39. [doi: 10.1109/mc.2019.2933192]
45. Shiffman S, Stone AA, Hufford MR. Ecological momentary assessment. Annu Rev Clin Psychol 2008;4:1-32. [doi: 10.1146/annurev.clinpsy.3.022806.091415] [Medline: 18509902]
46. Sandelowski M. Whatever happened to qualitative description? Res Nurs Health 2000 Aug;23(4):334-340. [doi: 10.1002/1098-240x(200008)23:4<334::aid-nur9>3.0.co;2-g] [Medline: 10940958]
47. Am I Rural? – Tool. Rural Health Information Hub. URL: https://www.ruralhealthinfo.org/am-i-rural [accessed 2020-08-17]
48. Oken MM, Creech RH, Tormey DC, Horton J, Davis TE, McFadden ET, et al. Toxicity and response criteria of the eastern cooperative oncology group. Am J Clin Oncol 1982 Dec;5(6):649-655. [Medline: 7165009]
49. Schoppee TM, Dyal BW, Scarton L, Ezenwa MO, Singh P, Yao Y, et al. Patients and caregivers rate the PainReportIt wireless internet-enabled tablet as a method for reporting pain during end-of-life cancer care. Cancer Nurs 2019 Sep 27:-epub ahead of print. [doi: 10.1097/NCC.0000000000000743] [Medline: 31517649]
50. Loh KP, Ramsdale E, Culakova E, Mendler JH, Liesveld JL, O'Dwyer KM, et al. Novel mhealth app to deliver geriatric assessment-driven interventions for older adults with cancer: pilot feasibility and usability study. JMIR Cancer 2018 Oct 29;4(2):e10296 [FREE Full text] [doi: 10.2196/10296] [Medline: 30373733]
51. Creswell J, Poth C. Qualitative Inquiry and Research Design: Choosing Among Five Approaches. Thousand Oaks, CA: Sage Publications; 2016.
52. Vasileiou K, Barnett J, Thorpe S, Young T. Characterising and justifying sample size sufficiency in interview-based studies: systematic analysis of qualitative health research over a 15-year period. BMC Med Res Methodol 2018 Nov 21;18(1):148 [FREE Full text] [doi: 10.1186/s12874-018-0594-7] [Medline: 30463515]
53. Børøsund E, Mirkovic J, Clark MM, Ehlers SL, Andrykowski MA, Bergland A, et al. A stress management app intervention for cancer survivors: design, development, and usability testing. JMIR Form Res 2018 Sep 6;2(2):e19 [FREE Full text] [doi: 10.2196/formative.9954] [Medline: 30684438]
54. Wolpin S, Halpenny B, Whitman G, McReynolds J, Stewart M, Lober W, et al. Development and usability testing of a web-based cancer symptom and quality-of-life support intervention. Health Informatics J 2015 Mar;21(1):10-23 [FREE Full text] [doi: 10.1177/1460458213495744] [Medline: 24406906]
55. Anderson K, Burford O, Emmerton L. Mobile health apps to facilitate self-care: a qualitative study of user experiences. PLoS One 2016;11(5):e0156164 [FREE Full text] [doi: 10.1371/journal.pone.0156164] [Medline: 27214203]
56. Morse JM. Determining sample size. Qual Health Res 2016 Jul;10(1):3-5. [doi: 10.1177/104973200129118183]

57. Langer SL, Ghosh N, Todd M, Randall AK, Romano JM, Bricker JB, et al. Usability and acceptability of a smartphone app to assess partner communication, closeness, mood, and relationship satisfaction: mixed methods study. JMIR Form Res 2020 Jul 6;4(7):e14161. [doi: 10.2196/14161] [Medline: 32628614]
58. Brick T, Mundie J, Weaver J, Fraleigh R, Oravecz Z. Low-burden mobile monitoring, intervention, and real-time analysis using the wear-it framework: example and usability study. JMIR Form Res 2020 Jun 17;4(6):e16072 [FREE Full text] [doi: 10.2196/16072] [Medline: 32554373]
59. Siemer L, Allouch SB, Pieterse M, Brusse-Keizer M, Sanderman R, Postel M. Patients' user experience of a blended face-to-face and web-based smoking cessation treatment: qualitative study. JMIR Form Res 2020 Jun 3;4(6):e14550 [FREE Full text] [doi: 10.2196/14550] [Medline: 32343245]
60. Sandelowski M. Sample size in qualitative research. Res Nurs Health 1995 Apr;18(2):179-183. [doi: 10.1002/nur.4770180211] [Medline: 7899572]
61. Adam R, de Bruin M, Burton CD, Bond CM, Giatsi Clausen M, Murchie P. What are the current challenges of managing cancer pain and could digital technologies help? BMJ Support Palliat Care 2018 Jun;8(2):204-212 [FREE Full text] [doi: 10.1136/bmjspcare-2016-001232] [Medline: 28554888]
62. Patel RA, Klasnja P, Hartzler A, Unruh KT, Pratt W. Probing the benefits of real-time tracking during cancer care. AMIA Annu Symp Proc 2012;2012:1340-1349 [FREE Full text] [Medline: 23304413]
63. Pain Management and the Opioid Epidemic. The National Academies Press. 2017. URL: http://nationalacademies.org/hmd/Reports/2017/pain-management-and-the-opioid-epidemic.aspx [accessed 2020-08-03]
64. Paice JA. Cancer pain management and the opioid crisis in America: how to preserve hard-earned gains in improving the quality of cancer pain management. Cancer 2018 Jun 15;124(12):2491-2497 [FREE Full text] [doi: 10.1002/cncr.31303] [Medline: 29499072]
65. Paice JA, Von Roenn JH. Under- or overtreatment of pain in the patient with cancer: how to achieve proper balance. J Clin Oncol 2014 Jun 1;32(16):1721-1726. [doi: 10.1200/JCO.2013.52.5196] [Medline: 24799468]
66. Joyce BT, Berman R, Lau DT. Formal and informal support of family caregivers managing medications for patients who receive end-of-life care at home: a cross-sectional survey of caregivers. Palliat Med 2014 Oct;28(9):1146-1155. [doi: 10.1177/0269216314535963] [Medline: 24854033]
67. Schumacher KL, Plano Clark VL, West CM, Dodd MJ, Rabow MW, Miaskowski C. Pain medication management processes used by oncology outpatients and family caregivers part II: home and lifestyle contexts. J Pain Symptom Manage 2014 Nov;48(5):784-796 [FREE Full text] [doi: 10.1016/j.jpainsymman.2013.12.247] [Medline: 24709364]
68. Tabi K, Randhawa AS, Choi F, Mithani Z, Albers F, Schnieder M, et al. Mobile apps for medication management: review and analysis. JMIR Mhealth Uhealth 2019 Sep 11;7(9):e13608 [FREE Full text] [doi: 10.2196/13608] [Medline: 31512580]
69. Han CJ, Chi N, Han S, Demiris G, Parker-Oliver D, Washington K, et al. Communicating caregivers' challenges with cancer pain management: an analysis of home hospice visits. J Pain Symptom Manage 2018 May;55(5):1296-1303 [FREE Full text] [doi: 10.1016/j.jpainsymman.2018.01.004] [Medline: 29360571]
70. Schumacher KL, Plano Clark VL, West CM, Dodd MJ, Rabow MW, Miaskowski C. Pain medication management processes used by oncology outpatients and family caregivers part I: health systems contexts. J Pain Symptom Manage 2014 Nov;48(5):770-783 [FREE Full text] [doi: 10.1016/j.jpainsymman.2013.12.242] [Medline: 24704800]
71. Gonella S, Sperlinga R, Sciannameo V, Dimonte V, Campagna S. Characteristics of breakthrough pain and its impact on quality of life in terminally ill cancer patients. Integr Cancer Ther 2019;18:1534735419859095 [FREE Full text] [doi: 10.1177/1534735419859095] [Medline: 31220961]
72. Tagami K, Okizaki A, Miura T, Watanabe YS, Matsumoto Y, Morita T, et al. Breakthrough cancer pain influences general activities and pain management: a comparison of patients with and without breakthrough cancer pain. J Palliat Med 2018 Nov;21(11):1636-1640. [doi: 10.1089/jpm.2017.0675] [Medline: 29975582]
73. Mercadante S, Marchetti P, Cuomo A, Caraceni A, Mediati RD, Vellucci R, IOPS-MS Study Group. Factors influencing the clinical presentation of breakthrough pain in cancer patients. Cancers (Basel) 2018 Jun 1;10(6):- [FREE Full text] [doi: 10.3390/cancers10060175] [Medline: 29865170]
74. Finan PH, Goodin BR, Smith MT. The association of sleep and pain: an update and a path forward. J Pain 2013 Dec;14(12):1539-1552 [FREE Full text] [doi: 10.1016/j.jpain.2013.08.007] [Medline: 24290442]
75. Glover J, Dibble SL, Dodd MJ, Miaskowski C. Mood states of oncology outpatients: does pain make a difference? J Pain Symptom Manage 1995 Feb;10(2):120-128 [FREE Full text] [doi: 10.1016/0885-3924(94)00073-t] [Medline: 7730684]
76. Amano K, Maeda I, Morita T, Okajima Y, Hama T, Aoyama M, et al. Eating-related distress and need for nutritional support of families of advanced cancer patients: a nationwide survey of bereaved family members. J Cachexia Sarcopenia Muscle 2016 Dec;7(5):527-534 [FREE Full text] [doi: 10.1002/jcsm.12102] [Medline: 27239421]

Abbreviations

BESI-C: Behavioral and Environmental Sensing and Intervention for Cancer
EMA: ecological momentary assessment
NIH PROMIS: National Institutes of Health Patient-Reported Outcomes Measurement Information System

Edited by G Eysenbach; submitted 04.06.20; peer-reviewed by K Schumacher, D Mayer; comments to author 25.06.20; revised version received 11.07.20; accepted 25.07.20; published 26.08.20.

Please cite as:
LeBaron V, Bennett R, Alam R, Blackhall L, Gordon K, Hayes J, Homdee N, Jones R, Martinez Y, Ogunjirin E, Thomas T, Lach J Understanding the Experience of Cancer Pain From the Perspective of Patients and Family Caregivers to Inform Design of an In-Home Smart Health System: Multimethod Approach
JMIR Form Res 2020;4(8):e20836
URL: http://formative.jmir.org/2020/8/e20836/
doi:10.2196/20836
PMID:32712581

Original Paper

Development of a Mobile Health Intervention with Personal Experiments for Smokers Who Are Ambivalent About Quitting: Formative Design and Testing

Jaimee L Heffner[1], PhD; Sheryl L Catz[2], PhD; Predrag Klasnja[3,4], PhD; Brooks Tiffany[3], MS; Jennifer B McClure[3], PhD

[1]Fred Hutchinson Cancer Research Center, Seattle, WA, United States

[2]Betty Irene Moore School of Nursing, University of California, Sacramento, CA, United States

[3]Kaiser Permanente Washington Health Research Institute, Seattle, WA, United States

[4]University of Michigan, Ann Arbor, MI, United States

Corresponding Author:
Jennifer B McClure, PhD
Kaiser Permanente Washington Health Research Institute
1730 Minor Ave
Suite 1600
Seattle, WA, 98101
United States
Phone: 1 206 287 2737
Fax: 1 206 287 2800
Email: Jennifer.B.McClure@kp.org

Abstract

Background: The majority of cigarette smokers want to quit someday but are not ready to commit to long-term abstinence. However, available smoking cessation treatments are not well-suited to meet the needs of these ambivalent smokers. Low-cost, high-reach mobile health (mHealth) interventions may be a cost-efficient means of offering assistance to ambivalent smokers, yet there are currently no evidence-based options available for this group.

Objective: The aim of this study was to develop and preliminarily evaluate the core content for an mHealth program targeting adult smokers who are ambivalent about quitting. The core content consisted of a series of “personal experiments” similar to those tested as part of a counseling intervention in prior work, including brief cognitive or behavioral tasks designed to boost readiness for changing smoking behavior.

Methods: We conducted individual user interviews (N=3) to refine program content, and then conducted a one-arm pilot study (N=25) to assess user receptivity and the potential impact of the experiments on motivation and self-efficacy to quit or reduce smoking.

Results: In user interviews, participants liked the concept of the personal experiments. Participants in the pilot study found a medium-fidelity prototype to be highly acceptable. After watching a brief orientation video that explained how the program works, most participants (80%, 20/25) indicated that it sounded interesting, primarily because it did not require any commitment to quit. All participants (100%, 25/25) completed all 7 experiments, including a 24-hour quit attempt, although not all were able to refrain from smoking for a full day based on qualitative feedback on the experiment. The mean rating of usefulness of the overall program was 4.12 (SD 1.09) out of 5, and the average rating of the difficulty of the experiments was 2.16 (SD 1.18) out of 5. At the last assessment point, 92% (23/25) of the participants indicated that they were more interested in either quitting or cutting back than when they began the program, and 72% (18/25) said that if the program had included a free trial of nicotine replacement therapy, they would have used it to try to quit smoking.

Conclusions: This formative work confirmed that ambivalent smokers are willing to use and will remain engaged with an mHealth intervention that employs the novel concept of personal experiments to enhance their motivation for and ability to quit smoking. The addition of action-oriented treatment (self-help and free nicotine replacement therapy, quitline referral) could further support users’ efforts to stop smoking and remain quit.

(JMIR Form Res 2020;4(8):e21784) doi:10.2196/21784

KEYWORDS

tobacco; nicotine; smoking; cessation; smartphone; motivation; mHealth; intervention; formative; development

Introduction

Background

According to the World Health Organization, tobacco kills approximately half of its users worldwide and is responsible for over 8 million deaths per year [1]. Helping people quit smoking is critical to reducing the human toll of tobacco use, but is also a difficult goal as most smokers are ambivalent about quitting. Approximately 70% of smokers in the United States report that they want to quit smoking someday, but they are not *currently* ready to quit nor actively seeking treatment [2]. This ambivalence makes it difficult to engage these individuals in nicotine dependence treatment programs. Consequently, reducing tobacco use on a population level will require new intervention strategies that can engage, motivate, and effectively assist ambivalent individuals in quitting smoking. Digital therapeutic interventions such as mobile health (mHealth) apps offer a potentially high-reach, effective strategy for achieving this important public health goal.

To date, no published trials have tested self-guided mHealth interventions designed for ambivalent smokers, despite strong rationale for doing so. The vast majority (81%) of US adults own a smartphone, including those with only a high school education (72%), those who make less than $30,000 a year (71%), and racial/ethnic minorities (Hispanic, 79%; Black, 80%) [3], which are also demographic groups with high smoking rates [4]. Further, rates of smartphone ownership are similar among smokers and nonsmokers [5]. Evidence from our preliminary work also suggests that smokers who are ambivalent about quitting—specifically, those who want to quit smoking someday, but not in the next 6 months—are quite receptive to mHealth interventions focused on smoking: 75% stated they would consider using a cessation app, 88% were interested in an app to help them reduce their smoking, and 91% were interested in an app that could help them decide "if , when, or how" to quit [6]. This suggests that, with appropriate intervention message-framing that takes their ambivalence into account, a self-guided mHealth intervention could reach and assist a sizable proportion of smokers who would not otherwise seek treatment. A particular advantage of mHealth in this context is that it does not require involvement of a treatment provider in contrast to other digital interventions that include ambivalent smokers as part of the target user group (eg, a provider-facilitated social media intervention targeting young adult smokers across all stages of change [7]). Thus, our aim was to create a novel, empirically validated, self-guided mHealth intervention that will be appealing to ambivalent smokers, able to keep them engaged over time, and ultimately assist them in quitting smoking.

Intervening With Ambivalent Smokers

Research suggests that ambivalent smokers may benefit from standard, evidence-based treatment approaches (ie, behavioral and pharmacological interventions), but that they may also need to be introduced to these methods in a softer, more gradual manner than would be used with smokers who are already committed to quitting [8]. For example, we found that ambivalent smokers are willing to enroll in clinical trials when it is clear that the goal is to help them explore their willingness to quit or to answer questions about the quitting process, as opposed to asking for a commitment to stop smoking, and many go on to successfully quit [9,10]. Other studies have found that ambivalent smokers are receptive to a goal of smoking reduction, which in turn can increase quit rates, particularly when the intervention is paired with a stop-smoking medication [11-15]. This concept—that ambivalent smokers can benefit from the same evidence-based strategies used with smokers ready to quit, if framed appropriately—is consistent with West's [16] PRIME theory, in which motivation is viewed as a dynamic, rapidly changing state rather than one that emerges slowly and in a staged manner. The implication is that interventions targeting ambivalent smokers should be responsive to rapid changes in motivation and able to support smokers' changing needs and interests, while still focusing on similar goals and strategies that have been found to be effective for smoking cessation.

Intervention Concept and Preliminary Work

We previously designed and pilot-tested a phone-based counseling program for smokers with depression, most of whom (69%) were ambivalent about quitting. A key component of the intervention was 9 weekly "experiments" that the participants were encouraged to try on their own between counseling sessions and to report back what they had learned. Each experiment was a short exercise designed to address cognitive restructuring and behavioral activation for mood management or to build self-efficacy for smoking cessation (eg, learning to delay smoking in response to urges, making a practice quit attempt) [17].

Findings from this study demonstrated the acceptability of using this approach to engage ambivalent smokers and support behavior change. We subsequently assessed ambivalent smokers' reactions using a similar concept as a component of an mHealth intervention during user-centered design workshops. Participants in our user-centered design workshops were strongly in favor of trying what they retermed "personal experiments" to guide them through short, discrete activities that could help them learn the skills needed to change their smoking habits or to explore their interest in quitting. They particularly liked the idea of accessing these "personal experiments" through an mHealth app and suggested that users have the opportunity to earn points or rewards for completing each experiment. Gamification is a common request from smokers in our design work [6,18,19].

Current Study

The goals of this study were to design and pretest a set of mHealth-delivered "personal experiments" for ambivalent smokers, using the experiments from our prior research as a guide, but modifying the topics and experiment structure (eg, duration) to work better as part of a self-guided program. Study

findings are currently being used to refine the experiment concept for subsequent testing of the efficacy of this intervention as part of a more comprehensive mHealth app.

Methods

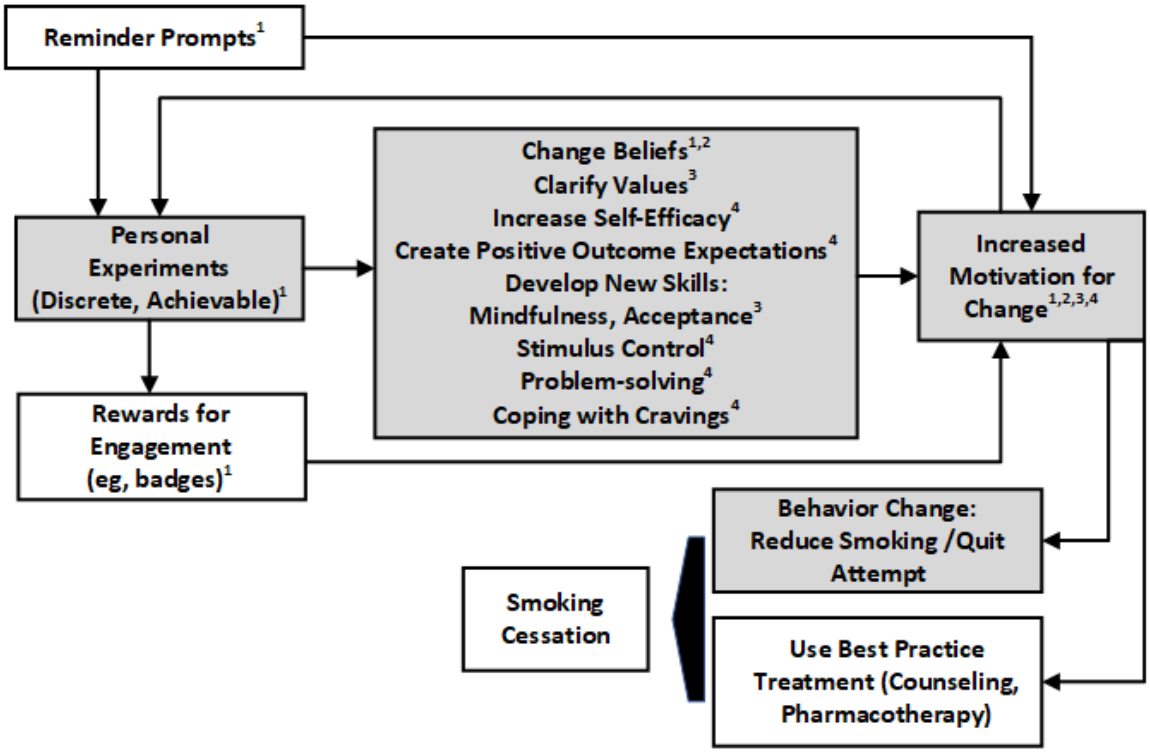
Prototype Development and Testing

Theoretical Foundation

The intervention concept was grounded in several prominent and complementary motivation and behavior change theories along with the empirically validated best-practice recommendations of the US Public Health Service (PHS) Guidelines for Treatment of Nicotine Dependence [20]. Consistent with the PRIME theory of motivation [16,21], the intervention acknowledges that motivation for behavior change is fluid and, in part, determined by one's situational beliefs. Thus, the intervention assumes that motivation will fluctuate over time and targets smokers' beliefs. In accordance with social cognitive theory [22-24], which forms the basis for cognitive behavioral therapy and many recommendations of the PHS Treatment Guidelines [20], we focus on promoting confidence (self-efficacy) and positive outcome expectations, since these are associated with success in quitting smoking. Toward this end, we used an approach to engage smokers in discrete experiments designed to shape their motivation and behavior change through traditional behavioral techniques such as successive approximation, reinforcement, and shaping. The experiments were also designed to teach the specific skills needed to cut back or quit smoking (eg, managing cravings) using techniques from both traditional cognitive behavioral therapy (eg, problem solving, stimulus control) and acceptance and commitment therapy (eg, values clarification, mindfulness, and other acceptance-based coping skills) [25-27]. We hypothesize that by engaging in the experiments, smokers will have successive mastery experiences that will build confidence and positive outcome expectations (ie, "I believe that I can control my smoking or I can quit when I am ready"), and in turn will encourage greater efforts for change, including making a quit attempt and ultimately quitting smoking. Drawing from Fogg's [28] model for persuasive design, the intervention also recognizes that when people have low motivation for change (as is expected for ambivalent smokers), it is important that the behaviors they are asked to engage in are fairly simple (ie, require low ability) and that these behaviors need to be coupled with extrinsic triggers to prompt engagement (ie, reminder prompts).

To ultimately achieve smoking cessation, individuals may also need to utilize other treatment aids such as counseling or pharmacotherapy [20], but we believe that engaging in the personal experiments will increase the likelihood that this will occur. Figure 1 shows the theoretically based conceptual model of the intervention; items in grey reflect the core intervention elements being developed and evaluated in the present work.

Figure 1. Conceptual model of the intervention.

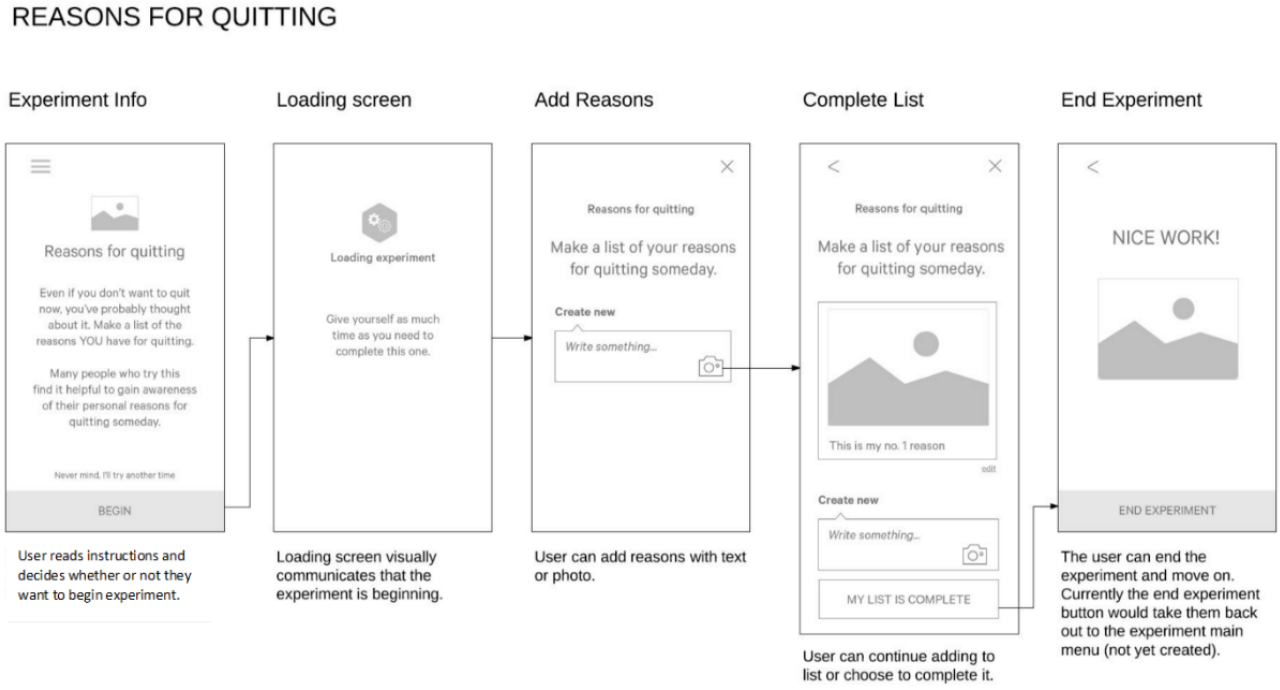


Preliminary User Interviews

As a first step, we wanted to assess individuals' reactions to the general experiment concept when presented as self-guided app-based exercises, and to collect feedback on several key issues that could help us refine the content and design of the experiments in the pilot study. For this purpose, we adapted 5

personal experiments from a prior study that used a similar experiment concept as part of a phone-based counseling program [17]. In that study, the counselor was able to explain each experiment to participants, who then had 1 week to complete each experiment. For an mHealth intervention, it is important that each experiment be self-explanatory, and we believe that a shorter time frame would be more optimal for keeping users engaged. Therefore, we created low-fidelity prototypes of 5 experiments that we believed were easy to understand and could be completed in less than 2 days (see Figure 2 for an example). We initially limited the number of experiments to 5 to reduce participant burden and maximize the time available during the interviews to more fully explore user reactions. The experiments included were specifically chosen to represent a range of goals and topics (ie, exploring motivation for quitting, tracking smoking behavior, changing smoking behavior or location), input options (ie, written comments, uploaded photos), duration (ie, experiments that could be completed in the moment and those that required taking action over a 24-48 hour period), and related design issues (eg, using progress indicators) that were considered pertinent to the final selection and design of the future experiments.

Figure 2. Example of a low-fidelity prototype of a personal experiment.



Five ambivalent smokers were screened as eligible to provide feedback on the initial prototype, but two cancelled and could not be rescheduled, leaving three final participants for this formative phase. Two of the three (67%) participants were women, all three were White, they were aged 40-46 years, reported smoking 5-10 cigarettes per day, and 2/3 (67%) reported household incomes under US $50,000/year (the other participant refused to answer). Each participant viewed a storyboard explaining the concept of the personal experiments as an mHealth intervention and were then walked through each of the experiment prototypes by a trained user-centered design researcher.

For each experiment, the participants were asked to identify any parts that were confusing, what they liked and disliked, whether they could see themselves trying the experiment, and what they anticipated would be the biggest challenges to their completing the experiment. They were also asked about the perceived helpfulness of a program that included these features; what, if anything, would help them stay engaged with the program over time; what other features they would like to see included; and whether they wanted to be able to share their experiment progress with others. The feedback from user interviews was then used to iteratively refine the basic intervention design and presentation prior to the pilot study.

Medium-Fidelity Prototypes

Using feedback collected from the user interviews, we designed a set of 7 personal experiments and then created a functional medium-fidelity prototype of the intervention program, including an initial program orientation and each personal experiment (see Table 1 for a summary of each and Figure 3 for sample content). Each experiment topic was chosen based on its theoretical or empirical utility for changing smoking-related attitudes, beliefs, or behaviors based on our prior research. The experiments were intentionally ordered based on the flow that we expect will maximize the intended therapeutic effects, starting with exercises designed to build or strengthen motivation for quitting, followed by experiments intended to teach coping or other behavioral skills needed to resist the urge to smoke in response to cravings and build self-confidence for making a quit attempt, and culminating in a 24-hour practice quit attempt.

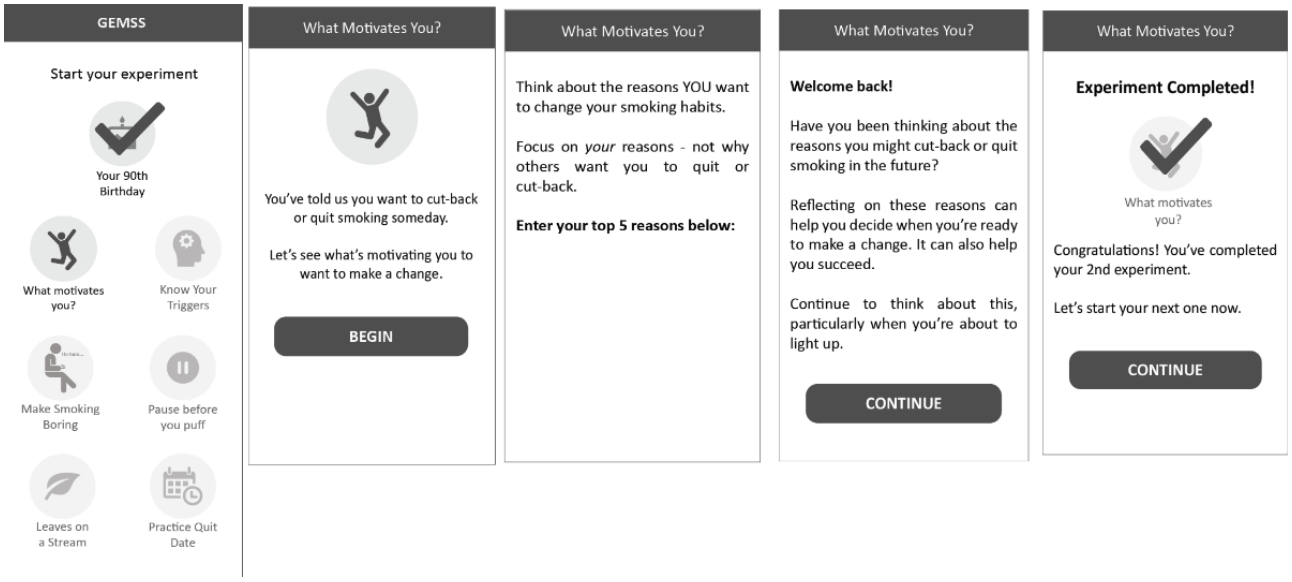
An overview of each experiment and its intended cognitive or behavioral target is provided in Table 1. All experiments were

designed to be completed either in the moment or within 48 hours.

Table 1. Personal experiments.

Experiment number	Title	Goal/skill targeted	Description
1	90th Birthday	Clarify personal values.	User is asked to imagine giving a speech at their 90th birthday party about what have been the most meaningful aspects of their life and to consider whether they want to be remembered as a smoker.
2	What Motivates You?	Identify reasons for quitting. Build motivation.	User is asked to make a list of all of the reasons that they want to quit smoking and review this list in response to cravings.
3	Know Your Triggers	Identify high-risk situations for smoking. Aid future problem-solving and coping.	User is asked to write down what they're doing every time they have a craving to smoke.
4	Make Smoking Boring	Stimulus control. Make smoking less reinforcing.	User is asked to do no other activities (eg, no friends, coffee, TV) while smoking and to notice how it feels.
5	Pause Before You Puff	Learn to delay smoking in response to urges.	User is asked to wait 1 minute before each cigarette and to use that time to consider personal values or reasons for quitting, or to do nothing (ie, make smoking boring).
6	Leaves on a Stream	Learn mindfulness-based coping strategy. Learn to let urges pass without smoking.	User is asked to visualize thoughts as leaves on a moving stream and to practice nonjudgmentally observing these thoughts as they come and go (including thoughts about smoking) without acting on them.
7	Practice Quit Attempt	Make a practice quit attempt. No smoking for 24 hours.	User is asked to attempt to stop smoking for 24 hours and is encouraged to use nicotine replacement to assist with this challenge.

Figure 3. Medium-fidelity prototype of a personal experiment.



Single-Arm Pilot Test

Participants

Twenty-five adult smokers were recruited to participate in a one-arm pilot test of the prototype intervention. Typically, randomized pilot trials of similar behavioral interventions include 25-30 participants per arm [29]; thus, 25 was viewed as an adequate sample size to assess the outcomes of interest in this formative study. Individuals were eligible if they were at least 18 years of age; smoked at least 100 cigarettes in their lifetime; reported any smoking in the last 7 days; were interested in quitting someday, but not currently trying to quit or planning to quit in the next month (ie, ambivalent about quitting); were comfortable reading and speaking in English; owned a smartphone; used apps on their smartphone at least once a week; and had a personal email account.

Procedures

Participants were recruited between September and November 2017 via advertisements on Craigslist in the following cities: Seattle, WA; Baltimore, MD; Columbus, OH; Atlanta, GA; and Oakland, CA. The cities were purposely chosen to obtain a geographically, racially, and socioeconomically diverse sample.

To determine eligibility, potential participants were screened by telephone, following verbal consent to participate.

Eligible participants completed a baseline survey via telephone, either immediately following the telephone screening or in a later scheduled call. The survey assessed demographics, tobacco use, motivation to quit smoking, smartphone use, and interest in and experience with mHealth apps.

Participants were emailed a URL that allowed them to access the initial orientation and the first experiment from their smartphone. Study staff walked individuals through this content while on the phone to assess their reactions in real time and to ensure that they understood how to access and use the online prototype.

Starting on day 2, the remaining experiments were pushed to smokers one at a time via an emailed link. When the link was opened, smokers were taken to a mobile-optimized website, which mimicked the appearance and functionality of an mHealth app. Each experiment used a similar format, starting with a brief explanation of the experiment's purpose, 1-2 screens explaining the action to be completed and encouragement to try it, and instructions that we would check back in 1-2 days to see how it went.

One to two days after each experiment was viewed online (which was monitored remotely), smokers were emailed a link that allowed them to complete a brief survey and then begin their next experiment. Surveys assessed motivation and self-efficacy for both quitting and reducing smoking, measured on a 5-point Likert-type scale where 1="not at all" and 5="very." In addition to being more feasible to administer in a brief, repeated assessment protocol, the predictive validity of single-item measures of motivation and self-efficacy has been supported in studies focused on change in smoking and other substance use behaviors [30,31]. Participants also rated how helpful and difficult each experiment was using the same 5-point scale. With the exception of the first two experiments that involved a separate pre-experiment survey, the timing of the surveys allowed each to serve both as a postexperiment assessment and a pre-experiment baseline for the next experiment, enabling assessments of change over time in relation to each experiment.

The final survey contained additional questions about overall perceptions and impact of the program, including whether the program caused them to think differently about quitting or cutting back on smoking (yes/no) and, if yes, whether it made them more interested in quitting, more interested in cutting back, less interested in quitting, or less interested in cutting back. Participants were also asked if the program had provided a free 2-week supply of nicotine replacement therapy (NRT), whether they would have used it to try to quit. Response options were: "yes," "unsure," and "no, I would have saved them until I am ready to quit smoking."

If participants failed to view an experiment within the planned 48-hour window, they were sent up to 4 email reminders. For the purpose of this pilot study, the entire series of experiments was designed to be completed in 2 weeks. Participants received US $75 for completing all experiments and a final follow-up survey.

The project was reviewed by the Kaiser Permanente Washington Human Subjects Review Board and deemed exempt from review due to its formative nature (ie, designed to develop a program rather than to produce generalizable knowledge).

Analyses

The majority of analyses for this formative work are descriptive, including means (SD) for continuous variables and frequencies and percentages for categorical variables. To assess the change in ratings of motivation and confidence, we used paired-sample *t* tests to compare pre- and postratings for each experiment, and we report the change score and 95% CIs for each comparison.

Results

User Interviews

Interviewees (N=3) responded positively to the intervention concept. They identified some aspects of the design that were confusing, including specific wording and iconography (eg, using an image of a camera to indicate the ability to upload photos). They also recommended several additional program features for future consideration, including allowing users to save their reflections about each exercise (eg, a journal), adding testimonials from other smokers, including statistics and information about smoking, and using gamification features to make the program more engaging and fun (eg, badges, challenges). Two of three participants were not interested in adding a social feature that would share their progress with others, citing a strong desire for privacy. Some of this feedback was incorporated into the medium-fidelity prototype (eg, dropping the ability to upload photographs). Other feedback that was out of the scope for the experiment concept (eg, adding a journal, testimonials, and reward badges) is being implemented and tested in an ongoing randomized pilot trial of the intervention.

Single-Arm Pilot Test

Participant Characteristics

Among the 25 participants, 15/24 (63%) reported their race/ethnicity as nonwhite (11 Black, 1 Asian, 1 Mexican American, 2 with multiple responses, and 1 invalid response); 12% (3/25) of the participants were Hispanic, 64% (16/25) were men, 24% (6/25) had an education of high school or less, 68% (17/25) were employed, and 48% (12/25) had an annual household income less than US $45,000. Regarding tobacco product use, 56% (14/25) smoked cigarettes only and 44% (11/25) used another form of tobacco or nicotine product (eg, 7 used electronic cigarettes, 4 smoked cigars, and 3 used other tobacco products) in addition to cigarettes. The average number of cigarettes smoked per day was 17 (SD 11). More participants had Android phones (68%, 17/25) than iPhones (32%, 8/25). Only 2 of the 25 (8%) participants had ever used a smoking cessation app, although most (84%, 21/25) said that they would consider using one. All 25 participants said they would consider using an app that helped them decide if, when, or how to quit smoking. Nearly half (12/25, 48%) reported having experience

using some other type of health app, with the most common being a physical activity app (9/25, 36%).

Receptivity

After watching the brief program orientation that explained how the program works, most participants (80%, 20/25) indicated that it sounded interesting, primarily because it did not require a commitment to quit. All participants (100%, 25/25) completed all 7 experiments; 80% (20/25) completed these within 2 weeks, as planned, and 100% (25/25) within 1 month. Most of the participants (88%, 22/25) liked the order of the experiments presented. All participants (100%, 25/25) tried the 24-hour quit attempt, although not all were able to stay quit for a full day. The mean rating of usefulness of the overall program was 4.12 (SD 1.09) out of 5, and the average rating of the difficulty of the experiments was 2.16 (SD 1.18) out of 5. Regarding difficulty, feedback on the Practice Quit Date exercise highlighted the difficulty that some participants experienced in trying to go 24 hours without smoking, with some being unable to do so. For example, one participant, when asked what they disliked about the experiment, stated, "I did not like the fact that I was not able to quit for one day." Another participant noted, "I tried using the tools that I learned from this study and to a certain degree it worked. I smoked less but I still smoked. I held out for most of one day, then I caved." Participants' comments also indicated that some were successful at the 24-hour abstinence goal. One participant stated, "I liked that I could go a whole day without smoking a cigarette. I thought I would have more withdrawal symptoms but I did not." Another participant said, "Although it was very hard I did it!"

Helpfulness ratings (on a 1 to 5 scale where 1=not at all and 5=very helpful) for individual experiments are shown in Table 2 and ranged from 3.44 to 3.96, indicating a net positive rating for all experiments. The three experiments rated highest on helpfulness were Make Smoking Boring, Know Your Triggers, and Pause Before You Puff.

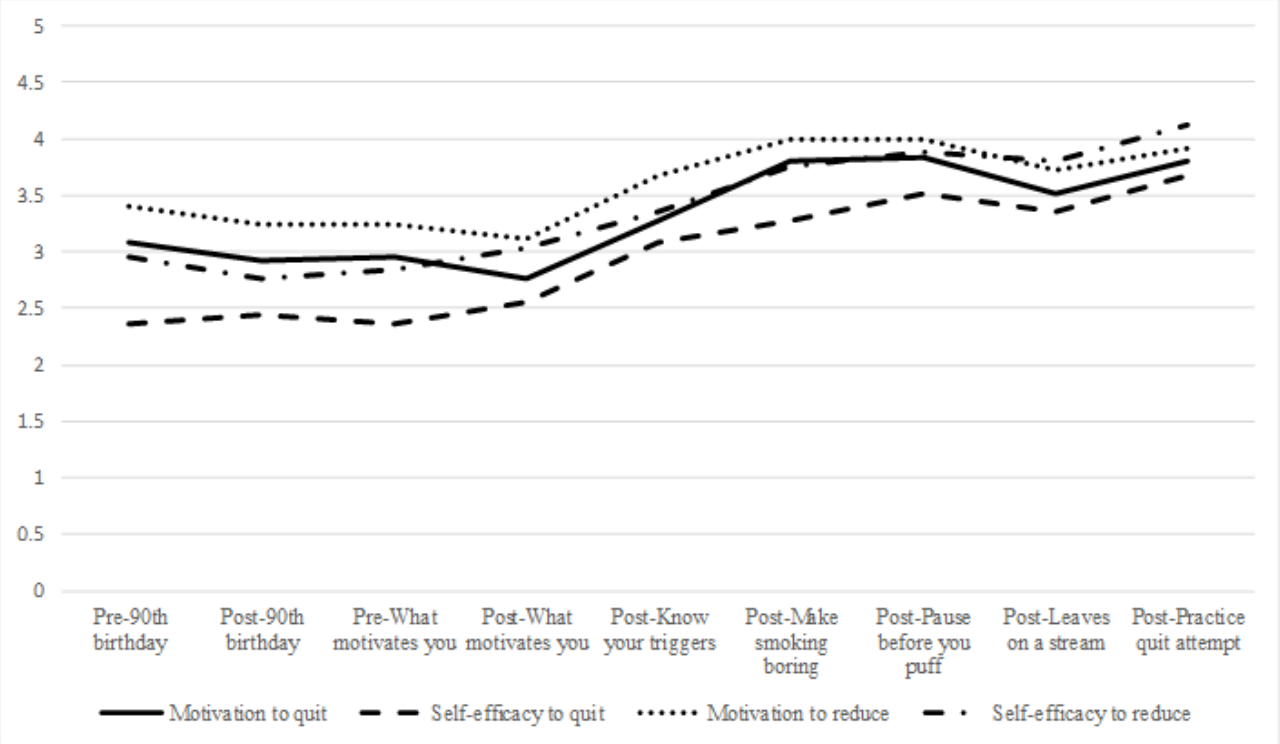
Table 2. Personal experiments and user receptivity outcomes.

Experiment	Title	Helpfulness, mean (SD)	Change in motivation to quit, mean (95% CI)	Change in confidence to quit, mean(95% CI)	Change in motivation to cut back, mean (95% CI)	Change in confidence to cut back, mean (95% CI)
1	90th Birthday	3.44 (1.26)	–0.16 (–0.63, +0.31)	+0.08 (–0.35, +0.51)	–0.16 (–0.63, +0.31)	–0.20 (–0.63, +0.23)
2	What Motivates You?	3.60 (1.04)	–0.20 (–0.58, +0.18)	+0.20 (–0.18, +0.58)	–0.12 (–0.34, +0.10)	+0.20 (–0.31, +0.71)
3	Know Your Triggers	3.75 (1.23)	+0.52 (+0.20, +0.84)	+0.52 (+0.09, +0.95)	+0.56 (+0.07, +1.05)	+0.32 (–0.15, +0.79)
4	Make Smoking Boring	3.96 (1.17)	+0.52 (+0.23, +0.82)	+0.20 (–0.12, +0.52)	+0.32 (–0.03, +0.67)	+0.40 (+0.11, +0.69)
5	Pause Before You Puff	3.75 (1.29)	+0.04 (–0.21, +0.29)	+0.24 (–0.08, +0.56)	+0.00 (–0.32, +0.32)	+0.12 (–0.16, +0.40)
6	Leaves on a Stream	3.46 (1.18)	–0.32 (–0.71, +0.07)	-0.16 (–0.51, +0.19)	–0.28 (–0.61, +0.05)	+0.08 (–0.44, +0.28)
7	Practice Quit Attempt	3.68 (1.31)	+0.28 (–0.07, +0.63)	+0.32 (+0.06, +0.58)	+0.20 (–0.14, +0.54)	+0.29 (–0.11, +0.70)

Motivation to Quit

At the last assessment point, 92% of the respondents (23/25) reported that trying the experiments made them think differently about quitting or cutting back, with a roughly even split between those who indicated that they were more interested in cutting back (11/25, 44%) and those who indicated that they were more interested in quitting completely (12/25, 48%). There was an average increase of 0.72 points on the 5-point motivation scale (95% CI +0.22 to +1.22) between the first pre-experiment assessment and the last postexperiment assessment. The impact of each experiment on motivation to quit is provided in Table 2, and Figure 4 shows the increases across experiments. The three experiments with the largest positive change in motivation to quit were Know Your Triggers, Make Smoking Boring, and Practice Quit Attempt.

Figure 4. Change in motivation and self-efficacy across experiments.



Quitting Self-Efficacy

Confidence in ability to quit also showed an increase across the experiments (see Figure 4), increasing by over one point (+1.32, 95% CI +0.78 to +1.86) from the first to the last assessment. The three exercises with the largest increase in confidence to quit were Know Your Triggers, Practice Quit Attempt, and Pause Before You Puff (Table 2).

Motivation to Reduce Smoking

Motivation to reduce smoking increased across the experiments by one-half of a point (mean +0.52, 95% CI +0.12 to +0.92) from the first to the last assessment (Figure 4). The three exercises with the greatest increase in motivation to cut back were Know Your Triggers, Make Smoking Boring, and Practice Quit Attempt (Table 2).

Self-Efficacy to Reduce Smoking

Confidence in ability to reduce smoking increased across experiments (Figure 4) by an average of 1.17 points (95% CI +0.61 to +1.72) from the first to the last assessment. The three experiments associated with the greatest increases in confidence to cut back were Make Smoking Boring, Know Your Triggers, and Practice Quit Attempt (Table 2).

User Feedback on Adding Pharmacotherapy

Overall, 72% (18/25) of the participants said that if the program had included a free trial of NRT, they would have used it to try to quit; 20% (2/25) said that they were unsure if they would use NRT, and 8% (2/25) said they would save it until they were ready to stop smoking.

Discussion

Principal Findings

Results of this formative work provide proof-of-concept evidence that ambivalent smokers are willing to use and will remain engaged with a self-guided mHealth intervention using the concept of personal experiments to enhance their ability to quit smoking. These findings expand on our prior mixed methods study to assess the preferences and behavioral intentions of ambivalent smokers, in which we found a high rate of interest in using a digital health program where the messaging was framed specifically for ambivalent smokers (ie, smoking reduction or decision support to help them decide if, how, and when to give quitting a try) [6].

Motivation and self-efficacy, both for quitting and for reducing smoking, increased across the period of use, suggesting that the program impacted key cognitive targets. Because self-efficacy and motivation to quit are predictive of quit attempts and quit success [32,33], these findings are an encouraging signal of potential efficacy for supporting cessation among ambivalent smokers. Taken together with the strong indications of acceptability based on high perceived usefulness and high engagement with the program, these findings warrant continued development of the program as a novel method for engaging and assisting ambivalent smokers in a cessation program.

Next Steps

Out of the 7 experiments, Leaves on a Stream was the only experiment that was identified as confusing based on the title and iconography (4/25), and after trying it, participants gave this experiment the lowest overall ratings. Based on this feedback, we plan to drop this experiment and replace it with an alternative mindfulness-based coping exercise. In addition,

we plan to add more experiments designed to help users resist the urge to smoke in response to cravings (a critical skill for smoking cessation), include gamification features (ie, badge rewards), a journal feature to allow users to record lessons learned, and testimonials, in addition to pairing the program with best-practice treatment (self-help advice, NRT, and access to quitline counseling). These refinements are responsive to feedback from participants in the preliminary user interviews and in our other prior design work [6], as well as a body of research suggesting that providing active treatment to unmotivated smokers encourages quit attempts and improves cessation rates [34]. Inclusion of these components is also aligned with the PRIME theory of motivation [21]. Since motivation is dynamic, people may convert from ambivalence to readiness for action at any time. Thus, providing these resources will ensure that they have action-oriented support when it is needed.

Limitations and Strengths

The small sample sizes in this research limit our ability to draw any conclusions about the generalizability of our findings. The lack of a control arm or long-term follow up in the pilot also prevents us from making strong assertions about the program's impact on motivation, self-efficacy, and behavior change. Self-report data are also subject to social desirability bias. Although these limitations should be considered when interpreting the results, the methods are appropriate for this formative stage of design.

Strengths of this work include a demographically diverse participant sample, a rigorous assessment strategy that included pre-post evaluation on key constructs of interest immediately prior to and following each experiment, and a high pilot retention rate. Given the dearth of knowledge about how best to utilize digital health technologies to support behavior change among people ambivalent about quitting, our target population is also a study strength.

Conclusions

This work highlights both the need and the promise for interventions targeted to smokers who are ambivalent about quitting. If found to be effective in future work, the planned intervention could provide an attractive new option for ambivalent smokers as well as for employers, quitlines, and health care organizations, none of whom currently has evidence-based options available to offer this group. The results may also be relevant when designing mHealth interventions for people not yet ready to commit to other types of health behavior change (eg, physical activity, dietary intake, alcohol use), as the concept of personal experiments may be a useful strategy for engaging users and promoting action without requiring a commitment to change.

Acknowledgments

This work was supported by grants from the National Cancer Institute (R21CA234003; principal investigator JM) and Kaiser Permanente Washington Health Research Institute (principal investigator JM). JH was supported by the Fred Hutchinson Cancer Research Center. PK was supported, in part, by the University of Michigan. Early design work and publication fees for this article were supported by the Betty Irene Moore School of Nursing at University of California Davis. We are grateful to Ella Thompson for her assistance managing this project.

Conflicts of Interest

JH has received research support from Pfizer. The remaining authors declare no conflict of interest.

References

1. WHO global report on trends in prevalence of tobacco use 2000–2025, third edition. Geneva: World Health Organization; 2019. URL: https://www.who.int/publications/i/item/who-global-report-on-trends-in-prevalence-of-tobacco-use-2000-2025-third-edition [accessed 2020-08-20]
2. Centers for Disease ControlPrevention (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]
3. Mobile Fact Sheet 2019. Pew Research Center. Washington, DC: Pew Research Center; 2019 Jun 12. URL: https://www.pewresearch.org/internet/fact-sheet/mobile/ [accessed 2020-08-14]
4. Phillips E, Wang TW, Husten CG, Corey CG, Apelberg BJ, Jamal A, et al. Tobacco Product Use Among Adults - United States, 2015. MMWR Morb Mortal Wkly Rep 2017 Nov 10;66(44):1209-1215. [doi: 10.15585/mmwr.mm6644a2] [Medline: 29121001]
5. 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]
6. McClure JB, Heffner J, Hohl S, Klasnja P, Catz SL. Design Considerations for mHealth Programs Targeting Smokers Not Yet Ready to Quit: Results of a Sequential Mixed-Methods Study. JMIR Mhealth Uhealth 2017 Mar 10;5(3):e31 [FREE Full text] [doi: 10.2196/mhealth.6845] [Medline: 28283465]
7. Ramo DE, Thrul J, Delucchi KL, Hall S, Ling PM, Belohlavek A, et al. A randomized controlled evaluation of the tobacco status project, a Facebook intervention for young adults. Addiction 2018 May 24;113(9):1683-1695 [FREE Full text] [doi: 10.1111/add.14245] [Medline: 29797621]

8. Fagerström KO. Can reduced smoking be a way for smokers not interested in quitting to actually quit? Respiration 2005;72(2):216-220. [doi: 10.1159/000084057] [Medline: 15824536]
9. McClure JB, Westbrook E, Curry SJ, Wetter DW. Proactive, motivationally enhanced smoking cessation counseling among women with elevated cervical cancer risk. Nicotine Tob Res 2005 Dec;7(6):881-889. [doi: 10.1080/14622200500266080] [Medline: 16298723]
10. McClure JB, Peterson D, Derry H, Riggs K, Saint-Johnson J, Nair V, et al. Exploring the "active ingredients" of an online smoking intervention: a randomized factorial trial. Nicotine Tob Res 2014 Aug;16(8):1129-1139 [FREE Full text] [doi: 10.1093/ntr/ntu057] [Medline: 24727369]
11. Carpenter MJ, Hughes JR, Solomon LJ, Callas PW. Both smoking reduction with nicotine replacement therapy and motivational advice increase future cessation among smokers unmotivated to quit. J Consult Clin Psychol 2004 Jun;72(3):371-381. [doi: 10.1037/0022-006X.72.3.371] [Medline: 15279521]
12. Chan SSC, Leung DYP, Abdullah ASM, Wong VT, Hedley AJ, Lam T. A randomized controlled trial of a smoking reduction plus nicotine replacement therapy intervention for smokers not willing to quit smoking. Addiction 2011 Jun;106(6):1155-1163. [doi: 10.1111/j.1360-0443.2011.03363.x] [Medline: 21226883]
13. Ebbert JO, Hughes JR, West RJ, Rennard SI, Russ C, McRae TD, et al. Effect of varenicline on smoking cessation through smoking reduction: a randomized clinical trial. JAMA 2015 Feb 17;313(7):687-694 [FREE Full text] [doi: 10.1001/jama.2015.280] [Medline: 25688780]
14. Wu L, Sun S, He Y, Zeng J. Effect of Smoking Reduction Therapy on Smoking Cessation for Smokers without an Intention to Quit: An Updated Systematic Review and Meta-Analysis of Randomized Controlled. Int J Environ Res Public Health 2015 Aug 25;12(9):10235-10253 [FREE Full text] [doi: 10.3390/ijerph120910235] [Medline: 26308034]
15. Moore D, Aveyard P, Connock M, Wang D, Fry-Smith A, Barton P. Effectiveness and safety of nicotine replacement therapy assisted reduction to stop smoking: systematic review and meta-analysis. BMJ 2009 Apr 02;338:b1024 [FREE Full text] [doi: 10.1136/bmj.b1024] [Medline: 19342408]
16. West R. Time for a change: putting the Transtheoretical (Stages of Change) Model to rest. Addiction 2005 Aug;100(8):1036-1039. [doi: 10.1111/j.1360-0443.2005.01139.x] [Medline: 16042624]
17. McClure JB, Catz SL, Ludman EJ, Richards J, Riggs K, Grothaus L. Feasibility and acceptability of a multiple risk factor intervention: the Step Up randomized pilot trial. BMC Public Health 2011 Mar 17;11:167 [FREE Full text] [doi: 10.1186/1471-2458-11-167] [Medline: 21414216]
18. McClure JB, Hartzler AL, Catz SL. Design Considerations for Smoking Cessation Apps: Feedback From Nicotine Dependence Treatment Providers and Smokers. JMIR Mhealth Uhealth 2016 Feb 12;4(1):e17 [FREE Full text] [doi: 10.2196/mhealth.5181] [Medline: 26872940]
19. Hartzler AL, BlueSpruce J, Catz SL, McClure JB. Prioritizing the mHealth Design Space: A Mixed-Methods Analysis of Smokers' Perspectives. JMIR Mhealth Uhealth 2016 Aug 05;4(3):e95 [FREE Full text] [doi: 10.2196/mhealth.5742] [Medline: 27496593]
20. Fiore M, Jaén C, Baker T, Bailey W, Benowitz N, Curry S. Treating tobacco use and dependence update. Clinical practice guideline. Rockville, MD: U.S. Department of Health and Human Services. Public Health Service; 2008.
21. West R. The PRIME Theory of motivation as a possible foundation for addiction treatment. In: Henningfield J, Santora P, editors. Drug Addiction Treatment in the 21st Century: Science and Policy Issues. Baltimore: Johns Hopkins University Press; 2007.
22. Bandura A. Self-efficacy: Toward a unifying theory of behavioral change. Psychol Rev 1977;84(2):191-215. [doi: 10.1037/0033-295x.84.2.191]
23. Bandura A. Social foundations of thought and action: A social cognitive theory. Englewood Cliffs, NJ: Prentice Hall; 1985.
24. Bandura A. Self-efficacy: The exercise of control. New York: WH and Freeman Company; 1997.
25. Hayes SC, Luoma JB, Bond FW, Masuda A, Lillis J. Acceptance and commitment therapy: model, processes and outcomes. Behav Res Ther 2006 Jan;44(1):1-25. [doi: 10.1016/j.brat.2005.06.006] [Medline: 16300724]
26. 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]
27. McClure JB, Bricker J, Mull K, Heffner JL. Comparative Effectiveness of Group-Delivered Acceptance and Commitment Therapy versus Cognitive Behavioral Therapy for Smoking Cessation: A Randomized Controlled Trial. Nicotine Tob Res 2020 Mar 16;22(3):354-362 [FREE Full text] [doi: 10.1093/ntr/nty268] [Medline: 30590810]
28. Fogg B. A behavior model for persuasive design. 2009 Presented at: Persuasive '09: Proceedings of the 4th International Conference on Persuasive Technology; April 2009; Claremont, CA p. 1-7. [doi: 10.1145/1541948.1541999]
29. Mekler LB. On the problem of oncogene of tumour viruses. Acta Virol 1975 Nov;19(6):501-508. [Medline: 2001]
30. Hoeppner BB, Kelly JF, Urbanoski KA, Slaymaker V. Comparative utility of a single-item versus multiple-item measure of self-efficacy in predicting relapse among young adults. J Subst Abuse Treat 2011 Oct;41(3):305-312 [FREE Full text] [doi: 10.1016/j.jsat.2011.04.005] [Medline: 21700411]

31. Boudreaux ED, Sullivan A, Abar B, Bernstein SL, Ginde AA, Camargo CA. Motivation rulers for smoking cessation: a prospective observational examination of construct and predictive validity. Addict Sci Clin Pract 2012 Jun 08;7:8 [FREE Full text] [doi: 10.1186/1940-0640-7-8] [Medline: 23186265]
32. Klemperer EM, Mermelstein R, Baker TB, Hughes JR, Fiore MC, Piper ME, et al. Predictors of smoking cessation attempts and success following motivation-phase interventions among people initially unwilling to quit smoking. Nicotine Tob Res 2020 Apr 01. [doi: 10.1093/ntr/ntaa051] [Medline: 32236417]
33. Castro Y, Cano MA, Businelle MS, Correa-Fernández V, Heppner WL, Mazas CA, et al. A cross-lagged path analysis of five intrapersonal determinants of smoking cessation. Drug Alcohol Depend 2014 Apr 01;137:98-105 [FREE Full text] [doi: 10.1016/j.drugalcdep.2014.01.013] [Medline: 24529688]
34. Jardin BF, Cropsey KL, Wahlquist AE, Gray KM, Silvestri GA, Cummings KM, et al. Evaluating the effect of access to free medication to quit smoking: a clinical trial testing the role of motivation. Nicotine Tob Res 2014 Jul;16(7):992-999 [FREE Full text] [doi: 10.1093/ntr/ntu025] [Medline: 24610399]

Abbreviations

mHealth: mobile health
NRT: nicotine replacement therapy
PHS: Public Health Service

Edited by G Eysenbach; submitted 25.06.20; peer-reviewed by D Kale, J Thrul; comments to author 17.07.20; accepted 03.08.20; published 27.08.20.

Please cite as:
Heffner JL, Catz SL, Klasnja P, Tiffany B, McClure JB
Development of a Mobile Health Intervention with Personal Experiments for Smokers Who Are Ambivalent About Quitting: Formative Design and Testing
JMIR Form Res 2020;4(8):e21784
URL: http://formative.jmir.org/2020/8/e21784/
doi:10.2196/21784
PMID:32852278

Original Paper

Investigation of a Mobile Health Texting Tool for Embedding Patient-Reported Data Into Diabetes Management (i-Matter): Development and Usability Study

Antoinette Schoenthaler[1*], EdD; Jocelyn Cruz[1*], MPH; Leydi Payano[1*], BS; Marina Rosado[1*], BS; Kristen Labbe[1*], BA; Chrystal Johnson[2*], BS; Javier Gonzalez[3*], BS; Melissa Patxot[4*], BS; Smit Patel[4*]; Eric Leven[4*], MBA, MSc; Devin Mann[5*], MD, MS

[1]NYU Langone Health, Department of Population Health, Center for Healthful Behavior Change, New York, NY, United States

[2]NYU Langone Health, Medical Center Information Technology Enterprise Project Management Office, New York, NY, United States

[3]NYU Langone Health, Department of Population Health, Digital DesignLab, New York, NY, United States

[4]Rip Road, Inc, New York, NY, United States

[5]NYU Langone Health, Department of Population Health, Healthcare Innovation Bridging Research, Informatics and Design Lab, New York, NY, United States

[*]all authors contributed equally

Corresponding Author:
Antoinette Schoenthaler, EdD
NYU Langone Health
Department of Population Health
Center for Healthful Behavior Change
180 Madison Ave
7th floor
New York, NY, 10016
United States
Phone: 1 6465013434
Email: antoinette.schoenthaler@nyumc.org

Abstract

Background: Patient-reported outcomes (PROs) are increasingly being used in the management of type 2 diabetes (T2D) to integrate data from patients' perspective into clinical care. To date, the majority of PRO tools have lacked patient and provider involvement in their development, thus failing to meet the unique needs of end users, and lack the technical infrastructure to be integrated into the clinic workflow.

Objective: This study aims to apply a systematic, user-centered design approach to develop i-Matter (investigating a mobile health [mHealth] texting tool for embedding patient-reported data into diabetes management), a theory-driven, mobile PRO system for patients with T2D and their primary care providers.

Methods: i-Matter combines text messaging with dynamic data visualizations that can be integrated into electronic health records (EHRs) and personalized patient reports. To build i-Matter, we conducted semistructured group and individual interviews with patients with T2D and providers, a design thinking workshop to refine initial ideas and design the prototype, and user testing sessions of prototypes using a rapid-cycle design (ie, design-test-modify-retest).

Results: Using an iterative user-centered process resulted in the identification of 6 PRO messages that were relevant to patients and providers: medication adherence, dietary behaviors, physical activity, sleep quality, quality of life, and healthy living goals. In user testing, patients recommended improvements to the wording and timing of the PRO text messages to increase clarity and response rates. Patients also recommended including motivational text messages to help sustain engagement with the program. The personalized report was regarded as a key tool for diabetes self-management by patients and providers because it aided in the identification of longitudinal patterns in the PRO data, which increased patient awareness of their need to adopt healthier behaviors. Patients recommended adding individualized tips to the journal on how they can improve their behaviors. Providers preferred having a separate tab built into the EHR that included the personalized report and highlighted key trends in patients' PRO data over the past 3 months.

Conclusions: PRO tools that capture patients' well-being and the behavioral aspects of T2D management are important to patients and providers. A clinical trial will test the efficacy of i-Matter in 282 patients with uncontrolled T2D.

Trial Registration: ClinicalTrials.gov NCT03652389; https://clinicaltrials.gov/ct2/show/NCT03652389

(JMIR Form Res 2020;4(8):e18554) doi:10.2196/18554

KEYWORDS

patient-reported outcome measures; mobile health; type 2 diabetes

Introduction

Background

Uncontrolled type 2 diabetes (T2D) is a significant public health problem in the United States, particularly among vulnerable populations (eg, low-income and racial and ethnic minorities) [1,2]. Annually, T2D incurs about US $250 billion in health care costs and lost productivity, representing a significant social and economic burden [3]. Despite recent improvements in the proportion of adults with T2D achieving hemoglobin A_{1c} (HbA_{1c}) targets <7%, only 50.9% achieved this level of control [4]. The number of patients who fail to meet these goals is even higher in resource-limited primary care practices—a place where most vulnerable populations receive their care [5,6].

Recognizing the central role patients play in the management of T2D (eg, being aware of its signs and symptoms and engaging in daily self-care behaviors), several national and local organizations have forged initiatives to support the development and use of patient-reported outcomes (PROs) in the evaluation of health and well-being of patients with T2D [7-11]. Measures of PROs are a *standardized and quantifiable* approach that allows for the collection and integration of data on patients' perspective of their chronic disease into its clinical management [12].

Much of the existing research that incorporates PROs in T2D has been limited to clinical drug trials examining patient tolerance to new treatments [13]. The few practice-based studies conducted in T2D have used long batteries of PRO measures, and patients report PROs only on a single occasion, most often immediately before clinic visits [14,15]. Such reporting introduces a recall bias because patients are asked to approximate changes in their symptoms and behaviors over several months. To address these challenges, a growing number of studies are utilizing mobile health (mHealth) platforms that enable real-time data collection to facilitate patient self-monitoring outside the clinic environment, enhance patient engagement in their care, and inform provider decision making [16-21].

Systematic reviews of mHealth interventions in patients with T2D have demonstrated positive, short-term benefits on HbA_{1c} levels and self-care behaviors [22-24]. However, these studies have several methodological shortcomings that limit their impact, including small sample sizes (24-180 patients), short study duration (mean 24 weeks), low patient compliance, limited integration with clinical practice, and exclusion of vulnerable populations that would benefit most from mHealth interventions [25]. More importantly, the PROs collected in the mHealth tools are researcher-driven and lack patient and provider involvement in the conceptualization of the intervention. As a result, the tools are not customized to address the complex and unique needs and preferences of patients and lack the technical infrastructure to support integration into the clinic workflow.

Objectives

The i-Matter (investigating an mHealth texting tool for embedding patient-reported data into diabetes management) trial aims to address this gap in the literature by evaluating the efficacy of an innovative mobile PRO system that incorporates patients' perspective of their disease into the management of T2D in primary care practices. The i-Matter intervention uses text messaging to capture patients' self-reported PROs in real time, enhances patient engagement through data-driven feedback and motivational messages, and creates dynamic visualizations of the PROs that can be shared in personalized reports and integrated into the clinical workflow. A future randomized controlled trial (RCT) will evaluate the efficacy of the i-Matter intervention versus usual care on reduction in HbA_{1c} and adherence to self-care behaviors at 12 months among 282 patients with uncontrolled T2D who receive care in resource-limited primary care practices. This paper discusses the iterative process of developing, integrating, and user testing the i-Matter intervention in the formative phase of the trial.

Methods

Theoretical Framework

The i-Matter intervention is a blend of 2 frameworks: technology acceptance model (TAM) and capability, opportunity, and motivation model of behavior (COM-B). The TAM is based on the theory of reasoned action and asserts that perceptions of usefulness and ease of use directly influence the intention to use a new technology, leading in turn to its adoption [26]. The TAM is considered a *gold standard* for characterizing the adoption and use of new health information technology [27,28]. COM-B is a parsimonious amalgamation of existing theories of behavior change [29], which states that interaction among 3 key components is necessary for successful behavior change: the person needs to feel *capable* (ie, the ability to engage in necessary physical and thought processes) of changing, needs to have the *opportunity* (ie, social and environmental factors) to change, and needs to feel *motivated* (ie, confidence and self-efficacy) to change [29]. The COM-B model has been proven effective for designing programs that help patients with T2D improve adherence to health behaviors [29,30]. The core components of the COM-B model are integrated into the design features of the i-Matter intervention to create a theoretically grounded technology solution (Table 1) [31-33].

Table 1. Application of capability, opportunity, and motivation model of behavior theoretical constructs to i-Matter design features.

i-Matter[a] components	COM-B[b] constructs	Mechanisms of action	Design features
PRO[c] assessments	• Capability (comprehension) • Motivation (habit formation)	• Rating PROs on a scale helps patients make more realistic assessments of their symptoms and behaviors • Daily ratings increase patients' awareness of their condition on their quality of life and daily functioning • Tracking PROs and observing patterns provides patients with reasons to adhere to their self-management regimen	• Daily text message questions • Asks patients to complete small doable actions at optimal times
Feedback messages (insights)	• Motivation (perceptions of illness and emotional response)	• Enables patients to identify changes in PROs that previously went undetected • Encourages self-reflection of PRO ratings and their impact on behavior	Data-driven insights based on PRO ratings, such as: • Correlational: association between PRO responses • Individual: comparisons of PRO responses across weeks
Motivational messages	• Motivation • Opportunity (perceived support)	• Uses text messages to maintain high levels of engagement in the program	Text messages that encourage journaling, such as: • Response-based: weekly supportive messages based on PRO responses • Activity-based: weekly messages based on response rates to the messages • Completion-based: messages based on patient duration in the study
Personalized reports	• Opportunity (patient-provider relationship) • Capability (comprehension and ability to plan)	• Facilitates informed discussions with provider • Provides provider with succinct and timely data on patient PROs • Motivates patients through the gradual completion of the personalized report, with landscape changes every 4 weeks • Enables patients to understand and identify patterns in their PROs and to develop behavioral changes to better manage PROs	• Patient reports will include pattern messaging, PRO data visualizations, reflective questions, and tips plus an area for notes • Monthly PRO patterns integrated into EHR[d], available during and between visits

[a]i-Matter: investigating an mHealth texting tool for embedding patient-reported data into diabetes management.

[b]COM-B: capability, opportunity, and motivation model of behavior.

[c]PRO: patient-reported outcome.

[d]EHR: electronic health record.

Overview of the Study Design

We used the evidence-based user-centered design (UCD) approach to conduct the formative phase of the trial [34-37]. The aims of this phase were to (1) systematically gather and incorporate feedback from patients and providers to develop and refine the i-Matter intervention and (2) optimize the design of the personalized report for patients and providers [38,39]. The formative phase consisted of 3 steps: (1) focus groups and semistructured interviews to *adapt* i-Matter to diverse patient and provider needs, including those of Spanish-speaking patients; (2) a design workshop to understand a *day in the life* of patients with T2D and provider workflow processes to *integrate* i-Matter into clinical practice; and (3) user testing to *evaluate* the usability and acceptability of i-Matter in patients with T2D and optimize the tool's performance and display of the personalized reports. The primary outcome of this phase was a refined, integrated, and well-tested mobile PRO system for T2D whose efficacy will be evaluated in the clinical trial.

Study Setting and Population

This study was conducted in a network of primary care practices of New York University Langone Health (NYULH). The practices comprised >1500 ambulatory physicians, nurse practitioners, and physician assistants who care for >800,000 patients in 235 facilities in New York City's 5 boroughs: Long Island, New Jersey, Westchester County, Putnam County, and Dutchess County. The participating sites include academic practices, many community-based practices, and federally qualified health centers, serving an ethnically diverse population.

All primary care practice sites share a single, integrated electronic health record (EHR; Epic).

The target enrollment for the formative phase was 36 patients and 14 providers. To be eligible, patients must (1) have had a diagnosis of T2D for ≥6 months; (2) have had uncontrolled T2D, defined as HbA_{1c} >7%, documented in the EHR at least twice in the past year; (3) be fluent in English or Spanish; (4) be willing to send and receive text messages; and (5) be aged ≥18 years. Patients were excluded if they (1) refused or were unable to provide informed consent; (2) had acute renal failure, end-stage renal disease (ESRD) or evidence of dialysis, renal transplantation, or other ESRD-related services documented in the EHR; (3) were participating in another T2D study; (4) had significant psychiatric comorbidity or reports of substance abuse (as documented in the EHR); (5) were pregnant or planning to become pregnant within 12 months; or (6) planned to discontinue care at the practice within the next 12 months. Providers were eligible if they (1) were a primary care provider (ie, medical doctor, nurse practitioner) practicing at the participating practices and (2) provided care to at least five patients with T2D. The NYULH Institutional Review Board approved this study.

Recruitment

We used 2 approaches to recruit patients and providers into the formative phase. First, potentially eligible patients were identified through a review of the EHR using the diagnosis-related group codes indicating the presence of T2D and receiving care from a primary care provider at one of NYULH practices. After retrieving a list of potentially eligible patients, research assistants (RAs) reviewed patients' EHR to determine if the patient met the eligibility criteria. Patients that met these criteria were contacted via telephone to confirm eligibility. During the telephone call, the RA gave patients a description of the study, including their role as participants in the study. If the patient remained interested in participating, they were given the option to either complete the focus groups or interviews in-person in a private room or via a remote session using the secure Webex conferencing platform. Providers were sent emails from the study principal investigators inviting them to provide feedback on the development of an interactive mHealth tool that could help enable patients with T2D to take a more active role in their diabetes management. All patients and providers provided written informed consent before participation in the study.

Development of the i-Matter Intervention

Table 2 provides an overview of the UCD process used to develop the PROs for i-Matter. A description of each step is also included below.

Table 2. Evidence-based user-centered design process for the development of patient-reported outcome text messages.

Steps	Methods	Outputs
1. Adapt	Patient focus groups and provider interviews	• Thematic analysis of patient and provider needs, preferences, and barriers and facilitators of tracking PROs[a] • Review of existing validated PRO questionnaires by study team based on thematic analysis • Initial list of PROs for i-Matter[b] comprised individual items extracted from existing questionnaires • Reduced list of PROs based on importance rankings from focus group participants
2. Integrate	Design workshop • Workflow mapping • Problem or opportunity analysis • Presentation of PRO list from step 1 • EHR[c] integration	• Refined list of PROs • Clinic workflow or patient journey maps • Essential features of i-Matter system • i-Matter[b] prototype: PRO text messages and personalized report
3. Evaluate	• 2 rounds of patient user testing sessions • Provider interviews	• Finalized PROs and personalized report • Fully functional i-Matter intervention

[a]PRO: patient-reported outcome.

[b]i-Matter: investigating an mHealth texting tool for embedding patient-reported data into diabetes management.

[c]EHR: electronic health record.

Step 1. Focus Groups and Interviews to Adapt i-Matter to Diverse T2D Patients and Primary Care Physician Needs

The goal of the focus groups was to select the PROs that would be integrated into the i-Matter intervention as it relates to patients' experiences living with T2D. A trained moderator used a semistructured guide to explore (1) patients' daily experiences living with T2D, (2) the barriers and facilitators to achieving their diabetes-specific goals, (3) descriptions of patient-provider conversations about T2D and goals for HbA_{1c}, and (4) interest in sharing PRO data with their provider to support treatment of T2D. A trained bilingual moderator also conducted separate focus groups with Spanish-speaking patients to inform the cultural and linguistic adaptation of i-Matter. Before starting each focus group, all patients completed questions about their

comfort with using technology and their interest in using mHealth tools for diabetes care.

A trained moderator conducted semistructured individual interviews with primary care providers at the participating practices. The goal of the interviews was to elicit provider feedback on the clinical relevance of the PROs discussed in the patient focus groups for the management of T2D. The interview guide also explored (1) providers' level of comfort with PRO data, (2) descriptions of patient-provider discussions about diabetes management, and (3) other important PROs not identified in the patient groups.

Results from the thematic analysis of the focus groups and interviews were used to develop a preliminary list of PROs for inclusion in i-Matter [40]. The PROs were individual items taken from existing validated PRO measures that assess the impact of T2D and its treatments on patients' psychosocial, physical, and behavioral functioning (eg, emotional distress, treatment and disease burden, adherence to medications, and lifestyle behaviors) [41]. The measures included the Problem Areas in Diabetes Questionnaire, Diabetes Treatment Satisfaction Questionnaire, Treatment Related Impact Measure-Diabetes, Audit of Diabetes-Dependent Quality of Life, Diabetes Impact Management Scale, and Diabetes Distress Screening Scale [42-47]. General items from the National Institutes of Health Patient-Reported Outcome Measurement Information System Global 10 measure, which assesses patients' physical, social, and emotional functioning [48], were also included on the list of candidate PROs.

The study team then recontacted patients from the focus groups to get their feedback on the candidate list and have them rank the perceived importance of each PRO for management of T2D on a 1 (least important) to 6 (most important) scale. The study team used patients' ratings in concert with the thematic analysis to narrow the list of PROs that would be presented to participants in the design workshop.

Step 2. Integrate i-Matter Into Provider Workflows and Patient Daily Lives

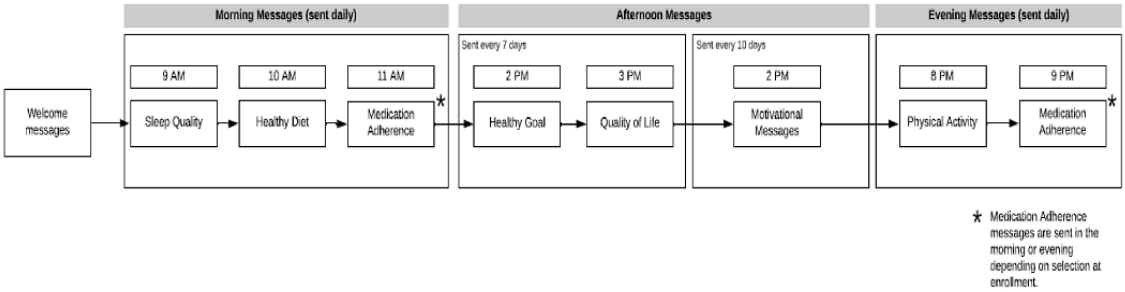
The design workshop comprised patients, providers, academic researchers with expertise in T2D and PROs; the digital health company Rip Road; and staff from the NYULH Medical Center Information Technology (MCIT) department. The design workshop used a UCD protocol adapted from the Agency for Healthcare Research and Quality [49] that sequentially led the group through a variety of activities (eg, story mapping and workflow or patient journey analysis) designed to further refine the i-Matter PRO content, stimulate ideas for the content and layout of the personalized report, discuss ideal workflow integration, and identify potential problems and opportunities to improve i-Matter for patients and providers.

Following steps 1 and 2, the study team collaborated with Rip Road to develop a prototype of i-Matter (ie, the beta version of the text message program and personalized report).

We wrote 2 variations of each PRO question to evaluate the wording and response formats that would yield the highest patient response rates and data quality. On the basis of our previous experiences and best practices for data collection via text message [50], all PRO questions were written so they require short, simple answer choices (eg, 0-10 rating or yes or no response), thereby minimizing the likelihood of missing and/or unanalyzable data that is common with open-ended (free text) response options. We also created 2 versions of the personalized report: a 1-month view and a 3-month view. All text messages and report content were translated to Spanish before user testing.

In addition to prototype development, we created decision rules that would drive the delivery of the text messages and report content. The rules, which were iteratively refined throughout the formative phase, outline the timing and order of the messages, the duration of time patients had to respond to each message (ie, response window), and the conditions that triggered specific motivational text messages and individualized insights displayed on the personalized report (Figure 1). As shown in Table 1, patients receive 3 types of motivational text messages over the course of the study: (1) response-based, (2) activity-based, and (3) completion-based. The personalized report displays 2 types of insights: (1) correlational, which compares associations between 2 different PROs, and (2) individual, which identify trends in patients' responses to the PROs over the past month (see Multimedia Appendix 1 for example messages).

Figure 1. i-Matter study flow.



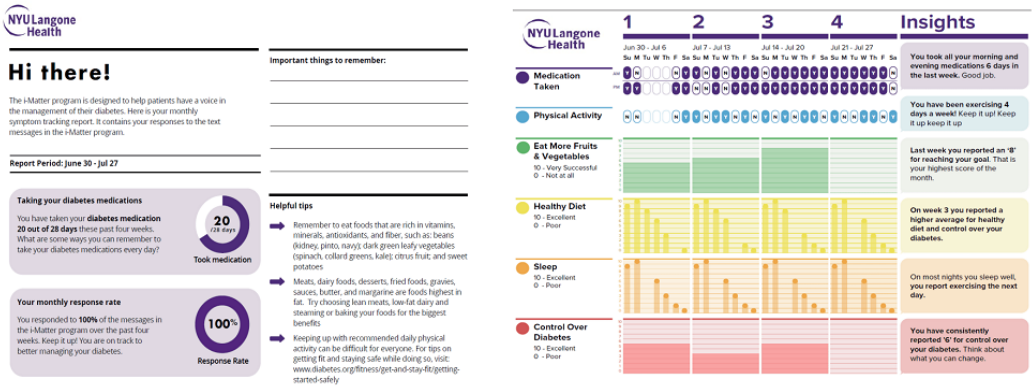
Step 3. User Testing of the i-Matter Prototype

User testing was conducted in a purposive sample of patients drawn from the focus groups and those who were naive to the tool (ie, did not participate in previous steps). A rapid-cycle design (ie, design-test-modify-retest in short intervals of time) was used to allow for iterative refinement of the i-Matter prototype between each user test. Patients participated in the user testing sessions for 2 weeks, during which time they

received and responded to the PRO questions sent via text message. At the end of the 2-week period, the study team sent patients a copy of the personalized report (Figure 2) and conducted an interview about their experiences. The interviews used a combination of think-aloud techniques and semistructured questions to collect patient feedback on the perceived ease of receiving and responding to the PRO questions; the clarity, timing, and frequency of the messages; and the perceived usefulness of systematically tracking the selected PROs for diabetes self-management. Patients also provided feedback on the personalized report, including the clarity of the data visualizations and data-driven feedback messages (herein called insights), content and layout of the report, and utility of the report for diabetes self-management. In addition to the interview, patients responded to questions derived from the TAM version 3 (TAM3) survey.

We also conducted interviews with providers to elicit their feedback on preferences for visual displays and placement of the report in the EHR and perceived barriers and facilitators to viewing the reports in clinical practice. The primary outcome of this step was the fully functional i-Matter intervention for testing in the RCT.

Figure 2. Example of a final personalized report after two rounds of user testing.



Measures

Participant demographics: a self-report instrument was used to collect patient sociodemographic data including gender, race or ethnicity, age, annual household income, education level, marital status, employment status, and current HbA_{1c} level.

Patient use of mobile technology: before the focus groups, patients completed a survey created for this study that assessed the frequency of mobile phone use, capabilities of their mobile phones (eg, Wi-Fi connection, Bluetooth, and mobile data plan), the most commonly used functions (eg, text messaging, phone calls, email, and apps), comfort with using their mobile phone to manage T2D, interest in enrolling in a text messaging diabetes program, and challenges to using their mobile phone for diabetes self-management.

Use behavior: these data were extracted from the i-Matter platform at the end of the user testing sessions and included the following metrics (described in the analysis section): number of mobile phone inputs, time-on-task, task success, number of missed responses to PRO questions, and number of responses by patients outside the response window.

TAM3 survey: following the 2-week user testing period, patients completed questions derived from the well-validated TAM3 survey that assessed the perceived ease of use, usefulness, and quality of i-Matter; the likelihood of using i-Matter in the future and recommending it to others (ie, behavioral intention); and perceived benefits of discussing i-Matter data with providers to help manage their diabetes (ie, communication). The internal consistency of this scale ranged from 0.86 (communication) to 0.94 (perceived usefulness).

Statistical Analysis

Sample size estimates for the formative phase were based on best practices for maximizing the information power of qualitative research, which recommends beginning with 6 to 8 participants per qualitative method and adding to the sample, as needed [51]. As with previous studies, user testing sessions were scheduled until data saturation was reached [36]. Our previous studies suggested that we would need 2 to 3 cycles of user testing to reach saturation [36].

Focus groups and interviews were audiotaped, translated where necessary, and transcribed verbatim. Both data sources were analyzed using the constant comparative method, in which text was categorized into themes with the use of codes developed iteratively to reflect the data [52,53]. The coding scheme was developed by the study investigators to focus on key themes identified both a priori (eg, from the interview protocols) and those that emerged during the interviews or focus group discussions. A trained qualitative researcher coded the transcripts independently, after which the research team met to discuss the coding and resolve any discrepancies.

After each round of user testing, the study team employed the best practices for instant data analysis of usability data for each PRO [54,55]. Task success was calculated as the percentage of PRO questions that were answered correctly without errors. Time-on-task was calculated as the average amount of time in minutes and seconds that patients took to respond to each PRO question. Mobile phone inputs were calculated as raw counts of PRO questions sent by the i-Matter platform and the number of responses received by patients. Missing data were calculated as the percentage of PRO questions that had no response by patients, and late responses were calculated as the percentage

of messages sent by patients that was outside a 1-hour window. In addition, we calculated frequencies for the TAM3 survey questions.

Following the analysis of use data, the research team categorized each issue with usability as either critical (abandon or remove), severe (significant delay or frustration in task completion requiring revision), or cosmetic (minor issue). Each of these issues were mapped onto the interview transcripts and survey responses to provide specific and detailed recommendations for refining i-Matter before proceeding to the next testing session.

Results

Step 1. Patient Focus Groups, Provider Interviews, and Ranking of Candidate PROs

We invited 55 patients with T2D (22 male and 33 female) to participate in the focus groups, of which 35 (64%) declined participation, leaving 20 potential participants. Reasons for declining participation included being too busy, limitations owing to other comorbid conditions, personal or family constraints, and lack of interest in participating in the research. Of the 20 people who agreed to participate, 12 (60%) attended one of the focus groups, 1 did not attend owing to a scheduling conflict with work, and 7 stopped responding to the RA's outreach calls. We held 4 focus groups: 2 for English-speaking patients (n=6) and 2 for Spanish-speaking patients (n=6). Table 3 describes the sociodemographic characteristics of the focus group participants and their comfort with technology.

Analysis of the focus groups identified 4 core themes: (1) patients felt as though their lives were controlled by their blood sugar values; (2) patients' greatest fear of having T2D were vision loss, kidney failure, or risk of amputation, and avoiding these consequences served as motivators for behavior change; (3) important goals for patients were being in control of their T2D, feeling well, living a long healthy life, and eventually not needing medications for T2D (owing to concerns about the negative long-term effects); and (4) forgetfulness, poor dietary adherence, physical inactivity, tiredness or fatigue, and poor emotional health were viewed as major barriers to keeping blood sugar in control. Patients in the Spanish-speaking focus groups also spoke about God being an important source of strength and motivation to improve their health.

We conducted 6 provider interviews (3/6, 50% female; 4 primary care providers, 1 endocrinologist, and 1 general surgeon and weight management specialist). Analysis of the interviews identified the central theme that providers want PRO data that are specific and actionable and can help them focus the clinic visit on what is most important for their T2D patients' care. All providers felt that an asset of a program like i-Matter would be having patients systematically track data such as dietary intake and medication adherence that they cannot reliably assess within the time constraints of a clinic visit. All providers liked the idea of showing correlations between PROs being tracked in i-Matter and clinical data that are already stored in the EHR, such as HbA_{1c} values. Providers varied on the importance of tracking patient functional status, quality of life, and psychosocial health, with two-thirds of the providers commenting that it was central to understanding patients' behaviors, whereas the remaining one-third felt they were *soft symptoms* that may be important for the patient but not for clinical management.

Next, the study team selected individual items from existing PRO measures that best represented themes derived from the focus groups and interviews. This resulted in the selection of items representing 8 categories of PROs: diabetes-related symptoms, quality of life, emotional health (eg, depression, mood, and distress), treatment-related symptoms, treatment satisfaction, diabetes-related functional status, medication adherence, and lifestyle behaviors. Patient ranking of the items further reduced the number of PRO categories to 5: diabetes-related symptoms, quality of life, emotional health, medication adherence, and lifestyle behaviors (Table 4).

Table 3. Sociodemographic characteristics and comfort with technology survey responses among focus group participants (n=12).

Sociodemographic characteristics	Values
Age (years), mean (SD)	62.5 (5.6)
HbA_{1c}[a], mean (SD)	7.95 (0.8)
Female, n (%)	8 (67)
Employed, n (%)	4 (33)
Retired, n (%)	4 (33)
Annual income <US $25,000, n (%)	7 (58)
Hispanic, n (%)	7 (58)
Race, n (%)	
White	5 (42)
Black	3 (25)
Asian	1 (8)
Other	4 (25)
Education, n (%)	
Less than high school	1 (8)
High school degree	4 (3)
Some college	2 (17)
College degree	5 (42)
Technology survey, n (%)	
Currently uses text messaging	7 (58)
Has an unlimited text messaging plan	12 (100)
Always has mobile phone with them	9 (75)
Comfortable downloading apps on their mobile phone	7 (58)
Comfortable receiving and responding to text messages about T2D[b]	8 (67)
Interested in using mobile phone to help keep track of T2D	7 (58)
Challenges to tracking T2D on mobile phone, n (%)	
Cost of receiving messages	2 (17)
Unreliable internet access	1 (8)
Do not use mobile phone regularly	3 (25)
Unsure of benefit	4 (33)
Concerns about privacy and security	2 (17)

[a]HbA_{1c}: hemoglobin A_{1c}.

[b]T2D: type 2 diabetes.

Table 4. Patient ranking of perceived importance of initial list of candidate patient-reported outcomes.

PRO[a] categories	PRO statements	Mean importance score, range 1 (low) to 6 (high)
Symptom	Over the past week, did you experience tingling or prickling sensations in hands or feet owing to your diabetes?	1.8
Symptom	Over the past week, did you experience dry mouth owing to your diabetes?	3.0
Symptom	Over the past week, how often were you bothered by blurred vision?	3.6
Symptom	Over the past week, how would you rate your level of fatigue owing to your diabetes?	4.3
Symptom	Over the past week, how often did you experience increased thirst and frequent urination owing to your diabetes?	4.1
Emotional health	Over the past week, how often were you been bothered by emotional problems such as feeling anxious, depressed, or irritable owing to your diabetes?	4.2
Emotional health	How often over the past month, did you feel overwhelmed by the demands of living with diabetes?	2.75
Lifestyle behavior	On average, how many days did you participate in at least 30 min of physical activity over the past 7 days?	3.13[b]
Lifestyle behavior	Over the past week, how often did you eat (favorite unhealthy food)?	4.6
Lifestyle behavior	Over the past week, how often did you eat (favorite healthy food)?	2.25
Lifestyle behavior	How would you rate your sleep quality over the past 7 days?	4.8
Quality of life	I was able to keep my blood sugar in good control today.	4.6
Medication adherence	Over the past week, how often were you able to take your diabetes medication on time?	4.9
Medication adherence	How many days did you miss or skip at least one dose of your diabetes medication over the past 7 days?	2.9

[a]PRO: patient-reported outcome.

[b]Despite the lower score, physical activity was added as a PRO after review of transcripts and notes from patient and provider interviews.

Step 2. Design Workshop

A total of 17 stakeholders participated in the design workshop. The following themes emerged when the group discussed the candidate list of PROs: (1) PROs should show variability in patients' responses over time and be actionable by both patients and providers, (2) PROs should be taken from validated questionnaires to increase provider confidence in the data patients report and be comparable with HbA_{1c} levels, (3) choosing fewer PROs would help increase patient response rates and reduce the burden on providers to view the data, (4) tracking PROs that focus on adherence to lifestyle behaviors were most appealing to patients, and (5) PRO content should be general (eg, "how are you feeling today?") as opposed to diabetes-specific ("how much does diabetes interfere with your life?"). The group reasoned that questions that were too specific may not be relevant to all patients and could lead to disengagement or missing data. Alternatively, a broader question could be used as a way to show care for patients' overall well-being and as an entry point for more diabetes-specific questions that may uncover new or different concerns the provider should be aware of.

On the basis of these discussions, the group generated several ideas for potential visualizations of PRO data. These included defining a threshold that patients' data can fall above or below and depicting it in a way that makes it easily detectible and actionable, using bar graphs to show directionality, including icons or coloring schemes in addition to PRO labels that enhance the readability of the report, and including summary data in percentages or raw numbers to show the patient's progress over time.

Applying the findings from steps 1 and 2, the study team reduced the number of PRO categories to 4. Diabetes-related fatigue (symptom category) was removed from the list because providers viewed it as too nebulous and not actionable, whereas patients felt sleep quality was a more meaningful PRO for their diabetes management. In addition, physical activity was added to the lifestyle category because many patients felt that physical inactivity was a major contributor to weight gain and poor diabetes control.

Step 3. User Testing of the i-Matter Prototype

Patients: Text Messages

We completed 2 rounds of user testing with patients: 7 patients completed the first round of testing (1 Spanish-speaking), and 3 patients completed the second round. Table 5 presents the results of the use behavior data for both rounds of user testing. The i-Matter platform sent 325 messages, and patients sent 256 responses (78.7%). The most common reason for invalid messages was the response being sent in the wrong format (eg,

sending free-text responses instead of a numerical response). The most common reasons for missing messages included problems with message filtering by the mobile carrier (which has been resolved by changing to short code messages), being busy during the response window, and not having their phone during the time the messages were sent. For the Spanish-speaking patient, the average response rate was 67.3 min (range 0-209.1 min). Overall, 59.7% (153/256) of the messages were answered within an hour, of which all (256/256, 100%) were answered within 1 min.

In the second round of user testing, the message protocol was modified to address the suboptimal percentage of missed responses. For example, to address the wide range of response times seen in the first round of testing (range 0-661.6 min), we restricted the patients' ability to respond to the morning and evening PRO questions to a 1-hour window (based on the median response time). Overall, the i-Matter platform sent 222 messages and received 188 responses (84.6%) from patients. The most frequently missed message was sleep quality (77/188, 40.9% of missed messages). The most common reason for an invalid message was the patient responding to a question outside the 1-hour response window.

In qualitative interviews, patients in both rounds of user testing described the program as easy to use, not intrusive to their daily life, and helpful for managing their T2D. Similar findings were seen in the TAM3 survey responses (Table 6). Patients also liked the consistency in message timing because it helped them build a habit to respond ("it becomes second nature"). Several patients commented that they felt as though a person was sending the messages to check up on them. Patients also felt that the number of messages sent was adequate, with 2 people commenting, "No number is too many because they want to get better." There were no differences in qualitative feedback or TAM3 responses between the English- and Spanish-speaking patients.

Table 5. Patient text messaging use behavior during user testing.

User behaviors	User testing round 1 (n=7)	User testing round 2 (n=3)
Time-on-task	44 min (range 0-661.6)	20 min (range 0.08-30)
Task success (messages), n (%)	232 (90.6)	175 (93.1)
Missed responses, n (%)	100 (39.2)	28 (15.0)
Late responses, n (%)	49 (19.3)	14 (7.5)
Invalid responses, n (%)	24 (9.4)	13 (6.9)

Table 6. Response to technology acceptance model version 3 survey questions.

Questions	Proportion of patients agreeing with statement, n (%)
PRO[a] (n=7)	
I would definitely use the i-Matter program in the future	5 (71)
The PRO questions are very helpful for managing T2D[b]	6 (86)
Receiving and responding to PRO questions was easy	7 (100)
I responded to the PRO questions all the time	5 (71)
I would recommend i-Matter to friends and family	7 (100)
My provider would be more effective managing T2D with my PRO data	5 (71)
Overall, the i-Matter program is great or excellent	6 (86)
Personalized report (n=9)	
I would definitely use the personalized report in the future	8 (89)
The personalized report is very helpful for managing T2D	7 (78)
The personalized report is easy to use	5 (56)
I would share the personalized report with friends or family	5 (56)
Showing my provider the personalized journal would help make clinic visits more effective	7 (78)
The charts and images are great	6 (67)
Overall, the personalized report is great	6 (67)

[a]PRO: patient-reported outcome.

[b]T2D: type 2 diabetes.

As shown in Table 7, patients recommended improvements to the wording and timing of several of the PROs (eg, sending the sleep message at 9 AM rather than 7 AM), which is reflective of the use data. Patients also recommended including

motivational messages to help sustain engagement with the program. After examining the data, the study team decided to remove the emotional health PRO (ie, labeled as critical). This was owing to the lack of variability in the use data (206/256, 80.6% of responses were 0-1 on a 10-point scale) and feedback from patients in the interviews that the PRO was not relevant to the management of their T2D.

Table 7. Recommended changes to patient-reported outcome text messages from user testing sessions.

PRO[a] categories and original messages	Original timing	Revised message	Revised timing
Medication adherence			
Have you taken all of your diabetes medications as prescribed today?	Daily at 7 AM	Retain as is	Allow patients to decide if they want the message in the AM or PM, or both (11 AM and 9 PM)
Lifestyle			
Reply with 1-4 to track ONE healthy living goal: • 1=Lose weight • 2=Eat more fruits and vegetables • 3=Eat less sweets and carbohydrates • 4=Have better portion control	Daily at 7 AM	Retain for all patients. Separate less carbs and sweets to 2 separate goals	Changes so patients choose healthy goal at baseline visit (with option to change goal every 3 months)
How successful were you in achieving your goal to (custom text healthy goal) yesterday? Response: 0 (not at all) to 10 (very successful)	Daily at 2 PM	How successful were you in achieving your goal to (custom text healthy goal) this *past week?*[b]	Change timing to weekly at 2 PM
In general, how healthy your overall diet was today?	Daily at 7 AM	Retain message but change timing to assessing overall diet *yesterday*[b]	Change to daily at 10 AM
Rate your sleep quality last night. Think how easily you fell asleep, how often you woke up and if sleep was refreshing. Response: 0 (poor) to 10 (excellent)	Daily at 7 AM	Reply with the number that best describes how well you slept last night	Change to daily at 9 AM
How many days in the past week did you do any physical activities like brisk walking where you breathed harder than normal?	Weekly at noon	Other than your regular job, did you do any physical activities like brisk walking for at least 30 min today?	Change to daily at 8 PM
Diabetes quality of life			
Reply with the number that best describes how much control you felt you had over your diabetes over the past 2 weeks	Biweekly at noon	Retain as is change timing to weekly	Change to weekly at 3 PM
Emotional health			
Reply with the number that best describes how irritable or moody you felt today owing to your diabetes	Daily at 7 PM	Remove	N/A[c]

[a]PRO: patient-reported outcome.

[b]Text in italics show the changes made to the PRO timing across user testing sessions.

[c]N/A: not applicable.

Patients: Personalized Reports

A total of 9 patients provided feedback on the 1-month and 3-month versions of the personalized report: 4 of these patients participated in the user testing (of which 2 were recruited from the focus groups), whereas 5 were naive to the program. Overall, the majority of patients (8/9, 89%) felt the report was easy to read, eye-catching, and comprehensive. There was a strong preference for the 1-month version of the report owing to the larger font size. Patients also felt that receiving the report more frequently would help sustain motivation. Patients preferred layouts that used darker fonts and lighter background colors to help make the text easier to read. All patients viewed the color-coded schema favorably because it helped draw attention to the most important aspects of the report and made the data easy to interpret.

Several patients had difficulty reading the bar graphs of PROs that were collected biweekly (eg, quality of life) and recommended changing the items to weekly measures to be consistent with other PROs. Finally, email was the preferred delivery method, and most patients would share the report with their family and friends (Table 6).

Benefits of using the personalized report for diabetes self-management included being able to analyze how well one is adhering to recommended diabetes behaviors (“being honest with yourself”), providing visual cues to take responsibility for one’s health (“a visual reminder of things I need to do but don’t do and how I can be better”), and providing support to stay on

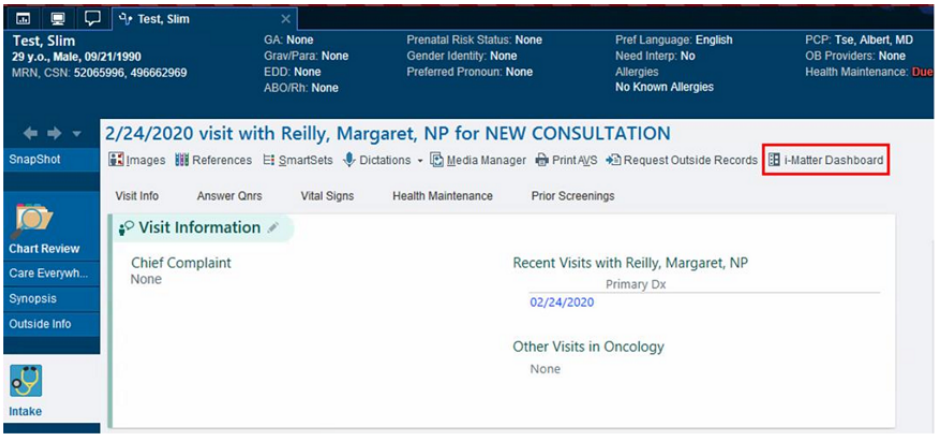
track to be successful with diabetes. In the first round of user testing, 3 critiques of the report included being of greater use to providers than patients, concerns about confidentiality, and being too limited because it did not include tips on how patients can improve unhealthy behaviors. From this feedback, we incorporated motivational text messages and insights into the i-Matter prototype and created a study website that included additional resources. Patients in the second round of testing regarded the inclusion of insight messages as a source of motivation to change their behaviors and to continue responding to text messages to monitor changes in behavior over time.

Provider Feedback

Overall, all (n=6) providers thought the report was a good tool to help patients manage their T2D. Similar to patients, they felt that the insight messages were helpful for interpreting the data and prompting behavioral changes. When reviewing the PRO content, providers cautioned that before starting the program, patients would need to be educated on the recommended dietary and physical activity guidelines for diabetic patients and the medications they are currently taking for their T2D to ensure they are reliably answering the questions. On the basis of this feedback, at the baseline visit, trained study staff provide a brief overview of evidence-based guidelines for healthy eating and physical activity for T2D using low-literacy and language-congruent patient education handouts from the American Diabetes Association and review the patient's current diabetes medication regimen.

To integrate the report into clinical practice, providers preferred having a separate tab built into the EHR, which included a summary of the personalized report and highlighted key trends in patients' PRO data over the past 3 months. All providers found value in discussing the report with patients during the clinic visit because the data complemented the questions that they had already asked about diabetes self-management. Finally, although they found value in the longitudinal trends displayed in the graphs, owing to time constraints, they felt that patients should bring up anything important that stood out in the detailed view. On the basis of this feedback, the study team is working in collaboration with NYULH MCIT to integrate the personalized report into Epic. This includes the development of security protocols that will link patients' encrypted research ID to their medical record number and integrate the report image into an Epic web integration record. Web integration records are used to visually integrate external apps with Epic. Providers will be able to access the i-Matter report via a button located within the patient's chart at the top of the Office Visit toolbar (Figure 3). The button will only be visible for patients randomized to the intervention arm.

Figure 3. Screenshot of i-Matter Epic integration.



Discussion

Strengths

Although achieving glycemic control is of clinical importance, it is the daily experiences of living with T2D that drive patients' decisions to adhere to treatment recommendations and become engaged in their care [56]. Even with the most efficacious treatments, failure to incorporate patients' perspective of their disease into clinical decision making will make achieving the outcomes desired by patients and providers unattainable. The i-Matter trial will assess whether a theory-driven mobile PRO system that incorporates a set of PROs that are meaningful to both patients and providers can lead to reductions in HbA_{1c} and improvements in patient adherence to self-care behaviors. Unlike existing programs, i-Matter is designed to collect real-time PRO data in the form of data-driven feedback, motivational messages, and dynamic data visualizations that are displayed in personalized reports for patients and providers.

This paper describes the design and refinement of i-Matter through an iterative user-centered approach that actively involved patients and providers throughout the process. Active involvement of end users in the development of the intervention can help to address the difficulties with protocol compliance, lack of clinical integration in the EHR, and provider skepticism about the utility of PROs in practice, which are hallmarks of previous trials, thus increasing the likelihood of developing a sustainable approach [57]. Findings from our formative phase resulted in several insights regarding issues with the design, usability, and workflow of i-Matter, which led to key changes in the content and delivery of the text messages and personalized report and the technical infrastructure to support the integration of i-Matter into the EHR to improve patient and provider acceptability. In addition to evaluating the clinical benefit of i-Matter, the RCT will provide much needed evidence on the

conditions under which mHealth interventions *work* in primary care settings and in patients' daily lives and the organizational, individual, and technical factors that are required to support their use.

Limitations

Although there are many strengths of our intervention approach, we note limitations that can be considered for future research. First, although our intervention is designed to target patients with T2D, it is more common for patients to have 2 or more chronic diseases (ie, multimorbidity) than 1 disease in isolation (89.3% vs 8.5%, respectively) [58]. In fact, recent research demonstrates the negative impact of multimorbidity on PROs such as quality life, psychosocial health, self-efficacy, physical function, and self-management behaviors (eg, physical activity and medication adherence) [59]. Thus, future research should examine whether adapting i-Matter for a multimorbid population would improve the integration and coordination of patient and provider management of co-occurring chronic diseases rather than using a single disease focus that can cause inefficiencies and fragmentation in care. Second, we did not perform psychometric testing of the final PROs before they were deployed in our intervention. We will use data collected in this study to assess the psychometric properties of our PRO questions and test their validity in subsequent research.

Finally, 2 (out of the 6) providers interviewed during the development of i-Matter indicated that they found less value in PROs that were not immediately actionable in primary care practice (eg, depression and quality of life). A key strength of the i-Matter study is the full EHR integration of the PROs with the health care team. Many previous PRO initiatives share the patient PRO data back with the providers in a workflow disruptive manner—asking providers to change their normal activities and make a special effort to review the PRO data. i-Matter overcomes these challenges by delivering the patient PRO data directly into the patient's chart in the EHR—presented as just another commonly viewed data visualization by the provider such as patients' lab and test results. Thus, our intervention will test the hypothesis that if actionable diabetes PRO data are delivered in the right context, it will influence patient-provider interactions. Early adopters of our intervention will also help to provide important data on the potential effectiveness and (time) efficiency of using PROs in clinical care. Sharing the outcomes of this work could provide providers who are hesitant to adopt such innovations with much needed information about the benefits of using these tools.

Acknowledgments

The authors would like to thank Aditya Verma and Sara Chokshi, DrPH, for their assistance with this project. This work was supported by a grant from Merck & Co, Inc, (principal investigator: AS) and the Agency for Healthcare Research and Quality, R01HS026522 (principal investigators: AS and DM).

Authors' Contributions

AS analyzed, interpreted data, and drafted the manuscript. JC, LP, KL and MR acquired data and critically reviewed the manuscript. CJ and JG oversaw the EHR integration and critically reviewed the manuscript. MP, SP, and EL developed the mHealth platform and contributed to all critical revisions of the manuscript. DM interpreted data and critically reviewed the manuscript.

Conflicts of Interest

DM, JC, LP, MR, KL, CJ, and JG have no competing interests or financial disclosures to declare. AS is a consultant for Rip Road, Inc. MP, SP, and EL were paid as consultants to develop the mHealth intervention for this project.

Multimedia Appendix 1
Example motivational and feedback messages.
[DOCX File , 16 KB - formative_v4i8e18554_app1.docx]

References

1. Fox CS, Golden SH, Anderson C, Bray GA, Burke LE, de Boer IH, American Heart Association Diabetes Committee of the Council on Lifestyle and Cardiometabolic Health, Council on Clinical Cardiology, Council on Cardiovascular and Stroke Nursing, Council on Cardiovascular SurgeryAnesthesia, Council on Quality of Care and Outcomes Research, American Diabetes Association. Update on prevention of cardiovascular disease in adults with type 2 diabetes mellitus in light of recent evidence: a scientific statement from the American heart association and the American diabetes association. Diabetes Care 2015 Sep;38(9):1777-1803 [FREE Full text] [doi: 10.2337/dci15-0012] [Medline: 26246459]
2. Geiss LS, Wang J, Cheng YJ, Thompson TJ, Barker L, Li Y, et al. Prevalence and incidence trends for diagnosed diabetes among adults aged 20 to 79 years, United States, 1980-2012. J Am Med Assoc 2014 Sep 24;312(12):1218-1226. [doi: 10.1001/jama.2014.11494] [Medline: 25247518]
3. National Diabetes Statistics Report: Estimates of Diabetes and Its Burden in the United States. Centers for Disease Control and Prevention. 2014. URL: https://www.cdc.gov/diabetes/data/statistics-report/index.html?CDC_AA_refVal=https%3A%2F%2Fwww.cdc.gov%2Fdiabetes%2Fdata%2Fstatistics%2Fstatistics-report.html [accessed 2020-08-12]

4. Carls G, Huynh J, Tuttle E, Yee J, Edelman SV. Achievement of glycated hemoglobin goals in the US remains unchanged through 2014. Diabetes Ther 2017 Aug;8(4):863-873 [FREE Full text] [doi: 10.1007/s13300-017-0280-5] [Medline: 28646411]
5. Boltri JM, Okosun IS, Davis-Smith M, Vogel RL. Hemoglobin A1c levels in diagnosed and undiagnosed black, Hispanic, and white persons with diabetes: results from NHANES 1999-2000. Ethn Dis 2005;15(4):562-567. [Medline: 16259477]
6. Ali MK, McKeever Bullard K, Imperatore G, Barker L, Gregg EW, Centers for Disease Control and Prevention (CDC). Characteristics associated with poor glycemic control among adults with self-reported diagnosed diabetes--national health and nutrition examination survey, United States, 2007-2010. MMWR Suppl 2012 Jun 15;61(2):32-37. [Medline: 22695461]
7. Cella D, Yount S, Rothrock N, Gershon R, Cook K, Reeve B, PROMIS Cooperative Group. The patient-reported outcomes measurement information system (PROMIS): progress of an NIH roadmap cooperative group during its first two years. Med Care 2007 May;45(5 Suppl 1):S3-11 [FREE Full text] [doi: 10.1097/01.mlr.0000258615.42478.55] [Medline: 17443116]
8. Snyder CF, Jensen RE, Segal JB, Wu AW. Patient-reported outcomes (PROs): putting the patient perspective in patient-centered outcomes research. Med Care 2013 Aug;51(8 Suppl 3):S73-S79 [FREE Full text] [doi: 10.1097/MLR.0b013e31829b1d84] [Medline: 23774513]
9. Speight J, Barendse SM. FDA guidance on patient reported outcomes. Br Med J 2010 Jun 21;340:c2921. [doi: 10.1136/bmj.c2921] [Medline: 20566597]
10. Funnell MM, Brown TL, Childs BP, Haas LB, Hosey GM, Jensen B, et al. National standards for diabetes self-management education. Diabetes Care 2007 Jun;30(6):1630-1637. [doi: 10.2337/dc07-9923] [Medline: 17526822]
11. Cho JH, Ha SJ, Kao LR, Megraw TL, Chae C. A novel DNA-binding protein bound to the mitochondrial inner membrane restores the null mutation of mitochondrial histone Abf2p in Saccharomyces cerevisiae. Mol Cell Biol 1998 Oct;18(10):5712-5723 [FREE Full text] [doi: 10.1128/mcb.18.10.5712] [Medline: 9742088]
12. Deutsch S, Smith L, Gage B, Kelleher C, Garfinkel D. Patient-Reported Outcomes in Performance Measurement. New York, USA: RTI Press; 2013.
13. Barsdorf A, Rubinstein E, Jaksa A. Patient-reported outcomes (Pros) in diabetes clinical trials. Value Health 2013 May;16(3):A168-A169. [doi: 10.1016/j.jval.2013.03.842]
14. Ridgeway JL, Beebe TJ, Chute CG, Eton DT, Hart LA, Frost MH, et al. A brief patient-reported outcomes quality of life (PROQOL) instrument to improve patient care. PLoS Med 2013 Nov;10(11):e1001548 [FREE Full text] [doi: 10.1371/journal.pmed.1001548] [Medline: 24265598]
15. Marshall S, Haywood K, Fitzpatrick R. Impact of patient-reported outcome measures on routine practice: a structured review. J Eval Clin Pract 2006 Oct;12(5):559-568. [doi: 10.1111/j.1365-2753.2006.00650.x] [Medline: 16987118]
16. Vieta A, Badia X, Sacristán JA. A systematic review of patient-reported and economic outcomes: value to stakeholders in the decision-making process in patients with type 2 diabetes mellitus. Clin Ther 2011 Sep;33(9):1225-1245. [doi: 10.1016/j.clinthera.2011.07.013] [Medline: 21856000]
17. Peters M, Crocker H, Jenkinson C, Doll H, Fitzpatrick R. The routine collection of patient-reported outcome measures (PROMs) for long-term conditions in primary care: a cohort survey. BMJ Open 2014 Feb 21;4(2):e003968 [FREE Full text] [doi: 10.1136/bmjopen-2013-003968] [Medline: 24561495]
18. Valderas JM, Kotzeva A, Espallargues M, Guyatt G, Ferrans CE, Halyard MY, et al. The impact of measuring patient-reported outcomes in clinical practice: a systematic review of the literature. Qual Life Res 2008 Mar;17(2):179-193. [doi: 10.1007/s11136-007-9295-0] [Medline: 18175207]
19. Maddigan SL, Majumdar SR, Guirguis LM, Lewanczuk RZ, Lee TK, Toth EL, et al. Improvements in patient-reported outcomes associated with an intervention to enhance quality of care for rural patients with type 2 diabetes: results of a controlled trial. Diabetes Care 2004 Jun;27(6):1306-1312. [doi: 10.2337/diacare.27.6.1306] [Medline: 15161780]
20. Baron JS, Hirani SP, Newman SP. Investigating the behavioural effects of a mobile-phone based home telehealth intervention in people with insulin-requiring diabetes: results of a randomized controlled trial with patient interviews. J Telemed Telecare 2017 Jun;23(5):503-512. [doi: 10.1177/1357633X16655911] [Medline: 27377790]
21. Årsand E, Frøisland DH, Skrøvseth SO, Chomutare T, Tatara N, Hartvigsen G, et al. Mobile health applications to assist patients with diabetes: lessons learned and design implications. J Diabetes Sci Technol 2012 Sep 1;6(5):1197-1206 [FREE Full text] [doi: 10.1177/193229681200600525] [Medline: 23063047]
22. Kitsiou S, Paré G, Jaana M, Gerber B. Effectiveness of mHealth interventions for patients with diabetes: an overview of systematic reviews. PLoS One 2017;12(3):e0173160 [FREE Full text] [doi: 10.1371/journal.pone.0173160] [Medline: 28249025]
23. Liang X, Wang Q, Yang X, Cao J, Chen J, Mo X, et al. Effect of mobile phone intervention for diabetes on glycaemic control: a meta-analysis. Diabet Med 2011 Apr;28(4):455-463. [doi: 10.1111/j.1464-5491.2010.03180.x] [Medline: 21392066]
24. Holtz B, Lauckner C. Diabetes management via mobile phones: a systematic review. Telemed J E Health 2012 Apr;18(3):175-184. [doi: 10.1089/tmj.2011.0119] [Medline: 22356525]
25. Baron JS, Hirani S, Newman SP. A randomised, controlled trial of the effects of a mobile telehealth intervention on clinical and patient-reported outcomes in people with poorly controlled diabetes. J Telemed Telecare 2017 Feb;23(2):207-216. [doi: 10.1177/1357633X16631628] [Medline: 26880694]

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. 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]
28. Yarbrough AK, Smith TB. Technology acceptance among physicians: a new take on TAM. Med Care Res Rev 2007 Dec;64(6):650-672. [doi: 10.1177/1077558707305942] [Medline: 17717378]
29. Michie S, van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. Implement Sci 2011 Apr 23;6:42 [FREE Full text] [doi: 10.1186/1748-5908-6-42] [Medline: 21513547]
30. Presseau J, Ivers NM, Newham JJ, Knittle K, Danko KJ, Grimshaw JM. Using a behaviour change techniques taxonomy to identify active ingredients within trials of implementation interventions for diabetes care. Implement Sci 2015 Apr 23;10:55 [FREE Full text] [doi: 10.1186/s13012-015-0248-7] [Medline: 25900104]
31. Michie S, Johnston M, Abraham C, Lawton R, Parker D, Walker A. Making psychological theory useful for implementing evidence based practice: a consensus approach. Qual Saf Health Care 2005 Feb;14(1):26-33 [FREE Full text] [doi: 10.1136/qshc.2004.011155] [Medline: 15692000]
32. 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]
33. Pal K, Eastwood SV, Michie S, Farmer AJ, Barnard ML, Peacock R, et al. Computer-based diabetes self-management interventions for adults with type 2 diabetes mellitus. Cochrane Database Syst Rev 2013 Mar 28(3):CD008776 [FREE Full text] [doi: 10.1002/14651858.CD008776.pub2] [Medline: 23543567]
34. Mann DM, Kannry JL, Edonyabo D, Li AC, Arciniega J, Stulman J, et al. Rationale, design, and implementation protocol of an electronic health record integrated clinical prediction rule (iCPR) randomized trial in primary care. Implement Sci 2011 Sep 19;6:109 [FREE Full text] [doi: 10.1186/1748-5908-6-109] [Medline: 21929769]
35. Mann DM, Lin JJ. Increasing efficacy of primary care-based counseling for diabetes prevention: rationale and design of the ADAPT (avoiding diabetes thru action plan targeting) trial. Implement Sci 2012 Jan 23;7:6 [FREE Full text] [doi: 10.1186/1748-5908-7-6] [Medline: 22269066]
36. Li AC, Kannry JL, Kushniruk A, Chrimes D, McGinn TG, Edonyabo D, et al. Integrating usability testing and think-aloud protocol analysis with 'near-live' clinical simulations in evaluating clinical decision support. Int J Med Inform 2012 Nov;81(11):761-772. [doi: 10.1016/j.ijmedinf.2012.02.009] [Medline: 22456088]
37. Chokshi SK, Mann DM. Innovating from within: a process model for user-centered digital development in academic medical centers. JMIR Hum Factors 2018 Dec 19;5(4):e11048 [FREE Full text] [doi: 10.2196/11048] [Medline: 30567688]
38. Arsand E, Demiris G. User-centered methods for designing patient-centric self-help tools. Inform Health Soc Care 2008 Sep;33(3):158-169. [doi: 10.1080/17538150802457562] [Medline: 18850399]
39. Hekler EB, Klasnja P, Riley WT, Buman MP, Huberty J, Rivera DE, et al. Agile science: creating useful products for behavior change in the real world. Transl Behav Med 2016 Jun;6(2):317-328 [FREE Full text] [doi: 10.1007/s13142-016-0395-7] [Medline: 27357001]
40. Reaney M, Black P, Gwaltney C. A systematic method for selecting patient-reported outcome measures in diabetes research. Diabetes Spectr 2014 Nov;27(4):229-232 [FREE Full text] [doi: 10.2337/diaspect.27.4.229] [Medline: 25647044]
41. Meadows KA, Abrams C, Sandbaek A. Adaptation of the diabetes health profile (DHP-1) for use with patients with type 2 diabetes mellitus: psychometric evaluation and cross-cultural comparison. Diabet Med 2000 Aug;17(8):572-580. [doi: 10.1046/j.1464-5491.2000.00322.x] [Medline: 11073178]
42. Brod M, Hammer M, Christensen T, Lessard S, Bushnell DM. Understanding and assessing the impact of treatment in diabetes: the treatment-related impact measures for diabetes and devices (TRIM-diabetes and TRIM-diabetes device). Health Qual Life Outcomes 2009 Sep 9;7:83 [FREE Full text] [doi: 10.1186/1477-7525-7-83] [Medline: 19740444]
43. Chawla A, Saha C, Marrero DG. A novel application of the problem areas in diabetes (PAID) instrument to improve glycemic control and patient satisfaction. Diabetes Educ 2010;36(2):337-344. [doi: 10.1177/0145721709354607] [Medline: 19959745]
44. Costa FA, Guerreiro JP, Duggan C. An audit of diabetes dependent quality of life (ADDQoL) for Portugal: exploring validity and reliability. Pharm Pract (Granada) 2006;4(3):123-128 [FREE Full text] [Medline: 25214898]
45. Hammond GS, Aoki TT. Measurement of health status in diabetic patients. Diabetes impact measurement scales. Diabetes Care 1992 Apr;15(4):469-477. [doi: 10.2337/diacare.15.4.469] [Medline: 1499460]
46. Polonsky WH, Fisher L, Earles J, Dudl RJ, Lees J, Mullan J, et al. Assessing psychosocial distress in diabetes: development of the diabetes distress scale. Diabetes Care 2005 Mar;28(3):626-631. [doi: 10.2337/diacare.28.3.626] [Medline: 15735199]
47. Saisho Y. Use of diabetes treatment satisfaction questionnaire in diabetes care: importance of patient-reported outcomes. Int J Environ Res Public Health 2018 May 9;15(5):947 [FREE Full text] [doi: 10.3390/ijerph15050947] [Medline: 29747423]
48. CMS Quality Measure Development Plan Supporting the Transition to the Merit-Based Incentive Payment System (MIPS) and Alternative Payment Models (APMs). CMS. 2016. URL: https://www.cms.gov/blog/cms-quality-measure-development-plan-supporting-transition-merit-based-incentive-payment-system-mips [accessed 2020-08-12]

49. What is Workflow? Agency for Healthcare Research and Quality. 2016. URL: https://healthit.ahrq.gov/health-it-tools-and-resources/workflow-assessment-health-it-toolkit/workflow [accessed 2016-09-30]
50. Shimoni N, Nippita S, Castaño PM. Best practices for collecting repeated measures data using text messages. BMC Med Res Methodol 2020 Jan 3;20(1):2 [FREE Full text] [doi: 10.1186/s12874-019-0891-9] [Medline: 31900108]
51. Malterud K, Siersma VD, Guassora AD. Sample size in qualitative interview studies: guided by information power. Qual Health Res 2016 Nov;26(13):1753-1760. [doi: 10.1177/1049732315617444] [Medline: 26613970]
52. Kolb S. Grounded theory and the constant comparitive method: valid research strategies for educators. J Emerg Trends Educ Res Policy Stud 2012;3:83-86 [FREE Full text]
53. Strauss A, Corbin J. Grounded theory methodology. In: Senzin N, Lincon YS, editors. The Landscape of Qualitative Research. Thousand Oaks, CA: Sage Publications; 1994.
54. Kjeldskov JS, Stage J. Instant data analysis: conducting usability evaluations in a day. Proc Third Nord Conf Human-Comput Interact 2004:- [FREE Full text] [doi: 10.1145/1028014.1028050]
55. Joe J, Chaudhuri S, Le T, Thompson H, Demiris G. The use of think-aloud and instant data analysis in evaluation research: exemplar and lessons learned. J Biomed Inform 2015 Aug;56:284-291 [FREE Full text] [doi: 10.1016/j.jbi.2015.06.001] [Medline: 26071683]
56. Engström MS, Leksell J, Johansson U, Gudbjörnsdottir S. What is important for you? A qualitative interview study of living with diabetes and experiences of diabetes care to establish a basis for a tailored patient-reported outcome measure for the Swedish national diabetes register. BMJ Open 2016 Mar 24;6(3):e010249 [FREE Full text] [doi: 10.1136/bmjopen-2015-010249] [Medline: 27013595]
57. McCurdie T, Taneva S, Casselman M, Yeung M, McDaniel C, Ho W, et al. mHealth consumer apps: the case for user-centered design. Biomed Instrum Technol 2012;Suppl:49-56. [doi: 10.2345/0899-8205-46.s2.49] [Medline: 23039777]
58. Fortin M, Bravo G, Hudon C, Vanasse A, Lapointe L. Prevalence of multimorbidity among adults seen in family practice. Ann Fam Med 2005;3(3):223-228 [FREE Full text] [doi: 10.1370/afm.272] [Medline: 15928225]
59. Sasseville M, Chouinard M, Fortin M. Patient-reported outcomes in multimorbidity intervention research: a scoping review. Int J Nurs Stud 2018 Jan;77:145-153. [doi: 10.1016/j.ijnurstu.2017.09.016] [Medline: 29080440]

Abbreviations

COM-B: capability, opportunity, and motivation model of behavior
EHR: electronic health record
ESRD: end-stage renal disease
HbA1c: hemoglobin A1c
i-Matter: investigating an mHealth texting tool for embedding patient-reported data into diabetes management
MCIT: Medical Center Information Technology
mHealth: mobile health
NYULH: New York University Langone Health
PRO: patient-reported outcome
RA: research assistant
RCT: randomized controlled trial
T2D: type 2 diabetes
TAM: technology acceptance model
UCD: user-centered design

Edited by G Eysenbach; submitted 04.03.20; peer-reviewed by H Barahimi, T Wieringa; comments to author 12.06.20; revised version received 09.07.20; accepted 26.07.20; published 31.08.20.

Please cite as:
Schoenthaler A, Cruz J, Payano L, Rosado M, Labbe K, Johnson C, Gonzalez J, Patxot M, Patel S, Leven E, Mann D
Investigation of a Mobile Health Texting Tool for Embedding Patient-Reported Data Into Diabetes Management (i-Matter): Development and Usability Study
JMIR Form Res 2020;4(8):e18554
URL: https://formative.jmir.org/2020/8/e18554
doi:10.2196/18554
PMID:32865505

Original Paper

How Health Care Organizations Approach Social Media Measurement: Qualitative Study

Chukwuma Ukoha[1], BSc, MSc, MIS

Centre for Informatics and Applied Optimisation, Federation University Australia, Ballarat, Australia

Corresponding Author:
Chukwuma Ukoha, BSc, MSc, MIS
Centre for Informatics and Applied Optimisation
Federation University Australia
Level 1, T Building
University Drive
Ballarat, VIC 3350
Australia
Phone: 61 353276435
Email: c.ukoha@federation.edu.au

Abstract

Background: Many health care organizations use social media to support a variety of activities. To ensure continuous improvement in social media performance, health care organizations must measure their social media.

Objective: The purpose of this study is to explore how health care organizations approach social media measurement and to elucidate the tools they employ.

Methods: In this exploratory qualitative research, Australian health care organizations that use social media, varying in size and locality, were invited to participate in the study. Data were collected through semistructured interviews, and the transcripts were analyzed using thematic analysis.

Results: The study identified health care organizations' approaches to social media measurement. While some measured their social media frequently, others used infrequent measurements, and a few did not measure theirs at all. Those that measured their social media used one or a combination of the following yardsticks: personal benchmarking, peer benchmarking, and metric benchmarking. The metrics tracked included one or more of the following: reach, engagement, and conversion rates. The tools employed to measure social media were either inbuilt or add-on analytics tools. Although many participants showed great interest in measuring their social media, they still had some unanswered questions.

Conclusions: The lack of a consensus approach to measurement suggests that, unlike other industries, social media measurement in health care settings is at a nascent stage. There is a need to improve knowledge, sophistication, and integration of social media strategy through the application of theoretical and analytical knowledge to help resolve the current challenge of effective social media measurement. This study calls for social media training in health care organizations. Such training must focus on how to use relevant tools and how to measure their use effectively.

(JMIR Form Res 2020;4(8):e18518) doi:10.2196/18518

KEYWORDS

health care organization; social media; measurement; benchmarking; metrics; analytics tools

Introduction

Background

There is a growing use of social media in health care settings, and an increasing number of health care organizations use social media [1-4]. Social media refers to internet-based communication and interactive tools that enable the capture, storage, and presentation of written, audio, and video communication [5]. Social media is the general term for internet-based applications underpinned by the ideological and technological foundations of Web 2.0, which enables and encourages user-generated content [6] such as texts, images, and videos [7]. Social media allows users to create a profile within a bounded system, identify other users with whom they have a connection, and view and access their list of connections and those made by others within the system [8]. The International Medical Informatics Association identified 13

types of social media platforms: social networks, professional networks, thematic networks, microblogs, blogs, Wikis, forums or Listserv, social photo and video-sharing tools, collaborative filtering tools, multiuser virtual environments, social apps and games, integration of social media with health information technologies, and others (eg, FriendFeed) [9]. Social media used in health care settings can be grouped broadly into two categories—general-purpose web-based social networks and online health communities that often serve as discussion forums [10].

The ubiquitous nature of social media makes it a convenient tool for health care organizations to connect with patients and colleagues, irrespective of their geographical locations [1]. It has been predicted that social media will become the second most important form of engagement with employees and customers, second only to face-to-face interactions [11]. Social media facilitates the expansion of professional networks and participation in various professional activities, thereby giving health care practitioners a platform to express their views about the health system, which in turn helps generate and inform health policy and public debate [1]. As many as 72% of internet users have sought health information online, and many individuals regard health care organizations as their primary source of health information, in preference to advice from other sources [12].

Objectives

Many studies [13-16] have highlighted the usefulness of social media in health care settings, and an increasing number of health care organizations are adopting the application [2-4,14]. Health care organizations use social media to support a variety of activities, including professional networking, harnessing patient feedback, public health promotion, professional education, patient education, organizational promotion, crowdsourcing, research, and patient collaboration [17].

Some health care organizations already use social media extensively, while others are relative neophytes, aiming to become “mature” social media users. Social media maturity entails a health care organization not only adopting and using the application but also possessing high levels of relevant knowledge and sophistication, along with an integrated social media strategy [18]. One of the indicators of a health care organization’s level of social media maturity is the ability to measure it [18].

Social media must be measured to ensure continuous improvement [17,19,20]. Thus, the ability to measure is critical to the success of health care social media initiatives [17,20,21]. As more forms of social media emerge, health care organizations must understand what tools to use, how to use them appropriately, and how to measure their effectiveness. The challenge for health care organizations is not just trying to find the best way to incorporate social media strategically, but also to find the best way to measure it. Against this background, this study explores how health care organizations approach social media measurement and the tools they employ.

Methods

Overview

This study is exploratory; thus, it follows a broadly interpretive [22] and inductive approach [23]. To solicit feedback relevant to the study’s objectives, health care organizations in Australia that use social media were invited to participate in the study.

Ethical Considerations

In line with Federation University Australia’s ethics procedure, an ethics review form was submitted, and approval was granted on August 23, 2017. Participants were interviewed between 2017 and 2019. All participants in the study were >18 years of age. Before the interviews, all potential interviewees were allowed to read about and consent to participate in the research. Thus, participation in the study was voluntary, and no financial rewards or incentives were offered.

Recruitment

Participant selection involved both purposive and snowball sampling. Purposive sampling involved the identification of major stakeholders [24] and ensured that initial participants were drawn from health care organizations that use social media. Initially, five hospitals that use social media were recruited to participate in the study through a combination of phone calls and emails. Apart from the initial participants, all but one of the participating organizations were recruited through snowball sampling—that is, participants suggested or helped to recruit other participants for the study. Finally, the participants were drawn from distinct types of health care organizations, varying in size and locality. Of the participating organizations, four were large hospitals that provide comprehensive health services, three were smaller hospitals that offer a wide range of medical and primary health services, and another was a medical research center. Other participants included a family practice and a clinic that promotes public health. Of the participating organizations, four were located in major cities, while the rest were located in regional areas.

Given that this study focused on health care organizations, all individuals that contribute to their organization’s social media were eligible to participate in the study. The final composition of participants was six medical doctors and five communications personnel (social media or communications managers).

Data Collection

Qualitative data were collected through semistructured interviews, as recommended by Walsham [25]. When developing the interview questions, the researcher initially outlined the broad areas of knowledge that were considered relevant to answering the larger research questions of the study. Questions were developed within each of these areas, adjusting the language of the interview to fit participants’ backgrounds so that clinicians and communications personnel could relate to questions. The goal was to tap into their experiences and expertise. The interview guide can be found in Multimedia Appendix 1.

Participants were interviewed at their preferred time and location. In line with the process of conducting semistructured

interviews, an interview guide was used flexibly [26], ensuring that conversations were free-flowing, yet focused. The flexible use of the researcher-developed interview questions enabled the interviewees to be probed further based on their responses [27]. Notes and probe questions in each interview were recorded and factored into the subsequent interviews.

Each interview was recorded with an audio recorder and then transcribed verbatim for analysis. The average duration of the interviews was approximately 50 minutes. Additional relevant information was obtained from publicly available literature about some participants in the study. After each interview transcription, the researcher carefully reviewed the transcripts and recordings to ensure that no relevant information had been missed.

The expectation was to conduct between 12 [28] and 15 interviews [29] to reach saturation of knowledge; however, after the seventh interview, the analysis of subsequent interview transcripts hardly yielded new themes. This redundancy signaled to the researcher that the data collection process might not produce additional information. In total, 11 in-depth interviews took place.

Data Analysis

The interview data were anonymized by removing content that could identify interviewees. Utmost care was taken to preserve the richness of the interview material wherever possible, while also protecting the privacy of participants [30].

The transcript was then uploaded to NVivo (QSR International) [31] in readiness for the analysis process. Relevant qualitative data were thematically analyzed until themes emerged that helped elucidate the phenomenon under investigation.

The analysis involved the inductive development of categories. The researcher first assigned summative or evocative attributes to different portions of the transcribed interviews, identifying similarities, patterns, and relationships. There were several initial codes, with some of them overlapping to some extent. A preliminary category system was applied to the interview data. Subsequently, codes with similar meanings were clustered, and a corresponding theme was formed. The researcher modified categories when the data showed additional and new information that required a new category. The researcher differentiated the resulting defined themes into main and subcategories and assigned relevant original statements in the transcripts to these categories.

Although the coding and analysis of interview transcripts were performed solely by the researcher, the findings were reviewed with peers, including experienced researchers. Furthermore, to ensure the validity of the results, the results were considered vis-à-vis explanations from relevant literature, in line with triangulation techniques [32].

Figure 1 presents the sequence of activities during data analysis.

Figure 1. The sequence of data analysis activities.

Results

Overview

Thematic analysis of interview transcripts yielded five main themes concerning health care organizations' approach to social media measurement: frequency of social media measurement, benchmarks used for social media measurement, metrics tracked in social media measurement, tools used in social media measurement, and challenging aspects of social media measurement.

The themes that emerged were spread over both medical doctors (MD) and communications personnel (CP).Hence, the contributions of MD and CP participants were blended and presented based on themes that emerged collectively rather than by group.

Frequency of Social Media Measurement

Three categories emerged based on how often participants measured their social media: frequently, infrequently, and never.

Frequently

Some participants reported that they measured their social media frequently. This study conceptualized frequent measurement as measuring one's social media at least once per week. In response to a question about how often they measure their social media, a participant had the following to say:

> *Every day. Every post that we do, we look to see how it went. So, we don't just look at them once a month, but we look at them every day.* [CP1]

According to another participant:

> *I monitor it on a day-to-day basis and then on a weekly basis, looking back, just to see that I am on track for where I need to be for the month. Because if you don't look at where you are until the end of the month, there's nothing you can do to fix it if you're not anywhere near your goal. So, it's continuous monitoring I would say.* [CP3]

Infrequently

It was observed that some health care organizations measured their social media infrequently. This study conceptualized infrequent measurement as measuring one's social media twice or fewer times per month. Regarding how frequently they measured their social media, a participant stated:

> *Not often. I give my results on a monthly report.* [CP5]

One reason for infrequent measurement could be that managers do not require employees who oversee the health care organization's social media to provide social media reports frequently. In the words of a participant:

> *I honestly haven't had to do that in a long while because they trust me.* [CP2]

Other possible reasons for not regularly measuring their social media are that some organizations do not sufficiently understand the information or are unable to afford the time, money, and effort needed.

> *I don't check it all that often because, number one, you need to understand the information, number two, you need to either have the time, the money, or [afford] the effort to change that.* [MD4]

Never

Interestingly, it was noted that some health care organizations did not ever measure their social media performance. As two participants explained:

> *I glance at the stats. I've never, in all my time, ever tried to calculate, like, the stats.* [MD5]

> *I don't measure anything. I am not sure these sorts of things can be measured.* [MD6]

The lack of measurement can be attributed to the practices of management staff in their organizations, who do not require them to provide periodic reports. As one participant put it:

> *I don't think I've been asked to measure once or to provide one of my monthly reports that I do. I don't think … I'm still yet to report that to anybody.* [CP4]

Benchmarks Used in Social Media Measurement

Three types of benchmarks were apparent in social media measurement: personal benchmarking, comparative benchmarking, and metric benchmarking.

Personal Benchmarking

At least one participant alluded to using personal benchmarks to track progress, evaluate performance, and determine areas for improvement. This study conceptualized personal benchmarking as using self-set targets to evaluate social media performance. The adoption of personal benchmarks appears to be a convenient way to make up for the absence of an official one:

> *I gave myself KPIs because nobody gave me any.* [CP2]

Comparative Benchmarking

An alternative to personal benchmarking is comparative benchmarking. It involves measuring social media against those of best-in-class peers. One of the participants stated that:

> *My whole team attended Mayo Clinic's Conference in Australia last year … in the next 5 years, we will like to be like Mayo clinic … We look at other organizations, for instance, the Royal Children's Hospital has very good social media … We use other organizations as benchmarks and try to do better.* [CP1]

Metric Benchmarking

Unlike other types of benchmarking, metric benchmarking enables the numerical measurement of performance levels and comparison with set targets. According to an interviewee:

> *We've got quite huge targets that they [management] want us to achieve within the next 5 years of growing the page and reaching more people ... And they are very keen, and they monitor those results in their quarterly board meetings. So, they get a presentation every 3 months of where we are versus where we should be, and then they make recommendations based on that … So, we have very defined targets that we want to reach on a yearly basis, but then we work it, obviously, back to a monthly basis.* [CP3]

Metrics Tracked in Social Media Measurement

Participants' responses revealed three areas that they considered relevant indicators of their social media performance. These are reach, engagement, and conversion rates.

Reach

To determine the size of the audience that has encountered the social media posts targeted at them, health care organizations use the reach metric.

Social media reach is based on the number of followers, fans, subscribers, connections, and visibility [33], as illustrated by the following comments:

> *… we specifically measure audience size. So, [on] Facebook and Twitter and Instagram, how many followers do we have, and how fast is that growing?* [CP3]

> *I've got a metric that's balanced towards being followed, which is good.* [MD3]

A reach metric allows health care organizations to estimate the proportion of an audience that sees a given social media message on a given social media platform.

> *If it's [social media posts] reaching 2000 people, or 200,000 people saw it, we know even though they might not have clicked like, they still saw it, and they might have gotten some benefit out of it.* [CP3]

Engagement

By using engagement metrics, health care organizations can gauge the degree of audience interaction with their social media efforts, using public shares, likes, retweets, check-ins, and comments as indicators. According to one of the participants:

> *Every month we look at our average brand impressions and make sure we exceed that of the previous month. So, we set targets for ourselves. We also look out for metrics on our engagement and check whether we are meeting the targets. We always aim to surpass that of the previous month.* [CP1]

Another participant shared a similar view:

> *I have a look at how many people look at my blog, my articles on my blog, which is an automated stats collection.* [MD5]

There are several ways to measure how much interest a health care organization's social media is generating, as noted below:

> *It could be the commentary, likes … it's derived from an algorithm.* [CP1]

Another way to measure engagement is to use yardsticks, such as the Alexa rank.

According to an interviewee:

> *I'll check my Alexa ranking every month or so just to make sure we're going in the right direction. And if that has a massive turn, then I know that something isn't right.* [MD4]

One of the alternatives to using yardsticks, such as the Alexa rank to measure engagement is to deploy an analytics tool to do the measurement. According to a participant:

> *…we have our Google Analytics running all the time on our website to be looking at traffic and so on. The regular posting on social media is designed to generate the traffic ... We track how successful they [social media] are at generating new and old traffic, how many clicks, how many people visit a certain number of pages after it, and so on. The bounce rate, things like that.* [MD3]

When health care organizations post relevant social media content, it encourages users to click through to the organization's website. To specifically identify traffic from social media platforms, health care organizations deploy relevant metrics. One participant alluded to this by saying:

> *Every time we run a little campaign or a post, we put a sticky label on it to see what's causing the traffic to rise or fall. So, we track which types of posts are most successful in driving more traffic through.* [MD3]

Another participant added that:

> *…we measure referrals back, so how many views on our website did we get from people who were using Facebook or Twitter or saw our stuff on Facebook or Twitter?* [CP3]

Measuring engagement also allows users to identify the platforms audiences are most interested in, the nature of their interaction with the platforms, and the geographical locations of the audience. In the words of a participant:

> *I measure the way in which people interact, including the platforms, the time zones, and the countries in which they interact. If there's increase in the number of users in a certain country, then I can start thinking about translating contents into the language of that country.* [MD4]

Conversion

The conversion rate metric enables health care organizations to determine the percentage of visitors to their social media platform who donated to their social crowdfunding initiatives. In that context, a higher conversion rate is an indication of the success of social media initiatives. As one participant stated:

> *I think the amount of money we have raised through social media campaigns is an indication of success.* [CP2]

Health care organizations can also use the conversion rate metric to measure the percentage of people who attend an event after learning of the event on their social media platform. According to some interviewees, if many people attend their events after interacting with social media posts about those events, this demonstrates that their social media initiative has been successful. In the words of a participant:

> *[We consider our social media initiative successful] when patients come in and say we are here because we saw you on Facebook.* [MD2]

Another added:

> *If we have a function … we can do a paid advert for A$2,000 in a local paper and get 5 people in. If we do a A$500 advert on Facebook, we'll get 50 people in, and the function is full, and we have to run additional events. So, from that, I suppose we can say we've been successful.* [CP5]

Tools Used in Social Media Measurement

The study found that health care organizations used two types of tools for social media measurement: inbuilt analytics tools and add-on analytics tools.

Inbuilt Analytics Tools

Inbuilt analytics tools are embedded in the social media platform. For instance, Facebook page analytics is an inbuilt tool used to track user interaction on a Facebook fan page to improve understanding of the page's performance. According to a participant:

> *Facebook Insights provides details on which posts have the most likes, comments, and shares, which means that page managers can see what content resonates with their audience and provide similar content to increase engagement with the page … Page Insights also provides basic demographic information about people who like your page, and this includes gender and age.* [MD2]

Another participant had the following to say about how they use inbuilt analytics tools for social media measurement:

> *... I will show them [management] monthly how many people were reached with the help of data obtained from Facebook Insights, and explain factors (humor, picture, videos, etc) that made the difference to audience engagement.* [CP2]

Add-on Analytics Tools

In contrast to inbuilt analytics tools, add-on analytics tools are not embedded in the social media platform. Rather, they are third-party software programs or scripts that are added to a social media platform to provide it with additional features and

abilities. The additional capabilities of add-on analytics tools appear to have made them popular among health care organizations that use social media. According to an interviewee:

> *I'll check my Alexa ranking every month or so just to make sure we're going in the right direction.* [MD4]

In the words of another interviewee:

> *… we have our Google Analytics running all the time on our website, to be looking at traffic and so on.* [MD3]

A participant had this to say about the benefits of using Google Analytics:

> *I think that using Google Analytics is useful from a geolocation point of view. ... If there's increase in number of users in a certain country, then I can start thinking about translating contents into the language of that country.* [MD4]

Challenging Aspects of Social Media Measurement

Some of the feedback from participants indicated that there are aspects of social media use that health care organizations would like to measure but are currently unable to. Those identified specifically were health care social media's conversion rate and its impact on both public health and patient satisfaction.

Conversion Rate

Although some of the interviewees alluded to measuring their social media conversion rate, it appears that many had more questions than answers. In the words of one particpant:

> *I wish there is a way to measure the follow-on. I can see how many people have gone to our website through Google Analytics, but the conversion rate is the problem. It's hard to measure. I wish I could find out if we got more people as a result of our social media post.* [CP2]

Echoing a similar sentiment, another interviewee commented:

> *… we want the hard facts, we want to know who are we convincing [through social media] to come here rather than a competitor.* [CP5]

Impact on Public Health

Another aspect identified by some participants as difficult to measure is the impact of their social media activities on public health. In the words of one participant:

> *Like everything in health, what you will like to measure is whether you are making an impact in people's health [through social media]… it's hard to know.* [MD2]

The difficulty inherent in measuring the impact of social media interventions is particularly obvious at the aggregate level. One interviewee had this to say regarding the issue of measuring the global impact of social media-based health interventions:

> *There are many anecdotal pieces of information [regarding the impact of social media interventions] ... But how can you measure the globality of impact, because a lot of it is subconscious? It is difficult.* [MD4]

Patient Satisfaction

Finally, it is difficult to measure the extent to which patients are satisfied with the information health care organizations share with them on social media. In their words:

> *So you can have an exchange with them [patients], on social media, but how satisfied were they with service that we provided? … You can't really measure the satisfaction that they got out of your exchange with them.* [CP03]

Discussion

Principal Findings

Five themes emerged from the analysis of health care organization participants' responses regarding their approaches to social media measurement: frequency of social media measurement, the benchmark used for social media measurement, metrics tracked in social media measurement, tools used in social media measurement, and challenging aspects of social media measurement. Table 1 presents a summary of the findings. The analysis presented in this section elucidates responses to the research question—how do health care organizations approach social media measurement?

Table 1. Principal findings: themes and categories.

Themes	Categories
Frequency of measurement	• Frequently • Infrequently • Never
Benchmarks used in measurement	• Personal benchmarking • Comparative benchmarking • Metric benchmarking
Metrics tracked in measurement	• Reach • Engagement • Conversion
Tools used in measurement	• Inbuilt analytics tools • Add-on analytics tools
Challenging aspects of measurement	• Conversion rate • Impact on health • Patient satisfaction

Frequency of Measurement

The study observed a discrepancy among participants in the frequency of social media measurement. While some health care organizations measured their social media frequently, others measured infrequently or not at all.

It appears that the frequency of social media measurement depended on the accountability expectations of management. Health care organizations whose management demand a formal report of their social media performance tend to measure their social media frequently. Whereas health care organizations whose management does not demand accountability from those who run their social media tend not to measure their social media at all, or they measure them infrequently.

Accountability expectations were influenced by whether the management trusted that their social media was performing a useful health service. Not measuring social media or infrequently measuring them could be an indication that managers of a health care organization are not fully aware of the usefulness of using social media in health care settings.

Further, it appears that medical doctors were less likely to measure social media regularly compared with communications personnel, perhaps due to their heavy workloads or limited experience using social media for business.

Social media should be tracked frequently, that is, weekly or monthly [34]. Health care organizations that frequently measure their social media can track how their social media initiatives are progressing vis-à-vis their target for the month, which allows them to address any issues that could impede their ability to meet set targets. Measurement needs to be an ongoing cycle [35]. Measuring social media on an ongoing basis would enable an understanding of the extent to which social media has supported the realization of set objectives [19,36-38].

Benchmarks Used in Measurement

Although many health care organizations use benchmarks to measure their social media performance, others do not. Benchmarks are standards against which performance is compared; thus, they enable social media users to know which areas require more attention [34]. The types of benchmarking that were apparent from the study are personal benchmarking, comparative benchmarking, and metric benchmarking.

Some health care organizations that use social media do not have an official benchmark against which they measure their performance; hence, some users have set personal benchmarks in the absence of an official one. Personal benchmarks are standards individuals set for themselves that allow them to track their progress and evaluate themselves to determine how they need to improve. The standards are set regarding what is important to individuals, thereby guiding them against irrational actions [39]. Setting a personal benchmark could help motivate health care organizations to perform better because it enables them to measure their performance.

An alternative approach used by some of the more ambitious health care organizations to measure their social media is comparative benchmarking. Health care organizations that used comparative benchmarking tracked their social media operations and compared their results with those of more established health care organizations. In this context, it can also be referred to as peer benchmarking. In health care settings, peer benchmarking is used when there is a need to raise performance levels to be on par with the performance of leaders in the field [40]. In doing so, individual results may be used to compare with peer results [41]. Comparative benchmarking has the potential to support health care organizations that are neophytes in their efforts to improve their social media performance. It encourages them to strive to be like more established health care social media users in the industry or region.

Health care organizations that have clear social media goals may prefer metric benchmarking to both personal benchmarking and comparative benchmarking. Metric benchmarking involves the use of statistical procedures to evaluate performance against set targets [42]. Consequently, social media use is data-driven, with quantitative performance improvement objectives that are

predictable and align with the needs of the health care organization, thereby ensuring a more objective appraisal of social media performance.

Metrics Tracked in Measurement

It was noted that metrics tracked by health care organizations that use social media include reach, engagement, and conversion rates.

Many health care organizations use reach metric to measure their social media performance because it is easy to calculate. Social media reach is an estimate of the number of users that could have contact with a social media post [33]. It is the aggregate of the audience of a social media platform, including subscribers and visitors [33]. The reach metric is informative, given that it allows a health care organization to obtain a better understanding of their audience and the geographical regions that their posts or content reach. An expanding audience would indicate to a health care organization that their online presence is growing.

A wide reach does not necessarily translate to deep engagement. In other words, it is possible for a health care organization's reach via social media to be wide yet only able to engage a small proportion of its target audience. Consequently, there is a need for the engagement metric. The engagement metric measures the ability of a social media user to establish dialogue and interaction with other users [37,43]. Measuring engagement allows health care organizations to know who is reading their social media posts, the content that interests users, and the platform that is popular with users. Social media data, such as the number of "likes," "fans," or "shares" for Facebook, or the number of "tweets," "retweets," or "replies" for Twitter are used to compute engagement [43]. The type of indicator required depends on the specific social media platform. For instance, engagement can be calculated by counting the number of replies on Twitter, the number of comments on Facebook, and the number of subscribers on YouTube [37,43]. More shares, likes, or comments for a health care organization's social media posts would indicate that their message resonates with their audience.

Although a high level of engagement indicates that the audience finds social media efforts interesting, it is not a confirmation that they are taking the desired action. To be sure that their social media posts are influencing the behavior of their audience, health care organizations use the conversion metric. Social media conversion rate is a measure of the percentage of the audience who take the desired action after interacting with social media content [44]. Health care organizations use conversion metrics to measure the number of people that respond to their call-to-action on social media. When a health care organization is seeking to motivate action on the part of its audience, tracking the conversion rate allows them to monitor the percentage of users who take the recommended action. Participants reported that the conversion metrics helped them to know the proportion of people who visited their social media pages that donated to their crowdfunding initiatives. Ultimately, the indicators used to measure conversion rates vary depending on the recommended action. For instance, if the recommended action is for the audience to donate to a crowdfunding initiative, the indicator would be the number of people that donated after finding the campaign on social media. Similarly, if the recommended action is that the audience attends an event organized by the health care organization, the indicator would be the number of people that arrive at the event after viewing the invitation on social media.

Tools Used in Measurement

It was noted that when it comes to measuring health care social media, both inbuilt analytics tools and add-on analytics tools are useful.

Inbuilt analytics tools are measurement tools built into most social media platforms [45]. For example, Facebook Analytics provides a general overview of a user's Facebook page, their audience, and the performance of their posts [45]. Data taken from one day, the previous week, or the last month can be drilled-down to reveal more high-level statistics [45]. It shows an organization the performance of their posts and the behavior of their followers, thus allowing them to identify the best time of day to post, the best day of the week to post, and the most popular type of content to post [45]. Other examples of inbuilt analytics tools include Pinterest Analytics, Twitter Analytics, Instagram Analytics, and YouTube Analytics [46].

Inbuilt analytics tools are popular because there are no acquisition costs [45,47]. Hence, they are suitable for health care organizations that have a relatively small social media budget and health care organizations that do not use social media extensively or use only one social media platform. The main limitations of inbuilt social media analytics tools are that they are usually only able to support individuals or small brands [47] and specific social media accounts [45].

Health care organizations that are mature social media users and own multiple social media platforms may prefer to use add-on social media analytics tools to appraise their social media. Add-on social media analytics tools include computer software services such as Google Analytics and Alexa ranking that can be added to social media to enable the tracking of relevant metrics. Users can employ the Google Analytics tool to sift and sort visitors to social media platforms with dimensions such as location. As an analytics tool, it can globally track social media and other online activities [45]. Google Analytics could work well with social media platforms. To optimize usage, the user must install both Google Analytics and Google Tag Manager before tagging the aspect of social media they would like to measure [48]. Google Tag Manager enables data and metrics from relevant social media websites to be sent to Google Analytics for analysis [48]. Alexa rank, on the other hand, is a measure of the popularity of a website or a social media site in terms of traffic [49]. It is calculated using a proprietary methodology that combines a site's estimated traffic and visitor engagement over some time [50].

Other examples of add-on analytics tools include BuzzSumo, Vizia, SumAll, and Quintly [46]. Add-ons are arguably easier to use than inbuilt analytics tools. They are cross-platform and are particularly relevant for managing multiple social media platforms and accounts [45-47]. Users can use them to track all their social media platforms simultaneously, thereby enhancing efficiency and ensuring more consistent and valuable results

[45]. The additional features mean that most of these tools entail additional costs. That notwithstanding, they are the most suitable analytics tools for health care organizations that use several social media platforms.

Challenging Aspects of Measurement

Information technology creates value but identifying where, how, and how much value can be problematic [51]. Social media has considerable potential to make a positive contribution to health care; however, results from its use are abstruse [52]. One of the most critical issues in the appraisal of information systems investments is the question of what to measure [53]. Although many participants alluded to measuring their social media, many did not appear confident that they were successful with certain aspects of measurement.

Participants identified conversion rate, impact on public health, and patient satisfaction as areas that were difficult to measure. Although some health care organizations found the conversion metrics useful for tracking audience responses to social crowdfunding initiatives, they appeared to have unanswered questions regarding using the metric in different contexts.

It was also noted that health care organizations would like to know the impact their social media initiatives are having on the broader society in terms of disease prevention, prolonging life, and promoting human health. However, according to the results of the study, participants do not know how to ascertain their impact. That is not surprising because the societal impact of information technology is difficult to conceptualize, and any conceptualization is likely to be subjective [54].

Participants also alluded to being interested in learning the level of patient satisfaction with their social media sites, but being unable to measure it. Patient satisfaction is an important and commonly used indicator for measuring the extent to which patients are content with the health care they have received from their health care organization [55]. Given that patient satisfaction affects clinical outcomes, patient retention, and medical malpractice claims [55], it is logical that health care organizations that use social media would be keen to know the extent to which patients are happy with their social media sites.

The ongoing social media measurement challenges highlight the need for a transparent, standardized, and flexible measurement framework [19]. The lack of a consensus approach concerning certain aspects of measurement suggests that social media measurement in health care settings is at a nascent stage. It seems that the health industry lags behind other industries in terms of knowledge, sophistication, and integration of social media strategy.

Strengths and Limitations

This study sheds light on the social media capabilities of health care organizations. It demonstrates the yardsticks that are effective for social media measurement as well as potential blind spots for which suitable yardsticks appear to be elusive. By exploring the merits and limitations of current techniques used to appraise social media in health care settings, this study provides information with which to revise and improve the existing measurement criteria.

A health care organization's ability to measure their social media, among other things, reflects their level of social media maturity [18]. By investigating how health care organizations in Australia approach social media measurement, this study reveals participants' level of social media maturity. Thus, the results of this study can help health care organizations take stock of their social media capabilities and determine which strategies are appropriate for their maturity level and for optimizing success [18].

Moreover, measurement is a critical success factor of social media initiatives [56]. Thus, by identifying the approach used by health care organizations in social media measurement, this study enables a deeper understanding of some of the tools and techniques required for successful social media campaigns.

Despite the contributions of this study to the growing body of research on the use of social software in health care settings, several important limitations need to be considered. First, the results of this study should be interpreted as indicative and not necessarily generalizable, considering that the study was restricted to Australia. Second, given that only medical doctors and the communications personnel of health care organizations were interviewed for this study, a research sample with more diverse profiles may suggest additional themes. Last, it is important to note the time frame of this study when considering its findings, since usage and attitude toward social media evolve rapidly.

Conclusions

This qualitative study provides insight into how health care organizations approach social media measurement. Although many participants showed great interest in using various tools and techniques to measure their social media, they still had some unanswered questions. Despite the availability of tools that enable users to track social followers and click-through, measuring the effectiveness of social media initiatives remains a challenge [21]. While many online activities can be appraised using defined quantitative metrics, social media, among other things, generates qualitative data, which traditional metrics alone cannot effectively measure [19]. This measurement problem is exacerbated by the lack of an overarching measurement approach, which causes difficulties for organizations wanting to prove the usefulness of their social media [19,57,58].

Without the ability to define and measure the use of social media, it will be difficult to derive value from them [21,59]. Therefore, the challenge for health care organizations that use social media is to determine what to measure and the data requirement of such measurement. A comprehensive and consistent measurement approach for social media would help improve its use.

Health care organizations should work towards improvements in how to use social media, and how to measure their effectiveness. This study calls for social media training in health care organizations. Such training must focus on both how to use relevant tools and how to effectively measure their use.

Acknowledgments

This research was supported by an Australian Government Research Training Program (RTP) Scholarship through Federation University Australia. RTP had no involvement in the study design, collection, analysis or interpretation of data, the writing of this paper or the decision to submit it for publication.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview Guide.
[PDF File (Adobe PDF File), 120 KB - formative_v4i8e18518_app1.pdf]

References

1. RACGP. Guide for the use of social media in general practice.: Royal Australian College of General Practitioners; 2015. URL: https://www.racgp.org.au/download/documents/e-health/social-media-guide-v5.pdf
2. Griffis HM, Kilaru AS, Werner RM, Asch DA, Hershey JC, Hill S, et al. Use of social media across US hospitals: descriptive analysis of adoption and utilization. J Med Internet Res 2014 Nov 27;16(11):e264 [FREE Full text] [doi: 10.2196/jmir.3758] [Medline: 25431831]
3. Brown J, Ryan C, Harris A. How doctors view and use social media: A national survey. J Med Internet Res 2014 Dec 02;16(12):e267 [FREE Full text] [doi: 10.2196/jmir.3589] [Medline: 25470407]
4. Panahi S, Watson J, Partridge H. Social media and physicians: Exploring the benefits and challenges. Health Informatics J 2016 Jun 18;22(2):99-112. [doi: 10.1177/1460458214540907] [Medline: 25038200]
5. Allen C. Tracing the evolution of social software.: Life with Alacrity; 2004. URL: http://www.lifewithalacrity.com/2004/10/tracing_the_evo.html
6. Kaplan AM, Haenlein M. Users of the world, unite! The challenges and opportunities of Social Media. Business Horizons 2010 Jan;53(1):59-68. [doi: 10.1016/j.bushor.2009.09.003]
7. Berthon PR, Pitt LF, Plangger K, Shapiro D. Marketing meets Web 2.0, social media, and creative consumers: Implications for international marketing strategy. Business Horizons 2012 May;55(3):261-271. [doi: 10.1016/j.bushor.2012.01.007]
8. van Dissel BMP. Social media and the employee's right to privacy in Australia. International Data Privacy Law 2014 Jul 02;4(3):222-234. [doi: 10.1093/idpl/ipu015]
9. Paton C, Luquel L. IMIA Social Media Working Group. 2015. URL: https://imiasocialmedia.wordpress.com/about/
10. Kordzadeh N. Social media in health care. In: Contemporary Consumer Health Informatics. Switzerland: Springer; 2016:101-123.
11. Williams DK, Scott MM. New research on why CEOs should use social media. Harvard Business Review. 2012. URL: https://hbr.org/2012/07/new-research-on-why-ceos-shoul
12. Fox S. The social life of health information. Pew Research Center. 2014. URL: https://www.pewresearch.org/fact-tank/2014/01/15/the-social-life-of-health-information/
13. Ferguson C. It's time for the nursing profession to leverage social media. J Adv Nurs 2013 Apr;69(4):745-747. [doi: 10.1111/jan.12036] [Medline: 23488814]
14. Ventola CL. Social media and health care professionals: benefits, risks, and best practices. P T 2014 Jul;39(7):491-520 [FREE Full text] [Medline: 25083128]
15. Ruddiman A. The value of social media to you and the profession. BC Medical Journal 2016;58(9):497.
16. Gibson CM. The democratization of medical research and education through social media. JAMA Cardiol 2017 Jan 01;2(1):9. [doi: 10.1001/jamacardio.2016.4933] [Medline: 27973665]
17. Ukoha C. On the value of healthcare social media: Exploring users' perspectives. 2018 Presented at: Pacific Asia Conference on Information Systems (PACIS); 2018; Yokohama, Japan.
18. Thomas L, Woodside JM. Social media maturity model. International Journal of Healthcare Management 2016 Jan 07;9(1):67-73. [doi: 10.1080/20479700.2015.1101940]
19. McCann M, Barlow A. Use and measurement of social media for SMEs. Jrnl of Small Bus Ente Dev 2015 May 18;22(2):273-287. [doi: 10.1108/JSBED-08-2012-0096]
20. Ukoha C, Stranieri A, Mehmood C. Deriving value from Health 2.0: A study of social media use in Australian healthcare organizations. 2017 Presented at: Pacific Asia Conference on Information Systems (PACIS); 2017; Langkawi, Malaysia.
21. Ukoha C, Stranieri A. Criteria to Measure Social Media Value in Health Care Settings: Narrative Literature Review. J Med Internet Res 2019 Dec 16;21(12):e14684 [FREE Full text] [doi: 10.2196/14684] [Medline: 31841114]
22. Green J. In: Thorogood N, editor. Qualitative methods for health research. London: Sage; 2013.

23. Thomas D. A general inductive approach for analyzing qualitative evaluation data. American Journal of Evaluation 2016 Jun 30;27(2):237-246. [doi: 10.1177/1098214005283748]
24. Palys T. Purposive sampling. The Sage Encyclopedia Of Qualitative Research Methods 2008;2(1):697. [doi: 10.4135/9781412963909.n349]
25. Walsham. Doing interpretive research. European Journal of Information Systems 2017 Dec 19;15(3):320-330. [doi: 10.1057/palgrave.ejis.3000589]
26. Dey I. Qualitative data analysis: A user-friendly guide for social scientists. London: Routledge; 1993.
27. LoflandL. H. Lofland, Analyzing social settings, Belmon, CA: Wadsworth Publishing Company, Inc 1984.
28. Guest A, Bunce A, Johnson L. How many interviews are enough? Field Methods 2016 Jul 21;18(1):59-82. [doi: 10.1177/1525822X05279903]
29. Kvale S, Brinkmann S. Interviews: Learning the craft of qualitative research interviewing. California: Thousand Oaks; 1996.
30. Saunders B, Kitzinger J, Kitzinger C. Anonymising interview data: Challenges and compromise in practice. Qual Res 2015 Oct;15(5):616-632 [FREE Full text] [doi: 10.1177/1468794114550439] [Medline: 26457066]
31. QSR International. URL: https://www.qsrinternational.com [accessed 2020-08-06]
32. Carter N, Bryant-Lukosius D, DiCenso A, Blythe J, Neville AJ. The use of triangulation in qualitative research. Oncol Nurs Forum 2014 Sep;41(5):545-547. [doi: 10.1188/14.ONF.545-547] [Medline: 25158659]
33. Fisher T. ROI in social media: A look at the arguments. J Database Mark Cust Strategy Manag 2009 Sep 14;16(3):189-195. [doi: 10.1057/dbm.2009.16]
34. Pritchard K. This is how often you really should track each of your marketing metrics in 2019.: Impact; 2019. URL: https://www.impactbnd.com/blog/how-often-track-marketing-metrics
35. Wilkerson D. How often should I post on Facebook or Twitter?.: Bounteous; 2013. URL: https://www.bounteous.com/insights/2013/08/19/how-often-should-i-post-facebook-or-twitter/?ns=l
36. Cooper C, Martin M, Kiernan T. Measuring the value of social software. Defining a measurement approach that maps activity to business value. Cambridge, MA: IBM; 2010.
37. Hoffman DL, Fodor M. Can you measure the ROI of your social media marketing? MIT Sloan Management Review 2010;52(1):41.
38. Nair M. Understanding and measuring the value of social media. J. Corp. Acct. Fin 2011 Feb 23;22(3):45-51. [doi: 10.1002/jcaf.20674]
39. Widger C. In: Crosby D, editor. Personal benchmark: integrating behavioral finance and investment management. New Jersey: John Wiley & Sons; 2014.
40. NHS Wales. Benchmarking. URL: http://www.wales.nhs.uk/technologymls/english/resources/pdf/tools/benchmarking%20final.pdf
41. White KR, Begun JW. Preceptor and employer evaluation of health administration student competencies. Journal of Health Administration Education 2006;23(1):1-16.
42. Ibrahim OM, Polk RE. Benchmarking antimicrobial drug use in hospitals. Expert Rev Anti Infect Ther 2012 Apr;10(4):445-457. [doi: 10.1586/eri.12.18] [Medline: 22512754]
43. Agostino D, Sidorova Y. A performance measurement system to quantify the contribution of social media: new requirements for metrics and methods. Measuring Business Excellence 2016 May 16;20(2):38-51. [doi: 10.1108/MBE-05-2015-0030]
44. Cothrel J. Measuring the success of an online community. Strategy & Leadership 2000 Apr;28(2):17-21. [doi: 10.1108/10878570010341609]
45. Meg. The complete social media analytics guide.: Talkwalker; 2019. URL: https://www.talkwalker.com/blog/social-media-analytics-guide
46. Smith K. The best free and paid social media analytics tools.: Brandwatch; 2019. URL: https://www.brandwatch.com/blog/social-media-analytics-tools/
47. Influencer Marketing Hub. 15 Social Media Monitoring Tools for 2020. 2020. URL: https://influencermarketinghub.com/social-media-monitoring-tools/
48. Tran T. A 4-step guide to tracking social media in Google Analytics.: Hootsuite; 2019. URL: https://blog.hootsuite.com/tracking-social-media-in-google-analytics/
49. Thakur A, L. Sangal A, Bindra H. Quantitative measurement and comparison of effects of various search engine optimization parameters on Alexa traffic rank. IJCA 2011 Jul 31;26(5):15-23. [doi: 10.5120/3100-4257]
50. Kosaka K. What is Alexa Rank?.: Alexa Blog; 2019. URL: https://blog.alexa.com/marketing-research/alexa-rank/
51. Tillquist J, Rodgers W. Using asset specificity and asset scope to measure the value of IT. Commun. ACM 2005 Jan 01;48(1):75-80. [doi: 10.1145/1039539.1039542]
52. Korda H, Itani Z. Harnessing social media for health promotion and behavior change. Health Promot Pract 2013 Jan;14(1):15-23. [doi: 10.1177/1524839911405850] [Medline: 21558472]
53. Schryen G. Revisiting IS business value research: what we already know, what we still need to know, and how we can get there. European Journal of Information Systems 2017 Dec 19;22(2):139-169. [doi: 10.1057/ejis.2012.45]

54. Bannister F, Remenyi D. The societal value of ICT: first steps towards an evaluation framework. Electronic Journal of Information Systems Evaluation 2003;6(2):197-206.
55. Prakash B. Patient satisfaction. Journal Cutan Aesthet Surg 2010;3(3):151-155.
56. Abedin B, Jafarzadeh H. Relationship development with customers on Facebook: A critical success factors model. 2015 Presented at: 48th Hawaii International Conference on Systems Science (HICSS); 2015; Hawaii. [doi: 10.1109/hicss.2015.227]
57. Mangiuc DM. Measuring Web 2.0 efficiency. Annales Universitatis Apulensis: Series Oeconomica 2009;11(1):74.
58. Flynn S, Hebert P, Korenstein D, Ryan M, Jordan WB, Keyhani S. Leveraging social media to promote evidence-based continuing medical education. PLoS One 2017;12(1):e0168962 [FREE Full text] [doi: 10.1371/journal.pone.0168962] [Medline: 28060854]
59. Culnan MJ, McHugh PJ, Zubillaga JI. How large US companies can use Twitter and other social media to gain business value. MIS Quarterly Executive 2010;9(4):243-259.

Abbreviations

CP: communications personnel
MD: medical doctor

Edited by G Eysenbach; submitted 02.03.20; peer-reviewed by R Lee, M Janodia; comments to author 23.03.20; revised version received 26.03.20; accepted 27.07.20; published 14.08.20.

Please cite as:
Ukoha C
How Health Care Organizations Approach Social Media Measurement: Qualitative Study
JMIR Form Res 2020;4(8):e18518
URL: http://formative.jmir.org/2020/8/e18518/
doi:10.2196/18518
PMID:32795994

Original Paper

Calibrating Wrist-Worn Accelerometers for Physical Activity Assessment in Preschoolers: Machine Learning Approaches

Shiyu Li[1], MSc; Jeffrey T Howard[2], PhD; Erica T Sosa[2], PhD; Alberto Cordova[3], PhD; Deborah Parra-Medina[4], PhD; Zenong Yin[2], PhD

[1]The University of Texas Health Science Center at San Antonio, San Antonio, TX, United States

[2]Department of Public Health, The University of Texas at San Antonio, San Antonio, TX, United States

[3]Department of Kinesiology, The University of Texas at San Antonio, San Antonio, TX, United States

[4]Department of Mexican American and Latina/o Studies, The University of Texas at Austin, Austin, TX, United States

Corresponding Author:
Shiyu Li, MSc
The University of Texas Health Science Center at San Antonio
7703 Floyd Curl Dr
San Antonio, TX, 78229
United States
Phone: 1 210 204 7621
Email: lis9@livemail.uthscsa.edu

Abstract

Background: Physical activity (PA) level is associated with multiple health benefits during early childhood. However, inconsistency in the methods for quantification of PA levels among preschoolers remains a problem.

Objective: This study aimed to develop PA intensity cut points for wrist-worn accelerometers by using machine learning (ML) approaches to assess PA in preschoolers.

Methods: Wrist- and hip-derived acceleration data were collected simultaneously from 34 preschoolers on 3 consecutive preschool days. Two supervised ML models, receiver operating characteristic curve (ROC) and ordinal logistic regression (OLR), and one unsupervised ML model, k-means cluster analysis, were applied to establish wrist-worn accelerometer vector magnitude (VM) cut points to classify accelerometer counts into sedentary behavior, light PA (LPA), moderate PA (MPA), and vigorous PA (VPA). Physical activity intensity levels identified by hip-worn accelerometer VM cut points were used as reference to train the supervised ML models. Vector magnitude counts were classified by intensity based on three newly established wrist methods and the hip reference to examine classification accuracy. Daily estimates of PA were compared to the hip-reference criterion.

Results: In total, 3600 epochs with matched hip- and wrist-worn accelerometer VM counts were analyzed. All ML approaches performed differently on developing PA intensity cut points for wrist-worn accelerometers. Among the three ML models, k-means cluster analysis derived the following cut points: ≤2556 counts per minute (cpm) for sedentary behavior, 2557-7064 cpm for LPA, 7065-14532 cpm for MPA, and ≥14533 cpm for VPA; in addition, k-means cluster analysis had the highest classification accuracy, with more than 70% of the total epochs being classified into the correct PA categories, as examined by the hip reference. Additionally, k-means cut points exhibited the most accurate estimates on sedentary behavior, LPA, and VPA as the hip reference. None of the three wrist methods were able to accurately assess MPA.

Conclusions: This study demonstrates the potential of ML approaches in establishing cut points for wrist-worn accelerometers to assess PA in preschoolers. However, the findings from this study warrant additional validation studies.

(JMIR Form Res 2020;4(8):e16727) doi:10.2196/16727

KEYWORDS

preschoolers; accelerometer; physical activity; obesity; machine learning

Introduction

Accelerometry has been widely accepted as the gold standard to measure physical activity (PA) in free-living settings [1,2] including preschools [3]. Triaxial accelerometers can record the magnitude of accelerations from three movement axes and convert accelerations to vector magnitude counts over a given user-specified cycling period (epoch) [4]. Counts are translated

into biologically meaningful PA volume and intensity levels using pre-established cut points for sedentary behavior, light physical activity (LPA), moderate physical activity (MPA), and vigorous physical activity (VPA) [3]. Although; traditionally, gold standard cut points are established using data derived from accelerometers placed on the right hip of the body [1,2], recent revolutions have focused on acceleration data from wrist-worn accelerometers for increased protocol compliance of study participants, better sensitivity to detect certain types of movements, and sleep measurement [5-9]. Fairclough et al [10] found that wrist-worn accelerometers had at least 10% higher compliance rate than hip-worn ones, regardless of the data processing criteria in school-age children. They also reported that wrist-worn accelerometers had a much lower study drop-off rate compared to hip-worn ones, regardless of the number of monitoring days. Thus, the wrist, instead of the hip, might be an ideal accelerometer placement site for preschoolers.

However, cut points from wrist-derived data are sparse for preschool-age children [11,12]. Johansson and colleagues [13,14] conducted the only studies that established wrist-referenced cut points for sedentary behavior and moderate and vigorous physical activity (MVPA) in preschoolers. Using direct observation of structured and free-play activities as the ground truth activities, hip- and wrist-derived cut points yielded comparable accuracy and validity of the observed activities. Nevertheless, Johansson [15] cut points did not differentiate between LPA, MPA, and VPA, and have not been replicated by others.

Machine learning (ML) algorithms are increasingly being used to translate accelerometer outputs to meaningful PA metrics [16]. Recent studies have applied part of the accelerometry data to ML algorithms as the training set to build statistical models that can predict PA intensities from a new set of accelerometry data without explicit instructions [17]. Research has demonstrated the promising performance of ML techniques in combination with the use of wrist-derived acceleration data in predicting the type and intensity of activities, as well as activity energy expenditure with comparable overall predictive accuracies in adult populations [16,18,19].

Combining ML techniques and wrist-worn accelerometers may help address the low compliance caused by the challenges in wearing hip-worn accelerometers [20,21] and the difficulty in measuring the various levels of activity intensity given the unique nature of the sporadic and short-burst activity patterns in preschool-age children [22,23]. Therefore, the purposes of this study were (1) to develop the cut point values for sedentary behavior, LPA, MPA, and VPA based on wrist-derived acceleration data using multiple ML algorithms, and (2) to examine classification accuracy of PA intensity in comparison to hip-reference cut points [24] and previously established wrist-referenced cut points [25] in preschool-age children.

Methods

Recruitment

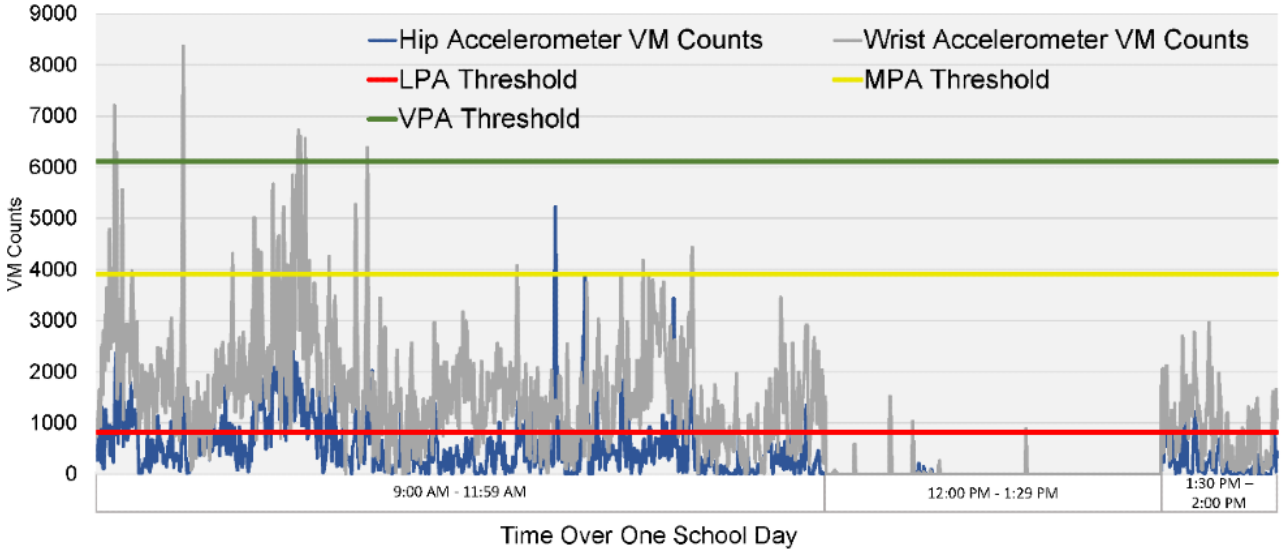
A total of 61 healthy children, aged 3-5 years, who were enrolled in a Head Start program (HS), were recruited to participate in the study in Fall 2018; in San Antonio, Texas. The HS is a federally funded program that serves children from low-income families through academic, health, nutrition, and family service programs [26]. Their parents/guardians signed an informed consent form before participation. Each participant received up to US $30 gift card for participating in the study. Children who were 3 years old at recruitment, were enrolled in the full-day HS program, and had no physical disabilities were eligible for the study.

This study was reviewed and approved by the Institutional Review Board of the University of Texas at San Antonio.

Accelerometer Data Collection

For 3 consecutive days, the children wore two triaxial accelerometers (ActiGraph model WGT3X-BT, ActiGraph, LLC) that collected raw accelerometry data at a sampling frequency of 30Hz (30 observations per second for each axis) from 9 AM to 2 PM, following a previously published protocol [27]. On the day of data collection, a trained research assistant placed one accelerometer on the nondominant wrist and the other one on the right hip of each child. For this study, raw accelerations were converted into 15-second epoch and thereafter collapsed to 60-second epochs. For this study, data were outputted as the vector magnitude (VM) counts, which is the square root of the sum of squares of each axis of acceleration data. Nonwear time was detected using the Choi wear time validation algorithm [28]. Participants with missing hip- or wrist-worn accelerometer epoch for more than 3 consecutive 5-hour days were excluded from the analysis. Accelerometer data processing was performed using ActiLife software (Version 6.13.3). Visual presentation of the accelerometer counts for one participant is shown in Figure 1.

Figure 1. Visual presentation of the changes of wrist- and hip-worn accelerometer vector magnitude counts for one participant throughout a school day, with previously established hip-based physical activity level thresholds as defined by Butte et al. [23].



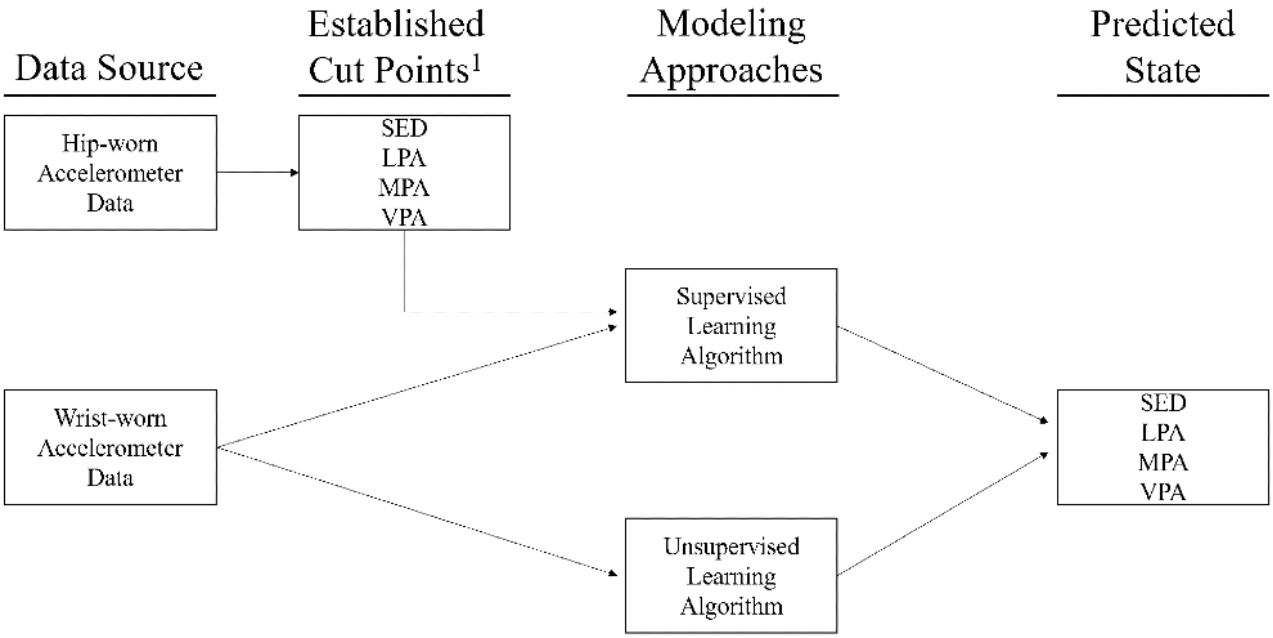
Hip- and Wrist-Reference Cut Points for Comparisons

The hip-reference cut points for assessing PA intensity were adopted from Butte et al [24] based on predicted energy expenditure from room calorimetry and doubly labeled water. The cut points were no more than 820 counts per minute (cpm) for sedentary behavior, 821-3908 cpm for LPA, 3909-6112 cpm for MPA, and greater than 6113 cpm for VPA for vector magnitude counts collected from ActiGraph hip-worn accelerometers during free-living activities in preschool-age children. The cut points reported by Butte and colleagues [13] are widely used and will be used as the gold-standard reference in this study. The cut points for sedentary behavior and MVPA from wrist-derived data captured during structured and free-play activities by Johansson et al [14] are the only available references for preschool-age children.

Applications of Machine Learning Techniques

Three ML models, two supervised and one unsupervised, were used to establish three different sets of wrist-worn accelerometer VM cut points as the new wrist methods to assess PA in preschoolers (Figure 2). Supervised ML models learn from hip-identified PA of each epoch and produce an inferred function that maps the wrist accelerometer count to a PA category; unsupervised models read the underlying structure of the wrist-worn accelerometer counts and detect the PA level of each count value [29]. PA intensity levels identified using hip-worn accelerometer cut points by Butte et al [24] were used as the hip reference to train the supervised ML models. The two supervised ML techniques were receiver operating characteristic (ROC) analysis and the ordinal logistic regression (OLR) model. Since ROC analysis was designed to predict binary outcomes, it was run three times to establish the upper threshold for sedentary behavior and the lower thresholds for LPA and VPA [30]. After ROC analysis calculated and compared sensitivity values for all possible threshold values, we selected thresholds based on the minimum difference between sensitivity and specificity [31]. For the OLR method, after being trained by the PA intensity levels as predicted by the hip reference, the newly constructed model calculated and compared the probability of each VM count value being classified into different PA intensity levels and assigned each count to the PA level with the highest predicted probability. K-means cluster analysis was the unsupervised learning approach used to separate each 15-second epoch for each participant into four distinct clusters, based solely on the input VM count data [32-34]. The number of clusters (k=4) was determined a priori because the four activity states of sedentary behavior, LPA, MPA, and VPA were known. Sedentary behavior cut points for OLR and k-means cluster analysis were determined based on the maximum count value within the sedentary behavior category; MPA and VPA cut points were determined based on the minimum count values within these two PA intensity categories; LPA was further determined based on the sedentary behavior and MPA cut points.

Figure 2. Modeling process diagram.



Statistical Analysis

The VM count for each epoch for each participant was categorized into a PA level based on each of the three sets of newly established wrist-worn accelerometer cut points and the hip reference, resulting in four separate PA level designations. Standard classification measures of sensitivity, specificity, false-negative rate, and false-positive rate were calculated to assess the discriminative ability of each method to correctly classify PA levels. Cohen kappa values were calculated to test the agreement between hip- and four wrist-derived measures, k-means, ROC, OLR, and Johansson's cut points. Daily amount of time in each PA level was also calculated and compared against PA estimates from the hip-worn accelerometer. Univariate analysis of covariance was conducted to compare mean daily time in each PA intensity level for each of the wrist-worn ML-based cut points versus the hip reference and the Johansson's cut points.

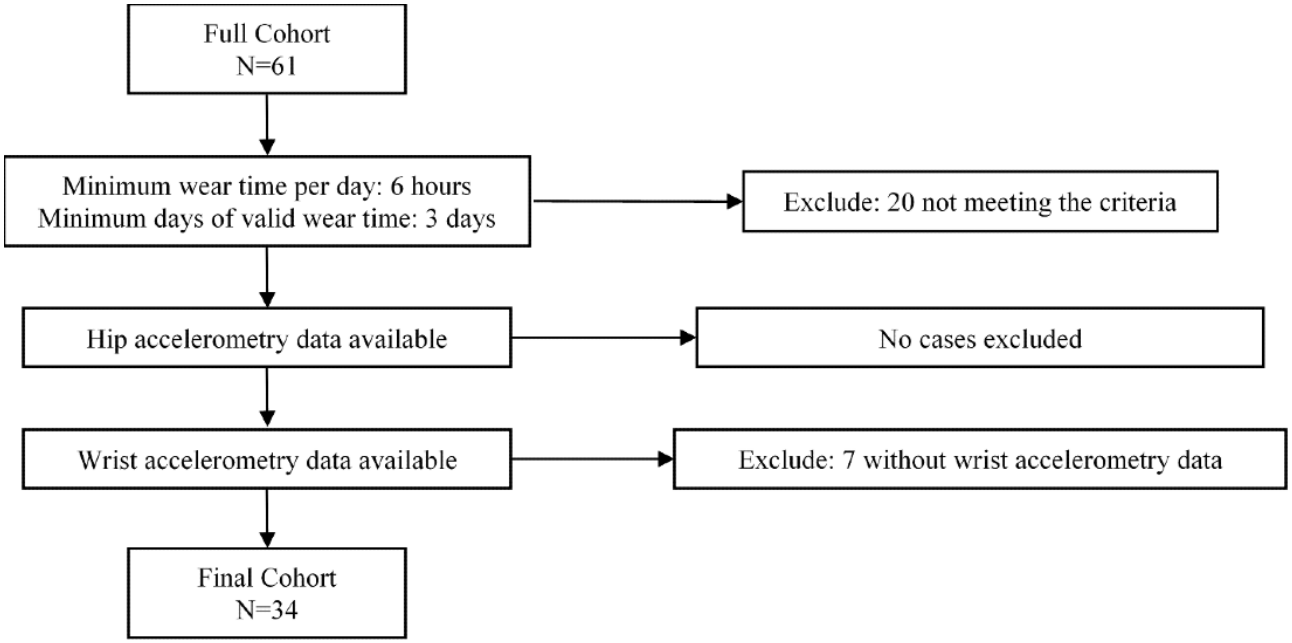
Results

All study participants were 3-5 years old, and more than 80% (29/34) of them were of Hispanic ethnicity (Table 1). Hip accelerometer identified 64.2% activity counts representing sedentary behavior and nearly 8.0% representing MVPA. Matched hip- and wrist-worn accelerometer data were collected and analyzed from 34 participants, yielding a total of 122,399 epochs (Figure 3).

Table 1. Descriptive characteristics of the study participants (N=34).

Variables	Total
Female, n (%)	20 (58.8)
Hispanic race, n (%)	29 (85.3)
Age (years), mean (SD)	3.97 (0.49)
Height (cm), mean (SD)	100.80 (4.70)
Weight (kg), mean (SD)	16.38 (2.157)
Hip-based activity counts in each physical activity level, n (%)	
Sedentary behavior	78,538 (64.2)
Light physical activity	34,243 (28.0)
Moderate physical activity	6784 (5.5)
Vigorous physical activity	2834 (2.3)
Moderate-to-vigorous physical activity	9618 (7.8)

Figure 3. Data flow diagram.



The three ML models grouped accelerometer counts differently and developed three sets of wrist-worn accelerometer VM cut points (lower and upper bounds; Table 2). For each ML model, mean count values increased as the PA intensity level increased, as expected.

Table 2. Wrist-worn accelerometer VM cut points established by each ML model and mean count value within each PA category.

Model	N	Cut points (cpm[a])		Vector magnitude counts (cpm)
		Lower Bound	Upper Bound	Mean (SD)
Receiver operating characteristic analysis				
Sedentary behavior	70,848	0	3406	744.40 (1077.67)
Light physical activity	20,194	3407	5690	4509.79 (652.33)
Moderate physical activity	4009	5691	6219	5952.98 (150.75)
Vigorous physical activity	27,348	6220	∞	10,054.39 (4267.24)
Ordinal logistic regression model				
Sedentary behavior	92,143	0	5837	1629.59 (1900.25)
Light physical activity	26,746	5838	14,020	8436.16 (2028.01)
Moderate physical activity	1688	14,021	17,432	15,481.18 (968.54)
Vigorous physical activity	1822	17,433	∞	22,350.31(4593.46)
K-means analysis				
Sedentary behavior	62,815	0	2556	457.57 (759.38)
Light physical activity	37,876	2557	7067	4655.99 (1270.28)
Moderate physical activity	18,559	7068	14,535	9474.95 (1895.16)
Vigorous physical activity	3149	14,536	∞	19,595.06 (4787.10)

[a]cpm: counts per minute.

Agreement Between Each Wrist Method and the Hip Reference

Agreement between the wrist methods and the hip reference at epoch level is presented at Table 3. When grouping the MPA and VPA, the prediction accuracy of each wrist method was more comparable with the hip reference, based on a higher classification rate and kappa value.

Table 3. Agreement of each wrist cut point compared to the hip reference.

Activity intensity	Johansson Wrist	ROC[a] analysis	OLR[b] model	K-means analysis
Sedentary behavior				
Sensitivity (%)	81.36	77.82	90.95	71.64
Specificity (%)	72.77	77.82	52.78	85.07
FPR (%)	27.23	22.18	47.22	14.93
FNR (%)	18.64	22.18	9.05	28.36
Kappa	0.53	0.54	0.47	0.53
Light physical activity				
Sensitivity (%)	67.34	26.80	43.27	50.90
Specificity (%)	75.69	87.50	86.47	76.81
FPR (%)	24.31	12.50	56.73	23.19
FNR (%)	32.66	73.20	13.53	49.10
Kappa	0.39	0.16	0.32	0.27
Moderate physical activity				
Sensitivity (%)	N/A[c]	4.47	8.83	48.42
Specificity (%)	N/A	96.79	99.06	86.79
FPR (%)	N/A	3.21	0.96	13.21
FNR (%)	N/A	95.53	76.11	51.58
Kappa	N/A	0.02	0.12	0.19
Vigorous physical activity				
Sensitivity (%)	N/A	79.00	23.89	33.94
Specificity (%)	N/A	79.00	99.04	98.17
FPR (%)	N/A	21.00	0.96	1.83
FNR (%)	N/A	21.00	76.11	66.06
Kappa	N/A	0.11	0.28	0.31
Moderate-to-vigorous physical activity				
Sensitivity (%)	16.02	78.89	24.27	68.37
Specificity (%)	99.53	78.92	98.96	86.58
FPR (%)	0.47	21.08	1.04	13.42
FNR (%)	83.98	21.11	75.73	31.63
Kappa	0.24	0.28	0.33	0.35
Overall agreement				
Correct classification (%)	N/A	59.51	71.51	63.68
Kappa	N/A	0.30	0.37	0.37
Overall agreement when MPA and VPA were grouped together				
Correct classification (%)	72.30	63.63	72.37	65.58
Kappa	0.45	0.35	0.39	0.40

[a]ROC: Receiver operating characteristic.

[b]OLR: Ordinal logistic regression.

[c]N/A: Not applicable.

According to the kappa values, all three ML models performed well on identifying sedentary behavior. When compared against the hip reference, ROC analysis and k-means cluster analysis derived cut points with acceptable sensitivity and specificity values (both >70%) and were comparable to the performance of Johansson et al [25] cut points.

In terms of classifying LPA, only the k-means LPA cut point resulted with sensitivity and specificity values greater than 50%. While the specificity value for k-means LPA cut point (76.81%) was similar to the Johansson cut point (75.69%), it had a much lower sensitivity value.

None of the three wrist methods were able to distinguish MPA as indicated by the low sensitivity and kappa values, although the k-means MPA cut point had the highest sensitivity (48.42%) and kappa value (0.19). OLR and k-means cut points also had low sensitivity values for identifying VPA (23.89% for OLR, 33.94 for k-means). Although the ROC cut point exhibited sensitivity and specificity values of 79%, the low kappa value (0.11) indicated that there was a low agreement between this method and the hip reference. When MPA and VPA were grouped together, the k-means cut points demonstrated higher sensitivity (68.37%), specificity (86.58%), and kappa values (0.35) than the other two wrist methods.

In general, k-means cut points resulted in sensitivity and specificity values above 50% for predicting sedentary behavior, LPA, and MVPA, with an acceptable kappa value for overall agreement (0.40).

Physical Activity Estimates by Method

Table 4 presents daily amount of time and percent of time in each PA intensity level as assessed by the hip reference, Johansson et al [25] wrist VM cut points, and the three newly developed wrist methods. ROC and k-means sedentary behavior cut points were close to the hip reference on estimating sedentary behavior time. LPA estimates were similar among the Johansson et al [25] cut point, the hip reference, and the k-means wrist method. None of the three wrist methods were comparable to the hip reference on estimating MPA and MVPA. Univariate ANOVA showed a similar VPA estimates for the k-means wrist method and the hip reference.

Table 4. Daily time in each Physical Activity intensity level (%) by different sets of cut points.

Activity intensity	Hip reference		Johansson wrist		ROC[a] analysis		OLR[b] model		K-means analysis	
	Mean (SD)	%	Mean (SD)	%	Mean (SD)	%	Mean (SD)	%	Mean (SD)	%
Sedentary behavior	148.07 (32.15)	57.8[c]	194.81 (41.58)	63.9	157.95 (37.88)	58.5[c]	208.61 (33.65)	76.1	138.13 (37.63)	51.6[c]
Light physical activity	83.13 (25.71)	30.4[d]	100.84 (39.53)	34.6[d]	47.47 (12.97)	16.7	60.11 (26.45)	21.2	89.25 (23.68)	31.3[d]
Moderate physical activity	19.27 (10.95)	7.1	N/A	N/A	9.20 (3.44)	3.3	3.73 (2.18)	1.3	41.46 (19.87)	14.7
Vigorous physical activity	13.24 (20.74)	4.7[e]	N/A	N/A	60.84 (28.09)	21.5	3.82 (2.88)	1.4	6.74 (4.42)	2.4[e]
Moderate-to-vigorous physical activity	32.61 (30.78)	11.8	4.35 (3.16)	1.5	70.98 (61.63)	24.8	7.62 (4.77)	2.7	48.87 (23.47)	17.1

[a]ROC: Receiver operating characteristic.

[b]OLR: Ordinal logistic regression.

[c] Indicates there is no statistical difference in PA estimate between the ML method and the hip-based reference for the Sedentary behavior intensity level.

[d] Indicates there is no statistical difference in PA estimate between the ML method and the hip-based reference for the Sedentary behavior intensity level.

[e] Indicates there is no statistical difference in PA estimate between the ML method and the hip-based reference for the Sedentary behavior intensity level.

Discussion

This study applied wrist derived VM data to assess PA in preschoolers based on three ML models. In the supervised ML models, the ROC analysis and OLR model were able to distinguish VM counts into each PA intensity level by reading the intensity label of each epoch as assigned by the gold standard hip reference. As an unsupervised ML model, k-means cluster analysis successfully grouped accelerometer count values into four PA clusters. When examining the agreement and comparing PA estimates from the wrist data compared with the hip reference, the k-means cluster analysis had the best performance among three ML models tested. Additionally, the k-means method of assigning PA levels had better agreement with the hip reference than the previously published cut points developed by Johansson et al [25].

Our results showed that cut points derived from the k-means cluster analysis produced better agreement with the gold standard hip reference than the other two ML approaches across all sedentary behavior and PA intensity levels, as indicated by sensitivity, specificity, and kappa values. K-means sedentary behavior, LPA, and VPA cut points showed the highest similarity to the hip reference on estimating PA time. Additionally, the estimated percent time in sedentary behavior, LPA, and MVPA by the k-means wrist reference method was comparable to the findings by Jones et al [35], who adopted a hip-worn accelerometry protocol for assessing PA in preschoolers similar to the one used in this study.

One reason that the k-means method had superior performance could be that the use of unlabeled data for calibration may result in less biased cut points for classifying epochs into PA levels [36]. During the data training process for the supervised learning

models, the hip reference-based PA levels were used as "labels" or "targets," which the models are trying to use as a basis to classify each epoch into the "true PA state." Since most of the epochs in our calibration data were observed in the sedentary state based on the hip reference, this could result in biasing the supervised model cut points toward lower activity levels. Thus, the lower activity levels might overly influence supervised learning models when developing cut points for wrist-worn accelerometers.

Similarly, the nature of the activities generating the calibration data may also influence the performance of the machine learning methods. For example, Butte et al [24] used 11 structured activities as the ground truth activity to calibrate the cut points for the hip reference [37]. However, others have found that using free-living activities for accelerometer calibration might generate higher counts per minute cut points than structured activities [38]. This might explain the higher MVPA cut point (≥16716 cpm) developed by Johansson et al [25], which incorporated free-play sessions during the calibration process, compared to our calibration data, which used only structured activities. Regardless, the k-means method showed superior performance in accurately assessing PA levels compared to the ROC, OLR and Johansson et al [25] methods. Thus, the k-means approach represents an improvement on existing methods for establishing wrist-based PA levels among pre-school aged children.

Our results showed that all wrist-based methods had difficulty in accurately assessing MPA. One possible explanation for this issue is that there is no consensus on what types of activities can represent MPA for this age group [25]. For example, the Johansson study [25] used a ball-toss activity, a 10-minute active video game, a 15-minute dancing activity, and an aerobic video activity to represent MPA, whereas Pate et al [39] defined the intensity between slow and brisk walking as the cut point for MPA and Sirard et al [40] used fast walking at 4.3 (SD 0.6) km/h as their criterion activity to represent MPA. Therefore, combined criterion-based and free-living activities for the generation of model training data may better reflect the full range of MPA, which could result in improved calibration processes for children in this age group. Future studies would also benefit from additional analysis of raw accelerometry data from wrist-worn devices with unsupervised learning methods.

Limitations

This study has several limitations. First, participants in this study were from low-income minority families and, therefore, accelerometer counts collected from the study sample might not represent the activity patterns of the general population. Previous studies have shown that young children from low-income minority families had a significantly lower motor performance and are less active during a preschool day compared to children from higher income families [41,42]. However, previous research found no difference in cut points in older children and adults from different income and socioeconomic background [13]. Second, this study chose a 30 Hertz sampling rate for both hip and wrist accelerometer placement sites to collect raw acceleration data to make the results comparable to other studies, but this approach might lead to an underestimate of activity intensity [13]. For instance, Clevenger et al [43] found that greater sampling rate resulted in a higher estimation of high-intensity PA in preschoolers for both hip- and wrist-worn accelerometers, even though the difference was not significant. Thus, the study should be replicated using higher sampling rate magnitude. Another limitation is that this study collected data during school hours on 3 days, which may not reflect the general PA patterns of preschool-aged children. Hesketh et al [44] found that children were more active in daycare than at home and were more active during weekdays than weekends days. In this case, collecting accelerometry data from both weekdays and weekend days based on a 24-hour accelerometry protocol would reflect the PA pattern in this age group more accurately. Finally, the cut points from ML models were not validated against ground truth activities that can substantiate the accuracy of the wrist-derived data [16].

Conclusion

This study demonstrated the potential of ML techniques to distinguish PA intensity levels, with the exception of MPA, in preschool-age children. Cut point established from k-means cluster analysis was comparable to the hip-reference criterion in predicting sedentary behavior. Although PA estimates from k-means cluster analysis of wrist-worn accelerometers were acceptable as compared to the hip reference, the finding needs to be replicated using ground truth activities in free-living setting.

Acknowledgments

This study is funded by the United States National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases (R01DK109323). Parent/Child Incorporated Of San Antonio And Bexar and Family Service Association of San Antonio, Inc, are the study collaborators who administer Head Start Programs in San Antonio, Texas. Our heartfelt appreciation goes to the children, parents, and staff of Head Start Programs who participated in our study. Finally, we want to thank our research staff, undergraduate intern students, and graduate students for their hard work and contribution to the study.

Authors' Contributions

ZY and SL contributed to the design and implementation of the study. JH and SL conducted the statistical analyses of this work. SL took the lead in writing manuscript. All authors participated in the implementation of the study protocol, helped interpret the results, contributed to the discussion and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Matthews CE, Hagströmer M, Pober DM, Bowles HR. Best practices for using physical activity monitors in population-based research. Med Sci Sports Exerc 2012 Jan;44(1 Suppl 1):S68-S76 [FREE Full text] [doi: 10.1249/MSS.0b013e3182399e5b] [Medline: 22157777]
2. Freedson P, Bowles HR, Troiano R, Haskell W. Assessment of physical activity using wearable monitors: recommendations for monitor calibration and use in the field. Med Sci Sports Exerc 2012 Jan;44(1 Suppl 1):S1-S4 [FREE Full text] [doi: 10.1249/MSS.0b013e3182399b7e] [Medline: 22157769]
3. Bornstein DB, Beets MW, Byun W, McIver K. Accelerometer-derived physical activity levels of preschoolers: a meta-analysis. J Sci Med Sport 2011 Nov;14(6):504-511. [doi: 10.1016/j.jsams.2011.05.007] [Medline: 21684809]
4. Chen KY, Bassett DR. The technology of accelerometry-based activity monitors: current and future. Med Sci Sports Exerc 2005 Nov;37(11 Suppl):S490-S500. [doi: 10.1249/01.mss.0000185571.49104.82] [Medline: 16294112]
5. Bochniewicz EM, Emmer G, McLeod A, Barth J, Dromerick AW, Lum P. Measuring Functional Arm Movement after Stroke Using a Single Wrist-Worn Sensor and Machine Learning. J Stroke Cerebrovasc Dis 2017 Dec;26(12):2880-2887. [doi: 10.1016/j.jstrokecerebrovasdis.2017.07.004] [Medline: 28781056]
6. Ellis K, Kerr J, Godbole S, Staudenmayer J, Lanckriet G. Hip and Wrist Accelerometer Algorithms for Free-Living Behavior Classification. Med Sci Sports Exerc 2016 May;48(5):933-940 [FREE Full text] [doi: 10.1249/MSS.0000000000000840] [Medline: 26673126]
7. Huberty J, Ehlers DK, Kurka J, Ainsworth B, Buman M. Feasibility of three wearable sensors for 24 hour monitoring in middle-aged women. BMC Womens Health 2015 Jul 30;15:55 [FREE Full text] [doi: 10.1186/s12905-015-0212-3] [Medline: 26223521]
8. Full KM, Kerr J, Grandner MA, Malhotra A, Moran K, Godoble S, et al. Validation of a physical activity accelerometer device worn on the hip and wrist against polysomnography. Sleep Health 2018 Apr;4(2):209-216 [FREE Full text] [doi: 10.1016/j.sleh.2017.12.007] [Medline: 29555136]
9. Kwon S, Zavos P, Nickele K, Sugianto A, Albert MV. Hip and Wrist-Worn Accelerometer Data Analysis for Toddler Activities. Int J Environ Res Public Health 2019 Jul 21;16(14) [FREE Full text] [doi: 10.3390/ijerph16142598] [Medline: 31330889]
10. Fairclough SJ, Noonan R, Rowlands AV, Van Hees V, Knowles Z, Boddy LM. Wear compliance and activity in children wearing wrist- and hip-mounted accelerometers. Med Sci Sports Exerc 2016 Feb;48(2):245-253. [doi: 10.1249/MSS.0000000000000771] [Medline: 26375253]
11. Kim Y, Beets MW, Welk GJ. Everything you wanted to know about selecting the "right" Actigraph accelerometer cut-points for youth, but…: a systematic review. J Sci Med Sport 2012 Jul;15(4):311-321. [doi: 10.1016/j.jsams.2011.12.001] [Medline: 22306372]
12. Cain KL, Sallis JF, Conway TL, Van Dyck D, Calhoon L. Using accelerometers in youth physical activity studies: a review of methods. J Phys Act Health 2013 Mar;10(3):437-450 [FREE Full text] [doi: 10.1123/jpah.10.3.437] [Medline: 23620392]
13. Migueles JH, Cadenas-Sanchez C, Ekelund U, Delisle Nyström C, Mora-Gonzalez J, Löf M, et al. Accelerometer Data Collection and Processing Criteria to Assess Physical Activity and Other Outcomes: A Systematic Review and Practical Considerations. Sports Med 2017 Sep;47(9):1821-1845 [FREE Full text] [doi: 10.1007/s40279-017-0716-0] [Medline: 28303543]
14. Johansson E, Ekelund U, Nero H, Marcus C, Hagströmer M. Calibration and cross-validation of a wrist-worn Actigraph in young preschoolers. Pediatr Obes 2015 Feb;10(1):1-6. [doi: 10.1111/j.2047-6310.2013.00213.x] [Medline: 24408275]
15. Hislop J, Palmer N, Anand P, Aldin T. Validity of wrist worn accelerometers and comparability between hip and wrist placement sites in estimating physical activity behaviour in preschool children. Physiol Meas 2016 Oct;37(10):1701-1714. [doi: 10.1088/0967-3334/37/10/1701] [Medline: 27653188]
16. Farrahi V, Niemelä M, Kangas M, Korpelainen R, Jämsä T. Calibration and validation of accelerometer-based activity monitors: A systematic review of machine-learning approaches. Gait Posture 2019 Feb;68:285-299 [FREE Full text] [doi: 10.1016/j.gaitpost.2018.12.003] [Medline: 30579037]
17. Koza J, Bennett III FH, Andre D, Keane MA. Automated design of both the topology and sizing of analog electrical circuits using genetic programming. In: Gero JS, Sudweeks F, editors. Artificial intelligence in design '96. Dordrecht: Kluwer Academic Publishers; 1996:151-170.
18. Ellis K, Kerr J, Godbole S, Lanckriet G, Wing D, Marshall S. A random forest classifier for the prediction of energy expenditure and type of physical activity from wrist and hip accelerometers. Physiol Meas 2014 Nov;35(11):2191-2203 [FREE Full text] [doi: 10.1088/0967-3334/35/11/2191] [Medline: 25340969]
19. Narayanan A, Desai F, Stewart T, Duncan S, Mackay L. Application of raw accelerometer data and machine-learning techniques to characterize human movement behavior: a systematic scoping review. J Phys Act Health 2020 Mar 01;17(3):360-383. [doi: 10.1123/jpah.2019-0088] [Medline: 32035416]

20. Scott JJ, Rowlands AV, Cliff DP, Morgan PJ, Plotnikoff RC, Lubans DR. Comparability and feasibility of wrist- and hip-worn accelerometers in free-living adolescents. J Sci Med Sport 2017 Dec;20(12):1101-1106. [doi: 10.1016/j.jsams.2017.04.017] [Medline: 28501418]
21. Tudor-Locke C, Barreira TV, Schuna JM, Mire EF, Chaput J, Fogelholm M, ISCOLE Research Group. Improving wear time compliance with a 24-hour waist-worn accelerometer protocol in the International Study of Childhood Obesity, Lifestyle and the Environment (ISCOLE). Int J Behav Nutr Phys Act 2015 Feb 11;12:11 [FREE Full text] [doi: 10.1186/s12966-015-0172-x] [Medline: 25881074]
22. Hinkley T, Crawford D, Salmon J, Okely AD, Hesketh K. Preschool children and physical activity: a review of correlates. Am J Prev Med 2008 May;34(5):435-441. [doi: 10.1016/j.amepre.2008.02.001] [Medline: 18407012]
23. Rowlands A. Moving forward with accelerometer-assessed physical activity: two strategies to ensure meaningful, interpretable, and comparable measures. Pediatr Exerc Sci 2018 Nov 01;30(4):450-456. [doi: 10.1123/pes.2018-0201] [Medline: 30304982]
24. Butte NF, Wong WW, Lee JS, Adolph AL, Puyau MR, Zakeri IF. Prediction of energy expenditure and physical activity in preschoolers. Med Sci Sports Exerc 2014 Jun;46(6):1216-1226 [FREE Full text] [doi: 10.1249/MSS.0000000000000209] [Medline: 24195866]
25. Johansson E, Larisch L, Marcus C, Hagströmer M. Calibration and validation of a wrist- and hip-worn actigraph accelerometer in 4-year-old children. PLoS One 2016;11(9):e0162436 [FREE Full text] [doi: 10.1371/journal.pone.0162436] [Medline: 27617962]
26. Kingsley MI, Nawaratne R, O'Halloran PD, Montoye AH, Alahakoon D, De Silva D, et al. Wrist-specific accelerometry methods for estimating free-living physical activity. J Sci Med Sport 2019 Jun;22(6):677-683. [doi: 10.1016/j.jsams.2018.12.003] [Medline: 30558904]
27. Cliff DP, Reilly JJ, Okely AD. Methodological considerations in using accelerometers to assess habitual physical activity in children aged 0-5 years. J Sci Med Sport 2009 Sep;12(5):557-567. [doi: 10.1016/j.jsams.2008.10.008] [Medline: 19147404]
28. Choi L, Liu Z, Matthews CE, Buchowski MS. Validation of accelerometer wear and nonwear time classification algorithm. Med Sci Sports Exerc 2011 Feb;43(2):357-364 [FREE Full text] [doi: 10.1249/MSS.0b013e3181ed61a3] [Medline: 20581716]
29. Alpaydin E. Introduction to machine learning. Cambridge: The MIT Press; 2020.
30. Zou KH, O'Malley AJ, Mauri L. Receiver-operating characteristic analysis for evaluating diagnostic tests and predictive models. Circulation 2007 Feb 06;115(5):654-657. [doi: 10.1161/CIRCULATIONAHA.105.594929] [Medline: 17283280]
31. Alhassan S, Robinson TN. Defining accelerometer thresholds for physical activity in girls using ROC analysis. J Phys Act Health 2010 Jan;7(1):45-53 [FREE Full text] [doi: 10.1123/jpah.7.1.45] [Medline: 20231754]
32. Fraley C. How many clusters? Which clustering method? Answers via model-based cluster analysis. Comput J 1998 Aug 01;41(8):578-588. [doi: 10.1093/comjnl/41.8.578]
33. Jain AK. Data clustering: 50 years beyond K-means. Pattern Recognit Lett 2010 Jun;31(8):651-666. [doi: 10.1016/j.patrec.2009.09.011]
34. Wilks DS. Statistical methods in the atmospheric sciences. Cambridge: Academic Press; 2011.
35. Jones RA, Okely AD, Hinkley T, Batterham M, Burke C. Promoting gross motor skills and physical activity in childcare: A translational randomized controlled trial. J Sci Med Sport 2016 Sep;19(9):744-749. [doi: 10.1016/j.jsams.2015.10.006] [Medline: 26774378]
36. van Kuppevelt D, Heywood J, Hamer M, Sabia S, Fitzsimons E, van Hees V. Segmenting accelerometer data from daily life with unsupervised machine learning. PLoS One 2019;14(1):e0208692 [FREE Full text] [doi: 10.1371/journal.pone.0208692] [Medline: 30625153]
37. Zakeri I, Adolph AL, Puyau MR, Vohra FA, Butte NF. Cross-sectional time series and multivariate adaptive regression splines models using accelerometry and heart rate predict energy expenditure of preschoolers. J Nutr 2013 Jan;143(1):114-122 [FREE Full text] [doi: 10.3945/jn.112.168542] [Medline: 23190760]
38. van Cauwenberghe E, Labarque V, Trost SG, de Bourdeaudhuij I, Cardon G. Calibration and comparison of accelerometer cut points in preschool children. Int J Pediatr Obes 2011 Jun;6(2-2):e582-e589. [doi: 10.3109/17477166.2010.526223] [Medline: 21121867]
39. Pate RR, Almeida MJ, McIver KL, Pfeiffer KA, Dowda M. Validation and calibration of an accelerometer in preschool children. Obesity (Silver Spring) 2006 Nov;14(11):2000-2006 [FREE Full text] [doi: 10.1038/oby.2006.234] [Medline: 17135617]
40. Sirard JR, Trost SG, Pfeiffer KA, Dowda M, Pate RR. Calibration and evaluation of an objective measure of physical activity in preschool children. J Phys Act Health 2005 Jul;2(3):345-357. [doi: 10.1123/jpah.2.3.345]
41. Iivonen S, Sääkslahti A, Mehtälä A, Villberg J, Soini A, Poskiparta M. Directly observed physical activity and fundamental motor skills in four-year-old children in day care. Eur Early Child Educ Res J 2016 Apr 21;24(3):398-413. [doi: 10.1080/1350293x.2016.1164398]
42. Innella N, Breitenstein S, Hamilton R, Reed M, McNaughton DB. Determinants of obesity in the hispanic preschool population: an integrative review. Public Health Nurs 2016 May;33(3):189-199. [doi: 10.1111/phn.12215] [Medline: 26118340]

43. Clevenger KA, Pfeiffer KA, Mackintosh KA, McNarry MA, Brønd J, Arvidsson D, et al. Effect of sampling rate on acceleration and counts of hip- and wrist-worn ActiGraph accelerometers in children. Physiol Meas 2019 Sep 30;40(9):095008. [doi: 10.1088/1361-6579/ab444b] [Medline: 31518999]
44. Hesketh KR, Griffin SJ, van Sluijs EMF. UK Preschool-aged children's physical activity levels in childcare and at home: a cross-sectional exploration. Int J Behav Nutr Phys Act 2015 Sep 26;12(1). [doi: 10.1186/s12966-015-0286-1]

Abbreviations

LPA: light physical activity
ML: machine learning
MPA: moderate physical activity
MVPA: moderate-to-vigorous physical activity
OLR: ordinal logistic regression
PA: physical activity
ROC: receiver operating characteristic
VPA: vigorous physical activity

Edited by G Eysenbach; submitted 17.10.19; peer-reviewed by K Tamura, J Migueles, M Chinapaw, E Shiroma; comments to author 30.03.20; revised version received 27.05.20; accepted 13.06.20; published 31.08.20.

Please cite as:
Li S, Howard JT, Sosa ET, Cordova A, Parra-Medina D, Yin Z
Calibrating Wrist-Worn Accelerometers for Physical Activity Assessment in Preschoolers: Machine Learning Approaches
JMIR Form Res 2020;4(8):e16727
URL: https://formative.jmir.org/2020/8/e16727
doi:10.2196/16727
PMID:32667893

Original Paper

Occupation Coding of Job Titles: Iterative Development of an Automated Coding Algorithm for the Canadian National Occupation Classification (ACA-NOC)

Hongchang Bao[1,2*], BSc, DA, MBA; Christopher J O Baker[1,3*], BSc (Hons), PhD, DUT; Anil Adisesh[4,5,6*], MB ChB, MSc, MD, FRCP, FFOM, FRCPC

[1]Department of Computer Science, Faculty of Science, Applied Science and Engineering, University of New Brunswick, Saint John, NB, Canada

[2]Department of Computing Science, University of Alberta, Edmonton, AB, Canada

[3]IPSNP Computing Inc, Saint John, NB, Canada

[4]Division of Occupational Medicine, Department of Medicine, University of Toronto, Toronto, ON, Canada

[5]Division of Occupational Medicine, St Michael's Hospital, Toronto, ON, Canada

[6]Faculty of Business, University of New Brunswick, Saint John, NB, Canada

[*]all authors contributed equally

Corresponding Author:
Christopher J O Baker, BSc (Hons), PhD, DUT
Department of Computer Science
Faculty of Science, Applied Science and Engineering
University of New Brunswick
100 Tucker Park Rd
Saint John, NB, E2K5E2
Canada
Phone: 1 (506) 648 2302
Email: bakerc@unb.ca

Abstract

Background: In many research studies, the identification of social determinants is an important activity, in particular, information about occupations is frequently added to existing patient data. Such information is usually solicited during interviews with open-ended questions such as "What is your job?" and "What industry sector do you work in?" Before being able to use this information for further analysis, the responses need to be categorized using a coding system, such as the Canadian National Occupational Classification (NOC). Manual coding is the usual method, which is a time-consuming and error-prone activity, suitable for automation.

Objective: This study aims to facilitate automated coding by introducing a rigorous algorithm that will be able to identify the NOC (2016) codes using only job title and industry information as input. Using manually coded data sets, we sought to benchmark and iteratively improve the performance of the algorithm.

Methods: We developed the ACA-NOC algorithm based on the NOC (2016), which allowed users to match NOC codes with job and industry titles. We employed several different search strategies in the ACA-NOC algorithm to find the best match, including exact search, minor exact search, like search, near (same order) search, near (different order) search, any search, and weak match search. In addition, a filtering step based on the hierarchical structure of the NOC data was applied to the algorithm to select the best matching codes.

Results: The ACA-NOC was applied to over 500 manually coded job and industry titles. The accuracy rate at the four-digit NOC code level was 58.7% (332/566) and improved when broader job categories were considered (65.0% at the three-digit NOC code level, 72.3% at the two-digit NOC code level, and 81.6% at the one-digit NOC code level).

Conclusions: The ACA-NOC is a rigorous algorithm for automatically coding the Canadian NOC system and has been evaluated using real-world data. It allows researchers to code moderate-sized data sets with occupation in a timely and cost-efficient manner such that further analytics are possible. Initial assessments indicate that it has state-of-the-art performance and is readily extensible upon further benchmarking on larger data sets.

(JMIR Form Res 2020;4(8):e16422) doi:10.2196/16422

KEYWORDS

occupation coding; automated coding; occupational health; job title

Introduction

In many research studies, and for governmental or other statistical purposes, data collection includes gathering information on occupation. Occupation is a widely used explanatory variable in health research, representing social status and class as well as exposure to environmental hazards [1]. Typically, such data are collected either by a self-completed questionnaire or by an interviewer. In either case, to be useful for secondary reuse in health care analytics, a coding system is applied to translate the job titles into meaningful categories and then match them to the appropriate standard codes. Manual coding is generally considered to be the most reliable approach; however, manual coding is very expensive and time-consuming and requires considerable expertise. It can, from our experience, take a manual coder 3-5 days to code 500 job titles.

The Canadian National Occupational Classification (NOC) 2016 is a 4-tiered hierarchical arrangement of occupational groups with successive levels of disaggregation [2]. It has 500 predefined occupational codes and unit groups that code >30,000 occupational job titles, although different unit groups may contain similar job titles. Occupational coding is therefore a complex task, with the possibility of different professional coders coding differently, sometimes developing individual coding preferences [3,4]. The agreement of manual occupation coders has been reported to range from 44% to 89% at the four-digit level [4]. Efforts have been made to automate such coding. Gweon [3] reported that researchers have implemented 2 kinds of automated coding approaches: a data-based approach and a rule-based approach.

A data-based approach involves using machine learning algorithms to create models from manually coded training data. Once the model has been trained, new data can be coded automatically [3]. Several different algorithms have been used for occupation coding in this manner. For example, Bethmann et al [5] employed 2 machine learning algorithms (Naive Bayes and Bayesian Multinomial) for occupation coding using a data set of 300,000 coded answers from the German National Educational Panel study. The approach used by Bethmann et al [5] was further deployed by Schierholz et al [6] to code 2 survey data sets from the German Institute for Employment Research. Gweon [3] proposed 3 methods for automatic coding: (1) combining separate models for detailed occupation codes into aggregate occupation codes, (2) a hybrid method that combines the duplicate-based approach with a machine learning algorithm, and (3) a modified nearest neighbor approach. These same authors [3] used data from the German General Social Survey (ALLBUS). In the United States, Russ et al [7] developed a stacked ensemble algorithm trained with 14,983 manually coded job titles to assign US SOC (Standard Occupational Classification) 2010 occupation codes based on job title, task, and industry input data. Nahoomi [8] used 65,962 SOC-coded job titles to report that support vector machines (SVMs) and convolutional neural networks (CNNs) have similar performance but perform better than Naive Bayes.

A rule-based approach involves building a classification algorithm based on several rules created by experts after analyzing manually coded data. Depending on their design, the application of rules can lead to the assignment of multiple classes to an input and a filter algorithm is applied to identify the best match. There exist several rule-based methods for automated matching of text in a hierarchical classification system. In one approach, an initial match between the input data and the text description of a top-level class triggers further matchmaking at lower levels in the hierarchy. Another method is to directly compare the input data with each of the hierarchical levels [9]. In the same study, Burstyn et al [9] applied an algorithm that mixes both methods to code against the SOC (2010) system. Schierholz [6] has since reported that this approach rarely coded >50% of records accurately.

A review of automated occupation coding system performance carried out by Nahoomi [8] considered the full range of approaches, including machine learning, hybrid, and rule-based approaches. It was found that occupation classification algorithms can deliver production rates of up to 100%, with accuracy levels ranging from 44% to 98%. The best dictionary-based approach produced a 43% production rate but a 94% accuracy, whereas machine learning approaches were able to deliver a 100% production rate and a maximum accuracy of 80%. The hybrid machine learning/rule-based approaches had a production rate of 73% to 80% and an accuracy rate of 98% to 100%, relying on large data sets and >3000 rules. Overall, the accuracy depends on the type of algorithms used and the manner in which accuracy is computed. The review did not identify any systems targeted or tested on the Canadian NOC.

In this study, we report on the development of a hybrid algorithm that uses a rule-based method for search and matchmaking and a filtering step to select the best match. The algorithm is rigorous in matching input text with the hierarchy textual descriptions of the NOC, meaning it is designed to accommodate poor-quality input data, specifically spelling and typing errors. It adjusts for inserted symbols such as hyphenation, forward slashes, or punctuation errors in the input strings. Multiple search strategies are performed sequentially until job titles are matched and filtered and the best-fitting NOC code is identified. This paper aimed to document the incremental design of the algorithm based on performance benchmarking and an analysis of uncoded or incorrectly coded occupation data.

Methods

Data Resources

The Canadian NOC is the national reference for occupations in Canada, providing a standard taxonomy and a Canadian

framework for collecting, analyzing, and disseminating occupational data for labor market information and employment-related program administration. It comprises >30,000 occupational titles gathered into 500 unit groups, organized according to 4 skill levels and 10 skill types. Unit groups are based on the similarity of skills, defined primarily by functions and employment requirements. Each unit group describes the main duties and employment requirements and details examples of occupational titles. Each unit group has a unique four-digit code. The first 3 digits of this code indicate the major and minor groups to which the unit group belongs [2].

Table 1 lists the first level of NOC (2016). NOC (2016) is organized in a four-level hierarchy, and there are 10 broad occupational categories in the first level, 46 major groups in the second level, 140 minor groups in the third level, and 500 unit groups in the fourth level [2]. NOC major group 00: senior management occupations showing minor and unit groups are shown in Textbox 1.

Table 1. Occupational categories of the Canadian National Occupational Classification (first level).

NOC[a] code	Occupational categories
0	Management occupations
1	Business, finance, and administration occupations
2	Natural and applied sciences and related occupations
3	Health occupations
4	Occupations in education, law and social, community, and government services
5	Occupations in art, culture, recreation, and sport
6	Sales and service occupations
7	Trades, transport, and equipment operators and related occupations
8	Natural resources, agriculture, and related production occupations
9	Occupations in manufacturing and utilities

[a]NOC: National Occupational Classification.

Textbox 1. National Occupational Classification major group 00: senior management occupations showing minor and unit groups.

- 0 - management occupations
- 00 - senior management occupations
- 001 - legislators and senior management
- 0011 - legislators
- 0012 - senior government managers and officials
- 0013 - senior managers: financial, communications, and other business services
- 0014 - senior managers: health, education, social and community services, and membership organizations
- 0015 - senior managers: trade, broadcasting, and other services, n.e.c (not elsewhere classified)
- 0016 - senior managers: construction, transportation, production, and utilities

Multiple versions (2001, 2006, 2011, and 2016) of the NOC can be accessed through a web browser [10]. StatCan also provides programmatic access to the 2016 classification via web services including GetNOCStructure, GetSkillType, GetSkillLevel, GetMajorGroup, GetMinorGroup, GetUnitGroup, GetNOCTitles, GetNOCTitlesAutoCoding, GetNOCDescription, and GetNOCDescriptionByKeyword. Furthermore, the NOC database dump web method, DumpNOCDatabase, was available, enabling local query access to the categories.

Data Set

The initial data set used to benchmark the algorithm consisted of 566 job and industry titles with manual coding to the NOC (2016) and to the North American Industrial Classification (NAICS), respectively. These data were gathered as part of the Canadian Immunization Research Network Community Acquired Pneumonia study [11] to investigate occupational associations. The data were coded by 1 researcher and underwent validation checks by a second coder. Table 2 lists the input data set of this project that contains the fields *Current Job Title*, *NOC Code Manually*, and *Current Industry*. The *NOC Code Manually* was manually coded and entered based on the current job title and industry. The data on job title and industry originate from free text entered in response to the questions “What is your job title?” and “In which type of industry do you work?” with additional questions asking about the last job held for those not currently working and enquiry about the longest held post.

Table 2. Example of the source data sets.

Current job titles	NOC[a] codes manually	Current industries
Owner of cleaning business	0651	Service industry
Managing website	0213	Computer technology
Manager	0311	Health care

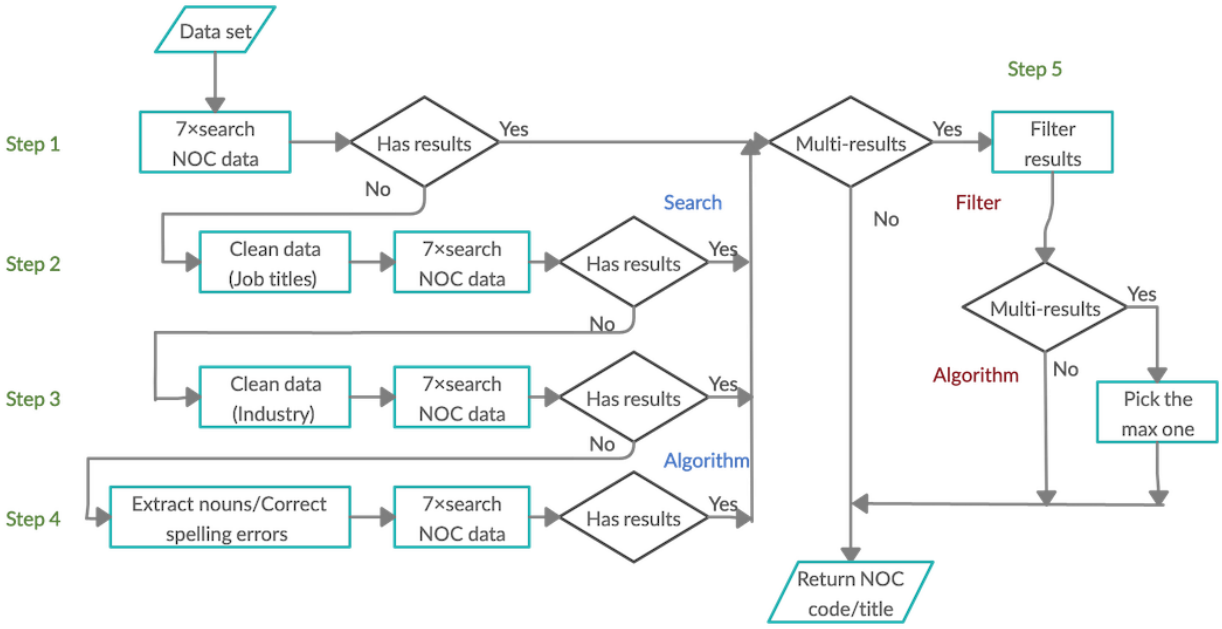
[a]NOC: National Occupational Classification.

The current job title and the current industry set are accessed from the data set, and the algorithm interrogates the local version of the NOC database to find the best matching NOC code. The NOC code in the data set is contrasted with the algorithmically derived code to arrive at a performance evaluation.

Algorithm

To develop an algorithm for automated occupation coding, direct access to the full NOC structure is required. Although manual coding relies on familiarization with a coding structure and computer-assisted keyword lookups through a web browser, we sought to use the NOC's web service access to look up codes for our sample data. In preliminary work, we assessed the accuracy with which we could identify the correct job codes using individual web service calls to the job title field only. Albeit limited, we recorded this as a baseline for subsequent performance comparisons. To improve the accuracy beyond using the web services, we initiated the design of the Automated Coding Algorithm (ACA)-NOC algorithm. To facilitate custom searches, the NOC data were downloaded and saved in a spreadsheet using the *NOC Database Dump Web Methods* provided by NOC. Thereafter, we designed a search algorithm, as shown in Figure 1, employing the 7 search strategies listed in the section *Search Strategies* to match text input strings with NOC titles and text descriptions in the locally stored data.

Figure 1. ACA-NOC algorithm. NOC: National Occupational Classification.



Search Algorithm

The search algorithm shown in Figure 1 comprises a sequence of steps that iteratively look for matches in the NOC until all the given job titles in the input data set are matched with one or more NOC codes. The initial step, step 1, runs multiple search strategies sequentially to match input data with the NOC data using either a job title or an industry description. Depending on the outcomes from these NOC data searches, matches to a single NOC code are archived or multiple NOC code matches are further filtered according to the NOC hierarchy and by industry, step 5. If no matches are generated, the input data are preprocessed (splitting job titles, removing stop words, stemming, and noun extraction from titles) in step 2 before NOC multisearches are reinitiated. If no NOC codes are matched at this stage, the input industry descriptions are preprocessed and NOC searches are reinitiated, step 3. In step 4, any input data that still fail to link to an NOC code are further processed to extract any nouns in the job title, and spelling checks and corrections are made. This is the final step in which NOC data searches are subsequently reinitiated.

ACA-NOC Algorithm

We applied multiple search strategies sequentially to match the input data with the NOC data.

1. Query the NOC job codes by exactly matching the input data with the NOC job titles in step 1.

2. If results are found, then initiate postprocessing in step 5. If not, then use the following procedure:
   - 2.1. Obtain results from the processed information (splitting job titles, removing stop words, stemming, and noun extraction from titles) in step 2.
   - 2.2. If results are found, then perform postprocessing in step 5. If not:
     - 2.2.1. Query NOC information with industry or split industry keywords in step 3.
     - 2.2.2. If results are found, then perform postprocessing in step 5. If not:
       - 2.2.2.1. Correct spelling errors for the input data. Then, query NOC data in step 2.2.2.2. If results are found, perform postprocessing in step 5. Otherwise, return none.
       - 2.2.2.2. If results are found, perform postprocessing in step 5. Otherwise, return none.

3. Postprocessing:
   - 3.1. Return the NOC code if a single result is found, else:
   - 3.2. Apply the filter algorithm to the multiple results and return the most pertinent NOC code in step 5.

Search Strategies

1. Exact: this search returns results that exactly match the input string(s) entered.
2. Minor exact: this search returns results that match exactly after correcting spelling errors in the input string(s).
3. Like: this search returns results that include every input string in the specific order as entered. The exact string may be included anywhere in the associated text.
4. Near (same order): this search returns results that include every input string matched in the same order.
5. Near (different order): this search returns results that include every input string matched in any order.
6. Any: this search returns results that include any/some of the string(s) entered.
7. Weak match: this search returns results that include any of the nouns found in the string(s) entered.

Filter Algorithm

The filter algorithm selects a single NOC code from a list of candidate NOC codes based on the frequency of the keyword (job title or industry title) in the NOC descriptions. The selection occurs in the following order:

1. Select the given job titles from the NOC skill-type names in the result sets.
2. Select the given job titles from the NOC major group names in the result sets.
3. Select the given job titles from the NOC minor group names in the result sets.
4. Select the given job titles from the NOC group titles in the result sets.
5. Select the industry title from the NOC job titles in the result sets.
6. Split the industry titles that have inserted symbols and search each of the job titles from the NOC job titles in the result sets. Then, return the data set with the highest frequency of the keyword.
7. Select the industry title from the NOC descriptions (lead statement, main duties, and employment requirements) in the result sets.

Results

Overview

The design of the ACA-NOC algorithm shown in Figure 1 was arrived at through an iterative process of design, deployment, and testing. Each generation of the algorithm was evaluated using the sample data set described in the section *Data Set*. In this section, we detail the different generations of the algorithm and the performance at each generation. Production rate and accuracy are metrics that were used for evaluating the performance, where the production rate is the proportion of observations that can be coded automatically. For a given production rate, accuracy is the proportion of codes that are coded correctly when compared with manual coding [3]. Accuracy can be assessed for each digit of the NOC code.

The ACA-NOC algorithm underwent 4 generations of design and evaluation. Table 3 lists the production rate and accuracy of each version of the algorithm compared with the simple web service lookup.

The first observation that could be made was that ACA-NOC generation 0 outperformed the web service on production rate, and a 100% production rate was achieved in ACA-NOC generation 1. Second, between ACA-NOC generation 0 and ACA-NOC generation 4, the accuracy in coding the NOC skill-type level (a one-digit code) improved from 67.3% to 73.3%, the NOC major group level (two-digit code) improved from 58.3% to 64.0%, the NOC minor group level (three-digit code) improved from 53.5% to 59.7%, and the NOC unit group level (four-digit code) improved from 51.2% to 55.5%. Expert review by a team member (author AA) familiar with occupational coding aided the development of the algorithm through identification of mismatched items and discussion within the development team of the likely reasons for error. These discussions led to modifications of the algorithm and further improved the performance. Finally, the expert review allowed the recognition that a number of mismatches did not necessarily indicate poor performance by the algorithm, but that, in some cases, the algorithm had identified a more accurate code. Specifically, human error caused data entry errors, and the algorithm code and manual code were both reasonable choices.

Table 3. Production rate and accuracy of occupation coding (NOC 2016) with the ACA-NOC algorithm by generation number (N=566).

Generation numbers	Production rate, %	Accuracy, %			
		1 digit	2 digits	3 digits	4 digits
Web service[a]	74.0	61.3	52.8	49.3	47.2
0	93.5	67.4	58.4	53.6	51.2
1	100	69.3	60.4	55.8	53.0
2	100	70.3	61.3	56.5	53.9
3	100	70.5	61.8	57.2	54.6
4	100	73.3	64.0	59.7	55.5
4 (following expert review of mismatches)	100	81.6	72.3	65.0	58.7
Mannetje and Kromhout [4] agreement rates (%) for occupation classification between manual coders	N/A[b]	75-97	61-92	56-80	44-89

[a]The method we used was *GetNOCTitlesAutoCoding*. This method returned a best match NOC code for the input titles. It queried the database titles tables first; if no result was found, then it queried database profile tables. The important parameters of the method are strTitle->job title, bytLang->0 (english), bytMultiTitles->0 (not required), bytGrouptitle->0 (not required).

[b]N/A: not applicable.

Generation 0

The first version of the algorithm included only steps 1 to 3. No hierarchical filtering steps were applied.

1. Step 1: if the entry of *Current Job Title* exists, then search for NOC job titles using the same. Otherwise, search for NOC job titles by industry keywords. If the search returns only 1 NOC code, then the search for this job title is completed.
2. Step 2: if the returned data set is *none*, then remove stop words or stem the job titles and search the NOC job titles using the modified job titles. If there are no results, search NOC descriptions using the job title provided. If the search returns only 1 NOC code, then the search is completed. If it returns >1 NOC code, proceed to step 3.
3. Step 3: if the returned data set contains >1 NOC code, search the returned data set using industry keywords. If there are no results, then search the lead statement related to the returned data set using industry keywords. Rank the results in alphabetical order and return the last one.

Search Strategies

1. Exact: this search returns results that exactly match the string(s) entered.
2. Like: this search returns results that include every string in the specific order as entered. The string may be included anywhere in the associated text.
3. Near: this search returns results that include every string entered in any order.
4. Any: this search returns results that include any of the strings entered.

Performance

The initial data set consisted of 566 job and industry titles with manual coding to the NOC (2016) and to the NAICS, respectively. Applying the ACA-NOC algorithm in its original form, Generation 0, to the input data set from step 1 to step 3 assigned NOC codes to 529 cases (93.5%) and resulted in 37 (6.5%) residual uncoded job titles. For 289 (51.2%) job titles, the assigned NOC codes matched the manual coding NOC codes at the fourth digit position, that is, identical matches to the manual coding NOC codes; for 303 (53.6%) job titles, the assigned NOC codes matched the manual coding NOC codes at the third digit position; for 330 (58.4%) job titles, the assigned NOC codes matched the manual coding NOC codes at the second digit position; for 381 (67.4%) job titles, the assigned NOC codes matched the manual coding NOC codes at the first digit position; and for 148 (26.2%) job titles, the assigned NOC codes differed from the manual coding NOC codes, with no matches on the 4 digits of the NOC code.

Generation 1

Following the evaluation of Generation 0, a review of cases in which the algorithm failed to code input job titles and industry descriptions was made. Table 4 lists cases showing that some current job titles contain a forward slash, such as *Interpretor/Translator*; some have spelling errors, such as *Constuction*; and some contain adjectives and spelling/typographical errors, such as *certified accontant*.

Consequently, the algorithm was updated as follows: any unmatched job titles are reviewed for (1) incorrect spelling. Correct the spelling errors before continuing NOC searches with the corrected job titles. (2) Inserted symbols such as hyphenation, forward slash, or punctuation in the job titles initiated a subroutine where job titles were split and coded separately to the NOC. (3) Selected target nouns were extracted and used to search the NOC data. (4) Any nouns found in the job titles were extracted and used to search the NOC data.

Table 4. Examples of unmatched data for Generation 0.

Current job titles	NOC[a] code by ACA[b]-NOC	NOC codes manually	Current industries
Interpretor/translator	N/A[c]	5121	N/A
CONSTUCTION	N/A	7611	CONSTUCTION
Certified accontant	N/A	1111	N/A

[a]NOC: National Occupational Classification.

[b]ACA: Automated Coding Algorithm.

[c]N/A: not applicable.

Performance

Compared with Generation 0, the improvements yielded 10 more job titles (58+20+5+5-48-18-9-3) accurately coded in the fourth position, 12 (4+2+1+3-2-1-5+10) more job titles accurately coded in the third position, 11 (4+2+9-6-4-1-5+12) more job titles accurately coded in second position, and 10 (8+4+1+18-8-2-2-20+11) more job titles accurately coded in the first position; all the current job titles matched with an NOC code. Table 5 lists the specific details of the improvement between the results of the previous Generation 0 and Generation 1 of the algorithm. Overall, 88 more job titles were accurately coded in the fourth position, although 78 job titles previously accurately coded in this position were lost. The reason for the lost matches was that these job titles were listed in alphabetical order where >1 job title was matched.

An example of an exact match that did not correlate with the manual code is the job title *Landscaper* in the industry *Landscaping*, where the algorithm assigned code 2225. This gave only a two-digit match with the manually assigned code 2215, which does not exist on investigation and represents a typographical source data input error. Therefore, the algorithm code was correct in finding the category *Landscape and horticulture technicians and specialists*, which includes the job title *landscaper*.

Table 5. Performance of Generation 1 compared with Generation 0.

G1[a]	G0[b]				
	No match	One-digit match	Two-digit match	Three-digit match	Four-digit match
No match	N/A[c]	8	4	4	58
One-digit match	8	N/A	2	2	20
Two-digit match	6	4	N/A	1	5
Three-digit match	2	1	0	N/A	5
Four-digit match	48	18	9	3	N/A

[a]G1: Generation 1.

[b]G0: Generation 0.

[c]N/A: not applicable.

Generation 2

Table 6 lists the unmatched sample data from Generation 0 of the algorithm. It was identified that some of the industry titles also contained forward slash, punctuation, and hyphenation. Additionally, it was noted that some current job titles would match with the correct NOC code if only the noun from the title was used for matching. For example, a match to an NOC code was possible with *producer* but not *agricultural producer*.

The algorithm was rewritten to include the following updates designed to address these issues:

The correction of spelling errors was moved to the end of the algorithm. Industry titles including hyphenation, forward slash, or punctuation were split, and both parts of the titles were used to search for NOC data. Compound words with spelling errors were corrected.

The description of the ACA-NOC Generation 2 algorithm is as follows:

1. Step 1: is the same as the original algorithm.
2. Step 2: if the returned data set is *none*, search the NOC data based on job titles after each of the operations on the input data set, as follows:
   - 2.1 Split job titles.
   - 2.2 Remove stopping words.
   - 2.3 Stem job titles
   - 2.4 Extract nouns from job titles.

   If the search returns only 1 NOC code, then finish the search; if it returns >1 NOC code, then repeat step 5.
3. Step 3: if the returned data set is still *none*, search the NOC data based on industry. If there are no results, then the industry is split and searched again. If the search returns only 1 NOC code, then finish the search; if it returns >1 NOC code, repeat step 5.
4. Step 4: If the returning data set is still *none*, search the NOC data based on any nouns from the job titles. If there are no results, then correct spellings and search again. If the search returns only 1 NOC code, then finish the search; if it returns

>1 NOC code, repeat step 5. If there are still no results, then return *none*.

5. Step 5: is the same as the original algorithm.

Table 6. Unmatched data for Generation 1.

Current job titles	NOC[a] codes by ACA[b]-NOC	NOC codes manually	Current industries
MARKETING IN MD OFFICE	2153	1123	MARKETING/HEALTH CARE
Property manager	0714	1224	Real estate/housing
Agricultural producer	8252	0821	Farm

[a]NOC: National Occupational Classification.

[b]ACA: Automated Coding Algorithm.

Performance

Compared with Generation 1 of the ACA-NOC algorithm in Table 7, the improvements yielded 5 (10+2+0+1-5-2-1-0) more job titles accurately coded in the fourth digit position, 4 (0+1+1+0-1-1-0-1+5) more job titles accurately coded in the third digit position, 5 (2+0+0+1-1-0-1-0+4) more job titles accurately coded in the second digit position, and 6 (4+0+1+2-3-0-1-2+5) more job titles accurately coded in the first digit position. A total of 13 additional job titles were accurately coded in the fourth digit position, but 8 job titles accurately coded in this position were lost because of the alphabetical selection order for multiple matches.

On reviewing apparent mismatches, the algorithm coded the job title *Addictions worker* with industry *hospital* to 4212 *Social and community service workers*, which gave a single digit match with the manually assigned code 4153 *Family, marriage and other related counsellors*. This latter category includes the job title *addictions counsellor*, whereas the code 4212 includes *Addictions worker*, thus the algorithm-assigned code was considered preferable, being more accurate.

Table 7. Performance of Generation 2 compared with Generation 1.

G2[a]	G1[b]				
	No match	One-digit match	Two-digit match	Three-digit match	Four-digit match
No match	N/A[c]	4	2	0	10
One-digit match	3	N/A	0	1	2
Two-digit match	1	0	N/A	1	0
Three-digit match	1	1	0	N/A	1
Four-digit match	5	2	1	0	N/A

[a]G2: Generation 2.

[b]G1: Generation 1.

[c]N/A: not applicable.

Generation 3

Table 8 lists the unmatched data. Some current job titles were identified as compounds of >1 occupation, although they may match the correct NOC code if run separately using the combination of each part of the job title and the industry title for searching. For Bartender/Waiter, we could use Bartender accommodation and food services to match the NOC code. The algorithm was updated as follows:

After splitting the job title, first search the NOC data by the combination of each job title and the industry title (the combination of the first job title and the industry title has the priority); if the returned data set is empty, then search the NOC data by each job title (the first job title has the priority); if the returned data set is empty, then search the NOC data by the industry title.

Table 8. Unmatched data example for Generation 2.

Current job titles	NOC[a] codes by ACA[b]-NOC	NOC codes manually	Current industries
Retail	0601	6421	Sales
Bartender/waiter	6513	6512	Accommodation and food services

[a]NOC: National Occupational Classification.

[b]ACA: Automated Coding Algorithm.

Performance

Compared with the previous algorithm in Table 9, the improvements yielded 4 (0+2+0+2-0-0-0-0) more job titles accurately coded in the fourth position, 4 (2+0+0+0-0-0-0-2+4) more job titles accurately coded in the third position, 3 (0+0+0+0-1-0-0-0+4) more job titles accurately coded in the second position, and 1 (0+0+0+0-0-0-0-2+3) more job title accurately coded in the first position. A total of 4 additional job titles were accurately coded in the fourth position, and no job titles accurately coded in this position were lost.

An example of apparent inaccuracy being overturned by expert review was the job title *Quality Control* with the industry *Food* coded 9465 *Testers and graders, food and beverage processing* by the algorithm. The manual code 2211 *Chemical technologists and technicians* was allocated presumably because of the included job title *quality control technician–food processing*. However, code 9465 includes a wide range of food-related grading and testing jobs, including *quality control checker–food and beverage processing*, and was considered preferable for accuracy because the assumption of technical expertise is not necessary.

Table 9. Performance of Generation 3 compared with Generation 2.

G3[a]	G2[b]				
	No match	One-digit match	Two-digit match	Three-digit match	Four-digit match
No match	N/A[c]	0	0	2	0
One-digit match	0	N/A	0	0	2
Two-digit match	1	0	N/A	0	0
Three-digit match	0	0	0	N/A	2
Four-digit match	0	0	0	0	N/A

[a]G3: Generation 3.

[b]G2: Generation 2.

[c]N/A: not applicable.

Generation 4

By checking the unmatched data in Table 10, it was identified that the term *teacher* was included in the major group title (Textbox 2). The algorithm was updated as follows:

1. Split the industry title and search each title separately; return the NOC data with the highest frequency of the industry title.
2. If >1 NOC code is returned, first filter the results hierarchically by given job title, then filter the results by industry title, and finally filter the results by given job title in the description of the NOC unit group level.

In this manner, the ACA-NOC algorithm respects the hierarchical structure of the NOC (2016) when producing results.

Table 10. Unmatched data for Generation 2.

Current job titles	NOC[a] codes by ACA[b]-NOC	NOC codes manually	Current industries
TEACHER	4021	4032	EDUCATION
LANDSCAPER	2225	2215	LANDSCAPING
Works in fashion show room	5231	5243	Fashion

[a]NOC: National Occupational Classification.

[b]ACA: Automated Coding Algorithm.

Textbox 2. National Occupational Classification (NOC) major group title from NOC (2016) contained teacher.

- Major Group 40 professional occupations in education services
- 403-secondary and elementary school teachers and educational counsellors
- 4031-secondary school teachers
- 4032–elementary school and kindergarten teachers
- 4033–educational counsellors

Performance

Compared with the previous algorithm, the improvements yielded 5 (3+2+5+1-5-0-1-0) more job titles accurately coded in the fourth position, 14 (4+0+6+0-0-0-0-1+5) more job titles accurately coded in the third position, 12 (7+2+0+1-0-1-6-5+14) more job titles accurately coded in the second position, and 16

(9+1+0+0-2-2-0-2+12) more job titles accurately coded in the first position. The difference between the results of ACA-NOC Generation 3 and Generation 4 is shown in Table 11. A total of 11 additional job titles were accurately coded in the fourth position, but 6 job titles accurately coded in this position were lost because of the alphabetical selection order.

The equivalence of an algorithm and manually assigned code that would otherwise be designated as a mismatch was sometimes adjudicated by expert review. The job title *Electrician* with industry given as *Trades* was allocated code 7241 *Electricians (except industrial and power system)* by the algorithm being a two-digit match for the manual code 7202 *Contractors and supervisors, electrical trades and telecommunications occupations*. It was considered on the basis of the information available that either code was acceptable, the main distinction being that for 7202 it includes those who own and operate their own businesses.

Table 11. Performance of Generation 4 compared with Generation 3.

G4[a]	G3[b]				
	No match	One-digit match	Two-digit match	Three-digit match	Four-digit match
No match	N/A[c]	9	7	4	3
One-digit match	2	N/A	2	0	2
Two-digit match	0	1	N/A	6	5
Three-digit match	0	0	0	N/A	1
Four-digit match	5	0	1	0	N/A

[a]G4: Generation 4.

[b]G3: Generation 3.

[c]N/A: not applicable.

Discussion

Principal Findings

We developed the ACA-NOC algorithm for automatically coding occupation from textual jobs and industry titles. Mannetje and Kromhout [4] discussed the major standard classification systems for occupations and the different methods for coding occupations, such as self-classification, clerical coding, and computer-assisted coding. They identified that the agreement rates for reliable occupation classification with manual coding is 44% to 89% for 4 to 5 digits, 56% to 80% for 3 digits, 61% to 92% for 2 digits, and 75% to 97% for 1 digit. On the basis of the agreement rates mentioned in their study, our method can be considered reliable because the accuracy of our method is in the range of the agreement rates. Overall, we describe the algorithm as rigorous, meaning that it is capable of allowing for spelling and typing errors as well as adjusting for inserted symbols such as hyphenation, forward slash, or punctuation in the search terms. It uses 7 search strategies sequentially to match job titles with the NOC codes. Through iterative development in successive generations of the algorithm, we have incorporated changes that enhance the ability to handle variations in the quality of input data. Additionally, we built a filtering mechanism that respects the hierarchical structure of the coding scheme in code selection. The system outputs a spreadsheet that includes the best-matched NOC codes for a given input, the type of search that found the match, for example, exact match, as well as the list of candidate matches from which the filter algorithm selected the best match. These are available for algorithm developers investigating mismatches on a benchmarking data set.

During development, we found that reviewing apparent coding mismatches was helpful in identifying opportunities for algorithm improvement and in determining that some automatically assigned codes were preferable to those allocated manually. We further tested the ACA-NOC algorithm with 2 other manually coded data sets of 218 and 186 cases, derived from the same question sets and answered by 2 different populations, with production rates and accuracy at least as good as for our development data.

After comparison with manual codes and with expert review, the accuracy of the algorithm is 58.7% at the four-digit level, 65.0% at the three-digit level, 72.3% at the two-digit level, and 81.6% at the one-digit level. We compared the performance of our approach side-by-side with the available web service algorithm coding with the NOC, gaining a 10% improvement with these early efforts. It is also possible to make some comparisons on the basis of algorithms performing the same task, although to different classification systems. A recent study, similarly reporting performance at each digit level of the occupation code, was conducted by Nahoomi [8], who measured the performance of multiple machine learning models (SVM, maximum entropy, and CNN) for coding job titles to the SOC hierarchy. The accuracy did not exceed 56% in the fourth level, 59% in the third level, 67% in the second level, and 77% in the first level. Our results are comparable, although marginally higher at each level.

Taking a somewhat broader comparison with other automated coding systems, highlighted in the introduction, higher accuracy levels at the four-digit level have been reported. These cases, involving other occupation classification systems, have relied on very large sets of training data and up to 3000 rules in hybrid rule–based systems. Investigation into the design of these systems will give us further insights to improve the system we have developed for coding to the NOC. Further pragmatic considerations for direct comparison of systems include (1) the

availability of manually curated test sets for performance benchmarking tasks, (2) the quality of input data, which is not only *very short text* but also full of spelling errors and irregularities, (3) design limitations of techniques used, and (4) the speed at which these algorithms can run on dedicated hardware platforms. Given that such information is not readily available, our comparisons are limited in scope.

Nonetheless, given the relatively modest scale of our project, the results are promising, particularly as there are no other available automated coding algorithms for the NOC (2016). We are currently negotiating access to larger sets of patient data manually coded to the NOC for further benchmarking.

Even at the current level of accuracy, significant time and cost savings are possible over manual coding by insurers or researchers in Canada. As identified by Burstyn et al [9], it can take manual coders days to months to code a few 100 occupations to a classification. They reported that it can take 2 months to code approximately 1600 free text descriptors of lifetime occupational histories to the 2010 SOC. The ACA-NOC algorithm will clearly be useful for efficient and cost-effective coding of data sets, allowing human expertise to be more focused on any residual uncoded cases and those with low confidence matches. With a very high production rate, a coding accuracy of approximately 60% to 4 digits, and being comparable with the intersubject performance of human coders [4], our software is both practicable and economically significant, particularly with large data sets becoming available from population health studies in Canada. One such study is the Canadian Partnership for Tomorrow Project, which has occupation information on >300,000 Canadian participants [12]. Additionally, the ACA-NOC may have utility for Workers' Compensation Boards and other insurers who collate occupational data.

Acknowledgments

The authors wish to thank Worksafe New Brunswick for providing funding through the WorksafeNB Chief Medical Officers Occupational Medicine Research Grant Competition 2017-2018 for the project *Using semantic coding of text-based occupation titles with the Canadian National Occupation Codes (NOC)*. The authors wish to acknowledge their collaborators in the Canadian Immunisation Research Network Community Acquired Pneumonia study, the Atlantic PATH Project, and the Canada East Spine Centre, who provided information on occupation and the industry. The authors are grateful to the Alberta Tomorrow Project for providing further data, allowing validation of the initial results. Thanks are also due to Mohammad Sadnan Al Manir for proofreading the manuscript and for his valuable suggestions.

Conflicts of Interest

None declared.

References

1. MacDonald LA, Cohen A, Baron S, Burchfiel CM. Occupation as socioeconomic status or environmental exposure? A survey of practice among population-based cardiovascular studies in the United States. Am J Epidemiol 2009 Jun 15;169(12):1411-1421. [doi: 10.1093/aje/kwp082] [Medline: 19429878]
2. The National Occupational Classification (NOC). URL: http://noc.esdc.gc.ca/English/NOC/AboutNOC.aspx?ver=16 [accessed 2019-03-28]
3. Gweon H. Three methods for occupation coding based on statistical learning. J Off Stat 1122;33(1):10. [doi: 10.1515/jos-2017-0006]
4. Mannetje A, Kromhout H. The use of occupation and industry classifications in general population studies. Int J Epidemiol 2003 Jun;32(3):419-428. [doi: 10.1093/ije/dyg080] [Medline: 12777430]
5. Bethmann A, Schierholz M, Wenzig K, Zielonka M. Automatic Coding of Occupations. Statistics Canada. 2014 Aug 31. URL: https://www150.statcan.gc.ca/n1/en/catalogue/11-522-X201300014291 [accessed 2020-07-03]
6. Schierholz M. Automating Survey Coding for Occupation. Open Access LMU. 2014. URL: https://epub.ub.uni-muenchen.de/21444/1/MA_Schierholz.pdf [accessed 2019-05-08]
7. Russ DE, Ho K, Colt JS, Armenti KR, Baris D, Chow W, et al. Computer-based coding of free-text job descriptions to efficiently identify occupations in epidemiological studies. Occup Environ Med 2016 Jun;73(6):417-424 [FREE Full text] [doi: 10.1136/oemed-2015-103152] [Medline: 27102331]
8. Negin N. Automatically Coding Occupation Titles to a Standard Occupation Classification. Semantic Scholar. 2018. URL: https://pdfs.semanticscholar.org/be53/193184dfd665525249568487bcef5455c93d.pdf [accessed 2019-05-08]
9. Burstyn I, Slutsky A, Lee D, Singer AB, An Y, Michael YL. Beyond crosswalks: reliability of exposure assessment following automated coding of free-text job descriptions for occupational epidemiology. Ann Occup Hyg 2014 May;58(4):482-492. [doi: 10.1093/annhyg/meu006] [Medline: 24504175]
10. Occupational Classifications. Statistics Canada. URL: https://www.statcan.gc.ca/eng/concepts/occupation [accessed 2020-04-30]

11. LeBlanc J, ElSherif M, Ye L, MacKinnon-Cameron D, Ambrose A, Hatchette TF, et al. Age-stratified burden of pneumococcal community acquired pneumonia in hospitalised Canadian adults from 2010 to 2015. BMJ Open Respir Res 2020 Mar;7(1) [FREE Full text] [doi: 10.1136/bmjresp-2019-000550] [Medline: 32188585]
12. Canadian Partnership for Tomorrow's Health: CanPath. URL: https://canpath.ca/ [accessed 2020-01-05]

Abbreviations

ACA: Automated Coding Algorithm
CNN: convolutional neural network
NAICS: North American Industrial Classification
NOC: National Occupational Classification
SOC: Standard Occupational Classification
SVM: support vector machine

Edited by C Lovis; submitted 27.09.19; peer-reviewed by M Friesen, C Lara; comments to author 10.11.19; revised version received 01.05.20; accepted 14.06.20; published 05.08.20.

Please cite as:
Bao H, Baker CJO, Adisesh A
Occupation Coding of Job Titles: Iterative Development of an Automated Coding Algorithm for the Canadian National Occupation Classification (ACA-NOC)
JMIR Form Res 2020;4(8):e16422
URL: https://formative.jmir.org/2020/8/e16422
doi:10.2196/16422
PMID:32755893

Original Paper

Recruiting Student Health Coaches to Improve Digital Blood Pressure Management: Randomized Controlled Pilot Study

Elena Vasti[1], MPH; Mark J Pletcher[2], MD, MPH

[1]University of California, San Francisco School of Medicine, Stanford, CA, United States

[2]Department of Biostatistics and Epidemiology, University of California, San Francisco, San Francisco, CA, United States

Corresponding Author:
Elena Vasti, MPH
University of California, San Francisco School of Medicine
Stanford Internal Medicine Residency
300 N Pasteur Dr s101
Stanford, CA, 94305
United States
Phone: 1 2094703263
Email: ecvasti@stanford.edu

Abstract

Background: Hypertension is a significant problem in the United States, affecting 1 in 3 adults aged above 18 years and is associated with a higher risk for cardiovascular disease and stroke. The prevalence of hypertension has increased in medically underserved areas (MUAs). Mobile health technologies, such as digital self-monitoring devices, have been shown to improve the management of chronic health conditions. However, patients from MUAs have reduced access to these devices because of limited resources and low health literacy. Health coaches and peer training programs are a potentially cost-effective solution for the shortage of physicians available to manage hypertension in MUAs. Activating young people as student health coaches (SHCs) is a promising strategy to improve community health.

Objective: This pilot study aims to assess (1) the feasibility of training high school students as health technology coaches in MUAs and (2) whether the addition of SHCs to digital home monitoring improves the frequency of self-monitoring and overall blood pressure (BP) control.

Methods: In total, 15 high school students completed 3-day health coach training. Patients who had a documented diagnosis of hypertension were randomly assigned to 1 of the 3 intervention arms. The QardioArm alone (Q) group was provided a QardioArm cuff only for convenience. The SHC alone (S) group was instructed to meet with a health coach for 30 min once a week for 5 weeks to create action plans for reducing BP. The student+QardioArm (S+Q) group received both interventions.

Results: Participants (n=27) were randomly assigned to 3 groups in a ratio of 9:9:9. All 15 students completed training, of which 40% (6/15) of students completed all the 5 meetings with their assigned patient. Barriers to feasibility included transportation and patient response drop-off at the end of the study. Overall, 92% (11/12) of the students rated their experience as very good or higher and 69% (9/13) reported that this experience made them more likely to go into the medical field. There was a statistically significant difference in the frequency of cuff use (S+Q vs Q groups: 37 vs 17; P<.001). Participants in the S+Q group reported better BP control after the intervention compared with the other groups. The average BP at the end of the intervention was 145/84 (SD 9/18) mm Hg, 150/85 (SD 18/12) mm Hg, and 128/69 (SD 20/14) mm Hg in the Q, S, and S+Q groups, respectively.

Conclusions: This pilot study demonstrates the feasibility of pairing technology with young student coaches, although challenges existed. The S+Q group used their cuff more than the Q group. Patients were more engaged in the S+Q group, reporting higher satisfaction with their SHC and better control of their BP.

(JMIR Form Res 2020;4(8):e13637) doi:10.2196/13637

KEYWORDS

mobile health; hypertension; coaching; health-related behavior; mobile phone

Introduction

Background

Despite advances in management, hypertension remains a significant public health challenge in the United States, affecting approximately 33% of adults aged above 18 years [1,2]. Hypertension is associated with a significantly higher risk for cardiovascular disease and stroke [3]. From 2004 to 2014, the death rate directly attributable to hypertension increased by 7.6% and the total number of deaths from hypertension increased by 34% [3]. Factors such as race and socioeconomic status (SES) are major social determinants that influence an individual's risk of hypertension. In one meta-analysis, lower occupational status and level of education were associated with odds ratios for hypertension of 1.31 and 2.02, respectively [4]. Another prospective trial demonstrated that incident hypertension was lower in participants from higher SES groups than in lower SES groups, suggesting that having limited social and economic resources plays a role in the disproportionate burden of hypertension seen in disadvantaged neighborhoods [5].

Mobile health (mHealth) offers a unique opportunity for improving blood pressure (BP) management [6,7]. Patient motivation, medication adherence, and diastolic BP have all been shown to improve with access to digital BP monitoring devices [8,9]. In rural areas, mHealth has been shown to improve outcomes. One study [10] demonstrated that the implementation of mHealth strategies, such as smart messaging systems, in rural health care centers across Lebanon significantly improved mean systolic BP in comparison with controls. These interventions have also been shown to improve adherence to BP medication regimens [11]. However, many of these studies require intensive resources to carry out mHealth interventions, including dedicated field researchers or community health workers. Although effective in the short term, it can be challenging to allocate the appropriate resources to these interventions. One qualitative study in rural Uganda discovered that providers and health care workers both articulated that a lack of patient education, limitations in time with patients, and funding for patient education were substantial barriers to implementing successful programs [12]. Health literacy, which was defined by Nielsen-Bohlman in 2004 as "the degree to which individuals have the capacity to obtain, process and understand basic health information and the services needed to make appropriate health decisions," is an important variable to consider in the intervention of health programs. Improving the health literacy of patients allows them to become agents of change in their own health trajectory, which can reduce the burden on the health care system. Lower health literacy is often associated with patients from underrepresented and vulnerable groups and can contribute to widening health disparities in chronic diseases, such as hypertension [13]. Health literacy is also related to mHealth adoption. It is strongly associated with patients' perceived ease of use and perceived usefulness of digital devices [14]. Thus, patients who are arguably the most likely to benefit from mHealth interventions potentially face barriers related to health literacy that limit their access to them. Health coaching has become an increasingly used strategy in low-income settings to manage BP outside of the clinic [15]. One study that examined health coaching as a possible strategy to prevent rehospitalization for chronic obstructive pulmonary disease (COPD) exacerbation randomized 215 patients hospitalized for COPD to either health coaching sessions with written action plans for exacerbations or usual care. The investigators trained volunteers as health coaches and evaluated the rate of COPD-related hospitalizations. There was a statistically significant difference of up to 9 months after discharge, and the investigators concluded that this is a potentially feasible approach for patients with chronic diseases that predispose them to hospitalization [16]. In addition, peer health coaching, where nonmedical community members are trained as health coaches, has been shown to be cost-effective and efficacious [17]. Social epidemiologists have begun to target youth civic engagement (YCE) as a strategy to promote community health in a social context [18], but to our knowledge, the impact of YCE *directly* on patient care has not been explored. In addition, bringing vibrant, young people into the health coaching intersection with digital technology is also an evidence-free zone. iPads have increasingly become a part of secondary education and have shown enhanced creativity and increased collaboration between students and teachers [19]. Our team felt that high school students' exposure to technology in schools would uniquely qualify them as health coaches at the intersection of mHealth and health coaching. Using our understanding of the benefits of health coaching and mHealth in reducing the burden of hypertension, the focus of this study was to demonstrate that these 2 entities could synergistically improve BP monitoring and health behaviors.

Objectives

This pilot study investigated an innovative approach for implementing a smartphone-based BP monitoring program in a medically underserved area (MUA) by engaging young people. First, we hypothesized that high school students could be trained as student health coaches (SHCs) to assist patients with smartphone devices for BP management. Second, we asked the question, "does the addition of weekly SHC visits to the use of digital BP monitoring devices improve patient engagement and overall BP control?"

Methods

Recruitment of Participants and Students

Participants aged above 18 years with a documented diagnosis of hypertension were recruited from a rural primary care clinic in Stockton, San Joaquin County, California, from June to July 2016. As recruitment took place at a small clinic with the primary care physician (PCP) available, patients who did not have chart documentation of hypertension but were verified by the PCP to be prescribed antihypertensive agents or currently attempting lifestyle changes for hypertension were included. This clinic was selected because it was the only primary care clinic located in South Stockton, San Joaquin County, California, where health disparities are most prevalent. All participants needed to own or have access to a Bluetooth-enabled smartphone. Patients were excluded if they were diagnosed with hypertension secondary to renal or endocrine disorders.

SHCs aged between 14 and 18 years were selected from a charter school in South Stockton, San Joaquin County, California. This charter school was a health academy, where the students who were enrolled had expressed an interest in careers in health care. The principal was actively involved in the development of this program as an extracurricular activity for students during the summer to build their resumes and skillsets. The University of California, San Francisco (UCSF) Institutional Review Board (IRB), who approved the protocol, required SHCs to complete research credentials. However, they did not require parental consent for participation, as the research activities were taking place in public places, such as the clinic itself or a coffee shop, if it was more convenient for the SHC. This was deemed to be of minimal risk for students in secondary school. They completed an extensive 3-day training on health coaching techniques. The curriculum included motivational interviewing; *ask-tell-ask* strategies for identifying barriers to patients' adherence to BP management; creating specific, measurable, achievable, realistic, and time-based (SMART) goals; and identifying hypertensive emergencies. SHCs were given a binder with reference sheets for use during their meetings. One day of the training was focused on using the QardioArm device itself. The health coaching curriculum was provided by the UCSF Center for Excellence in Primary Care. SHCs were all required to update an electronic tracking spreadsheet with deidentified notes from their meetings with patients. The purpose of the tracking spreadsheet was to ensure that SHCs were effective at guiding patients to create an action plan in the first visit and to document how they guided their patients through challenges and provided encouragement. Real-time feedback was provided by the research coordinator (EV) to the SHCs to improve the quality of the coaching sessions.

Study Design

Participants were randomized in a 1:1:1 ratio to 1 of the following 3 groups: *SHC+QardioArm (S+Q)*, *QardioArm alone (Q)*, or *SHC alone (S)*. Randomization was open, and a computerized randomization algorithm was used to assign patients to 1 of the 3 groups. We planned to recruit 30 participants. Each patient who met the inclusion criteria was given information on the trial and was assigned to an intervention arm after they signed a consent form. We followed this system so that participation would not be influenced by group assignment and, therefore, there would be minimal differences between the groups. The 5-week intervention consisted of (1) weekly meetings with the SHC (S+Q and S arms) and (2) provision of a QardioArm BP home monitoring device to be used at the patient's discretion (S+Q and S arms).

Qardio Inc, a digital monitoring device company based in San Francisco, California, donated QardioArm BP devices, which uploaded data to the patient's smartphone. Patients were enrolled concurrently in the UCSF Health eHeart (HeH) study, and the HeH technical team carried out BP data collection. This study was approved by the UCSF IRB.

Data Collection and Analysis

This pilot study used a mixed methods approach. SHCs completed postintervention surveys, which included both quantitative and qualitative data. Differences in pre- and postintervention BP across all 3 groups were reported. The frequency of QardioArm use and the number of active QardioArm users between the S+Q and Q groups were analyzed with a two-sided *t* test. Qualitative data reported by patients are quoted directly.

Results

A total of 27 eligible participants were enrolled in the study. The baseline characteristics are displayed for 89% (24/27) of the participants (Table 1). The remaining 3 participants never returned to the baseline surveys.

Table 1. Baseline characteristics of Health eHeart participants in the San Joaquin County, California, cohort (N=24).

Group	QardioArm alone	Student health coach alone	S+Q[a]
Number of respondents[b], n (%)	8 (89)	8 (89)	8 (89)
Age (years), mean (SD)	55 (15)	55 (6)	58 (17)
Race and ethnicity, n (%)			
Black	2 (25)	3 (37.5)	3 (37.5)
White	0 (0)	1 (12.5)	3 (37.5)
Asian	2 (25)	0 (0)	0 (0)
Hispanic	0 (0)	0 (0)	0 (0)
Other Pacific Islander	1 (12.5)	0 (0)	1 (12.5)
No answer	3 (37.5)	4 (50)	1 (12.5)
Health insurance, n (%)			
MediCal/MediCare	3 (37.5)	3 (37.5)	3 (37.5)
Health Plan of San Joaquin	2 (25)	1 (12.5)	2 (25)
Blue Cross	0 (0)	1 (12.5)	0 (0)
Care First	0 (0)	0 (0)	1 (12.5)
Alignment	0 (0)	0 (0)	1 (12.5)
None	3 (37.5)	3 (37.7)	1 (12.5)

[a]S+Q: student+QardioArm.

[b]One participant from each intervention arm did not return a baseline survey.

Feasibility of High School Students as Health Coaches

All 15 SHCs completed training and at least one meeting to create a lifestyle plan with their assigned patient. Patients were assigned to SHCs in a nonrandom way at an orientation event held at the clinic where SHCs introduced themselves to patients and signed up for the 5 meetings together. Any of the participants or SHCs who could not attend the orientation event were individually contacted by the study coordinator and assigned patients based on the patient's and SHC's availability. One hypertensive emergency was accurately identified by an SHC, who followed the protocol and advised the patient to go to the emergency room for treatment. Follow-up surveys were completed by 80% (12/15) of the students. These SHCs rated their overall experience favorably, with 92% (11/12) of the SHCs reporting their experience from very good (8) to exceptional (10) on a 10-point scale. When asked how this project influenced their future plans to go into the medical field, 69% (9/13) of SHCs reported that they were more likely to go into the medical field. Notably, all (10/10) the SHCs who were assigned to patients randomized into the S+Q group reported that the feedback from the QardioArm device enhanced patient motivation and improved BP management. SHCs noted that they enjoyed the autonomy of working with patients the most. For example:

> *I absolutely loved the fact that this project allowed me to become an actual health provider. It was like I was her doctor and I'm sure it will prepare me for the future.* [JT, 17 years]
>
> *I enjoyed meeting with patients and actually being involved with tackling health disparities.* [MS, 17 years]

Of the 15 SHCs, 6 (40%) completed all 5 meetings, with 5 (33%) of these SHCs assigned to the S+Q group. SHCs reported difficulty with patient retention in the program:

> *...the patients would sign up and not show up, and it was pretty sad for the mentors, because some of us were actually looking forward to meeting them.* [SN, 15 years]
>
> *My two patients specifically stopped coming to meet up with me after a few days.* [JT]

In addition, SHCs kept weekly logs of their meeting notes on a spreadsheet and submitted them to the research coordinator during each week of the intervention (Table 2). The SHCs were all able to create effective SMART goals with patients, as shown in the sample of two of the SHC responses in the table.

Table 2. Example of student health coach–logged visit summaries for student+QardioArm and student health coach–alone groups.

Week of intervention	S+Q[a]	Student health coach alone
Week 1	Health Goals made: Meditation videos on YouTube 3 times a week for 30 min a time. Goal is to reduce stress levels. Has Bell's Palsy and feels like she gets some eye pain when she gets anxious.	Patient made action plan goal to exercise 3 days a week for an hour. He plans on exercises Mondays, Wednesdays, and Fridays by walking around Victory Park or by going to the gym. Patient started diet about 2 weeks before starting the study. His diet includes having smoothies for breakfast as well as oatmeal. He has also cut carbohydrates, starches, and unhealthy sugars from his diet. Since he is already on a diet, he made another action plan goal to continue his diet to eat healthier.
Week 2	The meeting was a success. She completed and over exceeded her health Goal and even wanted to walk for 10 min, once a week, alone without her dog. He vision is still blurry because of her Bell's Palsy, but she has arranged an eye doctor's appointment. Meeting dates have been rearranged also.	Patient feels very confident with his action plan. He has exercise from Monday-Friday in the mornings for an hour, which is more than his original [*sic*][b] action plan states. Patient has maintained his diet and is sleeping better. Pateint [*sic*] has also lost 10 pounds and has set his own goal to lose weight. His goal weight is 170 pounds. The patient overall wants to be healthier and not have to take medication.
Week 3	In this meeting, my patient decided to add another health goal. Beside being already fantastic they wanted to cut down on the slts [*sic*]. We talked a little bit about her stress, and she was mentioning how her mother may contribute a little to her stress. She deffinately [*sic*] said that Bell's Palsy was one of her major problems that stress her out, but she also said she's been feeling a lot [*sic*] better. She had question about the program and what was going to happen after all five meeting were over and what we would do? She also experienced some technical issues with her Quardio [*sic*] Arm monitor, so I assured her I would ask. She also wnated to know what would happen with the Quardio [*sic*] arm after the five visits, and if she would return it? Beside her adding another [*sic*] health goal my patient has beeing [*sic*] following up with her other previous expectations and says she's been doing great, and that she certaianatly [*sic*] has seen a differece [*sic*] in her overall life style.	No show
Week 4	even though she ran a little late because of a traffic jam on the freeway, she was an excellent patient and came. She's been cutting down on her stress levels by listening to the meditation sound tracks, and she's been walking for 30 min with her grandchildren, and besides that she has also been cutting down on the salts. She also confesses she likes eating chips, so she will stay away from those too. I explained to her after the five sessions we could meet up outside [*sic*] the clinic but it was totally up to her. I also let her know to add any concern about the Quardio [*sic*] arm and its performance on the feedback sheet that will hopefully be given to her by me. She will also be keeping the Quardio [*sic*] Arm after her participation in the case study. But she wanted to know if other patients were experiencing anything wrong with their Quardio [*sic*] Arm? she also said she's pretty proud of her reading and significantly reduced them. Although she is a little scepticle [*sic*] because she has another High Blood Pressure monitor for her wrist and says that the wrist one, shows a different reading than the Quardio [*sic*] Arm.	Patient has been experiencing any problems and been doing well with his action plan. Patient was cleared to receive [*sic*] CPAP[c] machine. Patient stated that he feels great, better than ever. Patient has been getting better sleep and has more energy. Patient as exercised for the past week Monday-Friday walking at Victory Park for an hour in the mornings and has been keeping up with his diet.

Week of intervention	S+Q[a]	Student health coach alone
Week 5	No show	Patient appeared to be happier compared to the first meeting. Patient says he lost another 9 pounds. His current weight is 220 pounds, he started at 240. The patient feels good. He was been keeping up with his diet and exercising. He plans to keep these up in order to reach his goal weight, 170 pounds. He states that his energy has improved. He exercises for 5 days a for 1 hour and 30 min at a time. He has stated that the chanl-leges [*sic*] he has faced include not want to exercises, wanting to get unhealthy foods, and keeping everything consistent [*sic*]. He was able to not give into these challenges and has benefited from it. Patient and I said our farewells.

[a]S+Q: student+QardioArm.

[b]SHC responses reported exactly as written in the originally logged visit summaries and are not edited for spelling or grammar to preserve the primary data.

[c]CPAP: Continuous Positive Airway Pressure.

Patient Outcomes

There was a total of 97 and 264 distinct uses of QardioArm in the Q group and S+Q group, respectively. We found a statistically significant difference in the average total number of uses of the BP monitor during the 5-week intervention (average total number of uses for S+Q vs Q groups: 37 vs 17; *P*=.01). As shown in Figure 1, the number of active QardioArm users per week (defined as use at least once a week) in both the S+Q and Q groups decreased over time but was higher in the S+Q group (4.8 vs 2.4 users). Due to the small sample size and the risk of overinterpretation, statistical analysis is not included. In the Q group, the number of intervention days used by each active participant was 48, 14, 4, and 2 for each of the 4 active participants. The other 5 participants did not record QardioArm use. In the S+Q group, the number of intervention days used by each active participant was 57, 36, 19, 17, 3, and 2 for the 6 active participants. The other 3 participants did not record QardioArm use.

Participants were also asked about their perception of their BP control, which was elicited by querying how many days of the week they felt their BP was well controlled. *Well controlled* was intentionally left to the participant's interpretation to avoid introducing bias, given the variability in the health literacy of the participants. *Number of days of the week* was categorized as *all days of the week* for 7 days, *most days of the week* for 5 to 6 days, *half the days of the week* for 3 to 4 days, few days of the week for 1 to 2 days, and *no days of the week* for 0 days. Participants were also given the option to respond, "I'm not sure." As shown in Multimedia Appendix 1, 7 of 9 participants in the S+Q group reported their BP to be well controlled most days of the week after the intervention compared with those in the Q (3/9) and S (2/9) groups.

Participants' perceptions of the program, both the SHCs and QardioArm devices, varied across the 3 groups qualitatively. Participants in the S group reported less satisfaction with the SHCs compared with those in the S+Q group (4/9 vs 7/9 participants rated their SHC as *satisfactory*). The Q group expressed more frustration with their QardioArm and overall had a higher cessation rate of QardioArm use compared with the S+Q group. Although the S+Q group expressed the same challenges as the Q group with QardioArm, they were more engaged. For example, the S+Q group participants identified solutions to technical difficulties with the devices. Survey responses from this group also reflected the synergy between the QardioArm and SHC. Patients in the S+Q group reported:

> *The results are really inconsistent—I don't know if I trust it or not. I am thinking of getting my own machine, though. Nobody should be on blood pressure medication if they don't need to be so we should have a machine that works.* [P12]

> *Sometimes it works and sometimes it doesn't, but I still use it. My health coach was very involved, he was very good, would call me every week, and I would text him when it wasn't working and he was able to help me use it and he would see if it started working later.* [P20]

> *[SHC name] was incredibly supportive and as a result of the accountability, I was actually able to improve my diet...and has taught me how to better manage my blood pressure.* [P6]

In comparison, surveys returned from the Q group included:

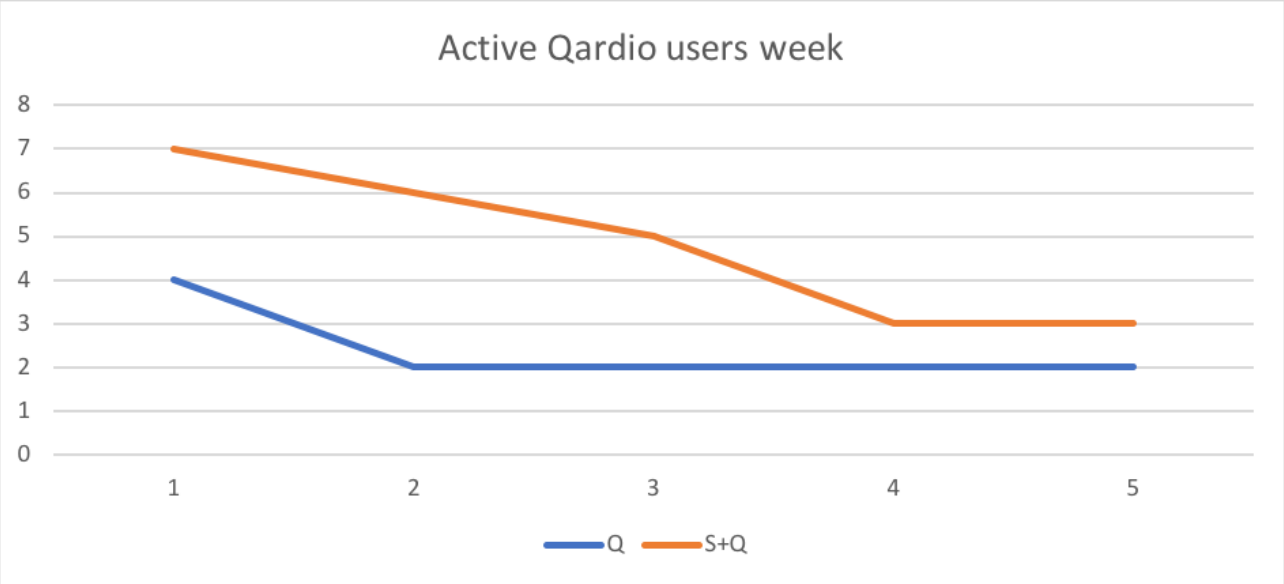
> *I was excited at first, but then it just went downhill. I got tired of messing with it.*" [P28]

> *Sometimes doesn't work properly. Had difficulty putting it on.* [P3]

BP values were obtained in the clinic during the first participant meeting. Each QardioArm distributed was calibrated with a clinic sphygmomanometer. Any QardioArm device that did not reliably yield a result within 5 mm Hg of the clinic sphygmomanometer was not used. Baseline values were obtained in 78% (7/9) of the participants in the S group, 56% (5/9) of the participants in the Q group, and 78% (7/9) of the participants in the S+Q group. Baseline BP values were obtained only for participants who were able to attend the first meeting with their SHC or, for Q group participants, if they were able to return during the first week of the intervention. This was done to avoid potential bias introduced by the intervention or

an uncalibrated QardioArm cuff. The follow-up BPs in the S+Q and Q groups were the final readings during the fifth week of the intervention from these calibrated QardioArm cuffs. These values represented 67% (6/9) of the participants assigned to the S+Q and Q groups. The S group was asked to come into the clinic to measure their BP. Only 44% (4/9) participants in this group returned to the clinic for a follow-up BP check. Given the small sample size in this pilot study, BPs reported should be interpreted with caution. At baseline, the mean BP was 149/85 (SD 15/10) mm Hg, 142/84 (SD 9/13) mm Hg, and 139/79 (SD 14/6) mm Hg in the Q, S, and S+Q groups, respectively. At the end of the intervention, the mean BP was 145/84 (SD 9/18) mm Hg, 150/85 (SD 18/12) mm Hg, and 128/69 (SD 20/14) mm Hg in the Q, S, and S+Q groups, respectively.

Figure 1. Active number of QardioArm users/week.



Discussion

Principal Findings

This proof-of-concept study showed that high school students are capable of learning health coaching skills that effectively facilitate patients' use of digital home monitoring devices to improve BP compared with either SHC or device alone. We found that it is feasible to train SHCs, as demonstrated by (1) 100% of training completion and (2) successful completion of at least one patient meeting where lifestyle modification plans were created and approved by the study coordinator. It is remarkable that even in this short 5-week intervention, 1 hypertensive emergency was detected by the QardioArm and the SHC was capable of properly referring the participant for emergent care. Participants in the S+Q group used their QardioArm more frequently across all weeks of the intervention. Compared with the Q and S groups, more participants in the S+Q group subjectively felt their BP was *well controlled* after the intervention. Thus, although this was reflective of participants' subjective experiences, the large increase in participants reporting their BP to be well controlled after the intervention suggests that the S+Q group had a larger impact from the intervention than the other two groups. The qualitative survey data also support these results. The S+Q group was more engaged with both the Q and SHC groups than the other groups. This was apparent in their willingness to troubleshoot obstacles, use their support systems, and recognize their self-agency over managing their hypertension. In addition, multiple participants in this group reached out to the study coordinator after the completion of the 5-week trial hoping to continue to meet with their SHC. Many reported that they felt responsible for cultivating the "future doctors of Stockton." In turn, SHCs felt accountable for their patients. In essence, they were the primary care providers for their patients, recognizing the complications of hypertension and encouraging participants to make lifestyle changes. Although the pilot study was not designed to detect a statistically significant difference in BP, the S+Q group showed a trend toward a clinically meaningful decrease in BP at the end of 5 weeks compared with the other groups. The SHC facilitated the use of the QardioArm to improve patient engagement in the S+Q group. This was reflected in both the frequency of active QardioArm use in this group as well as patient engagement in more regular BP monitoring despite challenges associated with the QardioArm.

Future Work

This study is promising because it highlights that although the use of smartphones is prevalent, the use of self-monitoring Bluetooth-enabled devices, such as QardioArm, is widely adopted by patients in MUAs who may have low health literacy. However, the assistance of a health coach can facilitate patient engagement in lifestyle changes related to hypertension, even when the health coach is a high school student without formal training or an advanced health degree. Health literacy in this study was measured by both patient use of QardioArm and participant responses of BP control most days of the week. Participants in the S+Q group had a higher frequency of use per week compared with the Q group, and overall, they had a higher proportion of participants reporting well-controlled BP on most days of the week. This demonstrates that the participants in this group had the information they needed after the intervention to

assess their health status, as it relates to their BP. This, in turn, was associated with higher rates of participant engagement.

The participants in this study were recruited from a clinic in South Stockton, a particularly underserved area in San Joaquin County, California, with a human development index score of 2.86 compared with the California average of 5.39 [21]. According to the community needs assessment conducted by the San Joaquin Department of Public Health in 2016, approximately 25% of people in this geographic region fall below the poverty line, and 1 in 10 workers cannot find employment [21]. In addition, the percentage of the population in San Joaquin County, California, insured by MediCal is 30.9%, similar to the reported MediCal and Medicare insurance coverage of the participants in this study. Hypertension is a prevalent and morbid condition affecting members of this community. For example, in the zip code 95202 which is located in south Stockton, San Joaquin County, California, there were 1749 visits and 722.5 hospitalizations related to hypertension in 2016. This is compared to 365.5 visits and 381 hospitalizations related to hypertension on average for the state of California [12]. Although we were unable to collect data on the other comorbidities of the participants because of inconsistent clinical documentation at the underresourced, rural clinic, we can use the community-level data to extrapolate the risk of serious hypertension-related outcomes to the patients studied. Therefore, this patient population is suitable for the assessment of the study interventions. Although the average initial diastolic BP was close to the goal, several studies have shown that isolated systolic hypertension is comparable with systolic-diastolic hypertension for the risk of incident heart failure and cardiovascular mortality [22]. In another study published in the *New England Journal of Medicine*, the authors showed that systolic hypertension had a greater effect on the composite outcome of myocardial infarction, hemorrhagic stroke, and ischemic stroke [23]. The study population clearly showed elevated systolic pressures in all groups at baseline, which showed a trend toward improvement in the S+Q group.

The students selected as SHCs were also from a charter school located in South Stockton, San Joaquin County, California, where 1 in 4 students dropped out of high school. This is twice the average dropout rate for the state of California. Demographic data were not collected on the SHCs, as this was not part of the IRB approval. However, all SHCs expressed a desire to go into the health field and were seeking mentorship that was not readily available within their school district.

Conclusions

This pilot study was designed as a proof-of-concept trial to explore the feasibility of a community-based, cost-effective integration of health coaching and digital home monitoring devices. There were several limitations to this study. Participation dropped off in week 4 of the intervention because of obstacles related to transportation and the start of the new school year for the students. However, one of the important aims of this project was to carry out the interventions in a low-resource setting with limited resources. We accepted donated QardioArm devices, and students' participation was completely voluntary. The results are promising in that they suggest that the need for health coaching in MUAs can be met by training and engaging motivated young people, who are looking for opportunities to work with patients. It would be worthwhile to explore health coaching as a possible avenue for both improving the health of communities and fostering positive youth development. Future exploration of this cost-effective, community-engaged approach is warranted.

Acknowledgments

The authors would like to acknowledge the UCSF Resource Allocation Program for Trainees for providing funding for this study, Qardio Inc, for donating QardioArm devices, and the HeH technical team for carrying out data collection.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Number of participants who answered that their blood pressure was well controlled on at least 5 days per week. Q: QardioArm alone; S: student health coach alone; S+Q: student+QardioArm.
[PNG File , 25 KB - formative_v4i8e13637_app1.png]

References

1. Fryar C, Ostchega Y, Hales CM, Zhang G, Kruszon-Moran D. Hypertension prevalence and control among adults: United States, 2015-2016. NCHS Data Brief 2017 Oct(289):1-8 [FREE Full text] [Medline: 29155682]
2. Whelton P, Carey R, Aronow W, Casey DE, Collins KJ, Himmelfarb CD, et al. 2017 ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA Guideline for the prevention, detection, evaluation, and management of high blood pressure in adults: a report of the American college of cardiology/American heart association task force on clinical practice guidelines. Hypertension 2018 Jun;71(6):e13-115. [doi: 10.1161/HYP.0000000000000065] [Medline: 29133356]

3. Benjamin EJ, Blaha MJ, Chiuve SE, Cushman M, Das SR, Deo R, American Heart Association Statistics Committee and Stroke Statistics Subcommittee. Heart disease and stroke statistics-2017 update: a report from the American Heart Association. Circulation 2017 Mar 7;135(10):e146-e603 [FREE Full text] [doi: 10.1161/CIR.0000000000000485] [Medline: 28122885]
4. Leng B, Jin Y, Li G, Chen L, Jin N. Socioeconomic status and hypertension: a meta-analysis. J Hypertens 2015 Feb;33(2):221-229. [doi: 10.1097/HJH.0000000000000428] [Medline: 25479029]
5. McDoom MM, Palta P, Vart P, Juraschek SP, Kucharska-Newton A, Diez Roux AV, et al. Late life socioeconomic status and hypertension in an aging cohort: the atherosclerosis risk in communities study. J Hypertens 2018 Jun;36(6):1382-1390 [FREE Full text] [doi: 10.1097/HJH.0000000000001696] [Medline: 29621068]
6. Logan AG, Irvine MJ, McIsaac WJ, Tisler A, Rossos PG, Easty A, et al. Effect of home blood pressure telemonitoring with self-care support on uncontrolled systolic hypertension in diabetics. Hypertension 2012 Jul;60(1):51-57. [doi: 10.1161/HYPERTENSIONAHA.111.188409] [Medline: 22615116]
7. Gagnon M, Ngangue P, Payne-Gagnon J, Desmartis M. m-Health adoption by healthcare professionals: a systematic review. J Am Med Inform Assoc 2015 Jun 15;23(1):212-220. [doi: 10.1093/jamia/ocv052] [Medline: 26078410]
8. Kim JY, Wineinger NE, Steinhubl SR. The influence of wireless self-monitoring program on the relationship between patient activation and health behaviors, medication adherence, and blood pressure levels in hypertensive patients: a substudy of a randomized controlled trial. J Med Internet Res 2016 Jun 22;18(6):e116 [FREE Full text] [doi: 10.2196/jmir.5429] [Medline: 27334418]
9. Fletcher BR, Hartmann-Boyce J, Hinton L, McManus RJ. The effect of self-monitoring of blood pressure on medication adherence and lifestyle factors: a systematic review and meta-analysis. Am J Hypertens 2015 Oct;28(10):1209-1221. [doi: 10.1093/ajh/hpv008] [Medline: 25725092]
10. Saleh S, Farah A, Dimassi H, El Arnaout N, Constantin J, Osman M, et al. Using mobile health to enhance outcomes of noncommunicable diseases care in rural settings and refugee camps: randomized controlled trial. JMIR Mhealth Uhealth 2018 Jul 13;6(7):e137 [FREE Full text] [doi: 10.2196/mhealth.8146] [Medline: 30006326]
11. Patel A, Praveen D, Maharani A, Oceandy D, Pilard Q, Kohli MP, et al. Association of multifaceted mobile technology-enabled primary care intervention with cardiovascular disease risk management in rural Indonesia. JAMA Cardiol 2019 Oct 1;4(10):978-986. [doi: 10.1001/jamacardio.2019.2974] [Medline: 31461123]
12. A!Community!Health!Needs!Assessment. Healthier San Joaquin County. 2013. URL: https://www.healthiersanjoaquin.org/pdfs/2013/CHNA_SJC_Final2.pdf [accessed 2020-06-26]
13. Paasche-Orlow MK, Wolf MS. Promoting health literacy research to reduce health disparities. J Health Commun 2010;15 Suppl 2(Suppl 2):34-41. [doi: 10.1080/10810730.2010.499994] [Medline: 20845191]
14. Mackert M, Mabry-Flynn A, Champlin S, Donovan EE, Pounders K. Health literacy and health information technology adoption: the potential for a new digital divide. J Med Internet Res 2016 Oct 4;18(10):e264 [FREE Full text] [doi: 10.2196/jmir.6349] [Medline: 27702738]
15. Margolius D, Bodenheimer T, Bennett H, Wong J, Ngo V, Padilla G, et al. Health coaching to improve hypertension treatment in a low-income, minority population. Ann Fam Med 2012;10(3):199-205 [FREE Full text] [doi: 10.1370/afm.1369] [Medline: 22585883]
16. Benzo R, Vickers K, Novotny PJ, Tucker S, Hoult J, Neuenfeldt P, et al. Health coaching and chronic obstructive pulmonary disease rehospitalization. A randomized study. Am J Respir Crit Care Med 2016 Sep 15;194(6):672-680 [FREE Full text] [doi: 10.1164/rccm.201512-2503OC] [Medline: 26953637]
17. Chang H, Hawley N, Kalyesubula R, Siddharthan T, Checkley W, Knauf F, et al. Challenges to hypertension and diabetes management in rural Uganda: a qualitative study with patients, village health team members, and health care professionals. Int J Equity Health 2019 Feb 28;18(1):38 [FREE Full text] [doi: 10.1186/s12939-019-0934-1] [Medline: 30819193]
18. Ballard PJ, Syme SL. Engaging youth in communities: a framework for promoting adolescent and community health. J Epidemiol Community Health 2016 Feb;70(2):202-206. [doi: 10.1136/jech-2015-206110] [Medline: 26443541]
19. Karsenti T, Fievez A. The Ipad in Education: Uses, Benefits and Challenges. A Survey of 6057 Students and 302 Teachers in Quebec, Canada. CRIFPE 2013:- [FREE Full text]
20. Thom DH, Ghorob A, Hessler D, de Vore D, Chen E, Bodenheimer TA. Impact of peer health coaching on glycemic control in low-income patients with diabetes: a randomized controlled trial. Ann Fam Med 2013;11(2):137-144 [FREE Full text] [doi: 10.1370/afm.1443] [Medline: 23508600]
21. San Joaquin County 2016 Community Health Needs Assessment. Healthier San Joaquin County. 2016. URL: https://healthiersanjoaquin.org/pdfs/2016/2016_CHNA_full_document-narrative_and_health_profiles.pdf [accessed 2020-06-26]
22. Tsimploulis A, Sheriff HM, Lam PH, Dooley DJ, Anker MS, Papademetriou V, et al. Corrigendum to 'systolic-diastolic hypertension versus isolated systolic hypertension and incident heart failure in older adults: insights from the cardiovascular health study' [int j cardiol 235 (2017) 11-16]. Int J Cardiol 2017 Jul 1;238:181. [doi: 10.1016/j.ijcard.2017.04.072] [Medline: 28487149]
23. Flint AC, Conell C, Ren X, Banki NM, Chan SL, Rao VA, et al. Effect of systolic and diastolic blood pressure on cardiovascular outcomes. N Engl J Med 2019 Jul 18;381(3):243-251. [doi: 10.1056/NEJMoa1803180] [Medline: 31314968]

Abbreviations

BP: blood pressure
COPD: chronic obstructive pulmonary disease
HeH: Health eHeart
IRB: institutional review board
mHealth: mobile health
MUA: medically underserved area
PCP: primary care physician
Q: QardioArm alone
S: student health coach alone
SES: socioeconomic status
SHC: student health coach
SMART goal: specific, measurable, achievable, realistic, and time-based goal
UCSF: University of California, San Francisco
YCE: youth civic engagement

Edited by G Eysenbach; submitted 05.02.19; peer-reviewed by J Edwards, J Volkman; comments to author 02.10.19; revised version received 25.02.20; accepted 14.05.20; published 25.08.20.

Please cite as:
Vasti E, Pletcher MJ
Recruiting Student Health Coaches to Improve Digital Blood Pressure Management: Randomized Controlled Pilot Study
JMIR Form Res 2020;4(8):e13637
URL: https://formative.jmir.org/2020/8/e13637
doi:10.2196/13637
PMID:32840489

Original Paper

Shared Decision Making and Patient-Centered Care in Israel, Jordan, and the United States: Exploratory and Comparative Survey Study of Physician Perceptions

Yaara Zisman-Ilani[1], MA, PhD; Rana Obeidat[2], PhD; Lauren Fang[3], BSc; Sarah Hsieh[3], BSc; Zackary Berger[3], MD, PhD

[1]Department of Social and Behavioral Sciences, College of Public Health, Temple University, Philadelphia, PA, United States

[2]Faculty of Nursing, Zarqa University, Zarqa, Jordan

[3]Johns Hopkins School of Medicine, Baltimore, MD, United States

Corresponding Author:
Yaara Zisman-Ilani, MA, PhD
Department of Social and Behavioral Sciences
College of Public Health
Temple University
1700 North Broad St
Philadelphia, PA, 19122
United States
Phone: 1 215 204 5618
Email: yaara@temple.edu

Abstract

Background: Shared decision making (SDM) is a health communication model that evolved in Europe and North America and largely reflects the values and medical practices dominant in these areas.

Objective: This study aims to understand the beliefs, perceptions, and practices related to SDM and patient-centered care (PCC) of physicians in Israel, Jordan, and the United States.

Methods: A hypothesis-generating comparative survey study was administered to physicians from Israel, Jordan, and the United States.

Results: A total of 36 surveys were collected via snowball sampling (Jordan: n=15; United States: n=12; Israel: n=9). SDM was perceived as a way to inform patients and allow them to participate in their care. Barriers to implementing SDM varied based on place of origin; physicians in the United States mentioned limited time, physicians in Jordan reported that a lack of patient education limits SDM practices, and physicians in Israel reported lack of communication training. Most US physicians defined PCC as a practice for prioritizing patient preferences, whereas both Jordanian and Israeli physicians defined PCC as a holistic approach to care and to prioritizing patient needs. Barriers to implementing PCC, as seen by US physicians, were mostly centered on limited appointment time and insurance coverage. In Jordan and Israel, staff shortage and a lack of resources in the system were seen as major barriers to PCC implementation.

Conclusions: The study adds to the limited, yet important, literature on SDM and PCC in areas of the world outside the United States, Canada, Australia, and Western Europe. The study suggests that perceptions of PCC might widely differ among these regions, whereas concepts of SDM might be shared. Future work should clarify these differences.

(JMIR Form Res 2020;4(8):e18223) doi:10.2196/18223

KEYWORDS

shared decision making; patient-centered care; Middle East; physicians; perceptions

Introduction

Shared decision making (SDM) is a central health communication model for supporting patient engagement in health care [1-3] and a recommended approach to increasing patient engagement and patient-centered care (PCC) in clinical decision making [4-6]. SDM evolved in Europe and North America [7] and largely reflects the values and medical practices

dominant in these areas [8,9]. Although SDM has become more widely discussed in recent years in non-Western countries (eg, China, Peru, Malaysia, Taiwan, Iran) [10], it has yet to be implemented on a wider scale, and less is known about how or whether attitudes, beliefs, and practices regarding PCC exist or differ in various other regions of the world.

The overall aim of the present exploratory study was to explore the factors that enable or impede SDM implementation in different geographical and political contexts. Specifically, we sought to conduct a hypothesis-generating study and to collect preliminary data to better understand SDM- and PCC-related beliefs, attitudes, and practices of physicians in four regions in the Middle East characterized by different health care systems, cultures, and political environments: Israel, Jordan, and the West Bank. As a point of reference, we conducted a similar survey among US physicians to serve as a benchmark for SDM- and PCC-related beliefs, attitudes, and practices. In addition, such a comparison may provide insights about the importance of a health care system that facilitates the practice of SDM and PCC. The study focused on the following specific questions: (1) What are physicians' common understandings and perceptions of the concepts of SDM and PCC? and (2) Do physicians find SDM and PCC to be feasible in their practice and in health care?

Methods

Settings: Context of Participating Countries

The survey was intended to be administered to physicians from different geographical and political contexts in the Middle East characterized by different health care systems: Israel, Jordan, and the West Bank. Israel is a democratic state with an efficient health care system that has been ranked among the top 10 health care systems for several years [11,12]. Israel's national health insurance system provides universal health coverage throughout the country, with a significant spread of hospital and clinics [13-15]. Residents can supplement the universal coverage with additional forms of private health insurance. Israel's health policy legislation is supportive of SDM principles, including the right to be informed of treatment options and risks [16]. Jordan is a constitutional monarchy state with a health care system characterized by diverse types of payers (public, private, and donors) [17]. The public health care sector is the largest in Jordan; however, only about 70% of residents have some form of public health insurance. Jordan has a ratio of 2.3 physicians to 1000 residents, and hospitals are mostly centralized in the larger urban areas. The massive influx of Syrian refugees due to the start of the Arab Spring in 2011 has further increased burdens on the Jordanian health care system, especially on public health facilities. The West Bank is an independent Palestinian territory governed by the Palestinian National Authority. The West Bank has low-functioning, inefficient health care systems that rely heavily on medical services to and referrals of patients to Israel (14% in 2011) or Jordan (13% in 2011) [17,18].

A parallel survey was planned among US physicians to serve as a benchmark. The United States is a representative democracy with a hybrid health care system but without universal health care coverage. In 2016, 48% of US health care spending came from private funds, with 28% coming from households and 20% coming from private businesses. The federal government accounted for 28% of spending, while state and local governments accounted for 17% of spending [19]. SDM in the United States is increasingly recognized as part of value-based care, and several federal initiatives have linked SDM to reimbursement [20].

Survey Development and Structure

Because the purpose of the present study was to explore aspects of SDM and PCC and to provide data for testing hypotheses, we chose to develop a survey as a research strategy [20]. As recommended by Enhancing the Quality and Transparency of Health Research (EQUATOR), we used "Good Practice in the Conduct and Reporting of Survey Research" as a reporting guideline [20]. We developed a short standardized survey form with 24 questions divided into 3 parts (see Multimedia Appendix 1): (1) demographic data, (2) qualitative evaluation, and (3) quantitative evaluation.

First, demographic data (eg, age and years in practice) was assessed with 9 questions.

Second, a qualitative evaluation was included. One part of this evaluation assessed the beliefs and attitudes of physicians around PCC and SDM in their health care setting (eg, "What do you see as barriers to implementing shared decision making in your practice?"). This comprised 5 open-ended questions and 1 multiple choice question. The second part of the qualitative evaluation assessed the understanding of physicians of their immediate environment of practice in the context of their larger health care system (eg, "What are the most important day-to-day problems in the practice of medicine or health care in your country or region?"). This included 3 open-ended questions.

Third, the quantitative evaluation of level of SDM practice was based on the Shared Decision Making Questionnaire, physician version (SDM-Q-DOC) scale [21]. This included 9 questions on a 6-point Likert scale ranging from "completely disagree" (0) to "completely agree" (5). The SDM-Q-DOC was developed to measure patients' and clinicians' agreement with steps and actions defined by a medically driven SDM model at the end of a medical consultation. It has been used in numerous studies to measure physicians' perspectives of SDM and was recommended for use in health policy survey responses targeting the implementation of SDM [22-24]

The qualitative part of the survey was developed through an iterative process based on SDM and PCC literature related to the delivery and perception of care and on the lead investigators' (YZI and ZB) knowledge [25-30]. The development process included discussions among the coinvestigators and piloting among colleagues. We conducted forward and backward translations based on accepted guidelines [31] to each new question in sections 1 and 2. We used the English version of the SDM-Q-DOC questionnaire [21] and the Arabic [32] and Hebrew [33] translations of the 9-item Shared Decision Making Questionnaire (SDM-Q-9), a parallel patient version [34], with the needed minor adaptions.

Procedure

A web-based survey developed for the study was emailed to physicians in Israel, the West Bank, and the United States using the Qualtrics platform (Qualtrics International Inc). In Jordan it was advised by one of the coauthors (RO) to administer the survey via face-to-face interviews, based on her previous experience conducting similar types of research in Jordan. A snowball sampling methodology was used to recruit physicians. Accordingly, participants were asked to identify and email the questionnaire to other colleagues. Surveys were administered in Hebrew, English, and Arabic, and all responses were anonymized. Data collection began in February 2017 and ended in June 2017. It was designed to stop after a sample of 15 in each country or after the maximum sample size closest to this threshold. As this was an exploratory study, this sample size target was based not on statistical considerations but on real-world experience with the number of respondents likely to provide a hypothesis-generating set of responses. At the end of the survey, participants were reimbursed via gift cards in the amount of US $10 or an equal value in the local currency.

Data Analysis

To summarize the qualitative results from the open-ended survey questions, we used an integrated approach that enabled both inductive (ie, data-driven) coding of participants' responses and deductive (ie, theory-driven) framework organization of codes [35]. Specifically, 2 coauthors (SH and ZB) and another research assistant read open-ended responses from participants and developed a draft of coding categories based on the responses' contents. These categories were reviewed by the lead author (YZI) and revised accordingly. Responses to open-ended questions were coded independently by all coauthors; differences and disagreements between the coders were resolved through discussions until consensus was achieved. The final coding of open-ended responses was double-checked for accuracy by the first and last authors (YZI and ZB) after finalization of the coding guide. Then, guided by SDM and PCC theories, all coauthors discussed the interrelationships between codes to finalize the grouping of the codes into themes and subthemes.

Chi-square and one-way ANOVA tests were conducted to describe demographic characteristics of the survey respondents. As recommended by the developers [21], multiplication of the raw score by 20/9 provided a transformed total score range from 0 to 100, where 0 indicates the lowest possible level and 100 indicates the highest possible level of SDM. A nonparametric Kruskal-Wallis test was conducted to compare the mean total score of SDM-Q-9 between the 3 countries. Results were considered significant below a *P* value of .05.

Ethics, Consent, and Permissions

Because the responses were anonymized and not identifiable and participation in the study was associated with minimal risk, the Institutional Review Board of the Johns Hopkins School of Medicine deemed this study exempt from requirements for approval (IRB00111847). All participants provided consent to participate through their responses to the survey.

Results

Participants

Eligible survey respondents were practicing physicians in the United States, Israel, Jordan, and the West Bank. A total of 36 survey responses were received from Israel (n=9), Jordan (n=15), and the United States (n=12). Throughout the study period, we received no responses from West Bank physicians to our emails or to our in-person attempts to contact them; therefore, we were unable to collect any data from that population. Most survey respondents were men (24/36, 67%), the mean age of survey respondents was 43.6 years (SD 11.2), and the mean years of clinical experience was 15.8 (SD 10.5). Comparison of demographic characteristics and clinical experience of clinicians in each country indicate similarity between US and Israeli respondents for gender distribution, age, and clinical experience (Table 1). The Jordanian respondents were significantly younger and less experienced, and almost all were men.

Table 1. Demographic and clinical experience characteristics of survey respondents.

Characteristics	Total sample (N=36)	United States (n=12)	Israel (n=9)	Jordan (n=15)	*P* value
Gender					.01[a]
Men, n (%)	5 (42)	24 (67)	5 (56)	14 (93)	—[b]
Women, n (%)	7 (58)	12 (33)	4 (44)	1 (7)	—
Age (years), mean (SD)	43.6 (11.2)	47.0 (8.6)	48.7 (13.4)	37.4 (9.2)	.02[c]
Years of clinical experience, mean (SD)	15.8 (10.5)	19.0 (9.8)	20.0 (14.2)	10.8 (6.0)	.05[d]

[a]Pearson χ^2_2=8.7.

[b]Not applicable.

[c]Analysis of variance *F* test ($F_{2,32}$=4.40).

[d]Analysis of variance *F* test ($F_{2,33}$=3.37).

Open-Ended Responses: Perception of SDM and PCC

We included in the analysis 34 survey responses with greater than 50% total completion (Jordan: n=15; United States: n=12; Israel: n=7). There were 12 responses to the open-ended questions from the US physicians, 7 from the Israeli physicians, and 14 from the Jordanian physicians.

Most respondents defined SDM as a process aimed at informing patients (United States: 8/12, 67%; Israel: 4/7, 57%; Jordan: 12/15, 80%). Whereas most US respondents also defined SDM as the participation of patients in their care (8/12, 67%), only a third of the Jordanian respondents defined it as patient participation (5/15, 33%), and 4 of the 7 (57%) Israeli respondents defined SDM also as collaboration between patient and physician.

> *[SDM is when] a patient makes decisions about medical tests and treatment that incorporate information about benefits and harms from the physician as well as the patient's own understanding of his or her values and priorities.* [US respondent]
>
> *[SDM is] an open conversation with the patient, in which I [the doctor] suggest/advise a variety of treatment options that fit the patient's medical condition, and together with the patient, choose the appropriate treatment method.* [Israeli respondent]
>
> *[SDM is] giving the patient information about his treatment options and his illness and giving him a chance to have a say in his treatment options.* [Jordanian respondent]

Most US and Israeli respondents indicated familiarity with the concept of PCC (Israel: 5/7, 71%; United States: 10/12, 83%), whereas only 6 of the 15 (40%) Jordanian respondents indicated their or their patients' familiarity with the concept. Prioritizing or meeting patient needs was commonly described as a feature of PCC by most respondents regardless of country of origin. In addition, US respondents commonly described PCC as accounting for patients' preferences, most Israeli respondents described PCC also as individualized care, and most Jordanian respondents also described PCC as a provision of holistic care.

> *[PCC refers to] care that balances the needs and desires of the person receiving care.* [US respondent]
>
> *[PCC aims] to provide the patient with all of the patient's needs and not just to solve a problem in the field, while maintaining proper communication and respect for the patient's values.* [Israeli respondent]
>
> *[PCC refers] to doing whatever is needed for the patient or referring him/her to someone who can.* [Jordanian respondent]

Respondents indicated several barriers affecting the provision of SDM and PCC systematically; however, the common barrier was related to the system itself. All US respondents (12/12, 100%) mentioned lack of time as a major barrier to SDM implementation, whereas only 5 of 12 (42%) mentioned it as a barrier to PCC implementation. The role of insurance companies and fragmentation of care were mentioned as additional possible barriers to PCC implementation. In Jordan, most respondents (9/15, 60%) mentioned patient-related barriers, low health literacy, and a lack of knowledge as barriers to SDM implementation, whereas system-related barriers, such as staff shortages and high patient loads, were identified as barriers to PCC.

> *Time, and often hard to do for many decisions: few are really straightforward. Would be nice to have tools readily available to do this & ways to facilitate it.* [US respondent, regarding SDM barriers]
>
> *Lack of knowledge among patients and patient unwillingness to be fully informed about his/her condition.* [Jordanian respondent, regarding SDM barriers]
>
> *The healthcare delivery system is still organized traditionally regarding appointment scheduling and how patients interact with doctors; short visits limit person-centered care.* [US respondent, regarding PCC barriers]
>
> *Bureaucracy of the Jordanian health care system and lack of medical specialties in the peripheral areas of the country.* [Jordanian respondent, regarding PCC barriers]

With respect to problems related to the patient-physician relationship, the responses of Israeli physicians emphasized a lack of time and training (eg, lack of time, lack of support for physicians during their work). Jordanian respondents emphasized disorganization of the health care system, and US physicians highlighted problems of cost, social determinants of health, and the role of insurance companies (eg, the payment system and its incentive structure, lack of universal health care, costs of pharmaceuticals).

SDM-Q-DOC Responses: Comparison of SDM Practice and PCC Behaviors

We included in the analysis 32 survey responses with greater than 50% total completion (Jordan: n=15; United States: n=10; Israel: n=7). Overall, physicians in our sample reported practicing SDM at a moderately high level (mean 76.6, SD 11.5; median 75.6), with a range of 53 to 100. This result is similar to findings in other studies [22]. The Kruskal-Wallis test results showed no significant difference in SDM-Q-DOC scores between the 3 countries (H_2=0.631, P=.73; United States: mean 74.9, SD 11.4; Israel: mean 78.7, SD 7.9; Jordan: mean 76.7, SD 13.4). Box plot diagrams of SDM-Q-DOC means and standard deviations imply that SDM practice and PCC behaviors vary more among the US and Jordanian respondents in our sample than among their Israeli counterparts, as seen in Figure 1.

In addition, we compared the individual item scores between respondents from the 3 countries. Only for the first item ("I make clear to my patient that a decision needs to be made") was a significant difference noticed, with higher scores for Jordanian physicians (Table 2). A nonsignificant but notable difference was noticed in the second item ("I want to know exactly from my patient how he/she wants to be involved in making the decision"), with a higher mean score for Jordanian physicians.

Figure 1. Box plot diagrams of SDM-Q-DOC score per country. SDM-Q-DOC: Shared Decision Making Questionnaire, physician version.

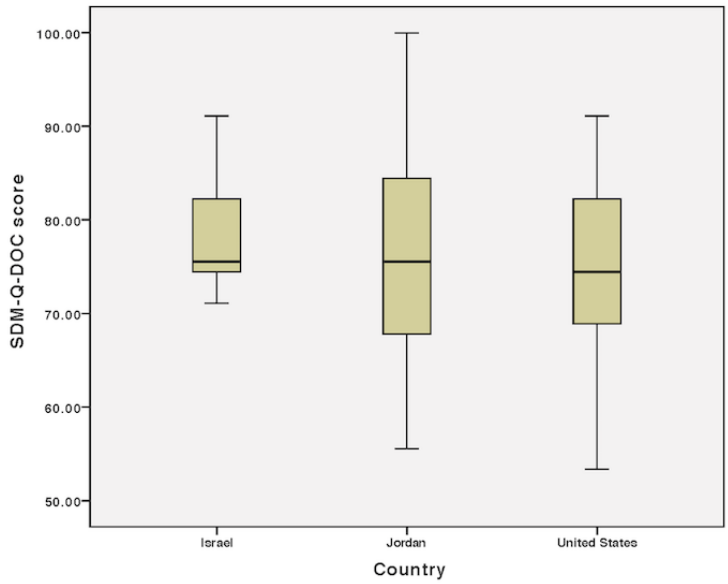


Table 2. Comparison of means and standard deviations of Shared Decision Making Questionnaire, physician version responses among respondents from the United States, Israel, and Jordan (n=32).

SDM-Q-DOC[a] item question	H statistic	*P* value	United States, mean (SD)[b]	Israel, mean (SD)[b]	Jordan, mean (SD)[b]
1. I make clear to my patient that a decision needs to be made.	9.55	.008	3.60 (0.84)	3.57 (0.98)	4.47 (0.52)
2. I want to know exactly from my patient how he/she wants to be involved in making the decision.	4.66	.10	3.00 (0.67)	3.29 (0.76)	3.73 (0.80)
3. I tell my patient that there are different options for treating his/her medical condition.	1.83	.40	4.50 (0.71)	4.43 (0.98)	4.21 (0.58)
4. I precisely explain the advantages and disadvantages of the treatment options to my patient.	2.59	.27	3.80 (0.92)	4.43 (0.53)	4.13 (0.64)
5. I help my patient understand all the information.	2.06	.36	3.80 (0.92)	4.29 (0.76)	4.27 (0.96)
6. I ask my patient which treatment option he/she prefers.	1.00	.61	4.10 (0.74)	4.00 (1.00)	3.73 (0.96)
7. My patient and I thoroughly weigh the different treatment options.	2.53	.28	3.40 (0.52)	4.00 (0.82)	3.47 (0.92)
8. My patient and I select a treatment option together.	0.23	.89	3.50 (0.53)	3.43 (0.79)	3.13 (1.46)
9. My patient and I reach an agreement on how to proceed.	1.47	.48	8.84 (0.76)	4.00 (5.80)	3.67 (0.90)

[a]SDM-Q-DOC: Shared Decision Making Questionnaire, physician version.

[b]Means represent level of agreement with each item on a 6-point Likert scale, where 0=completely disagree and 5=completely agree.

Discussion

Principal Results and Comparison With Prior Work

The present study describes findings of a small, exploratory hypothesis-generating survey [20] of Israeli, Jordanian, and US physicians' perceptions of SDM and PCC. Open-ended qualitative results suggest that respondents, regardless of country of origin, identify SDM as a process focused on providing information or as informed decision making, but PCC as a physician's effort to meet patients' individualized needs. These findings are aligned with the perception that SDM is "the pinnacle of PCC" [4] but also emphasize that SDM remains commonly perceived by physicians as a mean for delivering information rather than a collaborative discussion [36,37]. Barriers to implementing SDM and PCC were also identified and attributed to system- and patient-related factors [25]. The quantitative results of the total mean score of the SDM-Q-DOC

show medium to high levels of SDM-related behaviors among all respondents. Item-focused analysis showed that Jordanian respondents scored significantly higher on item 1 ("I make clear to my patient that a decision needs to be made"), with a nonsignificant but notable difference for item 2 ("I want to know exactly from my patient how he/she wants to be involved in making the decision"). These are interesting results for the psychometric quality of the SDM-Q-DOC. Although there is less literature on the psychometric qualities of the SDM-Q-DOC, ample literature exists on the psychometric characteristics of the SDM-Q-9, showing mixed results for item 1 and suggesting eliminating the item to improve the factorial structure [22]. In our small sample, item 1 served as a discriminate item.

Strengths and Limitations

The present study has several strengths. To the best of our knowledge, this is the first transnational comparison of the perceptions and practices of physicians in the United States, Israel, and Jordan. This survey study provided a conceptual overview of physicians' understanding of SDM and PCC as well as an evaluation of SDM- and PCC-related behaviors. Although SDM is a communication model and practice and PCC is considered the conceptual framework [6], the participating physicians' interpretation and understanding of the two were different. SDM was generally perceived as a means for delivering information to patients, whereas PCC was commonly perceived as a method for meeting a patient's individual needs. While US and Jordanian physicians in our sample interpreted SDM also as a patient-level process (ie, patient participation), Israeli physicians in our sample interpreted SDM as a dialogue- and dyadic-based process occurring between patient and physician. These insights can inform future research and education initiatives pertaining to SDM and PCC among physicians. Finally, a methodological strength is our ability to deliver surveys in Arabic, Hebrew, and English due to the multilingual expertise of our team.

There are also some limitations to note. Our sample is small and not random; thus, we are unable to make meaningful statistical inferences or to generalize our findings, decreasing the study's validity. However, because the purpose of this exploratory study was to generate hypotheses, our snapshot of the SDM landscape and PCC perceptions among physicians in Israel and Jordan is new and will inform future scaled-up surveys. We surveyed physicians, not other members of the health care team or patients. Clearly, a future comprehensive comparison of these health systems and practitioners' approaches to care requires surveying all stakeholders involved. The SDM-Q-DOC questionnaire was administered to explore physicians' perceptions of what is important in an SDM encounter, but it was not used to evaluate an actual encounter. Although the SDM-Q-DOC was developed to rate physicians' experiences of SDM in patient-physician encounters, recent literature indicates the feasibility of the Shared Decision Making Questionnaire family for use in surveys [23,28,38]. The final limitations are that we were unable to recruit and collect data from physicians in the West Bank and that the survey in Jordan was conducted using face-to-face interviews rather than web-based surveys. Using email to initiate communication with potential Jordanian and West Bank respondents was challenging, as was administering a web-based survey. In Jordan, we learned that recruitment by phone call or an offline survey methodology that does not require an internet connection could be better for collecting responses. Therefore, we employed face-to-face interviews successfully, but this might have caused a social desirability bias; that is, physicians might have overestimated their support for and use of the SDM approach. However, because the web-based platform was the tool rather than the purpose, we believe it was not critical and that the benefits of collecting data face to face instead of via a web-based survey were more important for this project.

Conclusions

The results of the present study add to the limited, yet important, literature on SDM and PCC in areas of the world outside the United States, Canada, Australia, and Western Europe [9,12,39-43]. They also add to the psychometric evaluation of SDM and PCC measures [44] and identify barriers to implementation. We hope this survey will motivate researchers and clinicians in Israel, Jordan, and other countries that are less represented in the SDM and PCC research and practice to encourage related discussions and practice and to facilitate implementation, measurements, and interventions.

Acknowledgments

The authors wish to thank Young Shin Kim for her help in this research.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Physician Perceptions of Shared Decision Making and Person-Centered Care Survey.
[DOCX File , 22 KB - formative_v4i8e18223_app1.docx]

References

1. Gordon JE, Leiman JM, Deland EL, Pardes H. Delivering value: provider efforts to improve the quality and reduce the cost of health care. Annu Rev Med 2014;65:447-458. [doi: 10.1146/annurev-med-100312-135931] [Medline: 24111890]

2. 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]
3. Charles C, Gafni A, Whelan T. Decision-making in the physician-patient encounter: revisiting the shared treatment decision-making model. Soc Sci Med 1999 Sep;49(5):651-661. [doi: 10.1016/s0277-9536(99)00145-8] [Medline: 10452420]
4. Barry MJ, Edgman-Levitan S. Shared decision making--pinnacle of patient-centered care. N Engl J Med 2012 Mar 01;366(9):780-781. [doi: 10.1056/NEJMp1109283] [Medline: 22375967]
5. Makoul G, Clayman ML. An integrative model of shared decision making in medical encounters. Patient Educ Couns 2006 Mar;60(3):301-312. [doi: 10.1016/j.pec.2005.06.010] [Medline: 16051459]
6. Salyers MP, Zisman-Ilani Y. Shared decision-making and self-directed care. In: Goldman HH, Frank RG, Morrissey JP, editors. The Palgrave Handbook of American Mental Health Policy. Cham, Switzerland: Palgrave Macmillan; 2020:197-228.
7. Hoving C, Visser A, Mullen P, van den Borne B. A History of Patient Education by Health Professionals in Europe and North America: From Authority to Shared Decision Making Education. Patient Education and Counseling 2010;78(3):275-281 [FREE Full text] [doi: 10.1016/j.pec.2010.01.015] [Medline: 20189746]
8. Hamann J, Bieber C, Elwyn G, Wartner E, Hörlein E, Kissling W, et al. How do patients from eastern and western Germany compare with regard to their preferences for shared decision making? Eur J Public Health 2012 Aug;22(4):469-473. [doi: 10.1093/eurpub/ckr112] [Medline: 21873278]
9. Obeidat RF, Homish GG, Lally RM. Shared decision making among individuals with cancer in non-Western cultures: a literature review. Oncol Nurs Forum 2013 Sep;40(5):454-463. [doi: 10.1188/13.ONF.454-463] [Medline: 23989019]
10. Härter M, Moumjid N, Cornuz J, Elwyn G, van der Weijden T. Shared decision making in 2017: International accomplishments in policy, research and implementation. Z Evid Fortbild Qual Gesundhwes 2017 Jun;123-124:1-5. [doi: 10.1016/j.zefq.2017.05.024] [Medline: 28546053]
11. Brammli-Greenberg S, Waitzberg R, Medina-Artom T, Adijes-Toren A. Low-budget policy tool to empower Israeli insureds to demand their rights in the healthcare system. Health Policy 2014 Dec;118(3):279-284 [FREE Full text] [doi: 10.1016/j.healthpol.2014.11.005] [Medline: 25467282]
12. Karnieli-Miller O, Miron-Shatz T, Siegal G, Zisman-Ilani Y. On the verge of shared decision making in Israel: Overview and future directions. Z Evid Fortbild Qual Gesundhwes 2017 Jun;123-124:56-60. [doi: 10.1016/j.zefq.2017.05.007] [Medline: 28529120]
13. Miron-Shatz T, Golan O, Brezis M, Siegal G, Doniger GM. Shared decision-making in Israel: status, barriers, and recommendations. Isr J Health Policy Res 2012 Jan 30;1(1):5 [FREE Full text] [doi: 10.1186/2045-4015-1-5] [Medline: 22913605]
14. Berenson A, Khalaila R. Patient's Rights Law and Culturally Competent Nursing Care: An Israeli Perspective. Med Law 2014 Oct;33(3):35-48. [Medline: 27359016]
15. Israel Ministry of Health. Inpatient Institutions and Day Care Units in Israel [report]. In: Israel Ministry of Health. Jerusalem, Israel: Israel Ministry of Health; 2015.
16. Nazer LH, Tuffaha H. Health Care and Pharmacy Practice in Jordan. Can J Hosp Pharm 2017;70(2):150-155 [FREE Full text] [doi: 10.4212/cjhp.v70i2.1649] [Medline: 28487583]
17. Vitullo A, Soboh A, Oskarsson J, Atatrah T, Lafi M, Laurance T. Barriers to the access to health services in the occupied Palestinian territory: a cohort study. The Lancet 2012 Oct;380:S18-S19. [doi: 10.1016/s0140-6736(13)60200-7]
18. Mahmoud A. Health Challenges in Palestine. Science & Diplomacy. 2013. URL: http://www.sciencediplomacy.org/perspective/2013/health-challenges-in-palestine [accessed 2020-07-03]
19. Agency for Healthcare Research and Quality. The SHARE Approach—Achieving Patient-Centered Care with Shared Decisionmaking: A Brief for Administrators and Practice Leaders. AHRQ. Rockville, MD: Agency for Healthcare Research and Quality; 2014. URL: https://www.ahrq.gov/health-literacy/curriculum-tools/shareddecisionmaking/tools/tool-9/index.html [accessed 2020-07-03]
20. Kelley K, Clark B, Brown V, Sitzia J. Good practice in the conduct and reporting of survey research. Int J Qual Health Care 2003 Jun;15(3):261-266. [doi: 10.1093/intqhc/mzg031] [Medline: 12803354]
21. Scholl I, Kriston L, Dirmaier J, Buchholz A, Härter M. Development and psychometric properties of the Shared Decision Making Questionnaire--physician version (SDM-Q-Doc). Patient Educ Couns 2012 Aug;88(2):284-290. [doi: 10.1016/j.pec.2012.03.005] [Medline: 22480628]
22. Doherr H, Christalle E, Kriston L, Härter M, Scholl I. Use of the 9-item Shared Decision Making Questionnaire (SDM-Q-9 and SDM-Q-Doc) in intervention studies-A systematic review. PLoS One 2017;12(3):e0173904 [FREE Full text] [doi: 10.1371/journal.pone.0173904] [Medline: 28358864]
23. Rencz F, Tamási B, Brodszky V, Gulácsi L, Weszl M, Péntek M. Validity and reliability of the 9-item Shared Decision Making Questionnaire (SDM-Q-9) in a national survey in Hungary. Eur J Health Econ 2019 Jun;20(Suppl 1):43-55 [FREE Full text] [doi: 10.1007/s10198-019-01061-2] [Medline: 31111402]
24. Calderon C, Ferrando PJ, Carmona-Bayonas A, Lorenzo-Seva U, Jara C, Beato C, et al. Validation of SDM-Q-Doc Questionnaire to measure shared decision-making physician's perspective in oncology practice. Clin Transl Oncol 2017 Nov;19(11):1312-1319. [doi: 10.1007/s12094-017-1671-9] [Medline: 28497424]

25. Légaré F, Witteman HO. Shared Decision Making: Examining Key Elements And Barriers To Adoption Into Routine Clinical Practice. Health Affairs 2013 Feb;32(2):276-284. [doi: 10.1377/hlthaff.2012.1078] [Medline: 23381520]
26. Hamann J, Mendel R, Bühner M, Kissling W, Cohen R, Knipfer E, et al. How should patients behave to facilitate shared decision making--the doctors' view. Health Expect 2012 Dec;15(4):360-366 [FREE Full text] [doi: 10.1111/j.1369-7625.2011.00682.x] [Medline: 21624024]
27. Kaminskiy E, Senner S, Hamann J. Attitudes towards shared decision making in mental health: a qualitative synthesis. Mental Health Review Journal 2017 Sep 11;22(3):233-256. [doi: 10.1108/mhrj-01-2017-0003]
28. Luxford K, Gelb Safran D, Delbanco T. Promoting patient-centered care: a qualitative study of facilitators and barriers in healthcare organizations with a reputation for improving the patient experience. Int J Qual Health Care 2011 Oct;23(5):510-515. [doi: 10.1093/intqhc/mzr024] [Medline: 21586433]
29. Epstein RM, Street RL. The values and value of patient-centered care. Ann Fam Med 2011;9(2):100-103 [FREE Full text] [doi: 10.1370/afm.1239] [Medline: 21403134]
30. Sepucha K, Breslin M, Graffeo C, Carpenter C, Hess E. State of the Science: Tools and Measurement for Shared Decision Making. Acad Emerg Med 2016 Dec;23(12):1325-1331 [FREE Full text] [doi: 10.1111/acem.13071] [Medline: 27770488]
31. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report measures. Spine (Phila Pa 1976) 2000 Dec 15;25(24):3186-3191. [doi: 10.1097/00007632-200012150-00014] [Medline: 11124735]
32. Alzubaidi H, Hussein A, Mc Namara K, Scholl I. Psychometric properties of the Arabic version of the 9-item Shared Decision-Making Questionnaire: the entire process from translation to validation. BMJ Open 2019 Apr 04;9(4):e026672 [FREE Full text] [doi: 10.1136/bmjopen-2018-026672] [Medline: 30948608]
33. Zisman-Ilani Y, Roe D, Scholl I, Härter M, Karnieli-Miller O. Shared Decision Making During Active Psychiatric Hospitalization: Assessment and Psychometric Properties. Health Commun 2017 Jan;32(1):126-130. [doi: 10.1080/10410236.2015.1099504] [Medline: 27168160]
34. Kriston L, Scholl I, Hölzel L, Simon D, Loh A, Härter M. The 9-item Shared Decision Making Questionnaire (SDM-Q-9). Development and psychometric properties in a primary care sample. Patient Educ Couns 2010 Jul;80(1):94-99. [doi: 10.1016/j.pec.2009.09.034] [Medline: 19879711]
35. Bradley EH, Curry LA, Devers KJ. Qualitative data analysis for health services research: developing taxonomy, themes, and theory. Health Serv Res 2007 Aug;42(4):1758-1772 [FREE Full text] [doi: 10.1111/j.1475-6773.2006.00684.x] [Medline: 17286625]
36. Hargraves I, LeBlanc A, Shah ND, Montori VM. Shared Decision Making: The Need For Patient-Clinician Conversation, Not Just Information. Health Aff (Millwood) 2016 Apr;35(4):627-629. [doi: 10.1377/hlthaff.2015.1354] [Medline: 27044962]
37. Coulter A. Shared decision making: everyone wants it, so why isn't it happening? World Psychiatry 2017 Jun;16(2):117-118 [FREE Full text] [doi: 10.1002/wps.20407] [Medline: 28498596]
38. Pollard S, Bansback N, Bryan S. Physician attitudes toward shared decision making: A systematic review. Patient Educ Couns 2015 Sep;98(9):1046-1057. [doi: 10.1016/j.pec.2015.05.004] [Medline: 26138158]
39. Rahimi SA, Alizadeh M, Légaré F. Shared decision making in Iran: Current and future trends. Z Evid Fortbild Qual Gesundhwes 2017 Jun;123-124:52-55. [doi: 10.1016/j.zefq.2017.05.018] [Medline: 28549749]
40. Liao H, Liang H, Chen H, Chang C, Wang P, Shih C. Shared decision making in Taiwan. Z Evid Fortbild Qual Gesundhwes 2017 Jun;123-124:95-98. [doi: 10.1016/j.zefq.2017.05.009] [Medline: 28526230]
41. Huang R, Gionfriddo MR, Zhang L, Leppin AL, Ting HH, Montori VM. Shared decision-making in the People's Republic of China: current status and future directions. Patient Prefer Adherence 2015;9:1129-1141 [FREE Full text] [doi: 10.2147/PPA.S82110] [Medline: 26273201]
42. Baicus C, Balanescu P, Zeh S, Oprisan E, Lapadatu R, Gurghean A, et al. Characteristics of shared decision making in Romania from the patient perspective: A cross-sectional multicentric study. J Eval Clin Pract 2019 Dec;25(6):1152-1159. [doi: 10.1111/jep.13257] [Medline: 31407420]
43. Scalia P, Elwyn G, Barr P, Song J, Zisman-Ilani Y, Lesniak M, et al. Exploring the use of Option Grid™ patient decision aids in a sample of clinics in Poland. Z Evid Fortbild Qual Gesundhwes 2018 Jul;134:1-8 [FREE Full text] [doi: 10.1016/j.zefq.2018.04.002] [Medline: 29858145]
44. Gärtner FR, Bomhof-Roordink H, Smith I, Scholl I, Stiggelbout A, Pieterse A. The quality of instruments to assess the process of shared decision making: A systematic review. PLoS One 2018;13(2):e0191747 [FREE Full text] [doi: 10.1371/journal.pone.0191747] [Medline: 29447193]

Abbreviations

EQUATOR: Enhancing the Quality and Transparency of Health Research
PCC: patient-centered care
SDM: shared decision making
SDM-Q-9: 9-item Shared Decision Making Questionnaire
SDM-Q-DOC: Shared Decision Making Questionnaire, physician version

Edited by G Eysenbach; submitted 12.02.20; peer-reviewed by N Diouf, T Wieringa, D Galasinski; comments to author 23.03.20; revised version received 21.04.20; accepted 13.05.20; published 03.08.20.

Please cite as:
Zisman-Ilani Y, Obeidat R, Fang L, Hsieh S, Berger Z
Shared Decision Making and Patient-Centered Care in Israel, Jordan, and the United States: Exploratory and Comparative Survey Study of Physician Perceptions
JMIR Form Res 2020;4(8):e18223
URL: https://formative.jmir.org/2020/8/e18223
doi:10.2196/18223
PMID:32744509

Review

What You Need to Know Before Implementing a Clinical Research Data Warehouse: Comparative Review of Integrated Data Repositories in Health Care Institutions

Kristina K Gagalova[1,2,3], MSc; M Angelica Leon Elizalde[3,4], MSc; Elodie Portales-Casamar[3,5], PhD; Matthias Görges[3,6], PhD

[1]Canada's Michael Smith Genome Sciences Centre, BC Cancer, Vancouver, BC, Canada

[2]Bioinformatics Graduate Program, University of British Columbia, Vancouver, BC, Canada

[3]Research Institute, BC Children's Hospital, Vancouver, BC, Canada

[4]School of Population and Public Health, University of British Columbia, Vancouver, BC, Canada

[5]Department of Pediatrics, University of British Columbia, Vancouver, BC, Canada

[6]Department of Anesthesiology, Pharmacology and Therapeutics, University of British Columbia, Vancouver, BC, Canada

Corresponding Author:
Matthias Görges, PhD
Department of Anesthesiology, Pharmacology and Therapeutics
University of British Columbia
Rm V3-324
950 West 28th Avenue
Vancouver, BC, V5Z 4H4
Canada
Phone: 1 875 2000 ext 5616
Email: mgorges@bcchr.ca

Abstract

Background: Integrated data repositories (IDRs), also referred to as clinical data warehouses, are platforms used for the integration of several data sources through specialized analytical tools that facilitate data processing and analysis. IDRs offer several opportunities for clinical data reuse, and the number of institutions implementing an IDR has grown steadily in the past decade.

Objective: The architectural choices of major IDRs are highly diverse and determining their differences can be overwhelming. This review aims to explore the underlying models and common features of IDRs, provide a high-level overview for those entering the field, and propose a set of guiding principles for small- to medium-sized health institutions embarking on IDR implementation.

Methods: We reviewed manuscripts published in peer-reviewed scientific literature between 2008 and 2020, and selected those that specifically describe IDR architectures. Of 255 shortlisted articles, we found 34 articles describing 29 different architectures. The different IDRs were analyzed for common features and classified according to their data processing and integration solution choices.

Results: Despite common trends in the selection of standard terminologies and data models, the IDRs examined showed heterogeneity in the underlying architecture design. We identified 4 common architecture models that use different approaches for data processing and integration. These different approaches were driven by a variety of features such as data sources, whether the IDR was for a single institution or a collaborative project, the intended primary data user, and purpose (research-only or including clinical or operational decision making).

Conclusions: IDR implementations are diverse and complex undertakings, which benefit from being preceded by an evaluation of requirements and definition of scope in the early planning stage. Factors such as data source diversity and intended users of the IDR influence data flow and synchronization, both of which are crucial factors in IDR architecture planning.

(JMIR Form Res 2020;4(8):e17687) doi:10.2196/17687

KEYWORDS

database; data warehousing; data aggregation; information storage and retrieval; data analytics; health informatics

Introduction

Background

An electronic health record (EHR) is a system for the input, processing, storage, and retrieval of digital health data. EHR systems have been increasingly adopted in the United States over the past 10 years [1], and their use is spreading worldwide in both hospital and outpatient care settings [2,3]. An EHR is typically organized in a patient-centric manner and has become a powerful tool to store data in a time-dependent and longitudinal structure. EHR data can also be integrated into an enterprise data warehouse or integrated data repository (IDR). IDRs collect heterogeneous data from multiple sources and present them to the user through a comprehensive view [4]. Unlike EHRs, IDRs offer specialized analytical tools for researchers or analysts to perform data analyses.

An IDR is a significant institutional investment in terms of both initial costs and maintenance, but it offers the advantage of clinical data reuse beyond direct clinical care, such as for research and quality improvement studies. Secondary use of clinical data is a rapidly growing field [5,6]; an increasing number of institutions have implemented in-house IDRs and several others are developing IDRs for future research endeavors.

Unlike clinical practice, which focuses on enhancing the well-being of current patients, the purpose of an IDR is to produce generalized knowledge that can be extended to future patients. Typical applications of IDRs include retrospective analysis and hypothesis generation [7]. Some IDRs also support clinical applications, such as clinical decision support systems (CDSSs), that work alongside clinical practice to estimate risk factors or predictive scores associated with clinical treatments. CDSSs help to avoid medical errors and deliver efficient and safer care by assisting the provider with diagnosis, therapy planning, and treatment evaluation decisions [8]. All these applications are valuable resources that have the potential to improve the quality of health care [9] and reduce health costs if implemented appropriately [10].

Objective

Our study is motivated by the need to develop a pediatric IDR at our institution and by the lack of literature providing practical recommendations to apply during the initial development stages. Reviews by Shin et al [11] and Huser et al [12] highlighted the recommended characteristics when designing an IDR; however, they include only a small set of examples and a limited number of example IDRs. Since 2014, the IDR landscape has evolved rapidly, and thus, we felt more recent developments needed to be better addressed as well. A 2018 review by Hamoud et al [13] provided a comprehensive description of most recent data warehouses, including information about their data content, processing, and main purpose; it also provides general recommendations for the implementation of an IDR, but no practical considerations to guide the planning stages.

This study compares the features of contemporary IDRs and presents some guiding principles for the design and implementation of a clinical research data warehouse. Our research objective was to identify the major features of contemporary IDRs and obtain a list of established architectures used in the field of health informatics. We expect that this review will be useful for other small- to medium-sized institutions that plan to implement an institutional IDR and have no extensive experience in the field.

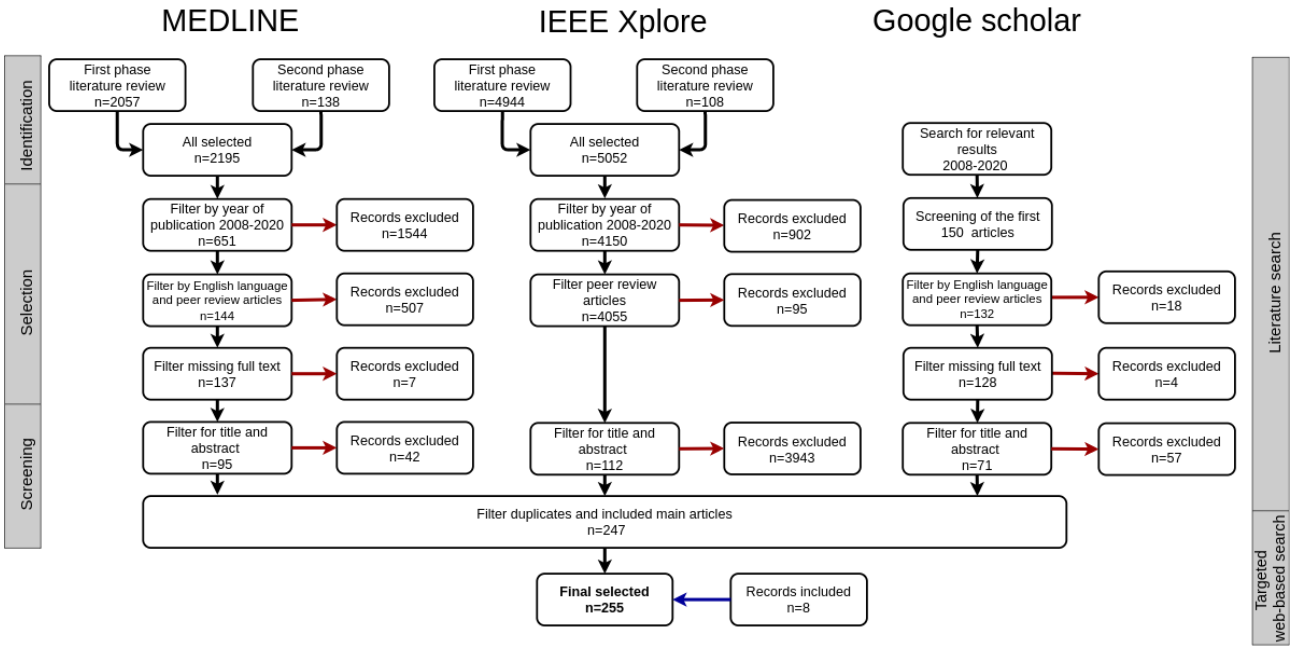
Methods

We conducted a literature review and a targeted web-based search to identify the major existing IDRs and synthesized the retrieved information around key themes.

Literature Review Search

We performed a narrative review following the procedure described below. First, a literature search was conducted using Ovid MEDLINE (Medical Literature Analysis and Retrieval System Online) and IEEE Xplore (Institute of Electrical and Electronics Engineers Xplore), queried in March 2020 (Figure 1). Articles were identified in 2 iterative phases. The first phase used an initial list of keywords querying for infrastructure purposes (data integration, such as linkage and harmonization) as well as infrastructure type and hospital setting (Multimedia Appendix 1: A1). The second phase search used additional keywords identified from the titles and abstracts of articles retrieved in the first phase (Multimedia Appendix 1: A1). Second, Google Scholar was queried for major article keywords (Integrated Data Repository) OR (Clinical Data Warehouse), and the first 150 retrieved hits were screened. The query was executed in a single search stage because the traditional search methods using Ovid MEDLINE and IEEE Xplore already produced exhaustive results.

Figure 1. Article selection process. The diagram shows the number of articles at each stage of selection for each of the 3 databases: MEDLINE (Medical Literature Analysis and Retrieval System Online), IEEE Xplore (Institute of Electrical and Electronics Engineers Xplore), and Google Scholar.



We selected peer-reviewed articles, published in the English language between January 2008 and March 2020, to include the most current data warehouse features. Non-English articles were excluded because of a lack of resources for translation. We retained articles for which the full text was available and removed duplicates. KG read the abstracts, and the articles describing specific data integration strategies, describing architecture structures, or providing more information about the data models were included. When it was unclear whether an article should be included, the authors EPC and MG were consulted. Duplicated articles were removed using EndNote reference management software (Clarivate Analytics). Additional articles providing the most up-to-date information about selected IDRs or cited by the selected articles were included in the selection process because they were considered relevant for the IDR definition. Targeted Web-Based Search of Known Institutional IDRs

We manually queried nonpublished resources with the goal of adding contemporary data warehousing practices implemented in large North American hospitals. A convenience sample of hospitals known to be leaders in these types of data warehousing was suggested by EPC and MG.

Additionally, we browsed publicly available information on each of the targeted institutional websites (Multimedia Appendix 1: A2). This was complemented with relevant peer-reviewed articles cited in these websites related to the design, implementation, and applications of such repositories.

Manual Shortlisting for a Comparative Review Analysis

For the comparative review analysis, we performed a manual selection to shortlist articles specifically describing IDR architectures. The shortlisting considered the major focus of the article and the presence of significant details describing data integration, data processing, or database services. The selected articles were searched for related IDR projects and further web-based resources (Table 1 and Multimedia Appendix 1: A3).

Table 1. Institutions and major features of the integrated data repositories.

IDR[a,b]	IDR scope	Architecture model	Standard common data model	Standard terminologies	Primary references
The National Institutes of Health Clinical Center-Small-sized institution					
Biomedical Translational Research Information System (BTRIS)	General care	General	N/A[c]	RED[d]	[14]
Deceased subjects (dsIDR)	Deceased subjects	General	N/A	RED	[15]
University of Kansas Medical Cent er-Medium-sized institution					
Healthcare Enterprise Repository for Ontological Narration (HERON)	General care	General	i2b2[e]	ICD[f]-9/ICD-10, CPT[g], RxNorm[h], SNOMED-CT[i], NDFRT[j], NCI[k], FDB[l]	[16,17]
Stanford University Medical Center- Medium-sized institution					
Stanford Translational Research Integrated Database Environment (STRIDE)	General care	General	i2b2, OMOP[m]	ICD-9, CPT, RxNorm, SNOMED-CT	[18,19]
STAnford Research Repository (STARR)	General care	General	i2b2, OMOP	ICD-9, CPT, RxNorm, SNOMED-CT	[20]
The Georges Pompidou University Hospital (HEGP)-Medium-sized institution					
HEGP CDW[n] platform	General care, cardiovascular, cancer	General	i2b2	ICD-10, LOINC[o], SNOMED-CT	[21-23]
Hanover Peter L. Reichertz Institute- Large-sized i nstitution					
Hanover Medical School Translational Research framework (HaMSTR)	General care	General	i2b2	ICD-10, LOINC	[24]
Erlangen University Hospital-Large-sized institution					
Clinical data warehouse	General care	General	I2b2	LOINC, NCI	[25]
Seoul St. Mary Hospital-Large-sized institution					
Prostate cancer research database	Cancer	General	N/A	N/A	[26]
Lead partner: Cincinnati Children's Hospital Medical Center-Collaborative project					
Maternal and Infant Data Hub (MIDH)	Perinatal	General	OMOP	ICD-9/ICD-10, SNOMED-CT	[27]
Georges Pompidou, Cochin and Necker Hospitals- Collaborative project					
CAncer Research for PErsonalized Medicine (CARPEM)	Cancer	General	Variant of i2b2 (tranSMART[p])	ICD-9/ICD-10, SNOMED-CT, ATC[q], GO[r], HPO[s]	[28]
Learning Healthcare System (LHS) across South Carolina-Collaborative project					
Health Science South Carolina (HSSC) clinical data warehouse	General care	General	i2b2	N/A	[29]
Windber Research Institute-Collaborative project					
Data Warehouse for Translational Research (DW4TR)	Cancer	General	N/A	MeSH[t], SNOMED-CT, NCI, caBIG VCDE[u]	[30,31]
Veterans' Health Administration-Collaborative project					
VA EHR (Veterans Administration's electronic health records)	General care	General with CDSS	N/A	ICD-9	[32]
Coordinated by Medtronic Iberica SA- Collaborative project					
Models and simulation techniques for discovering diabetes influence factors (MOSAIC)	Diabetes	General with CDSS	i2b2	ICD-9, DRG[v], ATC	[33]
National collaboration- Collaborative project					

IDR[a,b]	IDR scope	Architecture model	Standard common data model	Standard terminologies	Primary references
China Stroke Data Center (CSDC)	Cerebrovascular	General with CDSS	N/A	N/A	[34]
Houston Methodist Hospital-Large-sized institution					
Methodist Environment for Translational Enhancement and Outcomes Research (METEOR)	General care	General with CDSS	Extension of i2b2	ICD-9, CPT	[35]
Mayo Clinic- Large-sized institution					
Mayo Enterprise Data Trust (MEDT)	General care	General	i2b2	LexGrid[w]	[36]
Ovarian cancer registry	Cancer	General	i2b2	LexGrid	[37]
Translational Research Center (TRC)	Cancer	Biobank-driven	i2b2	LexGrid	[38]
Vanderbilt University Medical Center- Large-sized institution					
Synthetic Derivative	General care	General	N/A	FDB, ICD-9, CPT	[39]
BioVU	General care	Biobank-driven	N/A	FDB, ICD-9, CPT	[40]
The Children's Hospital of Philadelphia- Medium-sized institution					
Biorepository Portal (BRP)	Cancer, pediatric	Biobank-driven	Harvest	N/A	[41]
University of São Paulo-Large-sized institution					
BioBankWarden (BBW)	Cancer	Biobank-driven	N/A	ICD-10, SNOMED-CT, LOINC, GO	[42]
University of Pavia and Fondazione S. Maugeri- Large-sized institution					
onco-i2b2	Cancer	Biobank-driven	i2b2	SNOMED-CT	[43]
Leon Berard Cancer Center- Small-sized institution					
CLB-IT[x]	Cancer	User-controlled application layer	N/A	ADICAP[y], ICD-O	[44]
Lead partner: University of Utah-Collaborative project					
Federated Utah Research and Translational Health electronic Repository (FURTHeR)	Several	Federated	i2b2, OMOP, OpenMRS[z]	ICD-9/ICD-10, LOINC, SNOMED-CT, RxNorm	[45]
OpenFurther	Several	Federated	i2b2, OMOP, OpenMRS	ICD-9/ICD-10, LOINC, SNOMED-CT, RxNorm	[46]
Lead partner: University of Utah-Collaborative project					
Pediatric Health Information System (PHIS+)	Pediatric	Based on FURTHeR	i2b2	LOINC, SNOMED-CT	[47]
@neurIST European Project-Collaborative project					
@neurIST platform	Cerebrovascular	Federated as in FURTHeR	N/A	@neurIST ontology[aa]	[48]
University Clinics in Northern Germany- Collaborative project					

IDR[a,b]	IDR scope	Architecture model	Standard common data model	Standard terminologies	Primary references
Research Data Management System (RDMS)	Cancer	Federated as in FURTHeR	i2b2	ICD-10, SNOMED-CT	[49]

[a]IDR: Integrated Data Repository

[b]The IDRs are defined by their data scope, architecture model (as defined by the major design class represented in Figure 2), standard common data model, standard terminology, and primary reference.

[c]N/A: not applicable, n=4 in Standard Terminology.

[d]RED: Research Entities Dictionary, n=1.

[e]i2b2: Informatics for Integrating Biology and the Bedside

[f]ICD-9/ICD-10/ICD-O: International Classification for Diseases, version 9/10, O for oncology, n=14.

[g]CPT: Current Procedural Terminology, n=4.

[h]RxNorm: standardized nomenclature for clinical drugs, n=3.

[i]SNOMED-CT: Systematized Nomenclature of Medicine-clinical terms, n=11.

[j]NDFRT: National Drug File Reference Terminology, n=1.

[k]NCI: National Cancer Institute, n=2.

[l]FDB: First Databank, n=2.

[m]OMOP: Observational Medical Outcomes Partnership

[n]HEGP CDW: Hôpital Européen Georges Pompidou Clinical Data Warehouse, n=1.

[o]LOINC: Logical Observation Identifiers Names and Codes, n=5.

[p]tranSMART: Open-source data platform for translational research, n=1.

[q]ATC: Anatomical Therapeutic Chemical Classification, n=2.

[r]GO: Gene Ontology, n=2.

[s]HPO: Human Phenotype Ontology, n=1.

[t]MeSH: Medical Subject Headings; n=1.

[u]caBIG VCDE: the cancer Biomedical Informatics Grid Vocabulary and Data Elements Workspace, n=1.

[v]DRG: Diagnosis Related Group, n=1.

[w]LexGrid: Lexical Grid, n=1.

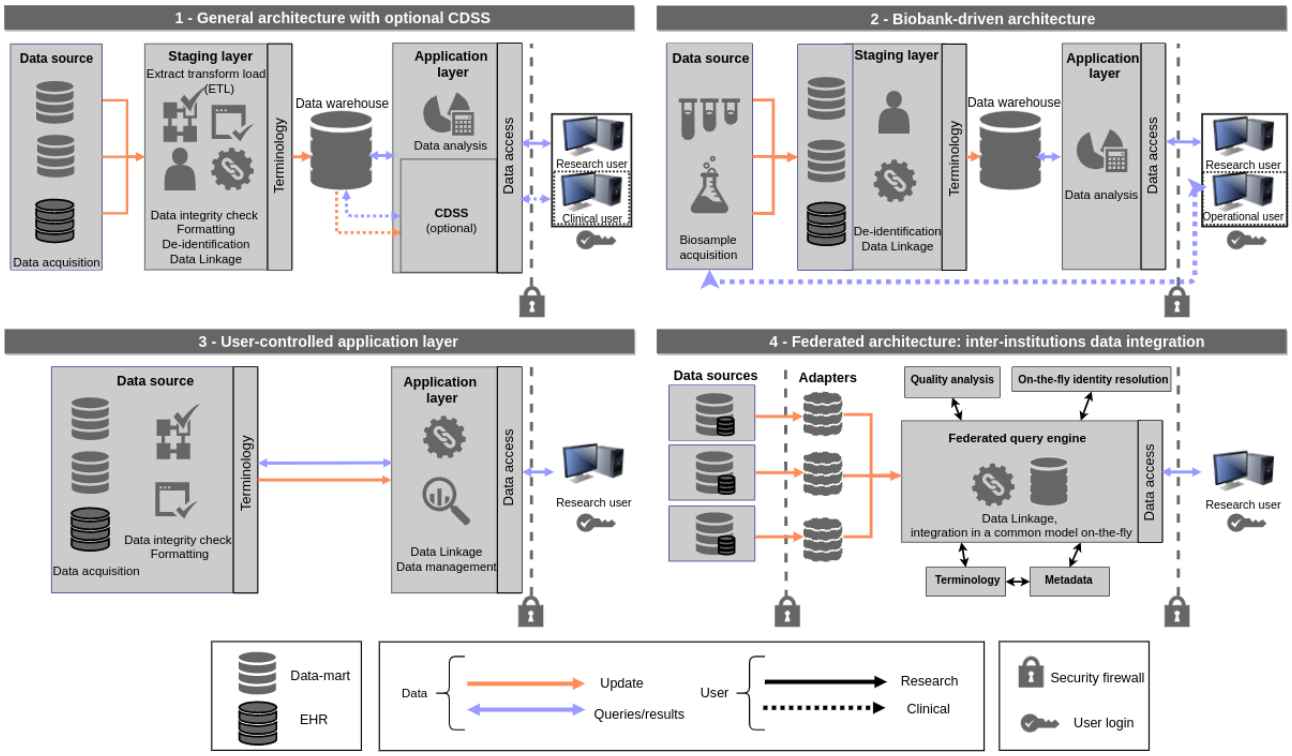
[x]CLB-IT: Léon Bérard Cancer Center-IT.

[y]ADICAP: Association pour le Développement de l'Informatique en Cytologie et en Anatomie Pathologique, n=1.

[z]OpenMRS: Open Medical Record System, n=1.

[aa]@neurIST ontology, n=1.

Figure 2. Architecture models identified from selected integrated data repositories (IDRs). Arrows indicate data output because of a query (blue) and data input (orange) because of data integration or update. Continuous lines show data query and integration applied by research users, whereas dashed lines are data queries performed by operational or clinical users.



Literature Synthesis and Institution Characterization

Information from the literature was aggregated through thematic analysis and collapsed into 4 classes of IDR architectures. We evaluated the main features of the identified IDRs, such as data processing components, data characteristics, common terminologies, and data models. Features were summarized, compared, and contrasted. We extracted information about host institutions and divided them into small (≤500 beds), medium (500-1000 beds), and large (>1000 beds) institutions based on the number of beds listed on the institution's websites.

Analysis of Word Content

Selected articles were uploaded into NVivo 12 (QSR International LLC) for qualitative analysis, specifically to count the word frequency in the selected papers. The words with a minimum length of 5 in the full text were counted, excluding stop words, and grouped by synonyms. The word frequency is represented as a word cloud, generated with R (R Foundation for Statistical Computing) and *wordcloud* package 2.6.

Citations Analysis

The references of the articles describing IDRs were downloaded in a semiautomated manner using Content Extractor and Miner software [50] to parse the full-text PDF files. References to web resources, video-cast meetings, and software were removed, and partial references were manually corrected. The references were grouped by *first author* and *year of publication* and loaded in R (R Foundation for Statistical Computing) and plotted with UpSetR [51].

Results

Overview

A total of 241 articles were identified in the literature search [11,13-19,21-29,31,33-35,37,43,44,47-49,52-264]; the largest number of articles were identified in IEEE Xplore (n=112), followed by MEDLINE (n=95), and Google Scholar (n=71). After removing duplicates (n=24), we added 3 articles that were frequently cited in the selected articles but were missing from our search results [30,36,42]. Three articles [38,40,45] were further added that provided additional details relevant to the review topic. Finally, 1 article was replaced by a more updated publication [265]. These 247 articles were combined with the targeted web-based search [32,39-41,266-269]; hence, we identified a total of 255 articles (Figure 1). The most frequent words in the articles were *system*,*information*, *study*, *project*, and *design* (Multimedia Appendix 1: A4.1). A total of 79 of these 255 articles were published between 2014 and 2016, and 34 were published in 2019; this date range covers the full range of initially identified articles in this domain area (Multimedia Appendix 1: A4.2 and A4.3).

A total of 116 articles were presented in proceedings of international scientific conferences, particularly those published in the book series Studies in Health Technology and Informatics (n=23); this included the World Congress of Medical and Health Informatics and Medical Informatics Europe. The second most frequent proceedings were the American Medical Informatics Association annual symposium and joint summits on translational science (n=12). The most frequently observed journals were the Journal of the American Medical Informatics Association (n=9) and BioMed Central (BMC; n=8), with BMC

Bioinformatics being the most common. More details about the individual conferences and journals can be found in Multimedia Appendix 1: A4.4 and A4.5.

For this review, we focused on the 34 articles describing 29 IDRs for which sufficient design details were presented. The additional web resources describing 2 IDRs, Stanford Translational Research Integrated Database Environment (STRIDE) and Federated Utah Research and Translational Health Electronic Repository (FURTHeR), referred to novel projects STAnford Research Repository (STARR) [20] and OpenFurther [46], respectively, which increased the number of IDRs to 31 from 25 different institutions or collaborative projects (Table 1). In reviewing the references in these 34 articles, we observed only a small overlap, with 1 reference [270] being found in common in a maximum of 11 articles (Multimedia Appendix 1: A5.1). The most frequently cited among the 34 are onco-Informatics for Integrating Biology and the Bedside (i2b2) [43,271], STRIDE [18], and the Mayo Clinic [36] IDRs, cited in 8, 5, and 4 articles, respectively (Multimedia Appendix 1: A5.2).

IDRs represent a variety of applications of health data warehousing for research. Although they share common characteristics, as described in detail below, they also demonstrate the many different purposes they can serve. For example, BioVU [40] and the Synthetic Derivative [39] at Vanderbilt University Medical Center are examples of a biobank-driven database that automatically couples patients' clinical information to biological samples (biosamples). The power of this system is its connection between genotype and phenotype and its large number of biosamples (>50,000), which allows a rich set of cohort research studies. The Maternal and Infant Data Hub (MIDH) at Cincinnati Children's Hospital Medical Center [27] is a regional perinatal data repository that integrates a large and diverse set of data from different institutions. The strength of the project is the combination of delivery and postdischarge hospital data and the linked mother and child data sets. The pilot database contains approximately 70,000 newborns and 42,000 pediatric postnatal visits. Another example is the Hanover Medical School Translational Research Framework (HaMSTR) framework at the Hanover Peter L. Reichertz Institute [24], which was developed to automatically load data from a clinical data repository into a standard data model that researchers can query; it is a successful example of fast data upload and query using data structures designed from standard data models available for clinical research.

Characteristics of the Institutions in the Selected IDR Sample

We identified 2 types of IDRs: those developed for use in a single institution (n=19) and those implemented for a collaborative project (n=12). The latter typically integrate patient data and provide project-specific tools. The median number of different institutional partners in a collaborative IDR is 6, with one of the partners acting as an organizational hub. The partners range from research institutes, laboratories, and private institutions to university medical centers.

The IDRs were further divided by their scope (Table 1), which were classified as general or specialized medical care (cancer, pediatrics, perinatal, cerebrovascular, or cardiovascular). Seven of the 10 IDRs containing specialized data were collaborative projects, likely indicating the need to pool data from several institutions when dealing with smaller but more focused patient populations.

Four Major Architecture Models Used in Our Selected IDR Sample

We identified 4 overarching conceptual architectures that summarize the data layers in the selected IDRs (Figure 2). Different institutions can implement multiple architectures for different purposes; we assigned each IDR to a category considering the major features of the IDR, as described in their respective articles.

The general architecture model is the most common model, with 19 identified IDRs structured around medical data mining (Figure 2, *General architecture with optional CDSS*). In outline, different data marts are transferred to a staging layer that harmonizes the input to a common data view; data are loaded into a common data warehouse and queried through an application layer that communicates with the user; a CDSS tool can provide added functionality. Hence, in this architecture, each data source is originally stored in an independent data mart, collecting data from a separate research or clinical source within the same institution. Data are processed in the staging layer, which reshapes the input to an integrated view through several steps of data linkage, transformation, and harmonization. The next stage of processing is loading the data into a single database connected to an application layer that provides the tools for end users, typically researchers, to access and analyze the data securely with different services. An example of an IDR providing multiple services is the STRIDE architecture stack [18], which includes several services for data analysis or research data management. The articles describing METEOR [35], CSDC [34], models and simulation techniques for discovering diabetes-related factors (MOSAIC) [33], and Veterans Administration's EHRs [32] provide further details about the integration of CDSS tools in the architecture. In these cases, the architecture model is divided into CDSS and data analysis modules, both of which communicate with the common database. The CDSS allows clinical staff to retrieve real-time individual patient records and to use analytical models to make risk prediction. The CDSS tools described by METEOR and MOSAIC, for example, learn from the clinical data stored in the data warehouse and estimate risk factors predicting hospital readmission or long-term complications.

The Health Science South Carolina (HSSC) [29] IDR gathers data from different clinical systems implemented in various institutions, all of which are party to a data collaboration agreement that authorizes data aggregation in a single data warehouse. This data warehouse contains a longitudinal record for each individual across all institutions. Data processing and terminology mapping occur in a conceptual staging layer, as in the case of the general architecture model.

In the case of the Erlanger University Hospital IDR [25], terminology is mapped using vocabularies that are manually curated and mapped through an automatic workflow that processes the raw data to the final data warehouse format. Other

IDRs that make use of multiple terminologies are health care enterprise repository for ontological narration [16], Research for PErsonalized Medicine (CARPEM) [28], and STRIDE [18], but further details of their mapping processes were not available.

The biobank-driven architecture model is built around a particular application, in this case, biobanking (Figure 2, *Biobank-driven architecture*). This model is similar to the general architecture model but, in this case, the IDR is built around the biosamples database. The biosample data integration occurs at the staging layer. The main feature is that the model allows the biosample operational user to access the raw and identified biobank data source for quality control and biosample management. An example of a biobank-driven structure is the biorepository portal (BRP) [41,266], which allows for the automatic integration of biosamples with clinical data, while maintaining unrestricted access to the biorepository for the operational team. The Mayo Clinic and Vanderbilt University adopt the general and biobank-driven architecture models in parallel.

The user-controlled application layer architecture model does not have a specific staging layer (Figure 2, *User-controlled application layer*). This architecture does not include a central data warehouse; the data are preprocessed and integrated from the original data sources only when the users query the data. Hence, data are processed in 2 stages: the first stage preprocesses the original data to a common format. The user query then carries out the final data integration function for the output delivery. In this architecture, a common data warehouse is not implemented, but rather the data are dynamically queried. An example is the text mining technology at the Léon Bérard Cancer Center (CLB) [44], which indexes text documents during the preprocessing stage and in which the users' queries return the exact documents matched.

The federated architecture is implemented for heterogeneous data retrieval and integration across multiple institutions (Figure 2, *Federated architecture*, adapted from OpenFurther). In this case, institutions selectively share their data through an adaptor system that applies common preprocessing, with data integrated on-the-fly in a virtual data warehouse. The FURTHeR federated query platform [45] builds a virtual IDR that responds to the needs of the user and calls several services for data resolution on-the-fly and upon query. The architecture model is flexible and operates using several services for data integration. An application of FURTHeR is the Pediatric Health Information System+ project [47], which combines data from 6 institutions. The IDR uses a federation component, which aggregates and stores translated query results in a temporary, in-memory database for presentation and analysis by the researcher for the duration of the user's session. Federated data integration was also proposed using a research data management system (RDMS) [49], which integrates clinical and biosample data from several institutions in Germany. The @neurIST [48] is a large IDR dedicated to translational research that includes data, computing resources, and tools for researchers and clinicians. Data are located across different sites and are securely shared with a grid infrastructure that allows federated data access.

The 4 types of architecture present different analytics tools, data presentation logic, and query interface based on the type of user they serve, which can be classified into 2 major groups: the first group, such as researchers and operational or business analysts, uses the IDR to identify important clinical features that occur at the level of patient cohorts. The second type of user, such as physicians and other health care professionals, uses the IDR to make decisions at an individual patient level, for example, to plan specific therapeutic interventions or predict risk. The first type of user is served by all the architecture models (*Research user* in Figure 2). The general architecture model that incorporates a CDSS presents a clear separation of both user types who have different applications for IDR data, with CDSS queries being made by clinical users (Figure 2, *General architecture with optional CDSS*). Similarly, the biobank-driven architecture model includes operational users who can directly query the information regarding patient biosamples for clinical applications (Figure 2, *Biobank-driven architecture*).

Data Retrieval and Update Are Influenced by the IDR Architecture Model

Both data update and integration schedules in an IDR are important features that define the timeliness of data. Here, we describe some of the key limiting steps and their occurrence in the different IDR architecture models.

Data Retrieval

The data processing involved in extraction, transformation, and loading (ETL) is described in detail in the articles of biomedical translational research information system (BTRIS) [14], HaMSTR [24], Mayo Translational Research Center (TRC) [38], CARPEM [28], onco-i2b2, Vanderbilt's Synthetic Derivative [39] and BioVU [40], and BRP [41]. These IDRs represent the general and biobank-driven architecture models, which implement a staging layer for the ETL process. A temporal sequence of the ETL steps is as follows:

1. Data extraction from source(s): The source data are extracted by an automatic (or manual) process.
2. Deidentification: Identifiable patient features, such as demographics or localization, are removed before loading into the IDR. The biobank-driven IDRs implement an automated process of this step without the need for extensive institutional reviews. In addition to the deidentified data, BTRIS [14] and Vanderbilt's Synthetic Derivative [39] maintain a parallel database with original identifiable patient entries for research purposes where appropriate.
3. Assignment of unique identifiers: Deidentified data are assigned unique patient identifiers that are used as a reference for linking.
4. Data transformation and standardization: Data are first checked for possible errors or missing values and are then transformed into a common format that is standard for all cohorts. Data may be subjected to transformation, such as the derivation of new values from the existing ones (pseudonymization) for maintaining privacy.
5. Standard terminology and ontology mapping: Data types are labeled with standard terminologies.

6. Data linkage: If the data are derived from multiple sources, they are linked and combined in the IDR.
7. Loading into the data warehouse: This is performed by either an update of existing data or a complete data re-import into the data warehouse.

The CLB [44] IDR (user-controlled application layer architecture model) uses specialized software to manipulate the content from unstructured data without using an ETL process. IDRs representing architecture model 4 do not provide additional information on the ETL process in their respective articles.

Data Update

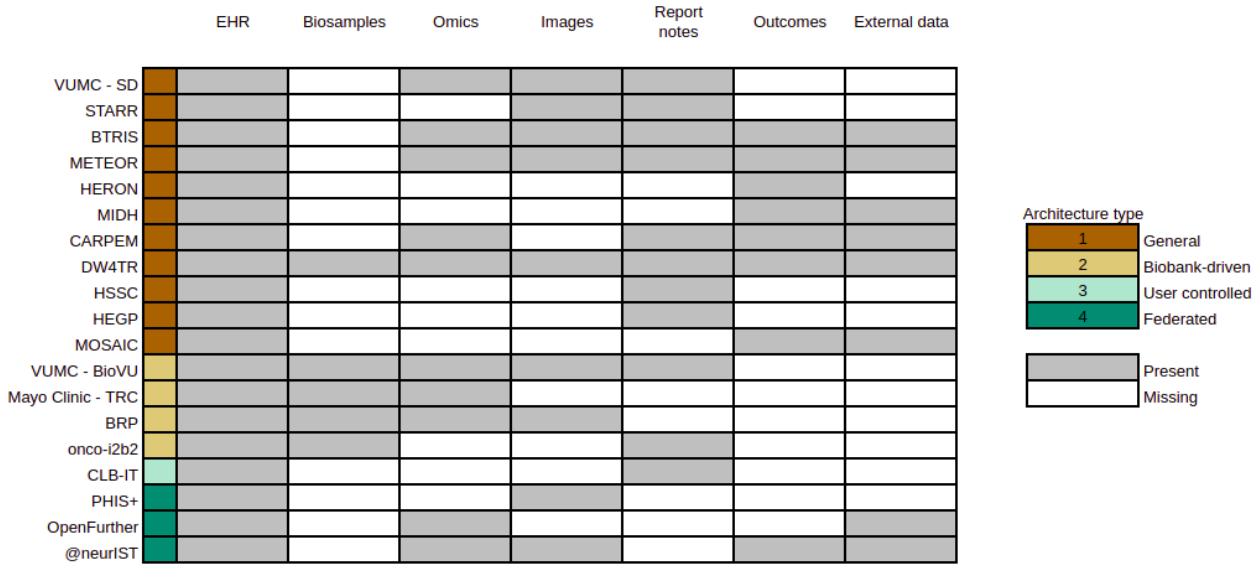
Five of the selected articles provide additional information about the frequency of data updates in their IDRs. BTRIS [14] and Vanderbilt's Synthetic Derivative [39] argue for daily IDR updates as new source data accumulate daily. Onco-i2b2 [43] performs more frequent data synchronization, as frequent as every 15 min. A real-time data update is presented by METEOR [35] and MOSAIC [33], which also integrate a CDSS in their architecture model and thus need this frequency to make actionable decisions. MOSAIC presents an example with asynchronous data update; although the CDSS is updated in real time, the demographics are synchronized only every 6 months. The general architecture model combined with a CDSS may require real-time data updates, whereas the general or the biobank-driven architecture models, without a CDSS, may have periodic updates that vary widely in frequency.

Major IDR Features: Data Type, Standard Terminology, and Common Data Model

Data Type

We have listed the data types in 19 of the selected IDRs based on information in the articles (Figure 3). The most common types of data are those extracted from EHR that include patient demographics, diagnoses, procedures, laboratory tests, and medications.

Figure 3. Common data types across IDRs. Columns show the main types of data collected in the selected IDRs. Gray-filled cells denote feature presence, with colors classifying the IDRs based on the examined architectures. Only 19 IDR articles contained enough information in their articles to be included in this figure. BRP: biorepository portal; BTRIS: biomedical translational research information system; CARPEM: cancer research for personalized medicine; CLB-IT: Léon Bérard Cancer Center Information Technology; DW4TR: Data Warehouse for Translational Research; EHR: electronic health record; HEGP: Hôpital Européen Georges Pompidou; HERON: health care enterprise repository for ontological narration; HSSC: Health Science, South Carolina; IDRs: integrated data repositories; Mayo Clinic-TRC: Mayo Clinic – Translational Research Center; METEOR: Methodist Environment for Translational Enhancement and Outcome Research; MIDH: Maternal and Infant Data Hub; MOSAIC: models and simulation techniques for discovering diabetes-related factors; Onco-i2b2; PHIS+: Pediatric Health Information System+; STARR: STAnford Research Repository; VUMC-BioVU: Vanderbilt University Medical Center–BioVU; VUMC-SD: Vanderbilt University Medical Center–Synthetic Derivative.



Several IDRs incorporate data from biosamples and their omics characterization, especially those based on the biobank-driven architecture model such as TRC [38], BRP [41,266], and BioVU [40]. Other examples of omics-based IDRs are CARPEM [28], Data Warehouse for Translational Research (DW4TR) [30], and @neurIST [48], which are dedicated to specific domains of research, namely cancer and cerebral aneurysm research.

Several types of images are part of modern IDRs, such as radiographic images in BTRIS [14] and document images in Methodist Environment for Translational Enhancement and Outcome Research (METEOR) [35]. In addition, medical reports are integrated in the IDRs. Clinical documents can be processed using natural language processing (NLP) algorithms to extract clinical conditions, medication types, and other features from common hospital procedures, which increases their utility through transformation into structured data. NLP modules are integrated in CLB-IT [44], which is specifically built for text processing entries, as well as BTRIS [14], METEOR [35], and onco-i2b2 [43].

IDRs including CDSS include outcome data types, which are relevant for calculating risk factors or predictive values in clinical domains. External data can also be integrated into the IDRs, including genomics data from disease model organisms (BTRIS) [14], patients from external sources (BTRIS [14] and DW4TR [30]), or environmental indices and geolocation (MIDH) [27].

Standard Terminology

Health information technology uses controlled terminologies to condense the information to a set of codes that can be manipulated more easily and automatically in data processing. We observed the adoption of both common [272,273] and specialized terminologies (eg, Anatomical Therapeutic Chemical Classification [274], human phenotype ontology [275], Gene Ontology [276]). The most broadly used were International Classification of Diseases (ICD)-9 and 10 for the classification of diseases, systematized nomenclature of medicine-clinical terms (SNOMED-CT) for a variety of medical domains, Logical Observation Identifiers, Names, and Codes for laboratory observations, and current procedural terminology for common procedures (Table 1). These terminologies were utilized within the EHR and further integrated into the IDRs.

Common Data Model

A common data model (CDM) is a standard data schema that enables data interoperability and sharing. Contemporary data warehouses propose an analytical platform built around the CDM that provides all the software components to construct and manage the data in a CDM. A few different CDMs have been developed and adopted by the wider clinical research community, although some institutions still favor using a custom data schema tailored to their specific needs. In our study, a standard CDM was adopted by 18 of the 29 IDRs. The most frequently applied CDM, found in 16 instances, is Informatics for Integrating Biology and the Bedside (i2b2) [277]. METEOR [35] applies i2b2 with an expanded schema, and CARPEM [28] applies tranSMART [278], which is a framework layered on top of i2b2, dedicated to integrating omics data with EHR data. Another popular CDM that has been used more frequently in recent years is the Observational Medical Outcomes Partnership (OMOP) [279], adopted by 3 IDRs, namely MIDH [27], OpenFurther [46], and STARR [20]. OpenFurther uses OpenMRS [280], which is an open-source software and CDM that delivers health care in low- and middle-income countries. The BRP [41] is the only example using Harvest as their CDM.

Discussion

Principal Findings

Our review identified several institutions of various sizes and scopes that utilize an IDR. These IDRs contain data used for both research and clinical decision-making purposes. The use of structured data from natural language processing of clinical notes, clinical imaging, and omics data are the most recent big data types to be integrated with standard clinical observations. Owing to the large heterogeneity, however, integration is complex and tailored to the specific needs during the IDR implementation and maintenance, as ETL necessitates a significant effort in both the initial modeling and the ongoing updates.

As a novel contribution, we proposed and classified IDR architectures into 4 major models that highlight the processing and integration steps. The most common architecture model employs a staging layer implemented before the data are loaded into the data warehouse.

A set of common features are applied across most IDRs: IDRs commonly use standard terminologies such as ICD-9/10 and SNOMED-CT, which are often already part of the EHR data. Several IDRs use an open-source translational research framework to model their data, as described by Huser et al [12]. We observed extensive use of i2b2 CDM and the emergent adoption of OMOP CDM, which has the possibility to map additional domain-specific terminologies. Interestingly, PCORnet is one of the newest CDMs, but its application was not discussed in the sample of IDRs reviewed. The PCORnet is the most recently implemented CDM that borrows from several other CDMs and is organized around patient outcomes [261].

To safeguard the data in the IDR, data security and privacy need to be ensured from the initial steps of development. Data security is an important factor in all architecture types, with a particular need in collaborative projects that share data across jurisdictions. For example, in the general architecture of HSSC [29], data need to be stored in physically and logically secure facilities, where data management is extended to all the parties involved, and data need to be transmitted between the participating institutions through private high-speed networks. In the case of federated data warehouses, such as @neurIST [48], there is a tight control of data flow between different institutions and clinical and research domains, following policies aligned with recommendations from the Legal and Ethics Advisory Board. Privacy, referring to the protection of patient's personal information, emerged as an important feature, especially in the biobank-driven architecture; here, identifiable patient information is deleted from both the biosamples and the patient clinical data. Developers at the Children's Hospital of Philadelphia and the Children's Brain Tumor Tissue Consortium created an electronic Honest Broker (eHB) and Biorepository Portal (BRP) eHB [41], which provides a method for patient privacy protection by removing all the exposure of the research staff to patient identifiers and automating the deidentification process. Following a different privacy-preservation approach, Vanderbilt's Synthetic Derivative database [39] alters the patient data by obfuscating the true entries while preserving their time dependence.

Guiding Principles

The implementation of an IDR is subject to several factors that must be considered before development. We identified 2 major factors: (1) the data stored in the IDR and (2) the scope of the IDR, either being exclusively used for research purposes or in combination with clinical or operational purposes, as shown in the general and biobank-driven architecture models. Data types, heterogeneity, and volume greatly influence system load, update, and query of the database. The scope of the IDR influences its primary end users, researchers, clinical users, or operational users, who have different needs and, thus, need access to different sets of tools to extract, analyze, and visualize the data. All the features influence both the data latency and the data synchronization, which are major elements in the model architecture. Moreover, available funding plays an important role in architecture decisions, as are considerations for future expansions.

Among the set of selected IDRs, we observed a number of collaborative projects that work within specialized medical domains, such as cancer or pediatrics. Collaborative IDRs are likely to integrate their data to increase the number of patients, thus increasing the statistical power of their respective cohorts.

On the basis of our analysis, we highlight the following guiding principles for small- to medium-sized institutions planning to implement an IDR:

1. The general architecture model, with or without CDSS, is the most straightforward to implement; the data staging layer facilitates ETL and data processing before loading into the data warehouse.
2. Select a standard CDM already in use by other institutions; both i2b2 and OMOP provide server and client services in a single unique platform that serves the user with all the necessary tools to set up a structured IDR.
3. Wherever possible, adopt standard terminologies; we listed the most common terminologies derived from the integration with EHR data (Table 1). One promising approach is that common terminologies are applied in the first phases of the IDR development with other, more specialized terminologies, added later as the project scope expands.
4. Finally, the data update requirements and ETL process design should be carefully considered, the level of automation, as these are the limiting stages in data integration and update.

Commercial electronic medical record platforms such as Epic, Cerner, Meditech, and Allscripts are dominant in large institutions. However, although some information about how to query underlying databases and application programming interfaces to communicate with these systems are available, little information on transforming such data into IDR is available in the literature, most likely because of their proprietary nature. Most vendors also sell tools for analysts to query and make use of data from these clinical production systems; however, they are not IDRs themselves and are not targeted toward secondary use for research.

As for lessons learned in the field, Epstein et al [281] demonstrate the feasibility of transferring the development of a perioperative data warehouse (schemas and processes) built on top of Epic's database from one institution to another.

Comparison With Prior Work

In their review, Hamoud et al [13] provided general requirements for building a successful clinical data warehouse, recommending a top-down approach to the initial stages of development. They recommended considering all the individual components of the final system to decrease integration obstacles when dealing with heterogeneous data sources.

Three major factors contributing to the success of IDRs were identified by Baghal [231] when developing their in-house IDR: (1) organizational, enhancing the collaboration between different departments and researchers; (2) behavioral, building new professional relationships through frequent meetings and communication between departments; and (3) technical improvements to deploy new self-service tools that empower researchers. Collectively, these factors increase the utility and adoption of IDRs in clinical research.

In addition, the report by Rizi and Roudsari [282] on lessons and barriers from their development of a public health data warehouse, which IDR developers might want to consider, specifically, not to underestimate technical challenges such as those related to extracting data from other systems, difficulties in modeling and mapping of data, as well as data security and privacy. Other considerations include leveraging the IDR to improve data quality at the source, implementing a data governance framework from the beginning, and ensuring that key organizational stakeholders endorse the project early and strongly [282].

Limitations

Our search was not intended to be a systematic search; therefore, we may have missed some articles. An example of missing articles is those describing raw and unstructured data repositories, also referred to as *data lakes*, as these did not appear in our search results although we knew they exist. One of the *data lakes* was presented by Foran et al [207] as a file reservoir, integrated in the data warehouse schema. For researchers to access those data, it was necessary to use a *feeder* database before their upload to the final data warehouse.

Furthermore, we were able to report on the IDRs and IDR features described in the literature, possibly omitting smaller institutions that are not actively publishing in peer-reviewed journals. In an attempt to mitigate this issue, we searched the representative institutional websites to retrieve additional details about the IDR architectures. As shown in Multimedia Appendix 1 [283-288]: A2, several organizations provide further details about their architecture in GitHub repositories or institutional Wiki pages, which can be explored for additional information besides the published literature.

This review includes articles and web resources shortlisted according to aspects of the IDR architectures that were considered relevant. Providing an exhaustive coverage of all aspects of IDR implementation, such as tools designed to interact with the IDR, are better left for a dedicated review. An example of such tools is the *Green Button* project, which provides critical help in treating patients [289-292]. Examples of CDM-based tools, built around an application, are the @neurIST platform [48], @neurLink, and @neurFuse application suites that consist of research-oriented modules dedicated to knowledge discovery and image processing. CDSS tools such as Green Button, @neurIST applications, or many other existing frameworks are essential in providing sophisticated analyses to support clinicians, but are beyond the scope of our review.

Conclusions

There is significant potential in the implementation of IDRs in health institutions, and their importance is evident from the growing number of projects developed in the past 10 years. Despite the common trends in IDR implementation observed in this study, there are also many variations. There are 2 major design factors, namely data heterogeneity and IDR scope, which

need to be carefully considered before embarking on the IDR design and planning process.

Finally, we aim to apply the knowledge presented in this study for the implementation of a pediatric IDR at our institution. By sharing our experience of planning and designing our IDR with those joining the field or planning to implement an IDR for research purposes, we hope to contribute to future IDR endeavors.

Acknowledgments

The project was supported, in part, by an Evidence to Innovation (E2i) Research Theme seed grant through the BC Children's Hospital Research Institute. The authors wish to thank Colleen Pawliuk for her help with the literature search strategy development and execution and Nicholas West and Zoltan Bozoky for editorial assistance.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary methods and results.
[PDF File (Adobe PDF File), 1649 KB - formative_v4i8e17687_app1.pdf]

References

1. Adler-Milstein J, Holmgren AJ, Kralovec P, Worzala C, Searcy T, Patel V. Electronic health record adoption in US hospitals: the emergence of a digital 'advanced use' divide. J Am Med Inform Assoc 2017 Nov 1;24(6):1142-1148. [doi: 10.1093/jamia/ocx080] [Medline: 29016973]
2. Lau F, Price M, Boyd J, Partridge C, Bell H, Raworth R. Impact of electronic medical record on physician practice in office settings: a systematic review. BMC Med Inform Decis Mak 2012 Feb 24;12:10 [FREE Full text] [doi: 10.1186/1472-6947-12-10] [Medline: 22364529]
3. Schoen C, Osborn R, Doty MM, Squires D, Peugh J, Applebaum S. A survey of primary care physicians in eleven countries, 2009: perspectives on care, costs, and experiences. Health Aff (Millwood) 2009;28(6):w1171-w1183. [doi: 10.1377/hlthaff.28.6.w1171] [Medline: 19884491]
4. MacKenzie SL, Wyatt MC, Schuff R, Tenenbaum JD, Anderson N. Practices and perspectives on building integrated data repositories: results from a 2010 CTSA survey. J Am Med Inform Assoc 2012 Jun;19(e1):e119-e124 [FREE Full text] [doi: 10.1136/amiajnl-2011-000508] [Medline: 22437072]
5. Anderson N, Abend A, Mandel A, Geraghty E, Gabriel D, Wynden R, et al. Implementation of a deidentified federated data network for population-based cohort discovery. J Am Med Inform Assoc 2012 Jun;19(e1):e60-e67 [FREE Full text] [doi: 10.1136/amiajnl-2011-000133] [Medline: 21873473]
6. Meystre SM, Lovis C, Bürkle T, Tognola G, Budrionis A, Lehmann CU. Clinical data reuse or secondary use: current status and potential future progress. Yearb Med Inform 2017 Aug;26(1):38-52 [FREE Full text] [doi: 10.15265/IY-2017-007] [Medline: 28480475]
7. Murphy SN, Dubey A, Embi PJ, Harris PA, Richter BG, Turisco F, et al. Current state of information technologies for the clinical research enterprise across academic medical centers. Clin Transl Sci 2012 Jun;5(3):281-284 [FREE Full text] [doi: 10.1111/j.1752-8062.2011.00387.x] [Medline: 22686207]
8. Kawamoto K, Houlihan CA, Balas EA, Lobach DF. Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success. Br Med J 2005 Apr 2;330(7494):765 [FREE Full text] [doi: 10.1136/bmj.38398.500764.8F] [Medline: 15767266]
9. Buntin MB, Burke MF, Hoaglin MC, Blumenthal D. The benefits of health information technology: a review of the recent literature shows predominantly positive results. Health Aff (Millwood) 2011 Mar;30(3):464-471. [doi: 10.1377/hlthaff.2011.0178] [Medline: 21383365]
10. Belard A, Buchman T, Forsberg J, Potter BK, Dente CJ, Kirk A, et al. Precision diagnosis: a view of the clinical decision support systems (CDSS) landscape through the lens of critical care. J Clin Monit Comput 2017 Apr;31(2):261-271. [doi: 10.1007/s10877-016-9849-1] [Medline: 26902081]
11. Shin S, Kim WS, Lee J. Characteristics desired in clinical data warehouse for biomedical research. Healthc Inform Res 2014 Apr;20(2):109-116 [FREE Full text] [doi: 10.4258/hir.2014.20.2.109] [Medline: 24872909]
12. Huser V, Cimino JJ. Desiderata for healthcare integrated data repositories based on architectural comparison of three public repositories. AMIA Annu Symp Proc 2013;2013:648-656 [FREE Full text] [Medline: 24551366]
13. Hamoud A, Hashim A, Awadh W. Clinical data warehouse: a review. Iraqi J Comput Inf 2018 Dec 31;44(2):16-26. [doi: 10.25195/ijci.v44i2.53]

14. Cimino JJ, Ayres EJ, Remennik L, Rath S, Freedman R, Beri A, et al. The national institutes of health's biomedical translational research information system (BTRIS): design, contents, functionality and experience to date. J Biomed Inform 2014 Dec;52:11-27 [FREE Full text] [doi: 10.1016/j.jbi.2013.11.004] [Medline: 24262893]
15. Huser V, Kayaalp M, Dodd ZA, Cimino JJ. Piloting a deceased subject integrated data repository and protecting privacy of relatives. AMIA Annu Symp Proc 2014;2014:719-728 [FREE Full text] [Medline: 25954378]
16. Liu M, Melton BL, Ator G, Waitman LR. Integrating medication alert data into a clinical data repository to enable retrospective study of drug interaction alerts in clinical practice. AMIA Jt Summits Transl Sci Proc 2017;2017:213-220 [FREE Full text] [Medline: 28815131]
17. Adagarla B, Connolly DW, McMahon TM, Nair M, VanHoose LD, Sharma P, et al. SEINE: Methods for Electronic Data Capture and Integrated Data Repository Synthesis with Patient Registry Use Cases. KU ScholarWorks. 2015. URL: http://hdl.handle.net/2271/1303 [accessed 2020-08-09]
18. Lowe HJ, Ferris TA, Hernandez PM, Weber SC. STRIDE--an integrated standards-based translational research informatics platform. AMIA Annu Symp Proc 2009 Nov 14;2009:391-395 [FREE Full text] [Medline: 20351886]
19. Hernandez P, Podchiyska T, Weber S, Ferris T, Lowe H. Automated mapping of pharmacy orders from two electronic health record systems to RxNorm within the STRIDE clinical data warehouse. AMIA Annu Symp Proc 2009 Nov 14;2009:244-248 [FREE Full text] [Medline: 20351858]
20. STARR Tools COVID-19 Research Support. Stanford Medicine. 2019. URL: http://med.stanford.edu/starr-tools.html [accessed 2020-07-10]
21. Jannot A, Zapletal E, Avillach P, Mamzer M, Burgun A, Degoulet P. The Georges Pompidou university hospital clinical data warehouse: a 8-years follow-up experience. Int J Med Inform 2017 Jun;102:21-28. [doi: 10.1016/j.ijmedinf.2017.02.006] [Medline: 28495345]
22. Boussadi A, Caruba T, Zapletal E, Sabatier B, Durieux P, Degoulet P. A clinical data warehouse-based process for refining medication orders alerts. J Am Med Inform Assoc 2012;19(5):782-785 [FREE Full text] [doi: 10.1136/amiajnl-2012-000850] [Medline: 22523345]
23. Zapletal E, Rodon N, Grabar N, Degoulet P. Methodology of integration of a clinical data warehouse with a clinical information system: the HEGP case. Stud Health Technol Inform 2010;160(Pt 1):193-197. [Medline: 20841676]
24. Haarbrandt B, Tute E, Marschollek M. Automated population of an i2b2 clinical data warehouse from an openEHR-based data repository. J Biomed Inform 2016 Oct;63:277-294. [doi: 10.1016/j.jbi.2016.08.007] [Medline: 27507090]
25. Zunner C, Ganslandt T, Prokosch H, Bürkle T. A reference architecture for semantic interoperability and its practical application. Stud Health Technol Inform 2014;198:40-46. [Medline: 24825683]
26. Choi IY, Park S, Park B, Chung BH, Kim C, Lee HM, et al. Development of prostate cancer research database with the clinical data warehouse technology for direct linkage with electronic medical record system. Prostate Int 2013;1(2):59-64 [FREE Full text] [doi: 10.12954/PI.12015] [Medline: 24223403]
27. Hall ES, Greenberg JM, Muglia LJ, Divekar P, Zahner J, Gholap J, et al. Implementation of a regional perinatal data repository from clinical and billing records. Matern Child Health J 2018 Apr;22(4):485-493 [FREE Full text] [doi: 10.1007/s10995-017-2414-9] [Medline: 29275460]
28. Rance B, Canuel V, Countouris H, Laurent-Puig P, Burgun A. Integrating heterogeneous biomedical data for cancer research: the CARPEM infrastructure. Appl Clin Inform 2016;7(2):260-274 [FREE Full text] [doi: 10.4338/ACI-2015-09-RA-0125] [Medline: 27437039]
29. Turley CB, Obeid J, Larsen R, Fryar KM, Lenert L, Bjorn A, et al. Leveraging a statewide clinical data warehouse to expand boundaries of the learning health system. EGEMS (Wash DC) 2016;4(1):1245 [FREE Full text] [doi: 10.13063/2327-9214.1245] [Medline: 28154834]
30. Hu H, Correll M, Kvecher L, Osmond M, Clark J, Bekhash A, et al. DW4TR: a data warehouse for translational research. J Biomed Inform 2011 Dec;44(6):1004-1019 [FREE Full text] [doi: 10.1016/j.jbi.2011.08.003] [Medline: 21872681]
31. Maskery S, Bekhash A, Kvecher L, Correll M, Hooke JA, Kovatich AJ, et al. Aggregated biomedical-information browser (ABB): a graphical user interface for clinicians and scientists to access a clinical data warehouse. J Comput Sci Syst Biol 2014;7:20-27. [doi: 10.4172/jcsb.1000134]
32. Rajeevan N, Niehoff KM, Charpentier P, Levin FL, Justice A, Brandt CA, et al. Utilizing patient data from the veterans administration electronic health record to support web-based clinical decision support: informatics challenges and issues from three clinical domains. BMC Med Inform Decis Mak 2017 Jul 19;17(1):111 [FREE Full text] [doi: 10.1186/s12911-017-0501-x] [Medline: 28724368]
33. Dagliati A, Sacchi L, Tibollo V, Cogni G, Teliti M, Martinez-Millana A, et al. A dashboard-based system for supporting diabetes care. J Am Med Inform Assoc 2018 May 1;25(5):538-547. [doi: 10.1093/jamia/ocx159] [Medline: 29409033]
34. Yu J, Mao H, Li M, Ye D, Zhao D. CSDC: a nationwide screening platform for stroke control and prevention in China. Conf Proc IEEE Eng Med Biol Soc 2016 Aug;2016:2974-2977. [doi: 10.1109/EMBC.2016.7591354] [Medline: 28268937]
35. Puppala M, He T, Yu X, Chen S, Ogunti R, Wong S. Data Security and Privacy Management in Healthcare Applications and Clinical Data Warehouse Environment. In: IEEE-EMBS International Conference on Biomedical and Health Informatics. 2016 Presented at: BHI'16; February 24-27, 2016; Las Vegas, NV, USA. [doi: 10.1109/bhi.2016.7455821]

36. Chute CG, Beck SA, Fisk TB, Mohr DN. The enterprise data trust at mayo clinic: a semantically integrated warehouse of biomedical data. J Am Med Inform Assoc 2010;17(2):131-135 [FREE Full text] [doi: 10.1136/jamia.2009.002691] [Medline: 20190054]
37. Hong N, Li Z, Kiefer R, Robertson MS, Goode EL, Wang C, et al. Building an i2b2-Based Integrated Data Repository for Cancer Research: A Case Study of Ovarian Cancer Registry. In: Lecture Notes in Computer Science. 2017 Presented at: LNCS'17; September 2-6, 2017; New Delhi, India. [doi: 10.1007/978-3-319-57741-8_8]
38. Horton I, Lin Y, Reed G, Wiepert M, Hart S. Empowering mayo clinic individualized medicine with genomic data warehousing. J Pers Med 2017 Aug 22;7(3):- [FREE Full text] [doi: 10.3390/jpm7030007] [Medline: 28829408]
39. Danciu I, Cowan JD, Basford M, Wang X, Saip A, Osgood S, et al. Secondary use of clinical data: the Vanderbilt approach. J Biomed Inform 2014 Dec;52:28-35 [FREE Full text] [doi: 10.1016/j.jbi.2014.02.003] [Medline: 24534443]
40. Roden DM, Pulley JM, Basford MA, Bernard GR, Clayton EW, Balser JR, et al. Development of a large-scale de-identified DNA biobank to enable personalized medicine. Clin Pharmacol Ther 2008 Sep;84(3):362-369 [FREE Full text] [doi: 10.1038/clpt.2008.89] [Medline: 18500243]
41. Felmeister AS, Masino AJ, Rivera TJ, Resnick AC, Pennington JW. The biorepository portal toolkit: an honest brokered, modular service oriented software tool set for biospecimen-driven translational research. BMC Genomics 2016 Aug 18;17(Suppl 4):434 [FREE Full text] [doi: 10.1186/s12864-016-2797-9] [Medline: 27535360]
42. Ferretti Y, Miyoshi NS, Silva WA, Felipe JC. BioBankWarden: a web-based system to support translational cancer research by managing clinical and biomaterial data. Comput Biol Med 2017 May 1;84:254-261 [FREE Full text] [doi: 10.1016/j.compbiomed.2015.04.008] [Medline: 25959800]
43. Segagni D, Tibollo V, Dagliati A, Zambelli A, Priori SG, Bellazzi R. An ICT infrastructure to integrate clinical and molecular data in oncology research. BMC Bioinformatics 2012 Mar 28;13(Suppl 4):S5 [FREE Full text] [doi: 10.1186/1471-2105-13-S4-S5] [Medline: 22536972]
44. Biron P, Metzger MH, Pezet C, Sebban C, Barthuet E, Durand T. An information retrieval system for computerized patient records in the context of a daily hospital practice: the example of the Léon Bérard cancer center (France). Appl Clin Inform 2014;5(1):191-205 [FREE Full text] [doi: 10.4338/ACI-2013-08-CR-0065] [Medline: 24734133]
45. Livne OE, Schultz ND, Narus SP. Federated querying architecture with clinical & translational health IT application. J Med Syst 2011 Oct;35(5):1211-1224. [doi: 10.1007/s10916-011-9720-3] [Medline: 21537849]
46. OpenFurther. 2019. URL: http://openfurther.org/ [accessed 2019-08-10]
47. Narus SP, Srivastava R, Gouripeddi R, Livne OE, Mo P, Bickel JP, et al. Federating clinical data from six pediatric hospitals: process and initial results from the PHIS+ consortium. AMIA Annu Symp Proc 2011;2011:994-1003 [FREE Full text] [Medline: 22195159]
48. Benkner S, Arbona A, Berti G, Chiarini A, Dunlop R, Engelbrecht G, et al. NeurIST: infrastructure for advanced disease management through integration of heterogeneous data, computing, and complex processing services. IEEE Trans Inf Technol Biomed 2010 Nov;14(6):1365-1377. [doi: 10.1109/TITB.2010.2049268] [Medline: 20435543]
49. Ulrich H, Kock A, Duhm-Harbeck P, Habermann JK, Ingenerf J. Metadata repository for improved data sharing and reuse based on HL7 FHIR. Stud Health Technol Inform 2016;228:162-166. [Medline: 27577363]
50. Tkaczyk D, Szostek P, Fedoryszak M, Dendek PJ, Bolikowski A. CERMINE: automatic extraction of structured metadata from scientific literature. Int J Doc Anal Recognit 2015 Jul 3;18(4):317-335. [doi: 10.1007/s10032-015-0249-8]
51. Conway JR, Lex A, Gehlenborg N. UpSetR: an R package for the visualization of intersecting sets and their properties. Bioinformatics 2017 Sep 15;33(18):2938-2940 [FREE Full text] [doi: 10.1093/bioinformatics/btx364] [Medline: 28645171]
52. Bortis G. Experiences With Mirth: an Open Source Health Care Integration Engine. In: Proceedings of the 30th International Conference on Software Engineering. 2008 Presented at: ICSE'08; May 10-18, 2008; Leipzig, Germany. [doi: 10.1145/1368088.1368179]
53. Kang B, Kim D, Kim H. Two-Phase chief complaint mapping to the UMLS metathesaurus in Korean electronic medical records. IEEE Trans Inf Technol Biomed 2009 Jan;13(1):78-86. [doi: 10.1109/TITB.2008.2007103] [Medline: 19129026]
54. Lhotska L, Aubrecht P, Valls A, Gibert K. Security Recommendations for Implementation in Distributed Healthcare Systems. In: 42nd Annual IEEE International Carnahan Conference on Security Technology. 2008 Presented at: CCST'08; October 13-16, 2008; Prague, Czech Republic. [doi: 10.1109/ccst.2008.4751280]
55. Maragoudakis M, Lymberopoulos D, Fakotakis N, Spiropoulos K. A hierarchical, ontology-driven Bayesian concept for ubiquitous medical environments--a case study for pulmonary diseases. Conf Proc IEEE Eng Med Biol Soc 2008;2008:3807-3810. [doi: 10.1109/IEMBS.2008.4650038] [Medline: 19163541]
56. Pruulmann J, Willemson J. Implementing A Knowledge-Driven Hierarchical Context Model in a Medical Laboratory Information System. In: The Third International Multi-Conference on Computing in the Global Information Technology (iccgi 2008). 2008 Presented at: ICCGI'08; July 27-August 1, 2008; Athens, Greece. [doi: 10.1109/iccgi.2008.19]
57. Riedl B, Grascher V, Fenz S, Neubauer T. Pseudonymization for Improving the Privacy in E-Health Applications. In: Proceedings of the 41st Annual Hawaii International Conference on System Sciences. 2008 Presented at: HICSS'08; January 7-10, 2008; Waikoloa, HI, USA. [doi: 10.1109/hicss.2008.366]

58. Sung T, Hung F, Chiu H. Implementation of an integrated drug information system for inpatients to reduce medication errors in administering stage. Conf Proc IEEE Eng Med Biol Soc 2008;2008:743-746. [doi: 10.1109/IEMBS.2008.4649259] [Medline: 19162762]
59. Taddei A, Dalmiani S, Vellani A, Rocca E, Piccini G, Carducci T, et al. Data Integration in Cardiac Surgery Health Care Institution: Experience at G Pasquinucci Heart Hospital. In: Computers in Cardiology. 2008 Presented at: CIC'08; September 14-17, 2008; Bologna, Italy. [doi: 10.1109/CIC.2008.4749034]
60. Zamboulis L, Poulovassilis A, Roussos G. Flexible Data Integration and Ontology-Based Data Access to Medical Records. In: 8th IEEE International Conference on BioInformatics and BioEngineering. 2008 Presented at: BIBE'08; October 8-10, 2008; Athens, Greece. [doi: 10.1109/bibe.2008.4696774]
61. Agorastos T, Koutkias V, Falelakis M, Lekka I, Mikos T, Delopoulos A, et al. Semantic integration of cervical cancer data repositories to facilitate multicenter association studies: the ASSIST approach. Cancer Inform 2009 Feb 3;8:31-44 [FREE Full text] [doi: 10.4137/cin.s963] [Medline: 19458792]
62. Amoretti M, Zanichelli F. The Multi-Knowledge Service-Oriented Architecture: Enabling Collaborative Research for E-Health. In: 42nd Hawaii International Conference on System Sciences. 2009 Presented at: HICSS'09; January 5-8, 2009; Big Island, HI, USA. [doi: 10.1109/hicss.2009.441]
63. Archer N, Cocosila M. Improving EMR System Adoption in Canadian Medical Practice: A Research Model. In: World Congress on Privacy, Security, Trust and the Management of e-Business. 2009 Presented at: CONGRESS'09; August 25-27, 2009; Saint John, NB, Canada. [doi: 10.1109/congress.2009.15]
64. Bradshaw RL, Matney S, Livne OE, Bray BE, Mitchell JA, Narus SP. Architecture of a federated query engine for heterogeneous resources. AMIA Annu Symp Proc 2009 Nov 14;2009:70-74 [FREE Full text] [Medline: 20351825]
65. Chung P, Afzal F, Hsiao H. A Software System Development for Probabilistic Relational Database Applications for Biomedical Informatics. In: International Conference on Advanced Information Networking and Applications Workshops. 2009 Presented at: WAINA'09; May 26-29, 2009; Bradford, UK. [doi: 10.1109/waina.2009.99]
66. Costanzo D. Biomedical Data Acquisition and Processing in the Decision Support Services of HEARTFAID Platform. In: International Workshop on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications. 2009 Presented at: IDAACS'09; September 21-23, 2009; Rende, Italy. [doi: 10.1109/idaacs.2009.5342976]
67. Dong J, Zhou D, Hu X, Zhang Z, Jiang K. Analysis and Design on Standard System of Electronic Health Records. In: Proceedings of the 2009 First International Workshop on Education Technology and Computer Science - Volume 01. 2009 Presented at: ETCS'09; March 7-8, 2009; Italy. [doi: 10.1109/etcs.2009.223]
68. Milano F, Eijo J, Gomez A, de Quiros F, Risk M. MedSiGRe: Medical Signal Grid Repository, an Integration to Italica Project. In: Latin American Network Operations and Management Symposium. 2009 Presented at: LANOMS;09; October 19-21, 2009; Punta del Este, Uruguay. [doi: 10.1109/lanoms.2009.5338792]
69. Ongenae F, Dupont T, Kerckhove W. Design of ICU Medical Decision Support Applications by Integrating Service Oriented Applications With a Rule-Based System. In: 2nd International Symposium on Applied Sciences in Biomedical and Communication Technologies. 2009 Presented at: ISABEL'09; January 8, 2009; Bratislava, Slovakia. [doi: 10.1109/isabel.2009.5373663]
70. Patra D, Ray S, Mukhopadhyay J, Majumdar B, Majumdar A. Achieving E-Health Care in a Distributed EHR System. In: 11th Internation11th International Conference on e-Health Networking, Applications and Services al Conference on e-Health Networking, Applications and Services. 2009 Presented at: Healthcom'09; December 16-18, 2009; Sydney, NSW, Australia. [doi: 10.1109/health.2009.5406205]
71. Rusu M, Saplacan G, Sebestyen G. Distributed e-Health system with Smart Self-Care Units. In: 5th International Conference on Intelligent Computer Communication and Processing. 2009 Presented at: ICCP'09; August 27-29, 2009; Cluj-Napoca, Romania. [doi: 10.1109/iccp.2009.5284744]
72. Siddiqi J, Akhgar B, Rahman F. Towards an Integrated Platform for Improving Hospital Risk Management. In: Sixth International Conference on Information Technology: New Generations. 2009 Presented at: IRNG'09; April 27-29, 2009; Las Vegas, NV, USA. [doi: 10.1109/itng.2009.204]
73. Wah T, Sim O. Development of a data warehouse for lymphoma cancer diagnosis and treatment decision support. WSEAS Transactions on Information Science and Applications 2009;6(3):530-543. [doi: 10.37394/23209]
74. Wu F, Williams M, Kazanzides P, Brady K, Fackler J. A Modular Clinical Decision Support System Clinical Prototype Extensible Into Multiple Clinical Settings. In: 3rd International Conference on Pervasive Computing Technologies for Healthcare. 2009 Presented at: Pervasive Health'09; April 1-3, 2009; London, UK. [doi: 10.4108/icst.pervasivehealth2009.6078]
75. Yang L, Tuzel O, Chen W, Meer P, Salaru G, Goodell LA, et al. PathMiner: a web-based tool for computer-assisted diagnostics in pathology. IEEE Trans Inf Technol Biomed 2009 May;13(3):291-299 [FREE Full text] [doi: 10.1109/TITB.2008.2008801] [Medline: 19171530]
76. Yu J. Distributed Data Processing Framework for Oral Health Care Information Management Based on CSCWD Technology. In: First International Conference on Information Science and Engineering. 2009 Presented at: ICISE'09; December 26-28, 2009; Nanjing, China. [doi: 10.1109/icise.2009.511]

77. Ceusters W, Smith B. A unified framework for biomedical terminologies and ontologies. Stud Health Technol Inform 2010;160(Pt 2):1050-1054 [FREE Full text] [Medline: 20841844]
78. Chen P, Freg C, Hou T, Teng W. Implementing RAID-3 on Cloud Storage for EMR System. In: International Computer Symposium. 2010 Presented at: ICI'10; December 16-18, 2010; Tainan, Taiwan. [doi: 10.1109/compsym.2010.5685395]
79. Couderc J. A unique digital electrocardiographic repository for the development of quantitative electrocardiography and cardiac safety: the Telemetric and Holter ECG Warehouse (THEW). J Electrocardiol 2010;43(6):595-600 [FREE Full text] [doi: 10.1016/j.jelectrocard.2010.07.015] [Medline: 20863512]
80. Dangl A, Demiroglu SY, Gaedcke J, Helbing K, Jo P, Rakebrandt F, et al. The IT-infrastructure of a biobank for an academic medical center. Stud Health Technol Inform 2010;160(Pt 2):1334-1338. [Medline: 20841901]
81. Duennebeil S, Sunyaev A, Leimeister J, Krcmar H. Strategies for Development and Adoption of EHR in German Ambulatory Care. In: 4th International Conference on Pervasive Computing Technologies for Healthcare. 2010 Presented at: PervasiveHealth'10; March 22-25, 2010; Munich, Germany. [doi: 10.4108/icst.pervasivehealth2010.8887]
82. El Fadly A, Lucas N, Rance B, Verplancke P, Lastic P, Daniel C. The REUSE project: EHR as single datasource for biomedical research. Stud Health Technol Inform 2010;160(Pt 2):1324-1328. [Medline: 20841899]
83. Frank L, Andersen S. Evaluation of Different Database Designs for Integration of Heterogeneous Distributed Electronic Health Records. In: International Conference on Complex Medical Engineering. 2010 Presented at: ICCME'10; July 13-15, 2010; Gold Coast, QLD, Australia. [doi: 10.1109/iccme.2010.5558844]
84. Frize M, Bariciak E, Weyand S. Suggested Criteria for Successful Deployment of a Clinical Decision Support System (CDSS). In: International Workshop on Medical Measurements and Applications. 2010 Presented at: MEMEA'10; April 30-May 1, 2010; Ottawa, ON, Canada. [doi: 10.1109/memea.2010.5480227]
85. Jiang L, Cai H, Xu B. A Domain Ontology Approach in the ETL Process of Data Warehousing. In: 7th International Conference on E-Business Engineering. 2010 Presented at: ICEBE'10; November 10-12, 2010; Shanghai, China. [doi: 10.1109/icebe.2010.36]
86. Kataria P, Juric R. Sharing E-health Information Through Ontological Layering. In: 43rd Hawaii International Conference on System Sciences. 2010 Presented at: HICSS'10; January 5-8, 2010; Honolulu, HI, USA. [doi: 10.1109/hicss.2010.338]
87. Kildea J, Evans M, Parker W. A Framework for Comprehensive Electronic QA in Radiation Therapy. In: Ninth International Conference on Machine Learning and Applications. 2010 Presented at: ICMLA'10; December 10-14, 2010; Washington, DC, USA. [doi: 10.1109/icmla.2010.157]
88. Mougiakakou SG, Bartsocas CS, Bozas E, Chaniotakis N, Iliopoulou D, Kouris I, et al. SMARTDIAB: a communication and information technology approach for the intelligent monitoring, management and follow-up of type 1 diabetes patients. IEEE Trans Inf Technol Biomed 2010 May;14(3):622-633. [doi: 10.1109/TITB.2009.2039711] [Medline: 20123578]
89. Payne PR, Borlawsky TB, Stephens W, Barrett MC, Nguyen-Pham T, Greaves AW. The TRITON project: design and implementation of an integrative translational research information management platform. AMIA Annu Symp Proc 2010 Nov 13;2010:617-621 [FREE Full text] [Medline: 21347052]
90. Ping H, Xin-Lei W. Health Information System Grid Based on WSRF. In: Second International Conference on Information Technology and Computer Science. 2010 Presented at: ITCS'10; July 24-25, 2010; Kiev, Ukraine. [doi: 10.1109/itcs.2010.133]
91. Podchiyska T, Hernandez P, Ferris T, Weber S, Lowe HJ. Managing medical vocabulary updates in a clinical data warehouse: an RxNorm case study. AMIA Annu Symp Proc 2010 Nov 13;2010:477-481 [FREE Full text] [Medline: 21347024]
92. Roelofs E, Persoon L, Qamhiyeh S, Verhaegen F, de Ruysscher D, Scholz M, et al. Design of and technical challenges involved in a framework for multicentric radiotherapy treatment planning studies. Radiother Oncol 2010 Dec;97(3):567-571. [doi: 10.1016/j.radonc.2010.08.009] [Medline: 20864198]
93. Sachdeva S, Mchome S, Bhalla S. Web Services Security Issues in Healthcare Applications. In: 9th International Conference on Computer and Information Science. 2010 Presented at: ICIS'10; August 18-20, 2010; Yamagata, Japan. [doi: 10.1109/icis.2010.134]
94. Spitzer AR, Ellsbury DL, Handler D, Clark RH. The pediatrix babysteps data warehouse and the pediatrix qualitysteps improvement project system--tools for 'meaningful use' in continuous quality improvement. Clin Perinatol 2010 Mar;37(1):49-70. [doi: 10.1016/j.clp.2010.01.016] [Medline: 20363447]
95. Tohouri R, Asangansi I, Titlestad O, Braa J. The Change Strategy Towards an Integrated Health Information Infrastructure: Lessons From Sierra Leone. In: 43rd Hawaii International Conference on System Sciences. 2010 Presented at: HICSS'10; January 5-8, 2010; Honolulu, HI, USA. [doi: 10.1109/hicss.2010.367]
96. Vcelák P, Klecková J, Rohan V. Cerebrovascular Diseases Research Database. In: 3rd International Conference on Biomedical Engineering and Informatics. 2010 Presented at: BMEI'10; October 16-18, 2010; Yantai, China. [doi: 10.1109/bmei.2010.5639845]
97. Yaowen Z, Wei X, Yuwan H. Cerebrovascular Diseases Research Databaseresearch on Healthcare Integrating Model of Medical Information System Based on Agent. In: International Conference on Computational and Information Sciences. 2010 Presented at: ICISS'10; December 17-19, 2010; Chengdu, China. [doi: 10.1109/iccis.2010.157]
98. Zhao J, Wang T. A General Framework for Medical Data Mining. In: International Conference on Future Information Technology and Management Engineering. 2010 Presented at: FITME'10; October 9-10, 2010; Changzhou, China. [doi: 10.1109/fitme.2010.5654718]

99. Zhou X, Chen S, Liu B, Zhang R, Wang Y, Li P, et al. Development of traditional Chinese medicine clinical data warehouse for medical knowledge discovery and decision support. Artif Intell Med 2010;48(2-3):139-152. [doi: 10.1016/j.artmed.2009.07.012] [Medline: 20122820]
100. Apte M, Neidell M, Furuya EY, Caplan D, Glied S, Larson E. Using electronically available inpatient hospital data for research. Clin Transl Sci 2011 Oct;4(5):338-345 [FREE Full text] [doi: 10.1111/j.1752-8062.2011.00353.x] [Medline: 22029805]
101. Chazard E, Băceanu A, Ferret L, Ficheur G. The ADE scorecards: a tool for adverse drug event detection in electronic health records. Stud Health Technol Inform 2011;166:169-179. [Medline: 21685622]
102. Cossu M, Furfori P, Taddei A, Mangione M, del Sarto P. Anesthesia Information Management System in Cardiac Surgery. In: Computing in Cardiology. 2011 Presented at: CIC'11; September 8-11, 2011; Hangzhou, China URL: https://ieeexplore.ieee.org/document/6164631
103. Crichton D, Mattmann C, Hart A. An Informatics Architecture for the Virtual Pediatric Intensive Care Unit. In: 24th International Symposium on Computer-Based Medical Systems. 2011 Presented at: CBMS'11; June 27-30, 2011; Bristol, UK. [doi: 10.1109/cbms.2011.5999031]
104. Cuggia M, Garcelon N, Campillo-Gimenez B, Bernicot T, Laurent J, Garin E, et al. Roogle: an information retrieval engine for clinical data warehouse. Stud Health Technol Inform 2011;169:584-588. [Medline: 21893816]
105. Hatakeyama Y, Kataoka H, Nakajima N, Watabe T, Okuhara Y, Sagara Y. An Education Support System with Anonymized Medical Data Based on Thin Client System. In: International Conference on Internet of Things and 4th International Conference on Cyber, Physical and Social Computing. 2011 Presented at: CPSCom'11; October 19-22, 2011; Dalian, China. [doi: 10.1109/ithings/cpscom.2011.37]
106. Kanagaraj G, Sumathi A. Proposal of an Open-source Cloud Computing System for Exchanging Medical Images of a Hospital Information System. In: 3rd International Conference on Trendz in Information Sciences & Computing. 2011 Presented at: TICS'11; December 8-9, 2011; Chennai, India. [doi: 10.1109/tisc.2011.6169102]
107. Kiong Y, Palaniappan S, Yahaya N. Health Ontology System. In: 7th International Conference on Information Technology in Asia. 2011 Presented at: CITA'11; July 12-13, 2011; Kuching, Sarawak, Malaysia. [doi: 10.1109/cita.2011.5999506]
108. Murphy SN, Gainer V, Mendis M, Churchill S, Kohane I. Strategies for maintaining patient privacy in i2b2. J Am Med Inform Assoc 2011 Dec;18(Suppl 1):i103-i108 [FREE Full text] [doi: 10.1136/amiajnl-2011-000316] [Medline: 21984588]
109. Rajala T, Savio S, Penttinen J, Dastidar P, Kähönen M, Eskola H, et al. Development of a research dedicated archival system (TARAS) in a university hospital. J Digit Imaging 2011 Oct;24(5):864-873 [FREE Full text] [doi: 10.1007/s10278-010-9350-1] [Medline: 21042830]
110. Suapang P, Yimmun S, Puditkanawat A. Web-based Medical Image Archiving and Communication System for Teleimaging. In: 11th International Conference on Control, Automation and Systems. 2011 Presented at: CAS'11; October 26-29, 2011; Gyeonggi-do, South Korea.
111. Tanioka T, Osaka K, Chiba S. PSYCHOMS, an Electronic Nursing Management System to Facilitate Interdisciplinary Communication and Improve Patient Outcomes in Psychiatric Hospitals. In: 7th International Conference on Natural Language Processing and Knowledge Engineeringtional Conference on Natural Language Processing and Knowledge Engineering. 2011 Presented at: NLPKE'11; November 27-29, 2011; Tokushima, Japan. [doi: 10.1109/nlpke.2011.6138241]
112. Tenenbaum JD, Whetzel PL, Anderson K, Borromeo CD, Dinov ID, Gabriel D, et al. The biomedical resource ontology (BRO) to enable resource discovery in clinical and translational research. J Biomed Inform 2011 Feb;44(1):137-145 [FREE Full text] [doi: 10.1016/j.jbi.2010.10.003] [Medline: 20955817]
113. Teodoro D, Choquet R, Schober D, Mels G, Pasche E, Ruch P, et al. Interoperability driven integration of biomedical data sources. Stud Health Technol Inform 2011;169:185-189. [Medline: 21893739]
114. Zheng L, Wang L, Wang D, Deng N, Lu X, Duan H. A Clinical Omics Database Integrating Epidemiology, Clinical, and Omics Data for Colorectal Cancer Translational Research. In: 4th International Conference on Biomedical Engineering and Informatics (BMEI). 2011 Presented at: BMEI'11; October 15-17, 2011; Shanghai, China. [doi: 10.1109/bmei.2011.6098677]
115. Antoniades A, Georgousopoulos C, Forgo N. Linked2Safety: A Secure Linked Data Medical Information Space for Semantically-interconnecting EHRs Advancing Patients' Safety in Medical Research. In: 12th International Conference on Bioinformatics & Bioengineering. 2012 Presented at: BIBE'12; November 11-13, 2012; Larnaca, Cyprus. [doi: 10.1109/bibe.2012.6399767]
116. Armstrong A, Reddy S, Garg A. Novel approach to utilizing electronic health records for dermatologic research: developing a multi-institutional federated data network for clinical and translational research in psoriasis and psoriatic arthritis. Dermatol Online J 2012 May 15;18(5):2 [FREE Full text] [Medline: 22630572]
117. Bernal JG, Lopez DM, Blobel B. Architectural approach for semantic EHR systems development based on detailed clinical models. Stud Health Technol Inform 2012;177:164-169. [Medline: 22942049]
118. Blechner M, Saripalle R, Demurjian S. A Proposed Star Schema and Extraction Process to Enhance the Collection of Contextual & Semantic Information for Clinical Research Data Warehouses. In: International Conference on Bioinformatics and Biomedicine Workshops. 2012 Presented at: BIBMW'12; October 4-7, 2012; Philadelphia, PA, USA. [doi: 10.1109/bibmw.2012.6470242]

119. de Mul M, Alons P, van der Velde P, Konings I, Bakker J, Hazelzet J. Development of a clinical data warehouse from an intensive care clinical information system. Comput Methods Programs Biomed 2012 Jan;105(1):22-30. [doi: 10.1016/j.cmpb.2010.07.002] [Medline: 20728956]
120. Holford ME, McCusker JP, Cheung K, Krauthammer M. A semantic web framework to integrate cancer omics data with biological knowledge. BMC Bioinformatics 2012 Jan 25;13(Suppl 1):S10 [FREE Full text] [doi: 10.1186/1471-2105-13-S1-S10] [Medline: 22373303]
121. Hsu W, Taira RK, El-Saden S, Kangarloo H, Bui AA. Context-based electronic health record: toward patient specific healthcare. IEEE Trans Inf Technol Biomed 2012 Mar;16(2):228-234 [FREE Full text] [doi: 10.1109/TITB.2012.2186149] [Medline: 22395637]
122. Jayapandian CP, Zhao M, Ewing RM, Zhang G, Sahoo SS. A semantic proteomics dashboard (SemPoD) for data management in translational research. BMC Syst Biol 2012;6(Suppl 3):S20 [FREE Full text] [doi: 10.1186/1752-0509-6-S3-S20] [Medline: 23282161]
123. Liu D, Görges M, Jenkins SA. University of Queensland vital signs dataset: development of an accessible repository of anesthesia patient monitoring data for research. Anesth Analg 2012 Mar;114(3):584-589. [doi: 10.1213/ANE.0b013e318241f7c0] [Medline: 22190558]
124. Majeed RW, Röhrig R. Automated realtime data import for the i2b2 clinical data warehouse: introducing the HL7 ETL cell. Stud Health Technol Inform 2012;180:270-274. [Medline: 22874194]
125. Meyer J, Ostrzinski S, Fredrich D, Havemann C, Krafczyk J, Hoffmann W. Efficient data management in a large-scale epidemiology research project. Comput Methods Programs Biomed 2012 Sep;107(3):425-435. [doi: 10.1016/j.cmpb.2010.12.016] [Medline: 21256617]
126. Pan X, Zhou X, Song H, Zhang R, Zhang T. Enhanced Data Extraction, Transforming and Loading Processing for Traditional Chinese Medicine Clinical Data Warehouse. In: 14th International Conference on e-Health Networking, Applications and Services (Healthcom). 2012 Presented at: HealthCom'12; October 10-13, 2012; Beijing, China. [doi: 10.1109/HealthCom.2012.6380066]
127. Sfakianakis S, Sakkalis V, Marias K, Stamatakos G, McKeever S, Deisboeck TS, et al. An architecture for integrating cancer model repositories. Conf Proc IEEE Eng Med Biol Soc 2012;2012:6628-6631. [doi: 10.1109/EMBC.2012.6347514] [Medline: 23367449]
128. Tamersoy A, Loukides G, Nergiz ME, Saygin Y, Malin B. Anonymization of longitudinal electronic medical records. IEEE Trans Inf Technol Biomed 2012 May;16(3):413-423 [FREE Full text] [doi: 10.1109/TITB.2012.2185850] [Medline: 22287248]
129. Vcelak P, Kratochvil M, Kleckova J, Rohan V. Metamed-Medical Meta Data Extraction and Manipulation Tool Used in the Semantically Interoperable Research Information System. In: 5th International Conference on BioMedical Engineering and Informatics. 2012 Presented at: BMEI'12; October 16-18, 2012; Chongqing, China. [doi: 10.1109/bmei.2012.6513038]
130. da Silva KR, Costa R, Crevelari ES, Lacerda MS, de Moraes Albertini CM, Filho MM, et al. Glocal clinical registries: pacemaker registry design and implementation for global and local integration--methodology and case study. PLoS One 2013;8(7):e71090 [FREE Full text] [doi: 10.1371/journal.pone.0071090] [Medline: 23936257]
131. Farley T, Kiefer J, Lee P, von Hoff D, Trent JM, Colbourn C, et al. The biointelligence framework: a new computational platform for biomedical knowledge computing. J Am Med Inform Assoc 2013 Jan 1;20(1):128-133 [FREE Full text] [doi: 10.1136/amiajnl-2011-000646] [Medline: 22859646]
132. Fraccaro P, Dentone C, Fenoglio D, Giacomini M. Multicentre clinical trials' data management: a hybrid solution to exploit the strengths of electronic data capture and electronic health records systems. Inform Health Soc Care 2013 Dec;38(4):313-329. [doi: 10.3109/17538157.2013.812648] [Medline: 23957714]
133. Gokulakannan E, Venkatachalapathy K. Survey on Privacy Preserving Updates on Unidentified Database. In: Fourth International Conference on Computing, Communications and Networking Technologies. 2013 Presented at: ICCCNT'13; July 4-6, 2013; Tiruchengode, India. [doi: 10.1109/icccnt.2013.6726507]
134. Hamoud AK, Obaid T. Building data warehouse for diseases registry: first step for clinical data warehouse. SSRN J 2013;4(7):636-640. [doi: 10.2139/ssrn.3061599]
135. Heath J. A Privacy Framework for Secondary Use of Medical Data. In: International Symposium on Technology and Society: Social Implications of Wearable Computing and Augmediated Reality in Everyday Life. 2013 Presented at: ISTAS'13; June 27-29, 2013; Toronto, ON, Canada. [doi: 10.1109/istas.2013.6613116]
136. Kovacevic A, Dehghan A, Filannino M, Keane JA, Nenadic G. Combining rules and machine learning for extraction of temporal expressions and events from clinical narratives. J Am Med Inform Assoc 2013;20(5):859-866 [FREE Full text] [doi: 10.1136/amiajnl-2013-001625] [Medline: 23605114]
137. Miyoshi NS, Pinheiro DG, Silva WA, Felipe JC. Computational framework to support integration of biomolecular and clinical data within a translational approach. BMC Bioinformatics 2013 Jun 6;14:180 [FREE Full text] [doi: 10.1186/1471-2105-14-180] [Medline: 23742129]
138. Post AR, Kurc T, Cholleti S, Gao J, Lin X, Bornstein W, et al. The analytic information warehouse (AIW): a platform for analytics using electronic health record data. J Biomed Inform 2013 Jun;46(3):410-424 [FREE Full text] [doi: 10.1016/j.jbi.2013.01.005] [Medline: 23402960]

139. Rüping S, Anguita A, Bucur A, Cirstea T, Jacobs B, Torge A. Improving the Implementation of Clinical Decision Support Systems. In: 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. 2013 Presented at: EMBC'13; July 3-7, 2013; Osaka, Japan. [doi: 10.1109/embc.2013.6610225]
140. Seah B. An Application of a Healthcare Data Warehouse System. In: Third International Conference on Innovative Computing Technology. 2013 Presented at: INTECH'13; August 29-31, 2013; London, UK. [doi: 10.1109/intech.2013.6653698]
141. Spyropoulos B, Tzavaras A, Zogogianni D, Botsivaly M. Adapting the design of anesthesia information management systems to innovations depicted in industrial property documents. Conf Proc IEEE Eng Med Biol Soc 2013;2013:890-893. [doi: 10.1109/EMBC.2013.6609644] [Medline: 24109831]
142. Tsafara A, Tryfonopoulos C, Skiadopoulos S. CloudStudy: A Cloud-based System for Supporting Multi-centre Studies. In: 13th IEEE International Conference on BioInformatics and BioEngineering. 2013 Presented at: BIBE'13; November 10-13, 2013; Chania, Greece. [doi: 10.1109/bibe.2013.6701549]
143. Zhao X, Dong T. Design and experimental approach to the construction of a human signal-molecule-profiling database. Int J Environ Res Public Health 2013 Dec 9;10(12):6887-6908 [FREE Full text] [doi: 10.3390/ijerph10126887] [Medline: 24351788]
144. Bernsmed K, Cruzes D, Jaatun M, Haugset B, Gjære E. Healthcare Services in the Cloud--Obstacles to Adoption, and a Way Forward. In: Ninth International Conference on Availability, Reliability and Security. 2014 Presented at: ARES'14; September 8-12, 2014; Fribourg, Switzerland. [doi: 10.1109/ares.2014.28]
145. Cano I, Tényi A, Schueller C, Wolff M, Huertas Migueláñez MM, Gomez-Cabrero D, et al. The COPD knowledge base: enabling data analysis and computational simulation in translational COPD research. J Transl Med 2014 Nov 28;12(Suppl 2):S6 [FREE Full text] [doi: 10.1186/1479-5876-12-S2-S6] [Medline: 25471253]
146. Chen F, Wang S, Mohammed N, Cheng S, Jiang X. PRECISE:privacy-preserving cloud-assisted quality improvement service in healthcare. IEEE Int Conf Systems Biol 2014 Oct;2014:176-183 [FREE Full text] [doi: 10.1109/ISB.2014.6990752] [Medline: 26146645]
147. Chouvarda I, Philip N, Natsiavas P, Kilintzis V, Sobnath D, Kayyali R, et al. WELCOME – innovative integrated care platform using wearable sensing and smart cloud computing for COPD patients with comorbidities. Conf Proc IEEE Eng Med Biol Soc 2014;2014:3180-3183. [doi: 10.1109/EMBC.2014.6944298] [Medline: 25570666]
148. Dagliati A, Sacchi L, Bucalo M. A Data Gathering Framework to Collect Type 2 Diabetes Patients Data. In: International Conference on Biomedical and Health Informatics. 2014 Presented at: BHI'14; June 1-4, 2014; Valencia, Spain. [doi: 10.1109/bhi.2014.6864349]
149. Dalpé G, Joly Y. Opportunities and challenges provided by cloud repositories for bioinformatics-enabled drug discovery. Drug Dev Res 2014 Sep;75(6):393-401. [doi: 10.1002/ddr.21211] [Medline: 25195583]
150. Dietrich G, Fette G, Puppe F. A Comparison of Search Engine Technologies for a Clinical Data Warehouse. In: Proceedings of the 16th LWA Workshops. 2014 Presented at: LWA'14; September 8-10, 2014; Aachen, Germany.
151. Fiehe C, Litvina A, Tonn J. Building a Medical Research Cloud in the EASI-CLOUDS Project. In: 6th International Workshop on Science Gateways. 2014 Presented at: IWSG'14; June 3-5, 2014; Dublin, Ireland. [doi: 10.1109/iwsg.2014.13]
152. Gavrielov-Yusim N, Friger M. Use of administrative medical databases in population-based research. J Epidemiol Community Health 2014 Mar;68(3):283-287. [doi: 10.1136/jech-2013-202744] [Medline: 24248997]
153. Ghane K. Healthcare information exchange system based on a hybrid central/federated model. Conf Proc IEEE Eng Med Biol Soc 2014;2014:1362-1365. [doi: 10.1109/EMBC.2014.6943852] [Medline: 25570220]
154. Jaja BN, Attalla D, Macdonald RL, Schweizer TA, Cusimano MD, Etminan N, et al. The subarachnoid hemorrhage international trialists (SAHIT) repository: advancing clinical research in subarachnoid hemorrhage. Neurocrit Care 2014 Dec;21(3):551-559. [doi: 10.1007/s12028-014-9990-y] [Medline: 24865271]
155. Laohakangvalvit T, Achalakul T. Cloud-Based Data Exchange Framework for Healthcare Services. In: 11th International Joint Conference on Computer Science and Software Engineering. 2014 Presented at: JCSSE'14; May 14-16, 2014; Chon Buri, Thailand. [doi: 10.1109/jcsse.2014.6841874]
156. Mohammed RO, Talab SA. Clinical data warehouse issues and challenges. Int J u-e-Serv Sci Technol 2014 Oct 31;7(5):251-262. [doi: 10.14257/ijunesst.2014.7.5.22]
157. Parane K, Patil N, Poojara S, Kamble T. Cloud Based Intelligent Healthcare Monitoring System. In: International Conference on Issues and Challenges in Intelligent Computing Techniques. 2014 Presented at: ICICICT'14; February 7-8, 2014; Ghaziabad, India. [doi: 10.1109/icicict.2014.6781365]
158. Pecoraro F, Luzi D, Ricci F. A Clinical Data Warehouse Architecture based on the Electronic Healthcare Record Infrastructure. In: Proceedings of the International Conference on Health Informatics. 2014 Presented at: CHI;14; March 3-6, 2014; Angers, France. [doi: 10.5220/0004764502870294]
159. Pennington JW, Ruth B, Italia MJ, Miller J, Wrazien S, Loutrel JG, et al. Harvest: an open platform for developing web-based biomedical data discovery and reporting applications. J Am Med Inform Assoc 2014;21(2):379-383 [FREE Full text] [doi: 10.1136/amiajnl-2013-001825] [Medline: 24131510]
160. Rosenbloom ST, Harris P, Pulley J, Basford M, Grant J, DuBuisson A, et al. The mid-south clinical data research network. J Am Med Inform Assoc 2014;21(4):627-632 [FREE Full text] [doi: 10.1136/amiajnl-2014-002745] [Medline: 24821742]

161. Scalone L, Cesana G, Furneri G, Ciampichini R, Beck-Peccoz P, Chiodini V, et al. Burden of diabetes mellitus estimated with a longitudinal population-based study using administrative databases. PLoS One 2014;9(12):e113741 [FREE Full text] [doi: 10.1371/journal.pone.0113741] [Medline: 25470484]
162. Schreiweis B, Schneider G, Eichner T, Bergh B, Heinze O. Health information research platform (HIReP)--an architecture pattern. Stud Health Technol Inform 2014;205:773-777. [Medline: 25160292]
163. Setareh S, Rezaee A, Farahmandian V, Hajinazari P, Asosheh A. A Cloud-based Model for Hospital Information Systems Integration. In: 7th International Symposium on Telecommunications. 2014 Presented at: IST'14; September 9-11, 2014; Tehran, Iran. [doi: 10.1109/istel.2014.7000792]
164. Thorogood A, Joly Y, Knoppers BM, Nilsson T, Metrakos P, Lazaris A, et al. An implementation framework for the feedback of individual research results and incidental findings in research. BMC Med Ethics 2014 Dec 23;15:88 [FREE Full text] [doi: 10.1186/1472-6939-15-88] [Medline: 25539799]
165. Tsumoto S, Hirano S. Healthcare IT: Integration of Consumer Healthcare Data and Electronic Medical Records for Chronic Disease Management. In: International Conference on Granular Computing. 2014 Presented at: GrC'14; October 22-24, 2014; Noboribetsu, Japan. [doi: 10.1109/grc.2014.6982855]
166. Yamaguchi H, Ito Y. Improving the Effectiveness of Interprofessional Work Teams Using EHR-Based Data in the Treatment of Chronic Diseases: An Action Research Study. In: International Conference on Management of Engineering & Technology. 2007 Presented at: PICMET'07; August 5-9, 2014; Portland, OR, USA. [doi: 10.1109/picmet.2007.4349302]
167. Adler-Milstein J, DesRoches CM, Kralovec P, Foster G, Worzala C, Charles D, et al. Electronic health record adoption in US hospitals: progress continues, but challenges persist. Health Aff (Millwood) 2015 Dec;34(12):2174-2180. [doi: 10.1377/hlthaff.2015.0992] [Medline: 26561387]
168. Agbele K, Oriogun P, Seluwa A, Aruleba K. Towards a Model for Enhancing ICT4 Development and Information Security in Healthcare System. In: International Symposium on Technology and Society. 2015 Presented at: ISTAS'15; November 11-12, 2015; Dublin, Ireland. [doi: 10.1109/istas.2015.7439404]
169. Ahmadi M, Ghazisaeidi M, Bashiri A. Radiology reporting system data exchange with the electronic health record system: a case study in Iran. Glob J Health Sci 2015 Mar 18;7(5):208-214 [FREE Full text] [doi: 10.5539/gjhs.v7n5p208] [Medline: 26156904]
170. Amato F, Cozzolino G, Maisto A, Mazzeo A, Moscato V, Pelosi S, et al. ABC: A Knowledge Based Collaborative Framework for E-health. In: 2015 IEEE 1st International Forum on Research and Technologies for Society and Industry Leveraging a better tomorrow (RTSI). 2015 Presented at: RTSI'15; September 16-18, 2015; Turin, Italy. [doi: 10.1109/rtsi.2015.7325107]
171. Combi C, Pozzani G, Pozzi G. Design, Development, Deployment of a Telemedicine System in a Developing Country: Dealing With Organizational and Social Issues. In: International Conference on Healthcare Informatics. 2015 Presented at: ICHI'15; October 21-23, 2015; Dallas, TX, USA. [doi: 10.1109/ichi.2015.103]
172. Delamarre D, Bouzille G, Dalleau K, Courtel D, Cuggia M. Semantic integration of medication data into the EHOP clinical data warehouse. Stud Health Technol Inform 2015;210:702-706. [Medline: 25991243]
173. Girardi D, Dirnberger J, Giretzlehner M. An ontology‐based clinical data warehouse for scientific research. Saf Health 2015 Jul 20;1:6. [doi: 10.1186/2056-5917-1-6]
174. Huang Z, Duan H, Li H. TCGA4U: a web-based genomic analysis platform to explore and mine TCGA genomic data for translational research. Stud Health Technol Inform 2015;216:658-662. [Medline: 26262133]
175. Kämpgen B, Werner H, Deeb R, Bornhövd C. Towards a Semantic Clinical Data Warehouse: A Case Study of Discovering Similar Genes. In: Proceedings of the 4th Workshop on Knowledge Discovery and Data Mining Meets Linked Open Data co-located with 12th Extended Semantic Web Conference (ESWC 2015). 2015 Presented at: KDD'15; May 31, 2015; Portoroz, Slovenia.
176. Kantorovitch J, Giakoumaki A, Korakis A. Knowledge Modelling Framework. In: 2nd International Conference on Information and Communication Technologies for Disaster Management. 2015 Presented at: ICT-DM'15; November 30-December 2, 2015; Rennes, France. [doi: 10.1109/ict-dm.2015.7402037]
177. Kong G, Xiao Z. Protecting privacy in a clinical data warehouse. Health Informatics J 2015 Jun;21(2):93-106. [doi: 10.1177/1460458213504204] [Medline: 25301198]
178. Lee H, Kim H. Ehealth Recommendation Service System Using Ontology and Case-Based Reasoning. In: International Conference on Smart City/SocialCom/SustainCom. 2015 Presented at: SmartCity'15; December 19-21, 2015; Chengdu, China. [doi: 10.1109/smartcity.2015.217]
179. Lu J, Keech M. Emerging Technologies for Health Data Analytics Research: A Conceptual Architecture. In: Proceedings of the 2015 26th International Workshop on Database and Expert Systems Applications (DEXA). 2015 Presented at: DEXA'15; September 1-4, 2015; Washington, DC, USA. [doi: 10.1109/dexa.2015.58]
180. Maaroufi M, Choquet R, Landais P, Jaulent M. Towards data integration automation for the French rare disease registry. AMIA Annu Symp Proc 2015;2015:880-885 [FREE Full text] [Medline: 26958224]
181. Martin-Sanchez F, Turner M, Johnstone A, Heffer L, Rafael N, Advisory Group, et al. Personalised medicine possible with real-time integration of genomic and clinical data to inform clinical decision-making. Stud Health Technol Inform 2015;216:1052. [Medline: 26262351]

182. Mohyuddin. Bridging the gap from bench to bedside--an informatics infrastructure for integrating clinical, genomics and environmental data (ICGED). Stud Health Technol Inform 2015;216:1054. [Medline: 26262353]
183. Mou X, Wang X, Wu Z, Wang X, Zhou M. An Automatic Ehealth Platform for Cardiovascular and Cerebrovascular Disease Detection. In: International Symposium on Bioelectronics and Bioinformatics. 2015 Presented at: ISBB'15; October 14-17, 2015; Beijing, China. [doi: 10.1109/isbb.2015.7344924]
184. Puppala M, He T, Chen S, Ogunti R, Yu X, Li F, et al. METEOR: an enterprise health informatics environment to support evidence-based medicine. IEEE Trans Biomed Eng 2015 Dec;62(12):2776-2786. [doi: 10.1109/TBME.2015.2450181] [Medline: 26126271]
185. Sanz-Requena R, Mañas-García A, Cabrera-Ayala J, García-Martí G. A Cloud-Based Radiological Portal for the Patients: IT Contributing to Position the Patient as the Central Axis of the 21st Century Healthcare Cycles. In: Proceedings of the 2015 IEEE/ACM 1st International Workshop on TEchnical and LEgal aspects of data pRivacy and SEcurity. 2015 Presented at: TELERISE'15; May 18, 2015; Florence, Italy. [doi: 10.1109/telerise.2015.18]
186. Schnell R, Borgs C. Building a National Perinatal Data Base Without the Use of Unique Personal Identifiers. In: International Conference on Data Mining Workshop. 2015 Presented at: ICDMW'15; November 14-17, 2015; Atlantic City, NJ, USA. [doi: 10.1109/icdmw.2015.19]
187. Sharghi H, Ma W, Sartipi K. Federated Service-Based Authentication Provisioning for Distributed Diagnostic Imaging Systems. In: 28th International Symposium on Computer-Based Medical Systems. 2015 Presented at: CBMS'15; June 22-25, 2015; Sao Carlos, Brazil. [doi: 10.1109/cbms.2015.85]
188. Skripcak T, Just U, Simon M, Buttner D, Luhr A, Baumann M, et al. Toward distributed conduction of large-scale studies in radiation therapy and oncology: open-source system integration approach. IEEE J Biomed Health Inform 2016 Sep;20(5):1397-1403. [doi: 10.1109/jbhi.2015.2450833]
189. Soudris D, Xydis S, Baloukas C. AEGLE: A Big Bio-Data Analytics Framework for Integrated Health-Care Services. In: International Conference on Embedded Computer Systems: Architectures, Modeling, and Simulation. 2015 Presented at: SAMOS'15; July 19-23, 2015; Samos, Greece. [doi: 10.1109/samos.2015.7363682]
190. Weber GM. Federated queries of clinical data repositories: scaling to a national network. J Biomed Inform 2015 Jun;55:231-236 [FREE Full text] [doi: 10.1016/j.jbi.2015.04.012] [Medline: 25957825]
191. Yang C, Liu C, Tseng T. Design and Implementation of a Privacy Aware Framework for Sharing Electronic Health Records. In: International Conference on Healthcare Informatics. 2015 Presented at: ICHI'15; October 21-23, 2015; Dallas, TX, USA. [doi: 10.1109/ichi.2015.92]
192. Yao Q, Tian Y, Li P, Tian L, Qian Y, Li J. Design and development of a medical big data processing system based on Hadoop. J Med Syst 2015 Mar;39(3):23. [doi: 10.1007/s10916-015-0220-8] [Medline: 25666927]
193. Andrew NE, Sundararajan V, Thrift AG, Kilkenny MF, Katzenellenbogen J, Flack F, et al. Addressing the challenges of cross-jurisdictional data linkage between a national clinical quality registry and government-held health data. Aust N Z J Public Health 2016 Oct;40(5):436-442. [doi: 10.1111/1753-6405.12576] [Medline: 27625174]
194. Bauer CR, Ganslandt T, Baum B, Christoph J, Engel I, Löbe M, et al. Integrated data repository toolkit (IDRT). A suite of programs to facilitate health analytics on heterogeneous medical data. Methods Inf Med 2016;55(2):125-135. [doi: 10.3414/ME15-01-0082] [Medline: 26534843]
195. Chelico JD, Wilcox AB, Vawdrey DK, Kuperman GJ. Designing a clinical data warehouse architecture to support quality improvement initiatives. AMIA Annu Symp Proc 2016;2016:381-390 [FREE Full text] [Medline: 28269833]
196. Denney MJ, Long DM, Armistead MG, Anderson JL, Conway BN. Validating the extract, transform, load process used to populate a large clinical research database. Int J Med Inform 2016 Oct;94:271-274 [FREE Full text] [doi: 10.1016/j.ijmedinf.2016.07.009] [Medline: 27506144]
197. García-de-León-Chocano R, Muñoz-Soler V, Sáez C, García-de-León-González R, García-Gómez JM. Construction of quality-assured infant feeding process of care data repositories: construction of the perinatal repository (part 2). Comput Biol Med 2016 Apr 1;71:214-222. [doi: 10.1016/j.compbiomed.2016.01.007] [Medline: 26950399]
198. Gupta S, Tripathi P. Big Data Lakes Can Support Better Population Health for Rural India - Swastha Bharat. In: International Conference on Innovation and Challenges in Cyber Security. 2016 Presented at: ICICCS-INBUSH'16; February 3-5, 2016; Noida, India. [doi: 10.1109/iciccs.2016.7542361]
199. Koppad S, Kumar A. Application of Big Data Analytics in Healthcare System to Predict COPD. In: International Conference on Circuit, Power and Computing Technologies. 2016 Presented at: ICCPCT'16; March 18-19, 2016; Nagercoil, India. [doi: 10.1109/iccpct.2016.7530248]
200. Liaw S, De Lusignan S. An 'integrated health neighbourhood' framework to optimise the use of EHR data. J Innov Health Inform 2016 Oct 4;23(3):826 [FREE Full text] [doi: 10.14236/jhi.v23i3.826] [Medline: 28059689]
201. Olaronke I, Oluwaseun O. Big Data in Healthcare: Prospects, Challenges and Resolutions. In: Future Technologies Conference. 2016 Presented at: FTC'16; December 6-7, 2016; San Francisco, CA, USA. [doi: 10.1109/ftc.2016.7821747]
202. Padula WV, Blackshaw L, Brindle CT, Volchenboum SL. An approach to acquiring, normalizing, and managing EHR data from a clinical data repository for studying pressure ulcer outcomes. J Wound Ostomy Continence Nurs 2016;43(1):39-45. [doi: 10.1097/WON.0000000000000185] [Medline: 26727681]

203. Vishnyakova D, Gaudet-Blavignac C, Baumann P, Lovis C. Clinical data models at university hospitals of Geneva. Stud Health Technol Inform 2016;221:97-101. [Medline: 27071885]
204. Bellini F, Gutierrez-Zorrilla J, Anza L, Ferreira E, Deneault L, Vanerio G. MDi: acquisition, analysis and data visualization system in healthcare. In: 2017 IEEE URUCON. 2017 Presented at: URUCON; 23-25 Oct. 2017; Montevideo, Uruguay p. 1-4. [doi: 10.1109/urucon.2017.8171879]
205. Dogaru D, Dumitrache I. Holistic Perspective of Big Data in Healthcare. In: E-Health and Bioengineering Conference (EHB). 2017 Presented at: EHB'17; June 22-24, 2017; Sinaia, Romania. [doi: 10.1109/ehb.2017.7995450]
206. Ewing A, Rogus J, Chintagunta P, Kraus L, Sabol M, Kang H. A Systems Approach to Improving Patient Flow at UVA Cancer Center Using Real-time Locating System. In: Systems and Information Engineering Design Symposium (SIEDS). 2017 Presented at: SIEDS'17; April 28, 2017; Charlottesville, VA, USA. [doi: 10.1109/sieds.2017.7937727]
207. Foran DJ, Chen W, Chu H, Sadimin E, Loh D, Riedlinger G, et al. Roadmap to a comprehensive clinical data warehouse for precision medicine applications in oncology. Cancer Inform 2017;16:1176935117694349 [FREE Full text] [doi: 10.1177/1176935117694349] [Medline: 28469389]
208. Khan SI, Hoque AS. Development of National Health Data Warehouse Bangladesh: Privacy Issues and a Practical Solution. In: 18th International Conference on Computer and Information Technology. 2017 Presented at: ICCIT'17; December 21-23, 2017; Dhaka, Bangladesh. [doi: 10.1109/ICCITechn.2015.7488099]
209. Lai J, Lee D, Yang M. Constructing the Cloud Computing System for Advanced Data Analysis of Biomedical Research. In: International Conference on Machine Learning and Cybernetics. 2017 Presented at: ICMLC'17; July 9-12, 2017; Ningbo, China. [doi: 10.1109/icmlc.2017.8108969]
210. Poenaru C, Merezeanu D, Dobrescu R, Posdarascu E. Advanced Solutions for Medical Information Storing: Clinical Data Warehouse. In: E-Health and Bioengineering Conference. 2017 Presented at: EHB'17; June 22-24, 2017; Sinaia, Romania. [doi: 10.1109/ehb.2017.7995355]
211. Sáez C, Moner D, García-De-León-Chocano R, Muñoz-Soler V, García-De-León-González R, Maldonado JA, et al. A standardized and data quality assessed maternal-child care integrated data repository for research and monitoring of best practices: a pilot project in Spain. Stud Health Technol Inform 2017;235:539-543. [Medline: 28423851]
212. Vuppala S, Dinesh M, Viswanathan S, Ramachandran G, Bussa N, Geetha M. Cloud Based Big Data Platform for Image Analytics. In: International Conference on Cloud Computing in Emerging Markets. 2017 Presented at: CCEM'17; November 1-3, 2017; Bangalore, India. [doi: 10.1109/ccem.2017.11]
213. Weng W, Wagholikar KB, McCray AT, Szolovits P, Chueh HC. Medical subdomain classification of clinical notes using a machine learning-based natural language processing approach. BMC Med Inform Decis Mak 2017 Dec 1;17(1):155. [doi: 10.1186/s12911-017-0556-8] [Medline: 29191207]
214. Brisimi TS, Chen R, Mela T, Olshevsky A, Paschalidis IC, Shi W. Federated learning of predictive models from federated electronic health records. Int J Med Inform 2018 Apr;112:59-67. [doi: 10.1016/j.ijmedinf.2018.01.007] [Medline: 29500022]
215. Deshpande P, Rasin A, Brown E. Big Data Integration Case Study for Radiology Data Sources. In: Life Sciences Conference. 2018 Presented at: LSC'18; October 28-30, 2018; Montreal, QC, Canada. [doi: 10.1109/lsc.2018.8572185]
216. Fette G, Kaspar M, Liman L, Dietrich G, Ertl M, Krebs J, et al. Exporting data from a clinical data warehouse. Stud Health Technol Inform 2018;248:88-93. [Medline: 29726423]
217. Garcelon N, Neuraz A, Salomon R, Bahi-Buisson N, Amiel J, Picard C, et al. Next generation phenotyping using narrative reports in a rare disease clinical data warehouse. Orphanet J Rare Dis 2018 May 31;13(1):85 [FREE Full text] [doi: 10.1186/s13023-018-0830-6] [Medline: 29855327]
218. de Quirós FG, Otero C, Luna D. Terminology services: standard terminologies to control health vocabulary. Yearb Med Inform 2018 Aug;27(1):227-233 [FREE Full text] [doi: 10.1055/s-0038-1641200] [Medline: 29681027]
219. Guo S, Jin Z, Gotz D, Du F, Zha H, Cao N. Visual progression analysis of event sequence data. IEEE Trans Vis Comput Graph 2018 Aug 20;25(1):417-426. [doi: 10.1109/TVCG.2018.2864885] [Medline: 30136953]
220. Kundalwal M, Singh A, Chatterjee K. A Privacy Framework in Cloud Computing for Healthcare Data. In: International Conference on Advances in Computing, Communication Control and Networking. 2018 Presented at: ICACCCN'18; October 12-13, 2018; Greater Noida (UP), India, India. [doi: 10.1109/icacccn.2018.8748480]
221. Maji G, Debnath N, Sen S. Data Warehouse Based Analysis with Integrated Blood Donation Management System. In: 16th International Conference on Industrial Informatics. 2018 Presented at: INDIN'18; July 18-20, 2018; Porto, Portugal. [doi: 10.1109/indin.2018.8471988]
222. Mande R, JayaLakshmi G, Yelavarti K. Leveraging Distributed Data Over Big Data Analytics Platform for Healthcare Services. In: 2nd International Conference on Trends in Electronics and Informatics. 2018 Presented at: ICOEI'18; May 11-12, 2018; Tirunelveli, India. [doi: 10.1109/icoei.2018.8553827]
223. Mullin S, Zhao J, Sinha S, Lee R, Song B, Elkin PL. Clinical data warehouse query and learning tool using a human-centered participatory design process. Stud Health Technol Inform 2018;251:59-62 [FREE Full text] [Medline: 29968601]
224. Patiny L, Zasso M, Kostro D, Bernal A, Castillo AM, Bolaños A, et al. The C6H6 NMR repository: an integral solution to control the flow of your data from the magnet to the public. Magn Reson Chem 2018 Jun;56(6):520-528. [doi: 10.1002/mrc.4669] [Medline: 28981966]

225. Pletcher MJ, Forrest CB, Carton TW. PCORnet's collaborative research groups. Patient Relat Outcome Meas 2018 Feb;Volume 9:91-95. [doi: 10.2147/prom.s141630]
226. Raisaro JL, Troncoso-Pastoriza JR, Misbach M, Sousa JS, Pradervand S, Missiaglia E, et al. MedCo: enabling secure and privacy-preserving exploration of distributed clinical and genomic data. IEEE/ACM Trans Comput Biol Bioinform 2019;16(4):1328-1341. [doi: 10.1109/TCBB.2018.2854776] [Medline: 30010584]
227. Rinner C, Gezgin D, Wendl C, Gall W. A clinical data warehouse based on OMOP and i2b2 for Austrian health claims data. Stud Health Technol Inform 2018;248:94-99. [Medline: 29726424]
228. Solbrig HR, Hong N, Murphy SN, Jiang G. Automated population of an i2b2 clinical data warehouse using FHIR. AMIA Annu Symp Proc 2018;2018:979-988 [FREE Full text] [Medline: 30815141]
229. Sylvestre E, Bouzillé G, Chazard E, His-Mahier C, Riou C, Cuggia M. Combining information from a clinical data warehouse and a pharmaceutical database to generate a framework to detect comorbidities in electronic health records. BMC Med Inform Decis Mak 2018 Jan 24;18(1):9 [FREE Full text] [doi: 10.1186/s12911-018-0586-x] [Medline: 29368609]
230. Ye B, Basdekis I, Smyrlis M, Spanoudakis G, Koloutsou K. A Big Data Repository and Architecture for Managing Hearing Loss Related Data. In: International Conference on Biomedical & Health Informatics. 2018 Presented at: BHI'18; March 4-7, 2018; Las Vegas, NV, USA. [doi: 10.1109/bhi.2018.8333397]
231. Baghal A, Zozus M, Baghal A, Al-Shukri S, Prior F. Factors associated with increased adoption of a research data warehouse. Stud Health Technol Inform 2019;257:31-35. [Medline: 30741168]
232. Black M, Wallace J, Rankin D. Meaningful Integration of Data, Analytics and Services of Computer-Based Medical Systems: The MIDAS Touch. In: 32nd International Symposium on Computer-Based Medical Systems. 2019 Presented at: CBMS'19; June 5-7, 2019; Cordoba, Spain. [doi: 10.1109/cbms.2019.00031]
233. Boujdad F, Gaignard A, Südholt M, Garzón-Alfonso W, Navarro L, Redon R. On Distributed Collaboration for Biomedical Analyses. In: 19th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGRID). 2019 Presented at: CCGRID'19; May 14-17, 2019; Larnaca, Cyprus. [doi: 10.1109/ccgrid.2019.00079]
234. Conte B, Fortin E, Singh P. Health Information Management System for a Rural Medical Clinic in Nicaragua. In: 2019 IEEE 39th Central America and Panama Convention. 2019 Presented at: CAPC'19; November 20-22, 2019; Guatemala City, Guatemala. [doi: 10.1109/concapanxxxix47272.2019.8977099]
235. Daniel C, Kalra D, Section Editors for the IMIA Yearbook Section on Clinical Research Informatics. Clinical research informatics: contributions from 2018. Yearb Med Inform 2019 Aug;28(1):203-205 [FREE Full text] [doi: 10.1055/s-0039-1677921] [Medline: 31419833]
236. Deshpande P, Rasin A, Furst J, Raicu D, Antani S. DiiS: a biomedical data access framework for aiding data driven research supporting fair principles. Data 2019 Apr 20;4(2):54. [doi: 10.3390/data4020054]
237. Duncan D, Vespa P, Pitkänen A, Braimah A, Lapinlampi N, Toga AW. Big data sharing and analysis to advance research in post-traumatic epilepsy. Neurobiol Dis 2019 Mar;123:127-136 [FREE Full text] [doi: 10.1016/j.nbd.2018.05.026] [Medline: 29864492]
238. DuVall SL, Matheny ME, Ibragimov IR, Oats TD, Tucker JN, South BR, et al. A tale of two databases: the DoD and VA infrastructure for clinical intelligence (DaVINCI). Stud Health Technol Inform 2019 Aug 21;264:1660-1661. [doi: 10.3233/SHTI190584] [Medline: 31438280]
239. Fette G, Kaspar M, Liman L, Dietrich G, Ertl M, Krebs J, et al. Query translation between openEHR and i2b2. Stud Health Technol Inform 2019;258:16-20. [Medline: 30942705]
240. Gaies M, Anderson J, Kipps A, Lorts A, Madsen N, Marino B, Cardiac Networks United Executive Committee and Advisory Board. Cardiac networks united: an integrated paediatric and congenital cardiovascular research and improvement network. Cardiol Young 2019 Feb;29(2):111-118. [doi: 10.1017/S1047951118001683] [Medline: 30567622]
241. Gardner BJ, Pedersen JG, Campbell ME, McClay JC. Incorporating a location-based socioeconomic index into a de-identified i2b2 clinical data warehouse. J Am Med Inform Assoc 2019 Apr 1;26(4):286-293 [FREE Full text] [doi: 10.1093/jamia/ocy172] [Medline: 30715327]
242. Gerl A, Meier B. Privacy in the Future of Integrated Health Care Services - Are Privacy Languages the Key? In: International Conference on Wireless and Mobile Computing, Networking and Communication. 2019 Presented at: WiMob'19; October 21-23, 2019; Barcelona, Spain. [doi: 10.1109/wimob.2019.8923532]
243. Juárez D, Schmidt EE, Stahl-Toyota S, Ückert F, Lablans M. A generic method and implementation to evaluate and improve data quality in distributed research networks. Methods Inf Med 2019 Sep;58(2-03):86-93 [FREE Full text] [doi: 10.1055/s-0039-1693685] [Medline: 31514209]
244. Khalique F, Khan SA, Nosheen I. A framework for public health monitoring, analytics and research. IEEE Access 2019;7:101309-101326. [doi: 10.1109/access.2019.2930730]
245. Klangprapunt P, Seresangtakul P. An Information Integration System to Continuing of Care Case study Nongsung Hospital, Mukdahan THAILAND. In: 16th International Joint Conference on Computer Science and Software Engineering. 2019 Presented at: JCSSE'19; July 10-12, 2019; Chonburi, Thailand. [doi: 10.1109/jcsse.2019.8864214]
246. Krithara A, Aisopos F, Rentoumi V. iASiS: Towards Heterogeneous Big Data Analysis for Personalized Medicine. In: 32nd International Symposium on Computer-Based Medical Systems. 2019 Presented at: CBMS'19; June 5-7, 2019; Cordoba, Spain. [doi: 10.1109/cbms.2019.00032]

247. Lee JS, Darcy KM, Hu H, Casablanca Y, Conrads TP, Dalgard CL, et al. From discovery to practice and survivorship: building a national real-world data learning healthcare framework for military and veteran cancer patients. Clin Pharmacol Ther 2019 Jul;106(1):52-57 [FREE Full text] [doi: 10.1002/cpt.1425] [Medline: 30838639]
248. Liu T, Liu X, Fan Y. Constructing a Comprehensive Clinical Database Integrating Patients' Data from Intensive Care Units and General Wards. In: 12th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics. 2019 Presented at: CISP-BMEI'19; October 19-21, 2019; Suzhou, China. [doi: 10.1109/cisp-bmei48845.2019.8966055]
249. Phan-Vogtmann LA, Helhorn A, Kruse HM, Thomas E, Heidel AJ, Saleh K, et al. Approaching clinical data transformation from disparate healthcare IT systems through a modular framework. Stud Health Technol Inform 2019;258:85-89. [Medline: 30942720]
250. Miller JB. Big data and biomedical informatics: preparing for the modernization of clinical neuropsychology. Clin Neuropsychol 2019 Feb;33(2):287-304. [doi: 10.1080/13854046.2018.1523466] [Medline: 30513257]
251. Motema J, Appiah M. Factors Affecting the Adoption of Cloud Computing in a South African Hospital. In: International Conference on Advances in Big Data, Computing and Data Communication Systems. 2019 Presented at: icABCD'19; August 5-6, 2019; Winterton, South Africa. [doi: 10.1109/icabcd.2019.8851015]
252. Offia CE, Crowe M. A theoritical exploration of data management and integration in organisation sectors. Int J Database Manag Syst 2019 Feb 28;11(01):37-56. [doi: 10.5121/ijdms.2019.11103]
253. Opali?ski A, Regulski K, Mrzygóod B. Medical Data Exploration Based on the Heterogeneous Data Sources Aggregation System. In: Federated Conference on Computer Science and Information Systems. 2019 Presented at: FedCSIS'19; September 1-4, 2019; Leipzig, Germany. [doi: 10.15439/2019f258]
254. Post A, Chappidi N, Gunda D, Deshpande N. A method for EHR phenotype management in an i2b2 data warehouse. AMIA Jt Summits Transl Sci Proc 2019;2019:92-101 [FREE Full text] [Medline: 31258960]
255. Raebel MA, Quintana LM, Schroeder EB, Shetterly SM, Pieper LE, Epner PL, et al. Identifying preanalytic and postanalytic laboratory quality gaps using a data warehouse and structured multidisciplinary process. Arch Pathol Lab Med 2019 Apr;143(4):518-524 [FREE Full text] [doi: 10.5858/arpa.2018-0093-OA] [Medline: 30525932]
256. Safavi KC, Driscoll W, Wiener-Kronish JP. Remote surveillance technologies: realizing the aim of right patient, right data, right time. Anesth Analg 2019 Sep;129(3):726-734 [FREE Full text] [doi: 10.1213/ANE.0000000000003948] [Medline: 31425213]
257. Sekeres MA, Gore SD, Stablein DM, DiFronzo N, Abel GA, DeZern AE, et al. The national MDS natural history study: design of an integrated data and sample biorepository to promote research studies in myelodysplastic syndromes. Leuk Lymphoma 2019 Dec;60(13):3161-3171. [doi: 10.1080/10428194.2019.1616186] [Medline: 31111762]
258. Seneviratne M, Kahn M, Hernandez-Boussard T. Merging heterogeneous clinical data to enable knowledge discovery. Pac Symp Biocomput 2019;24:439-443 [FREE Full text] [Medline: 30864344]
259. Shaanika I, Nehemia M. Developing an Integration Architecture to Manage Heterogeneous Data by Private Healthcare Practitioners: A Case of Namibia. In: Open Innovations. 2019 Presented at: OI'19; October 2-4, 2019; Cape Town, South Africa. [doi: 10.1109/oi.2019.8908188]
260. Shah C, Shaikh M, Shah D, Samdani K. A Review on Big Data Practices in Healthcare. In: International Conference on System, Computation, Automation and Networking. 2019 Presented at: ICSCAN'19; March 29-30, 2019; Pondicherry, India. [doi: 10.1109/icscan.2019.8878687]
261. Weeks J, Pardee R. Learning to share health care data: a brief timeline of influential common data models and distributed health data networks in US health care research. EGEMS (Wash DC) 2019 Mar 25;7(1):4 [FREE Full text] [doi: 10.5334/egems.279] [Medline: 30937326]
262. Lian W, Xue T, Lu Y, Wang M, Deng W. Research on hierarchical data fusion of intelligent medical monitoring. IEEE Access 2020;8:38355-38367. [doi: 10.1109/access.2019.2958854]
263. Yau Y, Khethavath P, Figueroa J. Secure Pattern-Based Data Sensitivity Framework for Big Data in Healthcare. In: International Conference on Big Data, Cloud Computing, Data Science & Engineering. 2019 Presented at: BCD'19; May 29-31, 2019; Honolulu, HI, USA. [doi: 10.1109/bcd.2019.8885114]
264. Zhang C, Ma R, Sun S, Li Y, Wang Y, Yan Z. Optimizing the electronic health records through big data analytics: a knowledge-based view. IEEE Access 2019;7:136223-136231. [doi: 10.1109/access.2019.2939158]
265. Johnson AE, Pollard TJ, Shen L, Lehman LH, Feng M, Ghassemi M, et al. MIMIC-III, a freely accessible critical care database. Sci Data 2016 May 24;3:160035 [FREE Full text] [doi: 10.1038/sdata.2016.35] [Medline: 27219127]
266. Felmeister AR, Masino A, Resnick A, Pennington J. Scalable Biobanking: a Modular Electronic Honest Broker and Biorepository for Integrated Clinical, Specimen and Genomic Research. In: International Conference on Bioinformatics and Biomedicine. 2015 Presented at: BIBM'15; November 9-12, 2015; Washington, DC, USA. [doi: 10.1109/bibm.2015.7359732]
267. Kheterpal S, Woodrum DT, Tremper KK. Too much of a good thing is wonderful: observational data for perioperative research. Anesthesiology 2009 Dec;111(6):1183-1184. [doi: 10.1097/ALN.0b013e3181c1496c] [Medline: 19934857]
268. Raman PW, Storm P, Lilly J, Mason J, Heath A, Felmeister A, et al. Cavatica- a pediatric genomic cloud empowering data discovery through the pediatric brain tumor atlas. In: Neuro-Oncology. 2017 Presented at: 4th Biennial Conference on

Pediatric Neuro-Oncology Basic and Translational Research; June 15-16, 2017; New York City, NY, USA p. iv21. [doi: 10.1093/neuonc/nox083.086]
269. Raman P, Resnick AC, Storm PB, Mueller S, Schultz N, Cerami E, et al. PedcBioPortal: a Cancer Data Visualization Tool for Integrative Pediatric Cancer Analyses. In: 21st Annual Scientific Meeting and Education Day of the Society for Neuro-Oncology. 2016 Presented at: Neuro-Oncology'16; November 17-20, 2016; Scottsdale, Arizona. [doi: 10.1093/neuonc/now212.337]
270. Murphy SN, Weber G, Mendis M, Gainer V, Chueh HC, Churchill S, et al. Serving the enterprise and beyond with informatics for integrating biology and the bedside (i2b2). J Am Med Inform Assoc 2010;17(2):124-130 [FREE Full text] [doi: 10.1136/jamia.2009.000893] [Medline: 20190053]
271. Segagni D, Tibollo V, Dagliati A, Perinati L, Zambelli A, Priori S, et al. The ONCO-I2b2 project: integrating biobank information and clinical data to support translational research in oncology. Stud Health Technol Inform 2011;169:887-891. [Medline: 21893874]
272. Bodenreider O, Cornet R, Vreeman DJ. Recent developments in clinical terminologies - SNOMED CT, LOINC, and RxNorm. Yearb Med Inform 2018 Aug;27(1):129-139 [FREE Full text] [doi: 10.1055/s-0038-1667077] [Medline: 30157516]
273. Cimino JJ. High-quality, standard, controlled healthcare terminologies come of age. Methods Inf Med 2011;50(2):101-104 [FREE Full text] [Medline: 21416108]
274. Structure and Principles. WHO Collaborating Centre for Drug Statistics and Methodology. 2020. URL: https://www.whocc.no/atc/structure_and_principles/ [accessed 2020-01-01]
275. Robinson PN, Köhler S, Bauer S, Seelow D, Horn D, Mundlos S. The human phenotype ontology: a tool for annotating and analyzing human hereditary disease. Am J Hum Genet 2008 Nov;83(5):610-615 [FREE Full text] [doi: 10.1016/j.ajhg.2008.09.017] [Medline: 18950739]
276. The Gene Ontology Consortium. Expansion of the gene ontology knowledgebase and resources. Nucleic Acids Res 2017 Jan 4;45(D1):D331-D338 [FREE Full text] [doi: 10.1093/nar/gkw1108] [Medline: 27899567]
277. Murphy SN, Mendis M, Hackett K, Kuttan R, Pan W, Phillips LC, et al. Architecture of the open-source clinical research chart from informatics for integrating biology and the bedside. AMIA Annu Symp Proc 2007 Oct 11:548-552 [FREE Full text] [Medline: 18693896]
278. Scheufele E, Aronzon D, Coopersmith R, McDuffie MT, Kapoor M, Uhrich CA, et al. TranSMART: an open source knowledge management and high content data analytics platform. AMIA Jt Summits Transl Sci Proc 2014;2014:96-101 [FREE Full text] [Medline: 25717408]
279. Hripcsak G, Duke JD, Shah NH, Reich CG, Huser V, Schuemie MJ, et al. Observational health data sciences and informatics (OHDSI): opportunities for observational researchers. Stud Health Technol Inform 2015;216:574-578 [FREE Full text] [Medline: 26262116]
280. Mohammed-Rajput NA, Smith DC, Mamlin B, Biondich P, Doebbeling BN, Open MRS Collaborative Investigators. OpenMRS, a global medical records system collaborative: factors influencing successful implementation. AMIA Annu Symp Proc 2011;2011:960-968 [FREE Full text] [Medline: 22195155]
281. Epstein R, Hofer I, Salari V, Gabel E. Successful implementation of a perioperative data warehouse using another hospital's published specification from epic's electronic health record system. Anesth Analg 2020 Apr 22:- epub ahead of print. [doi: 10.1213/ANE.0000000000004806] [Medline: 32332291]
282. Rizi SA, Roudsari A. Development of a public health reporting data warehouse: lessons learned. Stud Health Technol Inform 2013;192:861-865. [Medline: 23920680]
283. STARR Tools. Stanford Medicine. URL: http://med.stanford.edu/starr-tools.html [accessed 2018-03-26]
284. Wharton Research Data Services. URL: https://wrds-www.wharton.upenn.edu/ [accessed 2018-03-27]
285. MPOG – Multicenter Perioperative Outcomes Group. URL: https://mpog.org/ [accessed 2018-04-03]
286. Boston University Medical Campus and Boston Medical Center. URL: http://www.bumc.bu.edu/ohra/using-bmc-and-chc-data-for-research-purposes/ [accessed 2018-04-12]
287. Children's Hospital of Philadelphia® Center for Data-Driven Discovery in Biomedicine. URL: https://d3b.center/aboutd3b/history/ [accessed 2018-04-17]
288. Lex A, Gehlenborg N, Strobelt H, Vuillemot R, Pfister H. UpSet: visualization of intersecting sets. IEEE Trans Vis Comput Graph 2014 Dec;20(12):1983-1992 [FREE Full text] [doi: 10.1109/TVCG.2014.2346248] [Medline: 26356912]
289. Gallego B, Walter SR, Day RO, Dunn AG, Sivaraman V, Shah N, et al. Bringing cohort studies to the bedside: framework for a 'green button' to support clinical decision-making. J Comp Eff Res 2015 May;4(3):191-197. [doi: 10.2217/cer.15.12] [Medline: 25959863]
290. Longhurst CA, Harrington RA, Shah NH. A 'green button' for using aggregate patient data at the point of care. Health Aff (Millwood) 2014 Jul;33(7):1229-1235. [doi: 10.1377/hlthaff.2014.0099] [Medline: 25006150]
291. Gombar S, Callahan A, Califf R, Harrington R, Shah NH. It is time to learn from patients like mine. NPJ Digit Med 2019;2:16 [FREE Full text] [doi: 10.1038/s41746-019-0091-3] [Medline: 31304364]
292. Schuler A, Callahan A, Jung K, Shah NH. Performing an informatics consult: methods and challenges. J Am Coll Radiol 2018 Mar;15(3 Pt B):563-568 [FREE Full text] [doi: 10.1016/j.jacr.2017.12.023] [Medline: 29396125]

Abbreviations

BMC: BioMed Central
BRP: biorepository portal
BTRIS: biomedical translational research information system
CARPEM: CAncer Research for PErsonalized Medicine
CDM: common data model
CDSS: clinical decision support system
CLB: Léon Bérard Cancer Center
DW4TR: Data Warehouse for Translational Research
EHR: electronic health record
ETL: extraction, transformation, and loading
FURTHeR: Federated Utah Research and Translational Health Electronic Repository
HaMSTR: Hanover Medical School Translational Research Framework
HSSC: Health Science, South Carolina
ICD: International Classification of Diseases
IDR: integrated data repository
IEEE Xplore: Institute of Electrical and Electronics Engineers Xplore
i2b2: Informatics for Integrating Biology and the Bedside
MEDLINE: Medical Literature Analysis and Retrieval System Online
METEOR: Methodist Environment for Translational Enhancement and Outcome Research
MIDH: Maternal and Infant Data Hub
MOSAIC: models and simulation techniques for discovering diabetes-related factors
NLP: natural language processing
OMOP: Observational Medical Outcomes Partnership
SNOMED-CT: systematized nomenclature of medicine-clinical terms
STARR: STAnford Research Repository
STRIDE: Stanford Translational Research Integrated Database Environment
TRC: Translational Research Center

Edited by C Lovis, G Eysenbach; submitted 03.01.20; peer-reviewed by V Huser, S Mussavi Rizi, H Ulrich; comments to author 01.03.20; revised version received 09.06.20; accepted 17.07.20; published 27.08.20.

Please cite as:
Gagalova KK, Leon Elizalde MA, Portales-Casamar E, Görges M
What You Need to Know Before Implementing a Clinical Research Data Warehouse: Comparative Review of Integrated Data Repositories in Health Care Institutions
JMIR Form Res 2020;4(8):e17687
URL: http://formative.jmir.org/2020/8/e17687/
doi:10.2196/17687
PMID:32852280

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org