JMIR mHealth and uHealth

Impact Factor (2024): 5.4
Volume 7 (2019), Issue 2 ISSN 2291-5222 Editor in Chief: Lorraine Buis, PhD, MSI

Contents

Original Papers

Reviews

Viewpoint

Corrigenda and Addenda

Original Paper

Perceptions and Acceptance of mHealth in Patients With Cardiovascular Diseases: A Cross-Sectional Study

Jie Jiang[1], MD; Qin Zhu[2], MS; Yimei Zheng[1], BA; Yajing Zhu[2], MS; Yuxi Li[1], MD; Yong Huo[1], MD, FACC

[1]Department of Cardiology, Peking University First Hospital, Beijing, China

[2]Health Systems, Philips Research China, Shanghai, China

Corresponding Author:
Yong Huo, MD, FACC
Department of Cardiology
Peking University First Hospital
8 Xishiku Street
Xicheng District
Beijing, 100034
China
Phone: 86 1083572283
Email: huoyong@263.net.cn

Abstract

Background: Mobile health (mHealth)—a method of assisting long-term care in patients with chronic cardiovascular diseases (CVDs)—is gaining popularity in China, mainly owing to the large number of patients and limited clinical resources. Patients of different ages have varying needs for CVD management. However, evidence regarding how age influences Chinese CVD patients' use and perceptions of mHealth is limited.

Objective: This study aimed to explore age-related differences among Chinese patients with CVD regarding their use and perceptions of mHealth and to determine the factors that influence this population's willingness to use mHealth technologies.

Methods: We conducted a cross-sectional study of patients with chronic CVDs in a tertiary hospital in Beijing using a new questionnaire designed by the investigators. Participants were sourced using nonproportional quota-sampling methods, being recruited consecutively in each sampling category (age 18-49, 50-64, 65-74, and ≥75 years, with at least 25 men and 25 women in each age group). The survey consisted of 5 parts, including sociodemographic profile and medical history; current disease management situation; self-evaluation of disease management; current usage of mobile and internet technology (IT); and willingness to use an mHealth solution to perform disease self-management. Responses were compared among the 4 age groups as well as between patients who were willing to use mHealth solutions and those who were not. Multivariate logistic regression model was used to identify predictors of willingness to use mHealth for self-management.

Results: Overall, 231 patients (124 men) completed the questionnaire; of these, 53 were aged 18-49 years, 66 were aged 50-64 years, 54 were aged 65-74 years, and 58 were aged ≥75 years. Patients in the older cohorts visited hospitals more often than did those in the younger cohorts (P<.001), and they also showed lower technology skills regarding the use of mobile or internet devices (P<.001) and searched for health-related information on the internet less often (P<.001). In addition, 68.0% (157/231) of the patients showed interest in using mHealth solution to manage their disease; of these, 40.8% (64/157) were aged ≥65 years. Patients who were more willing to use mHealth solution to manage their diseases were younger (P<.001), more educated (P<.001), still working (P=.001), possessed higher skill regarding mobile or internet device use (P<.001), and more frequently searched for health information on the internet (P<.001). Finally, multivariate logistic regression showed that IT skill was the single indicator (P=.003) of willingness to use mHealth, not age.

Conclusions: Although age is associated with the use of mobile or internet devices, the sole indicator of mHealth use for self-management was participants' IT skills. Education regarding the use of mobile devices and development of easy-to-use software might improve the acceptance of mHealth solutions among older patient populations.

(JMIR Mhealth Uhealth 2019;7(2):e10117) doi:10.2196/10117

KEYWORDS

age; cardiovascular disease; internet; mHealth; mobile phone; secondary prevention; self-management

Introduction

Cardiovascular disease (CVD) is a major cause of death and disability worldwide, and its prevalence is increasing, particularly in low- and middle-income countries [1]. Approximately 290 million people in China have CVD, and this number is projected to grow rapidly over the coming decade [2]. Secondary prevention of CVD is critical for improving health outcomes and quality of life, as well as for achieving cost-effective health care [1]; however, for patients with CVD, the adherence rate to secondary prevention is relatively low. For example, a meta-analysis of 376,162 patients showed that the overall adherence to drugs that prevent CVD is just 57% [3]. The magnitude and impact of poor adherence in developing countries, like China, are assumed to be even higher given the paucity of health resources and existing inequities concerning access to health care [4]. This is important because poor adherence in chronic patients has been directly correlated with a greater incidence of recurrent cardiovascular events and increases in direct and indirect health care costs [1].

There is considerable international interest in mobile health (mHealth) as a possible intervention for people with chronic diseases, including patients with CVD [5]. Previous studies demonstrated that the application of mHealth technologies among patients with CVD has significant positive effects on self-care awareness, confidence regarding disease management, adherence behavior, and clinical outcomes. In addition, it has been indicated that mHealth-based self-management programs might help patients actively control and improve their disease condition [6]. Studies have shown that a daily automated short message service text message reminder can effectively increase medication adherence among patients with acute coronary syndrome during the early postdischarge period [7,8]. After using an mHealth system consisting of an automatic sphygmomanometer, lipid meter, and mobile phone, the intervention group of acute coronary syndrome survivors showed lower body mass index and had more patients achieving treatment goals for blood pressure (BP) and hemoglobin A_{1c} than did the usual care group at 12-month follow-up [9]. Individuals with a parental history of dementia showed better adherence and decreased BP level after using mobile phone-assisted technology during a 6-month period [10]. An mHealth self-monitoring program motivated patients with hypertension to improve their health management and led to decreased cigarette smoking, alcohol use, and BP level after 6 months [11]. Thus, mHealth technologies for the secondary prevention of CVD have the potential to improve the psychological condition, medication adherence, clinical outcome, and finally the lives of patients.

Success in designing mHealth interventions requires an adequate understanding of the patients' level of acceptance of the relevant technologies. Although diagnoses of CVD in younger populations are becoming more common, the main patient population remains those aged >65 years. Age at the time of diagnosis is a meaningful factor for long-term disease management. Younger patients are typically still working and may not pay enough attention to their disease, even though they will live with the disease and treatment for a longer time. On the other hand, the elderly take their disease and health more seriously but may lack the ability to perform self-management. Ensuring appropriate intervention designs for an audience of diverse ages and life stages is important, given these considerations.

Although many studies have been conducted on the topic of mHealth in the CVD population, most are in the Western context. A small number of studies have been conducted among the population of China, but these have been limited to narrowly defined interventions [12]. Furthermore, although most studies consider age to be a meaningful factor for evaluating acceptance and willingness, few have considered age-related differences concerning the use and perceptions in this regard.

Thus, to deepen the understanding of mHealth perceptions and acceptance, this study aimed to explore the use and perceptions of mHealth among Chinese patients with CVD using age-specific cohorts and to examine the factors that may influence this population's willingness to use mHealth technologies.

Methods

Study Design and Sample Strategy

This study was conducted in the Cardiology Department of Peking University First Hospital (PKU1), which is a tertiary hospital in the city of Beijing, China. This study was approved by the Clinical Research Ethical Committee of PKU1. All respondents provided written informed consent before the study questionnaires were distributed.

Recruitment was conducted in 2016 in the Cardiology Department of PKU1, including both inpatients and outpatients, following a cross-sectional study design. These participants were sourced using nonproportional quota-sampling methods. Patients were recruited consecutively in each sampling category based on age and gender; the minimum number of sampled units in each category was 25.

The inclusion criteria were as follows: (1) patients aged >18 years; (2) those capable of providing informed consent; (3) those being inpatients or outpatients of the Cardiology Department; and (4) those requiring long-term medication therapy because of hypertension, coronary heart disease, heart failure, or arrhythmia.

Instrumentation

To evaluate the applicability and perceptions of CVD-related mHealth, a survey method was used. A paper-and-pencil questionnaire was designed that featured multiple types of response scales with closed-ended questions. The survey was divided into 5 parts as follows: (1) sociodemographic profile and medical history; (2) current situation of disease management; (3) self-evaluation for disease management; (4) current usage of mobile and internet technology (IT); and (5) willingness to use a telehealth solution with a self-management system. The full questionnaire, with measurement responses, is displayed in Multimedia Appendix 1.

We developed the questionnaire based on a review of related studies [12,13], which was then evaluated by 3 cardiologists

and 3 cardiology nurses. Next, the questionnaire was pretested in 10 patients, and minor adjustments were made depending on their responses.

Data Collection

Data were collected during the interviews with patients, which were conducted by several research nurses, using a case record form. All research nurses were trained by one of the authors to ensure they understood the questionnaire equally. A nurse reviewed the questionnaire immediately after a patient had finished it to confirm the completeness of the questionnaire.

Two people with experience in data entry independently entered all case record form data into separate EpiData (Version 3.1, EpiData Software, Odense, Denmark) databases. The 2 databases were compared, and discrepancies were resolved by checking the original questionnaire.

Statistical Analysis

We used descriptive statistics and chi-square statistics to present the sociodemographic and CVD characteristics. Frequencies and percentages were evaluated for categorical variables, and means and SD were determined for continuous variables. For missing data, no imputation was performed during the analysis, and the actual percentages were used for describing categorical variables. In this study, age was stratified into 4 groups as follows: 18-49 years; 50-64 years; 65-75 years; and ≥75 years. In addition, 4 groups of questions were used to create a combined score to illustrate the overall impression on the responses for each group. “Frequency of visiting hospitals” was a combined score of “frequency of visiting hospital clinic” and “frequency of visiting community clinic”; “self-management at home” was a combined score of “frequency of measuring blood pressure at home,” “recording blood pressure at home,” “using pill box at home,” and “using wearable devices as self-management at home”; “self-management confidence” was a combined score of “capable of controlling one’s health condition,” “managing one’s health behavior,” “care for oneself at home,” and “have adequate knowledge to be able to care for oneself”; finally, “IT skill” score was created from the score for using smartphone, tablet, and personal computer (PC), as well as the capability to use Wi-Fi and install apps. For each question included in the combined scores, a score of 1-5 was given to each of the original responses. If there were 5 stage responses, they were given a 1-5 sequentially; for example, *Never*=1, *1-3 per year*=2, *1-3 per month*=3, *1-3 per week*=4, and *Everyday*=5. If there were binary responses, a 1 was given to *no* responses, while a 5 was given to *yes* responses. The sum of the scores for each question formed the combined results.

Chi-square statistics and analysis of variance were performed to test whether there were associations of age with (1) the current situation of disease management, (2) self-evaluation of disease self-management, and (3) current usage of mobile or IT technology as well as to examine the associations between the willingness to apply mHealth solutions and other variables. In addition, a multivariate logistic regression analysis was performed to identify the indicators of the willingness to use mHealth to manage CVD. Gender and variables with $P<.10$ were included in the regression analysis. We regarded $P<.05$ as statistically significant. All analyses were conducted using RStudio (Version 0.99.903, 2009-2016 RStudio, Inc)

Results

Sociodemographic and Medical History

Overall, 231 patients with CVD were recruited and completed the survey. The patients’ sociodemographic and main medical history are summarized in Table 1. Mean body mass index was 26.7 (SD 4.6) kg/m^2, which was similar across the 4 age groups. Almost half (103/229, 45.0%) of the patients had a university or master’s degree, and the education level significantly differed within the 4 age groups ($P<.001$), with 73.1% (38/53) of patients in the 18-49 years group having a university or master’s degree, while this percentage was only 32%-44% in other groups. Meanwhile, 64.5% (147/228) of patients were retired, including all but 1 patient in the group over 64 years and 56.2% (36/66) of participants in the 50-64 years age group. Only 10.4% (24/230) of participants were living outside of Beijing, and these were mostly younger patients (10/53, 18.9%, and 11/66, 16.7%, in the 18-49 and 50-64 years age groups, respectively). More than half of the patients (157/229, 68.6%) reported that they lived with their spouse. Almost half of the patients (131/231, 56.7%) were diagnosed with coronary artery diseases, mostly in the 3 older groups. Only 32.1% (17/53) of patients in the 18-49 years group had coronary artery diseases. Meanwhile, the percentage or hypertension in this younger group was 69.8% (37/53), which was higher than the average (120/231, 51.9%).

Current Usage of Mobile and Internet Technologies

Overall, 67.4% (155/230) of participants reported that they used smartphone every day, while almost half had never used a tablet (52.2%, 119/228) or PC (48.0%, 109/227). Meanwhile, 69.9% (160/229) reported that they were able to connect to Wi-Fi using their smartphone or tablet. However, 60.5% (138/228) reported that they did not know how to install a new app on their smartphones or tablets. Table 2 displays patients’ current usage conditions of mobile or IT and its association with age. The older cohort used smartphones, tablets, and PCs less frequently than did the younger cohort ($P<.001$). Similarly, a smaller proportion of older patients were able to connect to Wi-Fi and install apps ($P<.001$).

Table 1. Sociodemographic and medical history of patients by age.

Characteristics	Total (N=231), n (%)	Age stratification in years, n (%)				Chi-square test	
		18-49 (n=53)	50-64 (n=66)	65-74 (n=54)	≥75 (n=58)	χ^2 (df^a)	*P* value
Gender, n (%)						1.2 (*3*)	.76
Male	124 (53.7)	28 (52.8)	39 (59.1)	28 (51.9)	29 (50.0)		
Female	107 (46.3)	25 (47.2)	27 (40.9)	26 (48.1)	29 (50.0)		
Body mass index stratification in kg/m^2, n (%)						4.6 (*3*)	.21
<24	61 (27.2)	13 (26.5)	12 (18.5)	16 (30.2)	20 (35.1)		
≥24	163 (72.8)	36 (73.5)	53 (81.5)	37 (69.8)	37 (64.9)		
Education level, n (%)						25.9 (*6*)	<.001
Primary or junior	67 (29.3)	9 (17.3)	20 (30.3)	19 (35.2)	19 (33.3)		
Senior	59 (25.8)	5 (9.6)	25 (37.9)	16 (29.6)	13 (22.8)		
University or postgraduate	103 (45.0)	38 (73.1)	21 (31.8)	19 (35.2)	25 (43.9)		
Employment status, n (%)						153.2 (*6*)	<.001
Retired	147 (64.5)	1 (1.9)	36 (56.2)	52 (98.1)	58 (100.0)		
Part time	21 (9.2)	11 (20.8)	9 (14.1)	1 (1.9)	0 (0.0)		
Full time	60 (26.3)	41 (77.4)	19 (29.7)	0 (0.0)	0 (0.0)		
Home location, n (%)						14.0 (*3*)	.003
Beijing	206 (89.6)	43 (81.1)	55 (83.3)	52 (96.3)	56 (98.2)		
Outside of Beijing	24 (10.4)	10 (18.9)	11 (16.7)	2 (3.7)	1 (1.8)		
Living status, n (%)						7.0 (*6*)	.32
Alone	14 (6.1)	3 (5.9)	1 (1.5)	5 (9.3)	5 (8.6)		
Only with spouse	157 (68.6)	35 (68.6)	52 (78.8)	35 (64.8)	35 (60.3)		
With children	58 (25.3)	13 (25.5)	13 (19.7)	14 (25.9)	18 (31.0)		
Enrollment location						32.1 (*3*)	<.001
Outpatient	145(61.7)	37(69.8)	20(29.9)	16(28.6)	17(28.8)		
Inpatient	90(38.3)	16(30.2)	47(70.1)	40(71.4)	42(71.2)		
Medical history							
Coronary heart disease, n (%)						25.0 (*3*)	<.001
No	100 (43.3)	36 (67.9)	22 (33.3)	13 (24.1)	29 (50.0)		
Yes	131 (56.7)	17 (32.1)	44 (66.7)	41 (75.9)	29 (50.0)		
Hypertension, n (%)						12.7 (*3*)	.005
No	111 (48.1)	16 (30.2)	37 (56.1)	33 (61.1)	25 (43.1)		
Yes	120 (51.9)	37 (69.8)	29 (43.9)	21 (38.9)	33 (56.9)		
Arrhythmia, n (%)						10.8 (*3*)	.01
No	169 (73.2)	44 (83.0)	53 (80.3)	38 (70.4)	34 (58.6)		
Yes	62 (26.8)	9 (17.0)	13 (19.7)	16 (29.6)	24 (41.4)		
Heart failure, n (%)						5.9 (*3*)	.12
No	212 (91.8)	47 (88.7)	64 (97.0)	51 (94.4)	50 (86.2)		
Yes	19 (8.2)	6 (11.3)	2 (3.0)	3 (5.6)	8 (13.8)		
Diabetes, n (%)						5.2 (*3*)	.16
No	189 (81.8)	40 (75.5)	51 (77.3)	46 (85.2)	52 (89.7)		
Yes	42 (18.2)	13 (24.5)	15 (22.7)	8 (14.8)	6 (10.3)		

Characteristics	Total (N=231), n (%)	Age stratification in years, n (%)				Chi-square test	
		18-49 (n=53)	50-64 (n=66)	65-74 (n=54)	≥75 (n=58)	χ^2 (df^a)	*P* value
Hyperlipidemia, n (%)						4.3 (*3*)	.23
No	156 (67.5)	33 (62.3)	40 (60.6)	40 (74.1)	43 (74.1)		
Yes	75 (32.5)	20 (37.7)	26 (39.4)	14 (25.9)	15 (25.9)		
Cerebrovascular disease, n (%)						3.4 (*3*)	.34
No	214 (92.6)	52 (98.1)	61 (92.4)	49 (90.7)	52 (89.7)		
Yes	17 (7.4)	1 (1.9)	5 (7.6)	5 (9.3)	6 (10.3)		
Peripheral vascular disease, n (%)						2.9 (*3*)	.41
No	211 (91.3)	51 (96.2)	58 (87.9)	50 (92.6)	52 (89.7)		
Yes	20 (8.7)	2 (3.8)	8 (12.1)	4 (7.4)	6 (10.3)		

[a]*df*: degrees of freedom.

Current Disease Management

The frequency of searching for health information using the internet was found to be significantly correlated with age (*P*<.001); approximately 52.8% (94/178) of patients aged >50 years never searched the health information on the internet, while this percentage was only 11.3% (6/53) of the youngest group. Meanwhile, 78.9% (179/227) of patients reported that they were familiar with the medication they were taking. The frequency of visiting hospital clinics showed a significant difference across the 4 age groups (*P*=.004); >66.1% (74/112) of patients aged >65 years visited hospitals monthly. Although it seems that more patients aged ≥75 years used a pill box, there was no statistically significant difference among groups in this regard. In addition, the combined score for self-management at home had no significant correlation with age (*P*=.28). The participants' situation regarding disease management in terms of age is displayed in Table 3.

Self-Evaluation of Disease Management

Most patients reported being capable of controlling their health conditions (146/223, 65.5%), managing their health behaviors (150/222, 67.6%), and caring for themselves at home (175/221, 79.2%), but only 57.0% (126/221) believed that they had adequate knowledge to be able to care for themselves. In addition, the combined score of "self-management confidence" had no significant association with age (*P*=.60). Although no significant associations between self-evaluation of disease management and age were found, numerically patients aged <50 years and ≥75 years were less confident in their knowledge of their disease. Table 4 displays the self-evaluation of disease management and its association with age concerning our respondents.

Willingness to Use an mHealth Solution With a Self-Management System

Overall, 68.0% (157/231) of the participants were willing to use mHealth technologies to manage their CVD, of which 63.2% (86/136) felt that it might benefit them and 26.5% (36/136) patients reported just wanting to try it. Of the 32.0% (74/231) who reported that they were not interested in an mHealth solution, the major reasons were that they were unable to use the devices (64.4%, 47/73) and that the devices were too complicated for them (16.4%, 12/73). Participants were divided into 2 groups according to their willingness to use mHealth technologies. Patients in the younger cohort, who had a higher education level, who were still working, and had higher IT skills, had a higher frequency of internet searching, had the habit of self-management at home, and showed more interest in mHealth solution. The associations between willingness and other variables are shown in Table 5.

Multivariate logistic regression showed that although the willingness to use an mHealth solution varied among different age groups, this difference vanished after performing the multivariate logistic regression. Consequently, the IT skill was determined to be the only independent indicator for the willingness to use an mHealth solutions (*P*=.02). Results of multivariate logistic regression analysis are summarized in Table 6.

Table 2. Current usage of mobile or internet technology (IT) among patients and its association with age.

Characteristics	Total (N=231), n (%)	Age stratification in years, n (%)				Chi-square test	
		18-49 (n=53)	50-64 (n=66)	65-74 (n=54)	≥75 (n=58)	χ^2 (*df*[a])	*P* value
Smartphone, n (%)						57.0 (*12*)	<.001
Never	55 (23.9)	2 (3.8)	8 (12.1)	19 (35.2)	26 (44.8)		
1-3 per year	3 (1.3)	0 (0.0)	2 (3.0)	0 (0.0)	1 (1.7)		
1-3 per month	5 (2.2)	0 (0.0)	2 (3.0)	2 (3.7)	1 (1.7)		
1-3 per week	12 (5.2)	1 (1.9)	0 (0.0)	4 (7.4)	7 (12.1)		
Everyday	155 (67.4)	49 (94.2)	54 (81.8)	29 (53.7)	23 (39.7)		
Tablet, n (%)						46.0 (*12*)	<.001
Never	119 (52.2)	10 (18.9)	37 (56.9)	32 (60.4)	40 (70.2)		
1-3 per year	6 (2.6)	3 (5.7)	0 (0.0)	0 (0.0)	3 (5.3)		
1-3 per month	10 (4.4)	7 (13.2)	1 (1.5)	1 (1.9)	1 (1.8)		
1-3 per week	29 (12.7)	12 (22.6)	8 (12.3)	6 (11.3)	3 (5.3)		
Everyday	64 (28.1)	21 (39.6)	19 (29.2)	14 (26.4)	10 (17.5)		
Personal computer, n (%)						45.2 (*12*)	<.001
Never	109 (48.0)	10 (18.9)	27 (41.5)	34 (65.4)	38 (66.7)		
1-3 per year	10 (4.4)	3 (5.7)	3 (4.6)	2 (3.8)	2 (3.5)		
1-3 per month	8 (3.5)	2 (3.8)	1 (1.5)	3 (5.8)	2 (3.5)		
1-3 per week	20 (8.8)	3 (5.7)	9 (13.8)	4 (7.7)	4 (7.0)		
Everyday	80 (35.2)	35 (66.0)	25 (38.5)	9 (17.3)	11 (19.3)		
Able to connect to Wi-Fi, n (%)						25.0 (*3*)	<.001
No	69 (30.1)	5 (9.4)	15 (22.7)	20 (38.5)	29 (50.0)		
Yes	160 (69.9)	48 (90.6)	51 (77.3)	32 (61.5)	29 (50.0)		
Able to install apps, n (%)						60.8 (*3*)	<.001
No	138 (60.5)	10 (18.9)	39 (59.1)	39 (75.0)	50 (87.7)		
Yes	90 (39.5)	43 (81.1)	27 (40.9)	13 (25.0)	7 (12.3)		
IT skills[b] (score), mean (SD)	15.57 (7.16)	21.23 (4.84)	16.63 (6.38)	13.12 (6.39)	11.30 (6.78)	77.5 (*3*)[c]	<.001[c]

[a]*df*: degrees of freedom.

[b]"IT skills" was a combined result of the 5 questions in this part.

[c]The result was obtained using analysis of variance.

Table 3. Current situation of disease management by age.

Characteristics	Total (N=231), n (%)	Age stratification in years, n (%)				Chi-square test	
		18-49 (n=53)	50-64 (n=66)	65-74 (n=54)	≥75 (n=58)	χ^2 (df^a)	*P* value
Frequency of using the internet to find information about disease or treatment, n (%)						43.2 (*9*)	<.001
Never	100 (43.5)	6 (11.3)	26 (40.0)	30 (55.6)	38 (65.5)		
Rarely	27 (11.7)	7 (13.2)	10 (15.4)	7 (13.0)	3 (5.2)		
Sometimes	49 (21.3)	18 (34.0)	11 (16.9)	9 (16.7)	11 (19.0)		
Often	54 (23.5)	22 (41.5)	18 (27.7)	8 (14.8)	6 (10.3)		
Know your medications, n (%)						6.9 (*3*)	.08
No	48 (21.1)	9 (17.3)	21 (32.3)	8 (15.4)	10 (17.2)		
Yes	179 (78.9)	43 (82.7)	44 (67.7)	44 (84.6)	48 (82.8)		
Frequency of visiting hospital clinic, n (%)						27.9 (*12*)	.004
Never	31 (13.6)	8 (15.4)	15 (22.7)	6 (11.3)	2 (3.5)		
Once per year	29 (12.7)	9 (17.3)	11 (16.7)	6 (11.3)	3 (5.3)		
1-2 per half of year	42 (18.4)	15 (28.8)	8 (12.1)	7 (13.2)	12 (21.1)		
1-2 per month	120 (52.6)	20 (38.5)	31 (47.0)	33 (62.3)	36 (63.2)		
Weekly	6 (2.6)	0 (0.0)	1 (1.5)	1 (1.9)	4 (7.0)		
Frequency of visiting community clinic, n (%)						18.1 (*12*)	.11
Never	97 (42.5)	29 (54.7)	31 (47.7)	16 (30.2)	21 (36.8)		
Once per year	20 (8.8)	5 (9.4)	5 (7.7)	3 (5.7)	7 (12.3)		
1-2 per half of year	34 (14.9)	8 (15.1)	10 (15.4)	7 (13.2)	9 (15.8)		
1-2 per month	72 (31.6)	11 (20.8)	19 (29.2)	25 (47.2)	17 (29.8)		
Weekly	5 (2.2)	0 (0.0)	0 (0.0)	2 (3.8)	3 (5.3)		
Frequency of taking lab test, n (%)						18.1 (*12*)	.11
Never	26 (11.4)	9 (17.3)	11 (16.7)	4 (7.5)	2 (3.4)		
Once per year	63 (27.5)	16 (30.8)	20 (30.3)	11 (20.8)	16 (27.6)		
1-2 per half of year	90 (39.3)	18 (34.6)	26 (39.4)	20 (37.7)	26 (44.8)		
1-2 per month	45 (19.7)	9 (17.3)	7 (10.6)	17 (32.1)	12 (20.7)		
Weekly	5 (2.2)	0 (0.0)	2 (3.0)	1 (1.9)	2 (3.4)		
Frequency of measuring blood pressure at home, n (%)						16.1 (*12*)	.19
Never	40 (17.3)	8 (15.1)	15 (22.7)	12 (22.2)	5 (8.6)		
1-3 per year	16 (6.9)	5 (9.4)	5 (7.6)	5 (9.3)	1 (1.7)		
1-3 per month	49 (21.2)	14 (26.4)	15 (22.7)	7 (13.0)	13 (22.4)		
1-3 per week	57 (24.7)	15 (28.3)	12 (18.2)	15 (27.8)	15 (25.9)		
Everyday	69 (29.9)	11 (20.8)	19 (28.8)	15 (27.8)	24 (41.4)		
Did you record your blood pressure at home, n (%)						6.5 (*3*)	.09
No	138 (60.3)	28 (53.8)	42 (63.6)	39 (72.2)	29 (50.9)		
Yes	91 (39.7)	24 (46.2)	24 (36.4)	15 (27.8)	28 (49.1)		
Did you use pill box at home, n (%)						3.5 (*3*)	.32
No	91 (39.6)	23 (43.4)	29 (43.9)	22 (41.5)	17 (29.3)		
Yes	139 (60.4)	30 (56.6)	37 (56.1)	31 (58.5)	41 (70.7)		
Did you use wearables, n (%)						3.5 (*3*)	.32
No	198 (88.4)	42 (82.4)	58 (87.9)	48 (94.1)	50 (89.3)		
Yes	26 (11.6)	9 (17.6)	8 (12.1)	3 (5.9)	6 (10.7)		

Characteristics	Total (N=231), n (%)	Age stratification in years, n (%)				Chi-square test	
		18-49 (n=53)	50-64 (n=66)	65-74 (n=54)	≥75 (n=58)	χ^2 (df^a)	*P* value
Frequency of visiting hospital (score)[b], mean (SD)	5.58 (1.93)	4.94 (1.82)	5.12 (1.98)	6.17 (1.89)	6.14 (1.70)	17.2 (*3*)[c]	<.001[c]
Self-management at home (score)[d], mean (SD)	10.88 (4.17)	11.10 (4.43)	10.41 (3.98)	10.00 (4.25)	12.05 (3.90)	1.2 (*3*)[c]	.28[c]

[a]*df*: degrees of freedom.

[b]"Frequency of visiting hospital" is a combined result of "Frequency of visiting hospital clinic" and "Frequency of visiting community clinic".

[c]The result was obtained using analysis of variance.

[d]"Self-management at home" was a combined result of "Frequency of measuring blood pressure at home", "Did you record your blood pressure at home", "Did you use pill box at home", and "Did you use wearables".

Table 4. Self-evaluation of disease management by age.

Characteristics	Total (N=231), n (%)	Age stratification in years, n (%)				Chi-square test	
		18-49 (n=53)	50-64 (n=66)	65-74 (n=54)	≥75 (n=58)	χ^2 (df^a)	*P* value
Can control your health condition, n (%)						5.2 (*12*)	.95
Disagree Strongly	2 (0.9)	1 (2.0)	0 (0.0)	0 (0.0)	1 (1.7)		
Disagree	27 (12.1)	5 (9.8)	8 (12.7)	6 (11.8)	8 (13.8)		
Neutral	48 (21.5)	11 (21.6)	13 (20.6)	11 (21.6)	13 (22.4)		
Agree	122 (54.7)	28 (54.9)	33 (52.4)	28 (54.9)	33 (56.9)		
Agree Strongly	24 (10.8)	6 (11.8)	9 (14.3)	6 (11.8)	3 (5.2)		
Can manage your health behaviors, n (%)						9.4 (*12*)	.67
Disagree Strongly	3 (1.4)	1 (2.0)	1 (1.6)	0 (0.0)	1 (1.8)		
Disagree	28 (12.6)	10 (19.6)	6 (9.4)	4 (8.0)	8 (14.0)		
Neutral	41 (18.5)	7 (13.7)	13 (20.3)	8 (16.0)	13 (22.8)		
Agree	121 (54.5)	28 (54.9)	32 (50.0)	32 (64.0)	29 (50.9)		
Agree Strongly	29 (13.1)	5 (9.8)	12 (18.8)	6 (12.0)	6 (10.5)		
Can care for yourself at home, n (%)						13.2 (*12*)	.35
Disagree Strongly	2 (0.9)	1 (2.0)	0 (0.0)	0 (0.0)	1 (1.7)		
Disagree	12 (5.4)	3 (5.9)	2 (3.2)	2 (4.1)	5 (8.6)		
Neutral	32 (14.5)	8 (15.7)	7 (11.1)	4 (8.2)	13 (22.4)		
Agree	141 (63.8)	31 (60.8)	43 (68.3)	32 (65.3)	35 (60.3)		
Agree Strongly	34 (15.4)	8 (15.7)	11 (17.5)	11 (22.4)	4 (6.9)		
Have enough health knowledge to be able to care for yourself, n (%)						19.4 (*12*)	.08
Disagree Strongly	5 (2.3)	1 (2.0)	0 (0.0)	0 (0.0)	4 (7.0)		
Disagree	33 (14.9)	10 (19.6)	7 (11.3)	7 (13.7)	9 (15.8)		
Neutral	57 (25.8)	15 (29.4)	18 (29.0)	8 (15.7)	16 (28.1)		
Agree	105 (47.5)	24 (47.1)	28 (45.2)	30 (58.8)	23 (40.4)		
Agree Strongly	21 (9.5)	1 (2.0)	9 (14.5)	6 (11.8)	5 (8.8)		
Confidence in self-management (score)[b], mean (SD)	14.61 (2.90)	14.25 (3.08)	15.02 (2.60)	15.22 (2.53)	13.96 (3.21)	0.3 (*3*)[c]	.60[c]

[a]*df*: degrees of freedom.

[b]Confidence in self-management is a combined result for the 4 questions in this part.

[c]The result was obtained using analysis of variance.

Table 5. Willingness of participants to use mobile health (mHealth) technologies to manage their cardiovascular diseases.

Characteristics	Total (N=231), n (%)	Willingness to use mHealth, n (%)		Chi-square test	
		Interest (n=157)	No interest (n=74)	χ^2 (*df*[a])	*P* value
Age in years, mean (SD)	61.07 (15.62)	58.29 (16.19)	66.99 (12.51)	16.7 (*3*)[b]	<.001[b]
Age stratification in years, n (%)				13.8 (*1*)	.003
18-49	53 (22.9)	45 (28.7)	8 (10.8)		
50-64	66 (28.6)	48 (30.6)	18 (24.3)		
65-74	54 (23.4)	30 (19.1)	24 (32.4)		
≥75	58 (25.1)	34 (21.7)	24 (32.4)		
Gender, n (%)				0.1 (*1*)	.73
Male	124 (53.7)	86 (54.8)	38 (51.4)		
Female	107 (46.3)	71 (45.2)	36 (48.6)		
Body mass index stratification in kg/m^2 , n (%)				0.00 (*1*)	.96
≤24	61 (27.2)	41 (26.8)	20 (28.2)		
>24	163 (72.8)	112 (73.2)	51 (71.8)		
Education level, n (%)				24.4 (*2*)	<.001
Primary or junior	67 (29.3)	32 (20.5)	35 (47.9)		
Senior	59 (25.8)	38 (24.4)	21 (28.8)		
University or postgraduate	103 (45.0)	86 (55.1)	17 (23.3)		
Employment status, n (%)				13.1 (*2*)	.001
Retired	147 (64.5)	90 (58.1)	57 (78.1)		
Part time	21 (9.2)	13 (8.4)	8 (11.0)		
Full time	60 (26.3)	52 (33.5)	8 (11.0)		
Home location, n (%)				0.00 (*1*)	>.99
Beijing	206 (89.6)	141 (89.8)	65 (89.0)		
Outside of Beijing	24 (10.4)	16 (10.2)	8 (11.0)		
Living status, n (%)				4.1 (*2*)	.13
Alone	14 (6.1)	10 (6.5)	4 (5.4)		
Only with spouse	157 (68.6)	112 (72.3)	45 (60.8)		
With children	58 (25.3)	33 (21.3)	25 (33.8)		
Enrollment location				11.9 (*1*)	<.001
Outpatient	87 (37.7)	71 (45.2)	16 (21.6)		
Inpatient	144 (62.3)	86 (54.8)	58 (78.4)		
Frequency of using the internet to find information about disease or treatment, n (%)				36.1 (*3*)	<.001
Never	100 (43.5)	48 (30.8)	52 (70.3)		
Rarely	27 (11.7)	18 (11.5)	9 (12.2)		
Sometimes	49 (21.3)	43 (27.6)	6 (8.1)		
Often	54 (23.5)	47 (30.1)	7 (9.5)		
Did you know the medications you took well, n (%)				0.09 (*1*)	.77
No	48 (21.1)	31 (20.3)	17 (23.0)		
Yes	179 (78.9)	122 (79.7)	57 (77.0)		

Characteristics	Total (N=231), n (%)	Willingness to use mHealth, n (%)		Chi-square test	
		Interest (n=157)	No interest (n=74)	χ^2 (*df*[a])	*P* value
Frequency of taking lab test, n (%)				6.1 (*4*)	.19
Never	26 (11.4)	18 (11.5)	8 (11.0)		
Once per year	63 (27.5)	44 (28.2)	19 (26.0)		
1-2 per half of year	90 (39.3)	60 (38.5)	30 (41.1)		
1-2 per month	45 (19.7)	33 (21.2)	12 (16.4)		
Weekly	5 (2.2)	1 (0.6)	4 (5.5)		
Frequency of visiting hospital (score), mean (SD)	5.58 (1.93)	5.60 (1.88)	5.52 (2.03)	0.09 (*1*)[b]	.77[b]
Self-management at home (score), mean (SD)	10.88 (4.17)	11.24 (4.30)	10.09 (3.80)	3.7 (*1*)[b]	.05[b]
Confidence in self-management (score), mean (SD)	14.61 (2.90)	14.80 (2.77)	14.22 (3.13)	1.9 (*1*)[b]	.17[b]
Internet technology skills (score), mean (SD)	15.57 (7.16)	17.88 (6.37)	10.68 (6.24)	63.2 (*1*)[b]	<.001[b]

[a]*df*: degrees of freedom.

[b]The result was obtained using analysis of variance.

Table 6. Multivariate logistic regression concerning the willingness to use mobile health (mHealth) solutions.

Independent variables	Willingness to use mHealth		*P* value
	Coefficient	Odds ratio (95% CI)	
Age	–0.02	0.98 (0.94-1.02)	.35
Gender	–0.16	0.85 (0.42-1.73)	.65
Education level	0.18	1.12 (0.81-1.77)	.38
Employment status	–0.40	0.67 (0.34-1.34)	.26
Enrollment location	–0.63	0.53 (0.24-1.20)	.13
Internet technology skills	0.11	1.12 (1.03-1.21)	.006
Frequency of using the internet to search for health-related information	0.25	1.29 (0.87-1.90)	.20
Self-management at home	0.02	1.02 (0.94-1.11)	.65

Discussion

Principal Findings

The main findings of this study are that older patients visit hospitals more frequently than younger patients, but their frequency of searching for health-related information through the internet is lower. Older cohorts use mobile or IT less frequently and show lower technology skills than younger cohorts. More than half of the patients in the sample aged >65 years are interested in using mHealth solutions to manage their disease. Most significantly, experience in mobile technology use, not age, is the main indicator of the willingness to use mHealth.

Older patients visiting hospital more frequently showed that they have a higher motivation to manage their disease than younger patients. However, regarding their ability to manage their disease, the elderly less frequently searched for health-related information on the internet. Similar to the findings of this study, previous studies have shown that men in the older cohort (age >65 years) used the internet less frequently than younger men and also felt less comfortable using the internet [13]; meanwhile, younger patients (age <44 years) reportedly have more positive attitudes toward using mHealth apps for medication adherence than older participants [14]. However, in our research, we found that approximately 30% of our respondents aged >65 years used the internet to obtain health-related information. If a better solution for actively sending useful information to patients is developed, this may help the elderly obtain information more conveniently and improve their self-management capabilities.

As mentioned, older cohorts used the mobile or IT less frequently and showed lower technology skills than younger cohorts. These are not surprising findings. Importantly, however, a large proportion of patients used a smartphone almost every day; even in patients aged >75 years, this proportion is nearly 40%. Considering the large proportion of elderly with CVD in the population, there is huge potential for the use of smartphone apps in the self-management of CVD.

Previous studies showed that educated and younger individuals are more likely to download or use mHealth tools [12,13]. In

our study, we found that the difference between older and younger patients regarding the willingness to use mHealth is based on their mobile technology proficiency, not their age. More than half of the patients aged >65 years showed an interest in using an mHealth solution. Meanwhile, patients who were unwilling to use an mHealth solution reported that this mainly related to difficulties in using IT. This is consistent with the results of previous studies that older individuals are interested in mHealth but have concerns regarding their ability to use the devices [15]. Another study also indicated that clinical staff worried that mHealth devices might be too complicated for patients to use [16]. Therefore, it is essential to make mHealth tools easier-to-use and train elderly patients in technology-related skills, which might promote their willingness to use mHealth solutions.

Limitations

The study has several limitations. First, enrolled participants were sourced from a single tertiary hospital in Beijing, which compromises the generalizability of our results. Second, the limited sample size might constrain the power of statistical analysis. Third, based on the cross-sectional design of the study, we can only conclude that there is a significant correlation between IT skills and the willingness to use mHealth solution and cannot draw any conclusions regarding causality. Fourth, this study is based on a self-report survey of questionnaires. Using a newly designed questionnaire, even if pretested on a sample of 10 patients, does not bear the same weight as a validated questionnaire in prior studies. Thus, further validation of the instrument may be needed. Another limitation to note is that our items focused on the individuals' "willingness" to use mHealth, so we were examining their intent to use the technology rather than their actual use of it.

Conclusions

Although age is associated with the use of mobile or internet devices in the univariate analyses in this study, the sole indicator that we identified for mHealth use for self-management was the participants' IT skills. Education on the skills of using mobile devices and the invention of easy-to-use software might improve the acceptance of mHealth solutions among older patient populations.

Acknowledgments

This study was supported by the Chinese National Science and Technology Major Project for Major new drug creation (Project No. 2017ZX09304028-011), as well as a research fund from Philips. We thank Lifen Yu, Hongyu Chen, Jin Fang, Yuan Ma, Xiuying Wang, Juan Liu, and Xiaomeng Liu for administering the patients' questionnaire.

Authors' Contributions

JJ and JZ designed the questionnaire, oversaw the analysis, and wrote the manuscript. YMZ was the lead nurse who was in-charge of the study implementation. YJZ performed the statistical analysis. YXL helped with the designing of the questionnaire and the writing of the manuscript. YH was the Principal Investigator of the study who was responsible for all the results of the study, as well as the review and approval of the manuscript.

Conflicts of Interest

JZ and YJZ are employees of Philips Research. JJ, YMZ, YXL, and YH have no conflicts to declare.

Multimedia Appendix 1

Survey questionnaire.

[PDF File (Adobe PDF File), 51KB - mhealth_v7i2e10117_app1.pdf]

References

1. Bansilal S, Castellano JM, Fuster V. Global burden of CVD: focus on secondary prevention of cardiovascular disease. Int J Cardiol 2015 Dec;201 Suppl 1:S1-S7 [FREE Full text] [doi: 10.1016/S0167-5273(15)31026-3] [Medline: 26747389]
2. Chen W, Gao R, Liu L, Zhu M, Wang W, Wang Y, et al. Summary of Chinese cardiovascular disease report 2016. Chinese Circulation Journal 2017;32(6):521-530. [doi: 10.3969/j.issn.1000-3614.2017.06.001]
3. Naderi SH, Bestwick JP, Wald DS. Adherence to drugs that prevent cardiovascular disease: meta-analysis on 376,162 patients. Am J Med 2012 Sep;125(9):882-7.e1. [doi: 10.1016/j.amjmed.2011.12.013] [Medline: 22748400]
4. World Health Organization. Adherence to long-term therapies: evidence for action URL: http://www.who.int/chp/knowledge/publications/adherence_report/en/, [accessed 2018-11-20] [WebCite Cache ID 743vSVWgM]
5. Widmer RJ, Collins NM, Collins CS, West CP, Lerman LO, Lerman A. Digital health interventions for the prevention of cardiovascular disease: a systematic review and meta-analysis. Mayo Clin Proc 2015 Apr;90(4):469-480 [FREE Full text] [doi: 10.1016/j.mayocp.2014.12.026] [Medline: 25841251]
6. Lorig KR, Ritter PL, Laurent DD, Plant K. Internet-based chronic disease self-management: a randomized trial. Med Care 2006 Nov;44(11):964-971. [doi: 10.1097/01.mlr.0000233678.80203.c1] [Medline: 17063127]

7. Quilici J, Fugon L, Beguin S, Morange PE, Bonnet J, Alessi M, et al. Effect of motivational mobile phone short message service on aspirin adherence after coronary stenting for acute coronary syndrome. Int J Cardiol 2013 Sep 20;168(1):568-569. [doi: 10.1016/j.ijcard.2013.01.252] [Medline: 23462636]
8. Khonsari S, Subramanian P, Chinna K, Latif LA, Ling LW, Gholami O. Effect of a reminder system using an automated short message service on medication adherence following acute coronary syndrome. Eur J Cardiovasc Nurs 2015 Apr;14(2):170-179. [doi: 10.1177/1474515114521910] [Medline: 24491349]
9. Blasco A, Carmona M, Fernández-Lozano I, Salvador CH, Pascual M, Sagredo PG, et al. Evaluation of a telemedicine service for the secondary prevention of coronary artery disease. J Cardiopulm Rehabil Prev 2012;32(1):25-31. [doi: 10.1097/HCR.0b013e3182343aa7] [Medline: 22113368]
10. Wijsman LW, Richard E, Cachucho R, de Craen AJ, Jongstra S, Mooijaart SP. Evaluation of the Use of Home Blood Pressure Measurement Using Mobile Phone-Assisted Technology: The iVitality Proof-of-Principle Study. JMIR Mhealth Uhealth 2016 Jun 13;4(2):e67 [FREE Full text] [doi: 10.2196/mhealth.5485] [Medline: 27296628]
11. 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 Dec 22;18(6):e116 [FREE Full text] [doi: 10.2196/jmir.5429] [Medline: 27334418]
12. Sun L, Wang Y, Greene B, Xiao Q, Jiao C, Ji M, et al. Facilitators and barriers to using physical activity smartphone apps among Chinese patients with chronic diseases. BMC Med Inform Decis Mak 2017 Apr 19;17(1):44 [FREE Full text] [doi: 10.1186/s12911-017-0446-0] [Medline: 28420355]
13. Rising CJ, Bol N, Kreps GL. Age-Related Use and Perceptions of eHealth in Men With Prostate Cancer: A Web-Based Survey. JMIR Cancer 2015 Jun 25;1(1):e6 [FREE Full text] [doi: 10.2196/cancer.4178] [Medline: 28410165]
14. Davies MJ, Kotadia A, Mughal H, Hannan A, Alqarni H. The attitudes of pharmacists, students and the general public on mHealth applications for medication adherence. Pharm Pract (Granada) 2015;13(4):644 [FREE Full text] [doi: 10.18549/PharmPract.2015.04.644] [Medline: 26759619]
15. Lee J, Nguyen AL, Berg J, Amin A, Bachman M, Guo Y, et al. Attitudes and preferences on the use of mobile health technology and health games for self-management: interviews with older adults on anticoagulation therapy. JMIR Mhealth Uhealth 2014 Jul 23;2(3):e32 [FREE Full text] [doi: 10.2196/mhealth.3196] [Medline: 25098413]
16. Illiger K, Hupka M, von Jan U, Wichelhaus D, Albrecht U. Mobile technologies: expectancy, usage, and acceptance of clinical staff and patients at a university medical center. JMIR Mhealth Uhealth 2014 Oct 21;2(4):e42 [FREE Full text] [doi: 10.2196/mhealth.3799] [Medline: 25338094]

Abbreviations

BP: blood pressure
CVD: cardiovascular disease
IT: internet technology
mHealth: mobile health
PC: personal computer
PKU1: Peking University First Hospital

Edited by G Eysenbach; submitted 01.03.18; peer-reviewed by K Blondon, T Powell-Wiley; comments to author 05.08.18; revised version received 29.09.18; accepted 26.10.18; published 04.02.19.

Please cite as:
Jiang J, Zhu Q, Zheng Y, Zhu Y, Li Y, Huo Y
Perceptions and Acceptance of mHealth in Patients With Cardiovascular Diseases: A Cross-Sectional Study
JMIR Mhealth Uhealth 2019;7(2):e10117
URL: https://mhealth.jmir.org/2019/2/e10117/
doi:10.2196/10117
PMID:30714942

Original Paper

Quality Principles of App Description Texts and Their Significance in Deciding to Use Health Apps as Assessed by Medical Students: Survey Study

Urs-Vito Albrecht[1*], MD, MPH; Christin Malinka[1*]; Sarah Long[2*]; Tobias Raupach[2,3*], MD, MME; Gerd Hasenfuß[2*], MD; Ute von Jan[1*], Dr rer biol hum

[1]Peter L Reichertz Institute for Medical Informatics, Hannover Medical School, Hannover, Germany

[2]Department of Cardiology and Pneumology, University Medical Center Göttingen, Göttingen, Germany

[3]Division of Medical Education Research and Curriculum Development, University Medical Center Göttingen, Göttingen, Germany

[*]all authors contributed equally

Corresponding Author:
Urs-Vito Albrecht, MD, MPH
Peter L Reichertz Institute for Medical Informatics
Hannover Medical School
Carl-Neuberg-Straße 1
Hannover, 30625
Germany
Phone: 49 511 532 ext 3508
Fax: 49 511 532 2517
Email: albrecht.urs-vito@mh-hannover.de

Abstract

Background: Currently, there are no binding requirements for manufacturers prescribing which information must be included in the app descriptions of health apps.

Objective: The aim of this study was to investigate how medical students perceive a selection of quality principles, intended for usage decisions in the app context, and establish whether the information presented in a sample of app descriptions is perceived as sufficient for facilitating an informed usage decision.

Methods: A total of 123 students (mean age 24.2 years, SD 3.4) participating in a 6-week teaching module covering cardiology and pulmonology at the University of Göttingen (original enrollment 152 students, response rate 80.9%) were included. Students were asked to read 3 store description texts of cardiological or pneumological apps and initially assess whether the descriptions sufficed for a usage decision. Subsequently, they were queried on their perception of the relevance of 9 predefined quality principles, formulated for usage decisions. An appraisal of whether the app description texts contained sufficient information to satisfy these quality principles followed. By means of 20 guiding questions, participants were then asked to identify relevant information (or a lack thereof) within the descriptions. A reassessment of whether the description texts sufficed for making a usage decision ensued. A total of 343 complete datasets were obtained.

Results: A majority of the quality principles were described as “very important” and “important” for making a usage decision. When accessed via the predefined principles, students felt unable to identify sufficient information within the app descriptions in 68.81% (2124/3087) of cases. Notably, information regarding undesired effects (91.8%, 315/343), ethical soundness (90.1%, 309/343), measures taken to avert risks (89.2%, 306/343), conflicts of interest (88.3%, 303/343), and the location of data storage (87.8%, 301/343) was lacking. Following participants’ engagement with the quality principles, statistically significant changes in their assessment of whether the app descriptions sufficed for a usage decision can be seen—McNemar-Bowker test (3)=45.803919, P<.001, Cohen g=.295. In 34.1% (117/343) cases, the assessment was revised. About 3 quarters of changed assessments were seen more critically (76.9%, 90/117). Although, initially, 70% (240/343) had been considered “sufficient,” this rate was reduced to 54.2% (186/343) in the second assessment.

Conclusions: In a considerable number of app descriptions, participants were unable to locate the information necessary for making an informed usage decision. Participants’ sensitization to the quality principles led to changes in their assessment of app descriptions as a tool for usage decisions. Better transparency in app descriptions released by manufacturers and the exposure of users to quality principles could collectively form the basis for well-founded usage decisions.

(JMIR Mhealth Uhealth 2019;7(2):e13375) doi:10.2196/13375

KEYWORDS

mobile health; evaluation studies; mobile applications; quality criteria; usage decision

Introduction

Background

The market for health apps, that is, health-related apps running on mobile devices such as smartphones and tablet computers, is highly liberal and poorly regulated. This not only facilitates the creation of software, resulting in a large supply but also immensely influences user access and app usage. In this climate, we observe a flood of health apps, market dynamics typical for apps, and an associated lack of commitment to quality control [1]. From a government perspective, comprehensive (cross-border) monitoring of the market proves almost impossible [2,3]. Other entities (existing and emerging private and scientific testing or certification initiatives) [4-13] have yet to prove their efficiency and suitability for mapping the market [14]. The ultimate responsibility for deciding to utilize mobile apps rests with the users and cannot be transferred. In the context of health, this has even further-reaching implications than in other areas. Here, apps are used by laypersons as well as medical professionals in a highly sensitive environment. Apps, and the technology used to run them, are designed to be fully integrated into the user's everyday life. This aspect offers the greatest possible user comfort in both private and professional settings. Despite this unique advantage, it is important to recognize and respect certain legal boundaries, particularly addressing laws concerning medical practitioners [15]. These boundaries exist to protect both doctors and their patients and apply to using or recommending apps. In Germany, for example, laws cover confidentiality, advertising regulation, and the patient's freedom of choice concerning methods in diagnostics and therapy, given that these are appropriate and correspond to the current state of technological and scientific progress. These factors must be guaranteed by the medical staff as guarantors for their patients [16]. If applicable, rules are not followed, leading to damage infliction, and this is facilitated by a recommended or utilized app, medical staff involved can be held liable [16,17]. Consequently, doctors and other health professionals must (ethically and legally) inform themselves, undertaking a case-by-case risk-benefit assessment before recommending, or themselves deciding to use health-related apps. At the outset, similar to users with other backgrounds, medical professionals will likely—at least initially—rely on App Store description texts when selecting an app. Other information or test results and quality seals and the like are not often readily and reliably available [14] without (greater and time-consuming) research effort, or their reliability may be questionable because of various reasons. For this to be effective, it is imperative that manufacturers provide transparent information about their apps. Such transparency can serve as a reasonable basis for usage decisions. Thus, high-quality and trustworthy software has a better chance of asserting itself, and the self-regulatory capacity of the market can be supported [18].

Ideally, decisions for or against the use of an app are made by the interested parties who know their individual requirements best [19] and base their decisions on comprehensive information from multiple sources. A wide variety of tools and guidelines have been and are being developed on the basis of this principle [6,20-32], all of which share the common goal of supporting users in the decision process. In particular, there is a focus on requirements in the precarious context of health and medicine [7,33-40], taking into account both possible benefits and potential risks [41]. Many of these, for example, are published in the form of checklists that users may apply to the apps they are interested in [8,39], usually after installing them. However, it is currently almost impossible to estimate the extent to which the information available in the stores (in the form of app descriptions) can be used to adequately assess the suitability of an app before use. Existing studies, which also investigate the role of app descriptions, tend to focus on facets other than usage decisions, such as aspects related to marketing (and thus turnover-relevant aspects), rather than attempt to examine the quality of the content in serving its purpose [42]. With regard to app security, store description texts are used by researchers to compare the actual behavior of apps, for example, in the context of data transfers or potentially harmful functions (integration of advertising networks, etc), with the information contained in the descriptions [43,44].

Objectives

Supplementing gaps in existing research, in this study, the following questions were investigated: (1) which quality principles students consider fundamentally relevant for making a usage decision? (2) Whether or not the information in the submitted app descriptions is perceived as sufficient for a usage decision, (3) whether or not quality aspects can be identified within the description texts using key questions, and (4) whether or not exposure to the quality principles provokes a change in the students' assessment.

Methods

Setting

The study took place in the autumn of 2018 as part of a 6-week teaching module in the clinical phase of the undergraduate medical education program at the University of Göttingen, Germany. Within this module, a 6-hour practical training module was introduced, in which fourth-year students had the opportunity to explore health-related apps. The students had the opportunity to volunteer their data for this study. Nonparticipation would not have had any effect on the successful completion of the course. The students were informed in advance and were asked for their consent. The study was approved by the local Ethics Committee (application number 18/9/18), and all participants provided written consent.

For the purpose of this study, the Web-based survey system EvaSys (version 7.1, Electric Paper Evaluationssysteme GmbH,

Lüneburg, Germany) was used; the surveys were pseudonymized. In the first step, students were asked to provide demographic information. Each participant was then randomly assigned 3 app descriptions from a pool of health apps from the fields of cardiology and pulmonology, which were compiled by applying the keyword-based Semiautomated Retrospective App Store Analysis (SARASA) filtering processes to a readout of apps listed in the "Medical" category of Apple's App Store in August 2018 [45]. A wide range of apps for both patients and medical professionals was selected for the study. Examples of these include reference and learning apps as well as health diaries, treatment plans, and calculators. During the seminar, each student independently examined the app descriptions assigned to him or her in a multistep process.

After having provided basic demographic information, the students were asked for their initial assessment (not yet influenced by discussions, explanations, or having explored the quality criteria) of whether the app descriptions provided sufficient information for a decision on use ("The app description is sufficient for me to make a decision on use," "The app description is not sufficient for me to make a decision on use," or "I don't know"; see question block Q1, Figure 1). Immediately after this evaluation, the students were asked to express their—still uninfluenced—opinion on the importance of 9 quality principles for their usage decision—see definitions in subsection "Quality Principles and Operationalization" below, predominantly based on International Organization for Standardization/International Electrotechnical Commission (ISO/IEC) 25010 [46], question Block Q2, Figure 1, stratified by "very important," "important," "part/part," "less important," "unimportant," and "do not know."

Again, without explanations, the students were then requested to indicate whether the app descriptions provided allowed an assessment of the individual quality principles (see Q3, Figure 1). Subsequently, they were presented with 20 questions to be answered with "yes," "no," or "do not know" on the actual content of the app descriptions (eg, information on the purpose of the app, fields of application, target groups of the apps, or the respective providers; see section Quality Principles and Operationalization). These were based on the items presented in other studies [35,47] (see Q4, Figure 1) and covered aspects related to the 9 quality principles in the hope that working with these questions would increase participants' awareness of aspects related to these quality principles. Unfortunately, the question "Is there information about the aptitude (qualification) of the authors/developers of the app?" was not incorporated in the electronic survey, but for the sake of completeness, it is still listed in the section Quality Principles and Operationalization. Finally, the students were again prompted to assess whether, in their opinion, the app descriptions contained sufficient information for a decision on use (see Q5, Figure 1).

Figure 1. Study design and procedure.

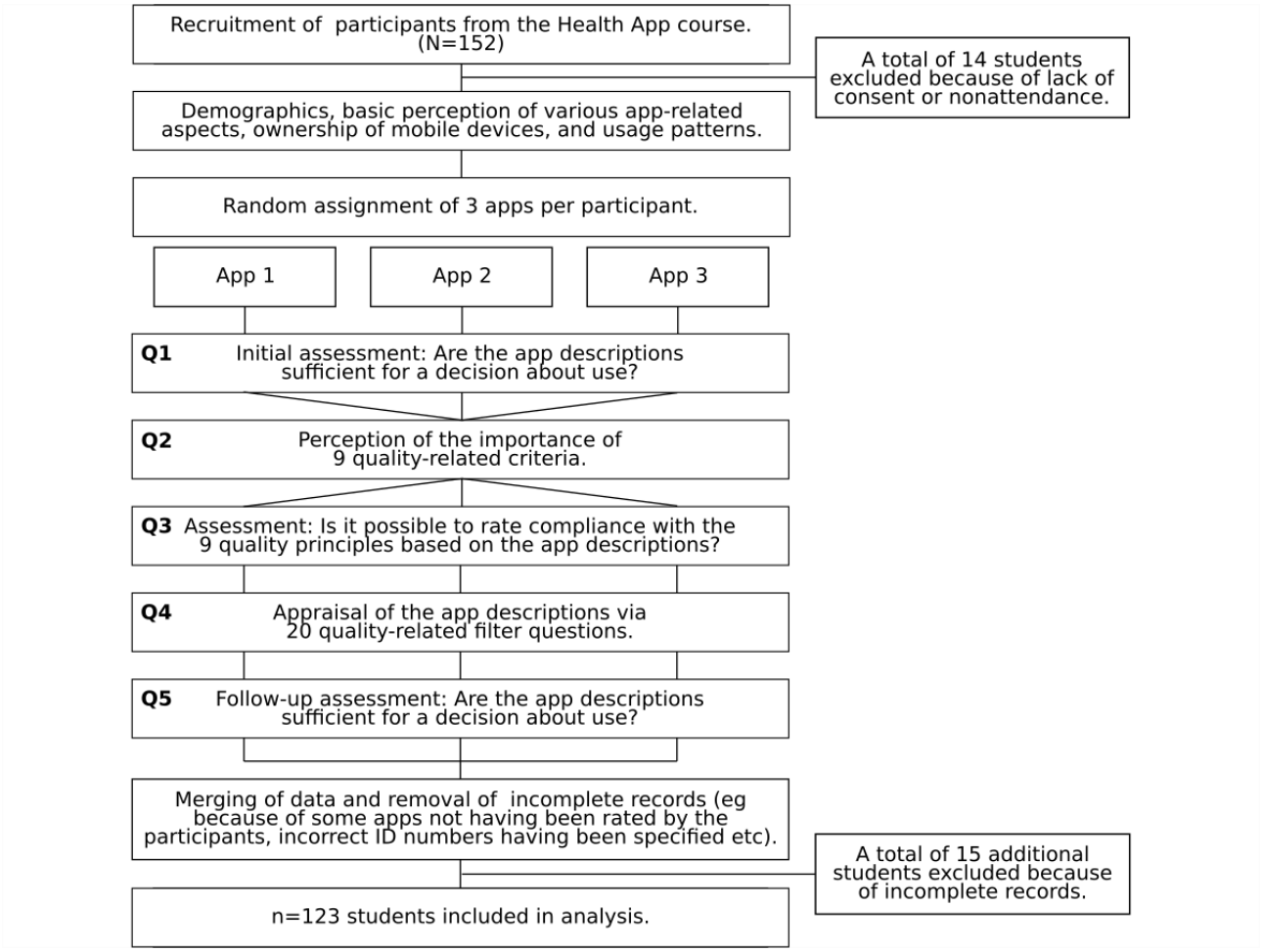


Table 1. Demographics for the participants.

Characteristics	Male (n=42)	Female (n=80)	Unspecified (n=1)	Total (n=123)
Age, mean (SD)	24.8 (3.2)	23.9 (3.4)	23 (—[a])	24.2 (3.4)
Years of study, mean (SD)	4.1 (0.4)	4.1 (0.4)	4 (—)	4.1 (0.4)
Mobile operating system (corresponding number of participants, n)				
iOS (tablet, smartphone, or iPod)	24	47	1	72
Android (tablet or smartphone)	21	36	0	57
Other (tablet or smartphone)	2	2	0	4
Several different OS (accumulated)	4	5	0	9
Use of apps in general (corresponding number of participants, n)				
No	4	2	1	7
Yes	38	78	0	117
Ratings submitted (total, N=343)	115	225	3	343
Ratings provided (per participant, mean [SD])	2.7 (0.6)	2.8 (0.4)	3 (—)	2.8 (0.5)
Apps assigned (n)	90	132	3	143

[a]Not applicable.

Study Population

Of a gross total of 152 medical students who had registered for the class, those who did not attend the course despite registration or did not give their consent (n=14) were not included in the study; thus, 138 participants in their fourth academic year remained (Table 1). The evaluation only included complete datasets. Participants' responses to the various parts of the survey (Figure 1) could be linked via their individual identification number and the name of the respective app. By answering all questionnaires for at least 1 of the 3 apps (selected from a set of 143 individual apps) assigned to them, participants qualified their data for inclusion in the analysis. Thus, a total of 343 app-related assessments (82.9% of 414 expected, dropout: 15 students) from 123 students (89.1%, 123/138) remained (42 males, 80 females, gender not indicated in 1 case, overall mean age 24.2 years, SD 3.4). Of these 123 students, not all completed all question blocks for all 3 of their assigned apps (or it was impossible to match these to a specific app or student, eg, because of errors typing identification numbers), resulting in an average of only 2.8 app evaluations per participant available for evaluation (SD 0.5).

Only 7 participants stated that they do not use any apps. Mainly iOS-based smartphones and tablets were in use (72/123 respectively 58.5% total, males: 24/42 respectively. 57.1%, females: 47/80 respectively. 58.8%), followed by Android-based mobile devices (57/123 mentions respectively. 46.3% total, males: 21/42 respectively. 50%, females: 36/80 respectively. 45%). With the exception of gender, the study population included in the evaluation is homogeneous. Approximately twice as many females were included, as opposed to males. This reflects the larger proportion of female students documented undertaking a medical degree at the University of Göttingen and at German universities in general [48]. A correlation between the evaluations of the app descriptions and participants' gender could not be shown—Pearson Chi-square χ^2_4=8.4, P=.77, n=123).

Quality Principles and Operationalization

The study was focused on 9 quality principles (see Table 2), predominantly modeled on ISO/IEC 25010 [46] for health software, that are currently being discussed in the context of coordinating interdisciplinary quality criteria in Germany, (eg, as compiled by Albrecht [49,50]). Although it could be argued that other criteria could also have been included in this study, we explicitly chose not to do so, as these proved to be too specific to be appropriate for the assessments we had planned. For example, although ISO/IEC 25023 [51] provides a "basic set of quality measures" for various quality aspects and "an explanation of how to apply software product and system quality measures," we chose not to include it, as the purpose of the part of the study presented here was not to measure app quality but rather to analyze participants' subjective perceptions of quality, via app descriptions.

Additional sources were used to support the compilation of the 9 principles; however, no single source was fully adopted, for reasons also noted by Nouri et al [40] in their study on quality principles in the app context.

There is hardly any agreement among different working groups or authors as to which quality categories and characteristics can be usefully applied to an assessment or which characteristics can be assigned to which quality categories and how it should be determined whether an app offers the desired characteristics. This can be illustrated exemplarily by the aspect of usability [3], but it can also be established in principle for all other areas relevant in the quality context. Differences exist, among other things, with regard to the assignment of different characteristics to the usability principle, but this may also be because of different objectives or target groups such as consumers or the restriction to selected application areas of the respective approaches. Objective as well as more subjective characteristics

are often included. Although Zapata et al [52], for example, included rather subjectively assessable aspects such as attractiveness, learnability, usability, and comprehensibility in their empirical analysis on usability, other authors approach the concept of usability from a technical and more objective point of view. Brown et al [53] did this by subdividing the usability of the "Health IT Usability Evaluation Model" presented in another study [54] into more detailed parts such as avoidance, completeness, memory, need for information, flexibility/adaptability, learnability, speed of performance, and competence. Nevertheless, in some cases, the various characteristics can be difficult to assess without in-depth technical and/or content-related knowledge or in some cases, time-consuming analyses. It is for this reason that, in our operationalization of the 9 quality principles, we tried to keep the questions the students were confronted with simple to comprehend and easy to answer, still addressing the quality principles without going into great technical detail. The operationalization itself (Table 3) was done by comparing the quality principles with existing question lists for self-assessments of health apps from the preliminary study conducted both internally and also in accordance with several other German initiatives [19,35,37].

Table 2. The 9 quality principles (predominantly based on ISO 25010, with supporting sources also listed).

Quality principle	Description	(Sub) Section of ISO/IEC[a] 25010 [46]	Supporting sources
Practicality	High-quality software must be flexible enough to be used for the intended purpose and, if possible, beyond it, to cover the widest possible range of use and application contexts.	4.1.3 satisfaction; 4.1.5 context coverage; 4.2.1 functional suitability; 4.4.11 stated purpose.	[35,37,38,55-64]
Risk adequacy	It must be possible to use software in a risk-appropriate manner without exposing the user or his or her environment to unreasonable health, social, or economic risks.	4.1.4 freedom from risk (economic, health and safety, and environmental risk mitigation).	[55,56,65,66]
Ethical soundness	Development, provision, operation, and use must be ethically innocuous to prevent discrimination and stigmatization and to provide fair access.	4.2.4.6 accessibility	[37,38,56,62,67,68]
Legal conformity	The legal conformity (eg, with regard to medical device law, professional codes of conduct, data protection laws, laws on the advertising of therapeutic products) for development, provision, operation, and use must be guaranteed for the protection of all parties involved (eg, providers, store operators, and users).	—[b]	[9,19,35,37,38,55,56,61,69-73]
Content validity	The content presented and used must be valid and trustworthy.	—[c]	[8,9,19,38,40,56,59-62,73-76]
Technical adequacy	Development, operation, and use need to be appropriately adapted to the capabilities of the technology and the current state-of-the-art to ensure sustainability in terms of maintainability, portability, interoperability, and compatibility.	4.2.3 compatibility; 4.2.5 reliability; 4.2.7 maintainability, and 4.2.8 portability	[40,66,73,77]
Usability	The software must have a high degree of usability appropriate for its target groups, that is, it must be user-friendly and easy to use, taking into account the relevant circumstances and conditions. This can facilitate fair and sustainable use that is also convenient and contributes to user satisfaction.	4.1.3.4 comfort; 4.2.4 usability; 4.2.8.1 adaptability.	[40,73,78]
Resource efficiency	Elements for resource-efficient operation and use should be taken into account during development.	4.1.2 efficiency and 4.2.2 performance efficiency (including time behavior, resource utilization, and capacity)	[65,73]
Transparency	Full transparency regarding the aforementioned criteria serves as a basis for software evaluations as well as for individual and collective usage decisions.	—[c]	[9,19,35,66,73]

[a]ISO/IEC: International Organization for Standardization/International Electrotechnical Commission.

[b]No longer covered in ISO/IEC 25010, but was part of ISO/IEC 9126-1:2001 [69], which 25010 revises.

[c]Not covered in ISO/IEC 25010.

Table 3. Operationalized quality aspects.

Question Number	Question	Affected quality principles
1	Has the purpose of the app been specified in the description text?	Practicality and transparency
2	Is there a description of the functions offered by the app (functionality)?	Practicality, usability, and transparency
3	Is there a description of the context and environment in which the app is to be used (application field)?	Practicality, usability, and transparency
4	Is the target group of the app (eg, doctors, students, and patients, or differently defined groups) described?	Practicality, usability, and transparency
5	Is there any indication as to whether feedback from the relevant user groups was incorporated into the design, development, or testing of the app?	Usability and transparency
6	Are there any details on where and how the app should not be used, where its limits lie (restrictions and limitations)?	Practicality, risk adequacy, and transparency
7	Are undesired effects that have already occurred been mentioned?	Risk adequacy and transparency
8	Is there a description of potential or actual risks (health, economic, and social) to which the user may be exposed when using the app?	Risk adequacy and transparency
9	Are precautions taken to avoid the above risks described?	Risk adequacy and transparency
10	Are authors or developers of the app named?	Content validity and transparency
11[a]	Is there information about the aptitude (qualification) of the authors or developers of the app?	Content validity and transparency
12	Are sources used for the app (eg, literature) named?	Content validity and transparency
13	Is it specified whether the app has been awarded certificates, quality seals or something similar by third parties?	Technical adequacy and content validity
14	Are details given with respect to quality assurance during development?	Technical adequacy and transparency
15	Is information given on whether the app is a medical device (keyword: CE label[b])?	Legal conformity, technical adequacy, risk adequacy, and transparency
16	Is there a description of how the app is financed or who is funding it?	Content validity and transparency
17	Are conflicts of interest named (eg, involvement of an author in the app company)?	Content validity and transparency
18	Are details provided on users' data protection rights in connection with the collection, storage, and deletion of data (eg, right to information, right of modification, right of revocation, and periods for deletion)?	Legal conformity, risk adequacy, and transparency
19	Are there any indications as to who the beneficiary(s) of the data is or are?	Legal conformity, risk adequacy, and transparency
20	Is the location where data are being stored (eg, in which country) named?	Legal conformity, risk adequacy, and transparency
21	Are there any indications of ethical innocuousness (eg, ethics vote for research apps)?	Ethical soundness, and transparency

[a]Unfortunately, question 11 was not included in the Web-based survey.

[b]Conformité Européenne. A CE labels indicates that a product sold within the European economic Area conforms to the required health, safety, and environmental protection standards.

Evaluation Strategy

A descriptive evaluation of the frequencies, mean values, and SDs was prepared.

The primary goal of the study was to detect a change in the assessment of sufficiency for usage decisions on the basis of app description texts. The hypothesis was tested that, after confrontation with the predefined quality principles, there would be no change in the students' assessment of the sufficiency of

app descriptions for the usage decision. Bowker test of symmetry (2-sided, alpha=.05, beta=.80) [79,80] was applied and for determining effect size, Cohen g [81,82] was calculated. The aforementioned symmetry test was chosen as it provides the opportunity to test multiple nominal characteristics in associated samples. In addition, in contrast to the McNemar test, the McNemar-Bowker test is able to consider more than 2 categories.

The following points acted as secondary aims within the study:

1. Assessment of the relevance of quality principles for the usage decision.
2. Evaluation of the sufficiency of the information provided in the app descriptions to assess compliance with the quality principles.
3. Frequency of mentioned aspects as identified by the key questions in the description texts.

IBM SPSS Statistics Subscription (Build 1.0.0.1118, IBM Corporation) and R (version 3.5.1, R Core Team) [83] were used for the evaluation.

Results

The 123 participating students regarded all 9 quality principles as "very important" or "important" (Table 4). In particular, they considered "content validity" (85.4%, 105/123) and "risk adequacy" (74%, 91/123), "legal conformity" (66.7%, 82/123) and "usability" (65%, 80/123) to be "very important." Furthermore, the principles of "ethical soundness" (55.3%, 68/123), "practicality" (50.4%, 62/123), and "transparency" (45.5%, 56/123) were regarded as "very important" but were weaker in terms of percentage for the decision on use. The quality principles "technical adequacy" (39%, 48/123) and "resource efficiency" (34.1%, 42/123) were classified as "important" (see Table 4).

After working with the quality principles, the students were asked if they were able to determine whether these principles were met on the basis of the app descriptions (Q3, Table 5). Affirmative answers to this question were given in 31.2% (943/3087) of the evaluations of the app descriptions, with "practicality" in 71.7% (246/343) and "usability" in 39.9% (137/343) assessed as fulfilled most frequently. The worst levels of fulfillment were found for "transparency" (16.9%, 58/343) and "resource efficiency" (19.8%, 68/343). In less than one third of the app descriptions, students were able to successfully determine compliance with the quality principles "content validity" (27.1%, 93/343), "ethical soundness" (26.8%, 92/343), and "legal conformity" (22.2%, 76/343; see Table 5).

On the basis of the total number of all individual answers, participating students were unable to identify the required information in the app descriptions in 70.4% (4831/6860) of the answers (see Q4, Table 6). In 5.9% (403/6860) of the answers, students were unsure as to whether the description texts contained suitable information ("do not know"). According to the students, the greatest deficits were the lack of information on "undesirable effects" (91.8%, 315/343), "ethical soundness" (90.1%, 309/343), "risk-avoidance" (89.2%, 306/343), "conflicts of interest" (88.3%, 303/343), and "naming the data storage location" (87.8%, 301/343). Sufficient information could be found via the filter questions on the "declaration of purpose" (93.6%, 321/343) and "description of functionalities" (86.9%, 298/343). In 76.7% (263/343) of the app descriptions, assessments of the field of application could be made. However, it should be noted that only 23.3% (1600/6860) of the answers given were positive (see Q4, Table 6), corresponding only to the presence of the information necessary to answer the question in the app description.

Table 4. Assessment of the relevance of the 9 quality principles (Q2) for one's own usage decision (for N=123 students).

Item	Very important, n (%)	Important, n (%)	Part/part, n (%)	Less important, n (%)	Unimportant, n (%)	Do not know, n (%)	No information, n (%)
Practicality	62 (50.4)	46 (37.4)	8 (6.5)	—[a]	2 (1.6)	5 (4.1)	—
Risk adequacy	91 (74.0)	11 (8.9)	9 (7.3)	1 (0.8)	2 (1.6)	7 (5.7)	2 (1.6)
Ethical soundness	68 (55.3)	37 (30.1)	11 (8.9)	3 (2.4)	2 (1.6)	1 (0.8)	1 (0.8)
Legal conformity	82 (66.7)	26 (21.1)	5 (4.1)	4 (3.3)	1 (0.8)	4 (3.3)	1 (0.8)
Content validity	105 (85.4)	12 (9.8)	2 (1.6)	1 (0.8)	1 (0.8)	2 (1.6)	—
Technical adequacy	3 (2.4)	48 (39.0)	47 (38.2)	3 (2.4)	—	3 (2.4)	3 (2.4)
Usability	80 (65.0)	34 (27.6)	8 (6.5)	—	—	1 (0.8)	—
Resource efficiency	36 (29.3)	42 (34.1)	24 (19.5)	10 (8.1)	3 (2.4)	8 (6.5)	—
Transparency	56 (45.5)	39 (31.7)	19 (15.4)	—	1 (0.8)	6 (4.9)	2 (1.6)

[a]No corresponding answer was given.

Table 5. Assessment as to whether compliance with the 9 quality principles could be determined on the basis of the available app descriptions (Q3, scale “yes,” “no,” and “do not know”), on the basis of N=343 assessments (3087 individual responses overall).

Item	Yes, n (%)	No, n (%)	Do not know, n (%)	No data, n (%)
Practicality	246 (71.7)	79 (23.0)	17 (5.0)	1 (0.3)
Risk adequacy	93 (27.1)	198 (57.7)	52 (15.2)	__[a]
Ethical soundness	92 (26.8)	211 (61.5)	40 (11.7)	__[a]
Legal conformity	76 (22.2)	231 (67.3)	36.0 (10.5)	__[a]
Content validity	93 (27.1)	210 (61.2)	37 (10.8)	3 (0.9)
Technical adequacy	100 (29.2)	199 (58.0)	41 (12.0)	3 (0.9)
Usability	137 (39.9)	179 (52.2)	25 (7.3)	2 (0.6)
Resource efficiency	68 (19.8)	205 (59.8)	69 (20.1)	1 (0.3)
Transparency	58 (16.9)	213 (62.1)	72 (21.0)	__[a]
Total number	963 (31.20)	1725 (55.88)	389 (12.60)	10 (0.32)

[a]Not applicable.

Table 6. Assessment of whether the 20 detailed questions could be answered on the basis of the available app descriptions (Q4, “yes”, “no”, “don't know”, based on N=343 evaluations with a total of 6860 individual answers).

Item	Yes, n (%)	No, n (%)	Do not know, n (%)	No data, n (%)
Indication of purpose	321 (93.6)	19 (5.5)	3 (0.9)	__[a]
Description of functionalities	298 (86.9)	38 (11.1)	7 (2.0)	__[a]
Information on the field of application	263 (76.7)	68 (19.8)	10 (2.9)	2 (0.6)
Information on the target group	233 (67.9)	96 (28.0)	13 (3.8)	1 (0.3)
Information on inclusion of feedback from the relevant user groups	40 (11.7)	273 (79.6)	28 (8.2)	2 (0.6)
Description of restrictions and limitations	43 (12.5)	284 (82.8)	15 (4.4)	1 (0.3)
Indication of undesired effects	8 (2.3)	315 (91.8)	18 (5.2)	2 (0.6)
Information on potential or actual risks	20 (5.8)	304 (88.6)	18 (5.2)	1 (0.3)
Information on the precautions taken to avoid the aforementioned risks	20 (5.8)	306 (89.2)	15 (4.4)	2 (0.6)
Authorship (authors or developers have been named)	67 (19.5)	249 (72.6)	25 (7.3)	2 (0.6)
Information on sources used	38 (11.1)	279 (81.3)	24 (7.0)	2 (0.6)
Information on certificates, quality seals, or something similar having been awarded	25 (7.3)	296 (86.3)	22 (6.4)	__[a]
Information on quality assured development	34 (9.9)	282 (82.2)	27 (7.9)	__[a]
Information on the medical device status	32 (9.3)	274 (79.9)	37 (10.8)	__[a]
Information on financing	45 (13.1)	280 (81.6)	16 (4.7)	2 (0.6)
Conflicts-of-interest-related information	10 (2.9)	303 (88.3)	27 (7.9)	3 (0.9)
Information about user privacy rights	41 (12.0)	277 (80.8)	23 (6.7)	2 (0.6)
Information on the beneficiary of the data	27 (7.9)	278 (81.0)	37 (10.8)	1 (0.3)
Specification of the data storage location	25 (7.3)	301 (87.8)	16 (4.7)	1 (0.3)
Information on ethical soundness	10 (2.9)	309 (90.1)	22 (6.4)	2 (0.6)
Total number of ratings	1600 (23.32)	4831 (70.42)	403 (5.87)	26 (0.38)

[a]Not applicable.

Initially, 70% (240/343) of the app descriptions were considered "sufficient" to make a decision on use. Following engagement with the quality principles, this rate was reduced to 54.2% (186/343). The proportion of app descriptions judged "insufficient" rose by an absolute value of 19.2% (66/343), from 22.7% (78/343) to 42% (144/343). The percentage of those who were undecided decreased from 7.3% (25/343) to 3.8% (13/343) and was thus almost halved (decline of 48%, 12/25; see Table 7). After the examination of quality aspects, significantly fewer assessments were considered "sufficient" than before—McNemar-Bowker Test (3)=45.803919, *P*<.001. The effect size according to Cohen was g=.295, which corresponded to a strong effect [81] (see Table 7). The calculated posthoc power was 0.99—Chi-square power calculation χ^2_4=0.3, *P*=.05, N=343.

Overall, 76 out of 123 students (61.8%) changed their opinion on the sufficiency of the app descriptions for a usage decision for at least 1 of the assigned apps. Of a total of 343 such assessments, 117 were revised (34.1%). A total of 90 of the 117 changes (76.9%) were corrected to a more critical assessment (changes to "insufficient" or "do not know"; see Table 8).

Table 7. Students' assessment as to whether the app description text is sufficient for the usage decision. Presentation of the contingency table (Q3 vs Q5) before and after the clarification of quality principles and the targeted search for these quality criteria (yes, no, and do not know) in 343 app evaluations from 123 students.

Before information and investigation	After information and investigation			
	"Insufficient"	"Do not know"	"Sufficient"	Total number
"Insufficient"	63	2	13	78
"Don't know"	11	2	12	25
"Sufficient"	70	9	161	240
Total	144	13	186	343

Table 8. Presentation of the directions of change in 117 out of 343 assessments of usage decisions based on information on quality principles and criteria by 76 (61.8%) of the 123 students.

Assessments	Changes in assessment, n (%)
From "do not know" to "sufficient"	12 (10.3)
From "do not know" to "insufficient"	11 (9.4)
From "sufficient" to "do not know"	9 (7.7)
From "insufficient" to "do not know"	2 (1.7)
From "sufficient" to "insufficient"	70 (59.8)
From "insufficient" to "sufficient"	13 (11.1)
Total	117 (100.0)

Discussion

Principal Findings

We conducted surveys before and after confrontation with quality principles and criteria. The students evaluated the same description in both surveys. Although we did not ask to what extent the students had previous knowledge on the subject or their assigned apps (and there were no indications for this in the free text comments they were allowed to make), on the basis of our design, we were nevertheless able to determine that, after having worked with quality principles, there were indeed changes in how the participants perceived the description texts with respect to whether these possibly suffice for initial usage decisions. We were also able to obtain insights into which elements can or cannot be commonly found in the descriptions.

The study showed that, following engagement with the 9 specified quality principles (Table 2), there was a statistically significant change in the students' assessment of the sufficiency of app descriptions for a decision on app use—McNemar-Bowker Test (3)=45.803919, *P*<.001, Cohen g=.295. In 34.1% (117/343) of the evaluations, the initial assessment was revised. Overall, more than 1 in 4 evaluations (or 3 in 4 changes of assessment) resulted in a more critical assessment. We assume that the following factors may have led to a sensitization, inciting further analytical thought when reassessing the initial question: First, the examination of app description quality by gauging the relevance of generic quality principles for the usage decision; second, the subsequent assessment of whether the description divulged the app's fulfillment of these principles; third, the search for specific information within the texts, guided by 20 filter questions. The students rated all quality principles as "very important" or "important" for their usage decisions. In particular, "content validity" (85.4%, 105/123) and "risk adequacy" (74%, 91/123) and "legal conformity" (66.7%, 82/123) and "usability" (65%, 80/123) were "very important." However, it was precisely these principles that the students were able to less identify with certainty in the app descriptions. It is for this reason that students were only able to assess the fulfillment of the quality criteria to a limited extent. The search for specific information in the app descriptions showed large deficits—for 16 of the 20

questions, more than 80% of the descriptions were found to contain insufficient information. In particular, statements on undesired effects (91.8%, 315/343), on the ethical harmlessness of the apps (90.1%, 309/343), on the measures taken to avoid risks (89.2%, 306/343), and regarding conflicts of interest (88.3%, 303/343) were lacking. An unspecified data storage location (87.8%, 301/343; Table 4) was also problematic.

The results allow the following conclusions to be drawn. First, when observing app descriptions, students were only able to identify a small amount of information on aspects relevant to the quality principles. This is in line with the work of other authors, in which the information content of store description texts was also evaluated as poor in terms of quality and content [84]. Second, it can be implicitly assumed that although awareness of quality principles exists, it is not generally transferred to descriptions of health apps. This is made apparent through the more critical assessment of the sufficiency of app descriptions after a sensitization to quality principles. Finally, it can be concluded that the abovementioned aspects represent essential elements for a well-founded user decision.

To form the basis for informed usage decisions, manufacturers need to provide relevant information on quality principles in an easy-to-understand manner, ideally following a universal, structured approach, easily comparable by interested parties [47]. The app description provides an ideal scope for this, as it is an obligatory requirement for all apps listed in stores on the major mobile platforms. In this study, we found that only a very small percentage of this information is made available. The specification of standardized information in the description [35] would help to solve this issue, especially if the users were to demand it. This can be achieved through the involvement of stakeholders, such as professional associations, industry associations, and consumer initiatives, that coordinate their activities across disciplines [49,50]. The message could be that manufacturers who do not include such content in the descriptions deny users the opportunity to make a well-founded decision on use. Recently, efforts have been made in various professional associations to consider compiling interdisciplinary quality criteria. Naturally, such processes are tedious because of the sheer quantity of opinions regarding the definition of the selection of criteria [50]. A process that could be concluded more quickly would be the agreement that transparency must be upheld on the part of app manufacturers and distributors.

Of course, transparency must also be appreciated and utilized by the user if a well-founded usage decision is to be made. To this end, users must become aware of their role and their individual responsibility in the (professional) use of this technology. The recognition of (professional) legal and ethical requirements of apps is not automatically conducted because of the general perception of smartphones and apps as “private matters.” In Weiser’s sense, mobile technology is already too “interwoven” with “our everyday life” [85] for it to be viewed in a differentiated way. However, the fact is that these technologies are used in professional contexts, even in health and medicine—with all their consequences. With small stakeholder campaigns and further training within the framework of the digitization debates, a great deal of sensitization could already be achieved, and a major contribution could be made in attaining the circumstances necessary for well-founded decisions on app use. Of all the solutions for evaluating apps, such as reviews, tests, certifications, and the preparation of scientific studies, app descriptions represent the first and fastest step taken by users.

Comparison With Other Approaches in the Quality Context

There are a number of helpful and validated tools available, aiming to support those interested in health-related apps and their quality [86]. Often in the form of a checklist, these tools address various user groups and application areas, for example, Mobile Application Rating Scale [7] and user version of the Mobile Application Rating Scale [39] as well as App Chronic Disease Checklist [8]. In addition to these tools, some third-party initiatives, such as national health bodies, assign quality seals to apps or compile lists of apps they have approved. The quality of such third-party evaluations is at times questionable. How well the quality assessment processes have been designed and implemented and the scope of the assessments that are performed (eg, assessments of whether the content is adequate vs also considering technical or security-related aspects) are critical aspects when making recommendations.

In terms of this study, it was not our aim to develop yet another assessment tool for determining whether an app is of high quality. Instead, we were interested in, first, whether potentially interested parties are aware of applicable quality criteria and are able to identify corresponding information in the app descriptions, second, whether for users who have previously been unfamiliar with such criteria, a familiarization can potentially lead to changes in how they assess quality aspects, on the basis of the app descriptions. In our analysis, we found strong indications for both of these aspects. We believe that this may facilitate future evaluations on the basis of the aforementioned quality assessment tools by enabling users to more easily apply these tools.

Limitations

App Selection

The inherent dependence of the quality of app selection on the quality of the search terms defined poses multiple limitations. While searching for suitable apps from the field of cardiology and pulmonology, it is possible that fitting search terms were not included or—especially with hits of partial terms—that some apps were incorrectly included. A complete (manual) screening of all apps available in the store categories “Medical” and “Health and Fitness” would not be possible because of the incredibly large volume of apps available. It is for this reason, despite limitations, that the keyword-based SARASA method [45] was used. Furthermore, it is possible that a sampling bias occurred during the selected search procedure in Apple's App Store. This is conceivable when considering the store’s category-based system, not recognizing apps falsely categorized by their manufacturers, and it may also be because of the limitation of the search to apps with German-language store descriptions, predetermined by the store front-end available for Germany. The situation may differ for App Stores available for other mobile platforms (eg, Android apps available from

Google's Play Store) or even for iOS-based apps from store front-ends in other countries or apps whose store descriptions are available in other languages, which should be taken into account in subsequent investigations.

In addition, the SARASA method led to a variable selection of apps that were probably not directly comparable because of their different application areas and target groups. Nevertheless, we believe that this variability was more a strength than a weakness of our evaluation, as we were not interested in the direct comparability of apps but rather in the evaluation of quality aspects in a typical setting. This is given as users are able to obtain apps using keyword-based searches in the store.

Study Population

It may also be argued that our participants' demographics are not fully representative of the German population, for example, with respect to their age, level of education, and smart device usage patterns, with almost 59% (72/123; Table 1) of the participants stating that they were using iOS-based devices versus only about 23% market share for such devices in the German population in December 2018 [87]. Despite these discrepancies, the study population reflects the often-mentioned greater popularity of the iOS platform among those working in the medical field [88], and thus the participants may well prove to be a representative sample, at least in comparison to their future colleagues. Platform-related effects on our results were probably negligible, as the students were requested to solely consider the provided store description texts, without platform specifics, and not the apps themselves. Moreover, it has been shown that there are only small differences among users of various mobile platforms, if sociodemographics are accounted for [89].

Another possible limitation regarding our choice of students as the study population may be the students' lack of experience in the medical field and their lack of exposure to the quality aspects investigated, potentially making it more difficult for them to assess the content of the app descriptions. Upon reflection, we believe this had little, if any, influence. As app descriptions are commonly written not to convey detailed, in-depth information, but rather to satisfy marketing requirements—after all, manufacturers hardly have a chance to restrict who has access to them—one would expect that only in rare circumstances would the information conveyed in descriptions require knowledge surpassing that of fourth year medical students. In addition, in a previous study [45], for a somewhat similar selection of apps, we applied automated algorithms for text complexity to the descriptions, with calculations based on sentence length, number of syllables, etc, to determine the level of education necessary for reading comprehension. In that case, for about 3 quarters of the apps, a level of high school education or less would have been sufficient for comprehension. We therefore believe that medical students, who are as far along in their studies as our fourth year participants, should have sufficient medical background and reading proficiency to perform basic checks of medically-oriented app descriptions. Moreover, an objection that students do not have the knowledge necessary for basic assessments of usability and information security can hardly be raised. For today's students, a majority of which have grown up with information technology and could therefore be considered "digital natives," at least a basic understanding of these aspects can be assumed. In any case, to be truly meaningful, expert-level assessments would require in-depth analyses of the apps themselves rather than an evaluation of store description texts.

Questionnaire Design

The questionnaires were pretested with 4 medical students from different semesters. It would have been sounder to test with a population comparable with the target group. Unfortunately, appropriate candidates could not have been recruited without provoking a bias (prospective course participants), which is why we refrained from doing so. The pretest was conducted without any evidence of comprehension problems when paraphrasing so that the authors saw no reason for any changes. Despite this, some of the questionnaires within the study were not fully completed. A dropout analysis was not carried out for reasons of capacity; however, it is planned for subsequent rounds.

Unfortunately, the filter question "Is there information about the aptitude (qualification) of the authors/developers of the app?" was not included in the Web-based survey, although this was planned. This will be done in a subsequent study, as determining whether the authors' and developers' qualifications befit the purpose of the app may be of interest—appropriate qualifications can be a surrogate parameter for the quality of the content. If those involved are experts in the respective field, be it because they obtained an academic degree or another type of suitable qualification, it is more likely that the content will be valid and of high quality than if it was written by others who are not similarly educated.

It would also have been desirable to discriminate between apps in general and health-related apps when asking participants to assess the importance of quality criteria. In addition to the general review of the quality principles, this would have made it possible to assess whether the participants' perceptions of quality criteria differ between general and particularly sensitive health contexts.

Outlook

Planned follow-up studies should aim to confirm and extend the results of this study. A more diverse study population (larger number of participants, other academic years, other health-related programs, and vocational training) should be included. On the whole, it is most important to facilitate analyses that can quantify the relevance of the individual quality principles and their contribution to the assessment process. This can be achieved by creating a larger database through experiment reproduction. Through this process, the isolation of a truly necessary and sufficient number of principles would be better possible. The operationalization of the quality principles will be examined in a separate paper. The aim is to identify potential candidates from the existing set of known criteria, to check their suitability and, if necessary, to synthesize new criteria. A time series, for example, through yearly evaluations in similar classes, possibly at other universities, could also be potentially used to determine whether, and if so, how, students'

awareness and perception of quality criteria in health-related app contexts change over time.

Conclusions

To provide users with orientation and to strengthen their decision-making competence, the app description texts must contain significantly more relevant information, for example, by including information compiled by following a standardized and comprehensive structure [19,35]. App stores should encourage this approach, as it would significantly aid in satisfying their users' need for information. However, whether (possibly mandatory) validations or cross checks of the provided information by independent experts, for example, before publication of a health-related app in an app store, would encourage trust and actually benefit users or would rather impede innovations seems questionable. Serious checks performed by experts in the respective field would—because of the steadily growing number of apps—require a significant number of experts to be able to perform these checks in a timely manner and would also introduce costs that many (at least smaller or startup) manufacturers would be unable or unwilling to bear. We therefore believe that sensitizing users to the importance of applying quality principles to any information available about an app, including app descriptions, will be much more effective.

Acknowledgments

The authors would like to thank the German Society for Internal Medicine (DGIM) e.V. for covering the publication costs.

Conflicts of Interest

None declared.

References

1. Duggal R, Brindle I, Bagenal J. Digital healthcare: regulating the revolution. Br Med J 2018 Dec 15;360:k6. [doi: 10.1136/bmj.k6] [Medline: 29335296]
2. Steinhubl SR, Muse ED, Topol EJ. The emerging field of mobile health. Sci Transl Med 2015 Apr 15;7(283):283rv3. [doi: 10.1126/scitranslmed.aaa3487] [Medline: 25877894]
3. Larson RS. A Path to Better-Quality mHealth Apps. JMIR Mhealth Uhealth 2018 Jul 30;6(7):e10414 [FREE Full text] [doi: 10.2196/10414] [Medline: 30061091]
4. Boulos MN, Brewer AC, Karimkhani C, Buller DB, Dellavalle RP. Mobile medical and health apps: state of the art, concerns, regulatory control and certification. Online J Public Health Inform 2014;5(3):229 [FREE Full text] [doi: 10.5210/ojphi.v5i3.4814] [Medline: 24683442]
5. Hale K, Capra S, Bauer J. A framework to assist health professionals in recommending high-quality apps for supporting chronic disease self-management: illustrative assessment of type 2 diabetes apps. JMIR Mhealth Uhealth 2015;3(3):e87 [FREE Full text] [doi: 10.2196/mhealth.4532] [Medline: 26369346]
6. Agencia de Calidad Sanitaria de Andalucía. Agencia de Calidad Sanitaria de Andalucía. 2012. Safety and quality strategies in mobile health apps: Complete list of recommendations on design, use and assessment of health Apps Internet URL: http://www.calidadappsalud.com/en/listado-completo-recomendaciones-app-salud/ [accessed 2019-02-17] [WebCite Cache ID 76G581Hy9]
7. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. JMIR Mhealth Uhealth 2015;3(1):e27 [FREE Full text] [doi: 10.2196/mhealth.3422] [Medline: 25760773]
8. Anderson K, Burford O, Emmerton L. App chronic disease checklist: protocol to evaluate mobile apps for chronic disease self-management. JMIR Res Protoc 2016 Nov 04;5(4):e204 [FREE Full text] [doi: 10.2196/resprot.6194] [Medline: 27815233]
9. Yasini M, Beranger J, Desmarais P, Perez L, Marchand G. mHealth quality: a process to seal the qualified mobile health apps. Stud Health Technol Inform 2016;228:205-209. [doi: 10.3233/978-1-61499-678-1-205] [Medline: 27577372]
10. Bradway M, Carrion C, Vallespin B, Saadatfard O, Puigdomènech E, Espallargues M, et al. mHealth assessment: conceptualization of a global framework. JMIR Mhealth Uhealth 2017 May 02;5(5):e60 [FREE Full text] [doi: 10.2196/mhealth.7291] [Medline: 28465282]
11. Fiore P. How to evaluate mobile health applications: a scoping review. Stud Health Technol Inform 2017;234:109-114. [doi: 10.3233/978-1-61499-742-9-109] [Medline: 28186025]
12. Leigh S, Ouyang J, Mimnagh C. Effective? Engaging? Secure? Applying the ORCHA-24 framework to evaluate apps for chronic insomnia disorder. Evid Based Ment Health 2017 Nov;20(4):e20. [doi: 10.1136/eb-2017-102751] [Medline: 28947676]
13. Ben-Mussa A, Paget AM. Popular apps on the medical category targeting patients and the general public in the United Kingdom: do they conform to the Health On the Net Foundation principles? Health Informatics J 2018 Sep;24(3):259-276. [doi: 10.1177/1460458216664193] [Medline: 27566750]

14. Albrecht UV, Hillebrand U, von Jan U. Relevance of trust marks and CE labels in German-language store descriptions of health apps: analysis. JMIR Mhealth Uhealth 2018 Apr 25;6(4):e10394 [FREE Full text] [doi: 10.2196/10394] [Medline: 29695374]
15. German Medical Association. German Medical Association. 2018 Dec 14. (Model) Professional Code for Physicians in Germany - MBO-Ä 1997 - The Resolutions of the 121st German Medical Assembly 2018 in Erfurt as amended by a Resolution of the Executive Board of the German Medical Association (English version) URL: http://www.bundesaerztekammer.de/fileadmin/user_upload/downloads/pdf-Ordner/MBO/MBO-AE_EN_2018.pdf [accessed 2019-02-17] [WebCite Cache ID 76G6m3ll0]
16. Albrecht UV, Pramann O. Liability issues when using health apps: nobody is perfect. Dtsch Arztebl 2018;115(122) [FREE Full text]
17. Palmer S. Swipe right for health care: how the state may decide the future of the mHealth app industry in the wake of FDA uncertainty. J Leg Med 2017;37(1-2):249-263. [doi: 10.1080/01947648.2017.1303289] [Medline: 28910227]
18. Albrecht UV. Secure medical apps: transparency through leaflets. Dtsch Arztebl 2013;110(44) [FREE Full text]
19. Albrecht UV. Transparency of health-apps for trust and decision making. J Med Internet Res 2013;15(12):e277 [FREE Full text] [doi: 10.2196/jmir.2981] [Medline: 24449711]
20. Canadian Medical Association. 2015. Guiding principles for physicians recommending mobile health applications to patients URL: https://www.cma.ca/Assets/assets-library/document/en/advocacy/cma_policy_guiding_principles_for_physicians_recommending_mobile_health_applications_to_patients_pd1-e.pdf [accessed 2019-02-17] [WebCite Cache ID 76G7YHIiT]
21. 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]
22. APPKRI. 2018. APPKRI criteria for health apps URL: https://ehealth-services.fokus.fraunhofer.de/BMG-APPS/ [accessed 2019-02-17] [WebCite Cache ID 76G8lDcs6]
23. Ranked. 2019. Curated Health Apps & Devices URL: http://www.rankedhealth.com/ [accessed 2019-02-17] [WebCite Cache ID 76GI8nFvL]
24. Haute Autorité de Santé. 2016. Good practice guidelines on health apps and smart devices (mobile health or mhealth) URL: https://www.has-sante.fr/portail/upload/docs/application/pdf/2017-03/dir1/good_practice_guidelines_on_health_apps_and_smart_devices_mobile_health_or_mhealth.pdf [accessed 2019-02-17] [WebCite Cache ID 76GIGMNhv]
25. Ministry of Health NZ. 2017. Health applications assessment guidance URL: https://www.health.govt.nz/our-work/ehealth/digital-health-sector-architecture-standards-and-governance/health-applications-assessment-guidance [accessed 2019-02-17] [WebCite Cache ID 76GIfFAki]
26. Myhealthapps. My health apps: tried and tested by people like you URL: http://myhealthapps.net/ [accessed 2019-02-17] [WebCite Cache ID 76GIx5uE8]
27. NHS. developer.nhs.uk. 2019. Digital Assessment Questions - Beta URL: https://developer.nhs.uk/digital-tools/daq/ [accessed 2019-02-17] [WebCite Cache ID 76GJ8NdwT]
28. NHS. 2019. Home - NHS Apps Library URL: https://www.nhs.uk/apps-library/ [accessed 2019-02-17] [WebCite Cache ID 76GJGy0jz]
29. Ontario Telemedicine Network. Practical Apps. 2019. URL: https://practicalapps.ca/ [accessed 2019-02-17] [WebCite Cache ID 76GJQbEFn]
30. Royal College of Physicians. Royal College of Physicians. 2015. Using apps in clinical practice guidance URL: https://www.rcplondon.ac.uk/guidelines-policy/using-apps-clinical-practice-guidance [accessed 2019-02-17] [WebCite Cache ID 76GJmjpwE]
31. West of England Academic Health Science Network. West of England Academic Health Science Network. 2016. Mobile Health Apps Resources URL: https://www.weahsn.net/work-with-us/working-with-businesses/resources/mobile-health-apps-resources/ [accessed 2019-02-17] [WebCite Cache ID 76GKepyfb]
32. Xcertia. 2019. mHealth App Guidelines URL: https://www.xcertia.org/the-guidelines/ [accessed 2019-02-17] [WebCite Cache ID 76GN5YQEN]
33. Albrecht UV, von Jan U, Jungnickel T, Pramann O. App-synopsis - standard reporting for medical apps. Stud Health Technol Inform 2013;192:1154. [doi: 10.3233/978-1-61499-289-9-1154] [Medline: 23920928]
34. Lewis TL. A systematic self-certification model for mobile medical apps. J Med Internet Res 2013;15(4):e89 [FREE Full text] [doi: 10.2196/jmir.2446] [Medline: 23615332]
35. Albrecht UV, Pramann O, von Jan U. Synopsis for health apps - transparency for trust and decision making. In: Househ M, Borycki E, Kushniruk A, editors. Social media and mobile technologies for healthcare. Hershey, Pennsylvania: IGI Global; 2014:94-108.
36. BinDhim NF, Hawkey A, Trevena L. A systematic review of quality assessment methods for smartphone health apps. Telemed J E Health 2015 Feb;21(2):97-104. [doi: 10.1089/tmj.2014.0088] [Medline: 25469795]

37. Albrecht UV. Chapter 13 - Orientation for users of health apps. In: Albrecht UV, editor. Chances and Risks of Health Apps (CHARISMHA). Hannover, Germany: Hannover Medical School; 2016:282-300.
38. Hartz T, von Jan U, Albrecht UV. Chapter 14 - Orientation for Professional Users of Health Apps. In: Albrecht UV, editor. Chances and Risks of Health Apps (CHARISMHA). Hannover, Germany: Hannover Medical School; 2016:302-319.
39. Stoyanov SR, Hides L, Kavanagh DJ, Wilson H. Development and validation of the user version of the mobile application rating scale (uMARS). JMIR Mhealth Uhealth 2016;4(2):e72 [FREE Full text] [doi: 10.2196/mhealth.5849] [Medline: 27287964]
40. Nouri R, R Niakan Kalhori S, Ghazisaeedi M, Marchand G, Yasini M. Criteria for assessing the quality of mHealth apps: a systematic review. J Am Med Inform Assoc 2018 Aug 01;25(8):1089-1098. [doi: 10.1093/jamia/ocy050] [Medline: 29788283]
41. Albrecht UV. Chapter 8 - Health Apps and Risks. In: Albrecht UV, editor. Chances and Risks of Health Apps (CHARISMHA). Hanover, Germany: Hannover Medical School; 2016:176-192.
42. Lee G, Raghu T, Park S. SSRN. 2015 Mar 28. Do App Descriptions Matter? Evidence from Mobile App Product Descriptions URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2581279
43. Ma S, Wang S, Lo D, Deng RH, Sun C. Active Semi-supervised Approach for Checking App Behavior against Its Description. 2015 Presented at: IEEE 39th Annual Computer Software and Applications Conference; 1-5 July, 2015; Taichung, Taiwan p. 179-184. [doi: 10.1109/COMPSAC.2015.93]
44. Gorla A, Tavecchia I, Gross F, Zeller A. Checking app behavior against app descriptions. In: Proceedings of the 36th International Conference on Software Engineering. New York, NY, USA: ACM; 2014 Presented at: ICSE 2014; May 31 - June 07, 2014; Hyderabad, India p. 1025-1035 URL: https://dl.acm.org/citation.cfm?id=2568276 [doi: 10.1145/2568225.2568276]
45. Albrecht UV, Hasenfuß G, von Jan U. Description of cardiological apps from the German app store: semiautomated retrospective app store analysis. JMIR Mhealth Uhealth 2018 Nov 20;6(11):e11753. [doi: 10.2196/11753] [Medline: 30459140]
46. ISO/IEC. BSI Standards Publication. 2011 Mar. ISO/IEC 25010 - Systems and software engineering - Systems and software Quality Requirements and Evaluation (SQuaRE) - System and software quality models URL: https://pdfs.semanticscholar.org/57a5/b99eceff9da205e244337c9f4678b5b23d25.pdf [accessed 2019-02-19] [WebCite Cache ID 76IbIAQPg]
47. Albrecht UV, Noll C, von Jan U. App-synopsis: self-assessment on trust or distrust of health-apps. Stud Health Technol Inform 2014;202:233-236. [doi: 10.3233/978-1-61499-423-7-233] [Medline: 25000059]
48. Ludwig S, Schuelper N, Brown J, Anders S, Raupach T. How can we teach medical students to choose wisely? A randomised controlled cross-over study of video- versus text-based case scenarios. BMC Med 2018 Jul 06;16(1):107 [FREE Full text] [doi: 10.1186/s12916-018-1090-y] [Medline: 29976211]
49. Albrecht UV. Health Apps: Interdisciplinary quality criteria are essential. Dtsch Arztebl 2018;115(3) [FREE Full text]
50. Albrecht UV. Warum so kompliziert? E-HEALTH-COM 2018(5):58-59.
51. ISO/IEC. webstore.iec.ch. 2016 Jun. ISO/IEC 25023 - Systems and software engineering - Systems and software Quality Requirements and Evaluation (SQuaRE) - Measurement of system and software product quality URL: https://webstore.iec.ch/preview/info_isoiec25023%7Bed1.0%7Den.pdf [accessed 2019-02-19] [WebCite Cache ID 76Ibdv4A1]
52. Zapata BC, Fernández-Alemán JL, Idri A, Toval A. Empirical studies on usability of mHealth apps: a systematic literature review. J Med Syst 2015 Feb;39(2):1. [doi: 10.1007/s10916-014-0182-2] [Medline: 25600193]
53. Brown W, Yen P, Rojas M, Schnall R. Assessment of the Health IT Usability Evaluation Model (Health-ITUEM) for evaluating mobile health (mHealth) technology. J Biomed Inform 2013 Dec;46(6):1080-1087 [FREE Full text] [doi: 10.1016/j.jbi.2013.08.001] [Medline: 23973872]
54. Yen PY. Health Information Technology Usability Evaluation: Methods, Models, and Measures. Dissertation, Columbia University. Ann Arbor, Michigan, USA: Proquest, Umi Dissertation Publishing; 2010.
55. Chan S, Torous J, Hinton L, Yellowlees P. Towards a framework for evaluating mobile mental health apps. Telemed J E Health 2015 Dec;21(12):1038-1041. [doi: 10.1089/tmj.2015.0002] [Medline: 26171663]
56. PLRI MedAppLab. 2015. app-synopsis - Checkliste für Nutzer URL: http://plrimedapplab.weebly.com/uploads/7/4/0/7/7407163/app-synopse-nutzer--dt-040315-1907.pdf [accessed 2019-02-17] [WebCite Cache ID 76GPOW28V]
57. Albrecht UV, von Jan U, Pramann O. Standard reporting for medical apps. Stud Health Technol Inform 2013;190:201-203. [doi: 10.3233/978-1-61499-276-9-201] [Medline: 23823422]
58. Leigh S, Ouyang J, Mimnagh C. Effective? Engaging? Secure? Applying the ORCHA-24 framework to evaluate apps for chronic insomnia disorder. Evid Based Ment Health 2017 Sep 25;20(4):e20. [doi: 10.1136/eb-2017-102751] [Medline: 28947676]
59. Albrecht UV, Matthies HK, Pramann O. In: Reiter H, Deussen S, editors. Mensch & Computer 2012: workshop band; 12th Interdisciplinary Conference for Interactive and Cooperative Media; interactively informed - omnipresent and all-encompassing !?. München: Oldenbourg Verlag; 2012:261-266.
60. afgis. 2012. Health Apps: Health App Fact Sheet URL: https://www.afgis.de/standards/gesundheits-app-fact-sheet [accessed 2019-02-17] [WebCite Cache ID 76GDdThkt]
61. APPcheck. 2018. URL: https://appcheck.de/ [accessed 2018-04-19] [WebCite Cache ID 6ynRmIcQE]

62. HealthOn. 2018. Online test for developers or providers of health apps URL: https://www.healthon.de/app-test_fuer_entwickler [accessed 2019-02-17] [WebCite Cache ID 76GEV7PVl]
63. Ben-Mussa A, Paget AM. Popular apps on the medical category targeting patients and the general public in the United Kingdom: do they conform to the Health On the Net Foundation principles? Health Informatics J 2018 Sep;24(3):259-276. [doi: 10.1177/1460458216664193] [Medline: 27566750]
64. Ali EE, Teo AK, Goh SX, Chew L, Yap KY. MedAd-AppQ: a quality assessment tool for medication adherence apps on iOS and android platforms. Res Social Adm Pharm 2018 Dec;14(12):1125-1133. [doi: 10.1016/j.sapharm.2018.01.006] [Medline: 29402755]
65. 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. [doi: 10.1016/j.jbi.2016.02.002] [Medline: 26903153]
66. ISO/IEC. ISO/IEC. 2016 Oct. ISO/IEC 82304-1 - Health software - Part 1: General requirements for product safety URL: https://www.iso.org/obp/ui/ [accessed 2019-02-19] [WebCite Cache ID 76IcXWk8w]
67. Fangerau H, Griemmert M, Albrecht UV. Chapter 9 - Health Apps and Ethics. In: Albrecht UV, editor. Chances and Risks of Health Apps (CHARISMHA). Hannover, Germany: Hannover Medical School; 2016:194-213.
68. Albrecht UV, Fangerau H. Do ethics need to be adapted to mHealth? A plea for developing a consistent framework. World Med J 2015 Jul;61(2):75-77 [FREE Full text]
69. ISO/IEC. ISO/IEC. 2001. ISO/IEC 9126-1 - Software engineering - Product quality - Part 1: Quality model URL: https://webstore.iec.ch/preview/info_isoiec9126-1%7Bed1.0%7Den.pdf [accessed 2019-02-19] [WebCite Cache ID 76Icp12Oq]
70. Pramann O. Chapter 11 - Health apps as medical devices. In: Albrecht UV, editor. Chances and Risks of Health Apps (CHARISMHA). Hannover, Germany: Hannover Medical School; 2016:228-243.
71. Pramann O. Chapter 10 - Health Apps and Privacy. In: Albrecht UV, editor. Chances and Risks of Health Apps (CHARISMHA). Hannover, Germany: Hannover Medical School; 2016:214-227.
72. Pramann O, Albrecht UV. Handling medical devices in the hospital and doctor's office. Düsseldorf: German Hospital Publishing Company; 2016.
73. Brönner M, Meister S, Breil B, Albrecht UV. Chapter 15 - Orientation for Health Apps Manufacturers. In: Albrecht UV, editor. Chances and Risks of Health Apps (CHARISMHA). Hannover, Germany: Hannover Medical School; 2016:320-340.
74. Böhme C, von Osthoff MB, Frey K, Hübner J. Qualitative evaluation of mobile cancer apps with particular attention to the target group, content, and advertising. J Cancer Res Clin Oncol 2018 Jan 19;144(1):173-181. [doi: 10.1007/s00432-017-2533-0] [Medline: 29052039]
75. Böhme C, von Osthoff M, Frey K, Hübner J. Development of a rating tool for mobile cancer apps: information analysis and formal and content-related evaluation of selected cancer apps. J Cancer Educ 2019 Feb;34(1):105-110. [doi: 10.1007/s13187-017-1273-9] [Medline: 28819928]
76. Martínez-Pérez B, de la Torre-Díez I, López-Coronado M. Experiences and results of applying tools for assessing the quality of a mHealth app named Heartkeeper. J Med Syst 2015 Nov;39(11):142. [doi: 10.1007/s10916-015-0303-6] [Medline: 26345452]
77. Loy JS, Ali EE, Yap KY. Quality assessment of medical apps that target medication-related problems. J Manag Care Spec Pharm 2016 Oct;22(10):1124-1140 [FREE Full text] [doi: 10.18553/jmcp.2016.22.10.1124] [Medline: 27668561]
78. Chen J, Cade JE, Allman-Farinelli M. The most popular smartphone apps for weight loss: a quality assessment. JMIR Mhealth Uhealth 2015;3(4):e104 [FREE Full text] [doi: 10.2196/mhealth.4334] [Medline: 26678569]
79. McNemar Q. Note on the sampling error of the difference between correlated proportions or percentages. Psychometrika 1947 Jun;12(2):153-157. [doi: 10.1007/BF02295996]
80. Bowker AH. A test for symmetry in contingency tables. J Am Stat Assoc 1948 Dec;43(244):572. [doi: 10.2307/2280710]
81. Cohen J. Statistical Power Analysis For The Behavioral Sciences (2nd Edition). New York: Academic Press; 1988.
82. Cohen J. A power primer. Psychol Bull 1992 Jul;112(1):155-159. [doi: 10.1037/0033-2909.112.1.155] [Medline: 19565683]
83. R-project. Vienna, Austria: R Foundation for Statistical Computing; 2018. The R Project for Statistical Computing URL: https://www.R-project.org [accessed 2018-09-23] [WebCite Cache ID 72eV2dBoz]
84. Larsen ME, Nicholas J, Christensen H. Quantifying app store dynamics: longitudinal tracking of mental health apps. JMIR Mhealth Uhealth 2016 Aug 09;4(3):e96 [FREE Full text] [doi: 10.2196/mhealth.6020] [Medline: 27507641]
85. Weiser M. Scientific American. Special Issue on Communications, Computers, and Networks. 1991 Sep. The computer for the 21st century URL: https://www.ics.uci.edu/~corps/phaseii/Weiser-Computer21stCentury-SciAm.pdf [accessed 2019-02-19] [WebCite Cache ID 76IddPBtm]
86. Albrecht UV. Transparenz schaffen und Orientierung bieten - Methoden und Werkzeuge als Entscheidungshilfe für die Nutzung von Gesundheits-Apps. Erstellung einer ersten Auslegeordnung zur Entwicklung eines Hilfsmittels für schweizerische Anwender (eHealth Suisse). Bern, Switzerland: ehealth Suisse; 2019. [doi: 10.26068/mhhrpm/20190116-000]
87. Statista. 2018. Android vs iOS: market share in smartphone sales in Germany 2018 | Statistic URL: https://www.statista.com/statistics/461900/android-vs-ios-market-share-in-smartphone-sales-germany/ [accessed 2019-02-18] [WebCite Cache ID 76H92N3ss]

88. Patel RK, Sayers AE, Patrick NL, Hughes K, Armitage J, Hunter IA. A UK perspective on smartphone use amongst doctors within the surgical profession. Ann Med Surg (Lond) 2015 Jun;4(2):107-112 [FREE Full text] [doi: 10.1016/j.amsu.2015.03.004] [Medline: 25905017]
89. Götz FM, Stieger S, Reips U. Users of the main smartphone operating systems (iOS, Android) differ only little in personality. PLoS One 2017;12(5):e0176921 [FREE Full text] [doi: 10.1371/journal.pone.0176921] [Medline: 28467473]

Abbreviations

CE: Conformité Européenne
ISO/IEC: International Organization for Standardization/International Electrotechnical Commission
SARASA: Semiautomated Retrospective App Store Analysis

Edited by G Eysenbach; submitted 11.01.19; peer-reviewed by B Breil, S Sullivan, K Anderson, R Eckhoff, M Kelley; comments to author 30.01.19; revised version received 09.02.19; accepted 10.02.19; published 27.02.19.

Please cite as:
Albrecht UV, Malinka C, Long S, Raupach T, Hasenfuß G, von Jan U
Quality Principles of App Description Texts and Their Significance in Deciding to Use Health Apps as Assessed by Medical Students: Survey Study
JMIR Mhealth Uhealth 2019;7(2):e13375
URL: http://mhealth.jmir.org/2019/2/e13375/
doi:10.2196/13375
PMID:30810534

Review

Efficacy and Effectiveness of Mobile Health Technologies for Facilitating Physical Activity in Adolescents: Scoping Review

Alexandra M Lee[1*], BS; Sarah Chavez[1], MPH; Jiang Bian[1], PhD; Lindsay A Thompson[2], MD, MS; Matthew J Gurka[1], PhD; Victoria G Williamson[3]; François Modave[4*], MS, PhD

[1]Department of Health Outcomes and Biomedical Informatics, College of Medicine, University of Florida, Gainesville, FL, United States

[2]Department of Pediatrics, Division of General Pediatrics, College of Medicine, University of Florida, Gainesville, FL, United States

[3]Department of Psychology, College of Liberal Arts and Sciences, University of Florida, Gainesville, FL, United States

[4]Center for Health Outcomes and Informatics Research, Health Sciences Division, Loyola University Chicago, Maywood, IL, United States

[*]these authors contributed equally

Corresponding Author:
François Modave, MS, PhD
Center for Health Outcomes and Informatics Research
Health Sciences Division
Loyola University Chicago
2160 S 1st Avenue
CTRE126
Maywood, IL, 60153
United States
Phone: 1 7082167962
Email: fmodave@luc.edu

Abstract

Background: Increasing physical activity (PA) levels in adolescents aged 12 to 18 years is associated with prevention of unhealthy weight gain and improvement in cardiovascular fitness. The widespread availability of mobile health (mHealth) and wearable devices offers self-monitoring and motivational features for increasing PA levels and improving adherence to exercise programs.

Objective: The aim of this scoping review was to identify the efficacy or effectiveness of mHealth intervention strategies for facilitating PA among adolescents aged 12 to 18 years.

Methods: We conducted a systematic search for peer-reviewed studies published between 2008 and 2018 in the following electronic databases: PubMed, Google Scholar, PsychINFO, or SportDiscus. The search terms used included mHealth or "mobile health" or apps, "physical activity" or exercise, children or adolescents or teens or "young adults" or kids, and efficacy or effectiveness. Articles published outside of the date range (July 2008 to October 2018) and non-English articles were removed before abstract review. Three reviewers assessed all abstracts against the inclusion and exclusion criteria. Any uncertainties or differences in opinion were discussed as a group. The inclusion criteria were that the studies should (1) have an mHealth component, (2) target participants aged between 12 and 18 years, (3) have results on efficacy or effectiveness, and (4) assess PA-related outcomes. Reviews, abstracts only, protocols without results, and short message service text messaging–only interventions were excluded. We also extracted potentially relevant papers from reviews. At least 2 reviewers examined all full articles for fit with the criteria and extracted data for analysis. Data extracted from selected studies included study population, study type, components of PA intervention, and PA outcome results.

Results: Overall, 126 articles were initially identified. Reviewers pulled 18 additional articles from excluded review papers. Only 18 articles were passed onto full review, and 16 were kept for analysis. The included studies differed in the sizes of the study populations (11-607 participants), locations of the study sites (7 countries), study setting, and study design. Overall, 5 mHealth intervention categories were identified: website, website+wearable, app, wearable+app, and website+wearable+app. The most common measures reported were subjective weekly PA (4/13) and objective daily moderate-to-vigorous PA (5/13) of the 19 different PA outcomes assessed. Furthermore, 5 of 13 studies with a control or comparison group showed a significant improvement in PA outcomes between the intervention group and the control or comparison group. Of those 5 studies, 3 permitted isolation of mHealth intervention components in the analysis.

Conclusions: PA outcomes for adolescents improved over time through mHealth intervention use; however, the lack of consistency in chosen PA outcome measures, paucity of significant outcomes via between-group analyses, and the various study designs that prevent separating the effects of intervention components calls into question their true effect.

(JMIR Mhealth Uhealth 2019;7(2):e11847) doi:10.2196/11847

KEYWORDS

review; mobile health; adolescent; exercise

Introduction

Background

Mobile health (mHealth) is the use of mobile or wireless devices to support medical and public health practice [1]. mHealth leverages the availability of technological innovations such as biological sensors, short message service (SMS), Global Positioning System, and accelerometry that are small enough to be embedded into wearable devices and smartphones. These novel technologies provide an easy way to collect health-related data and allow consumers to monitor their own health data. mHealth offers modalities that are easily accessible and low-cost to implement, permitting potential reach across socioeconomic gradients and into hard-to-reach populations [2]. The mHealth market was projected to reach US $23 billion in 2017 and it is estimated to grow more than 35% in the next 3 years [3]. A large proportion of growth in mHealth is in the native app and wearable activity device market. Native apps are apps developed specifically for a smartphone device that can be directly downloaded onto the device platform from the app marketplace (eg, Apple App Store, Google Play store). Wearable activity devices collect information about the movement activity of the individual wearing the device, and some are compatible with native or Web apps (eg, FitBit). In 2017, there were over 325,000 health-related apps available in major app stores (eg, Apple's App Store and Google's Play Store) equating to 3.7 billion app downloads [4]. Smartphone use is widespread in US adults with 77% owning a smartphone; however, in adolescents, it is even more prevalent, with 94% owning or having access to a smartphone and 89% indicating that they access the internet *almost constantly* or *several times a day* [5].

Behavior Change and Mobile Health

Emerging evidence shows that mHealth can aid in health behavior change resulting in better health outcomes, from smoking cessation and glucose monitoring to antiretroviral medication adherence and asthma control [6-8]. For successful prevention or management of many health conditions, health behavior changes are required, and a common recommendation has been to change physical activity (PA) behavior. Decreased PA levels are associated with several leading causes of death in the United States such as cardiovascular disease, cancer, and diabetes [9]. Previous reviews looking at the efficacy or effectiveness of mHealth for facilitating PA in adults have shown mixed evidence for the effectiveness of mHealth at increasing PA. One systematic review showed no impact on PA outcomes [10], and a meta-analysis presented a moderate effect on step counts and a nonsignificant effect on time in moderate-to-vigorous PA (MVPA) [11]. This lack of decisive findings could be attributed to low app quality and adherence to guidelines for exercise prescription [12] or a lack of variety in theories of behavior change employed by the app [13].

For the sake of clarity, in this paper, we use the term efficacy to discuss the performance of an intervention in a controlled environment, whereas effectiveness is understood as the performance of an intervention in a pragmatic setting. Nonetheless, these terms should not be viewed as binary attributes but rather as the range of performance as a continuous variable. Despite this mixed evidence as to the efficacy and effectiveness, understood in the context of continuity [14], of mHealth for facilitating PA in adults, there is a lack of evidence for an adolescent population despite its higher level of smartphone adoption and use. Adolescence is a sensitive period of neurocognitive development with effects on decision-making and behavior, making it a prime time for intervention regarding health-related behaviors. Though opportunities for prevention of chronic diseases can begin as early as the prenatal period, new health-related behaviors can arise in adolescence, making it a critical time point for prevention [15]. Vigorous PA levels decline by as much as 17.8% in boys and 11.0% in girls from middle to high school [16]. Increased knowledge about exercise, self-motivation, peer modeling and support, parental support, and availability of supplies or equipment are all positively associated with PA in adolescents [17,18]. mHealth modalities could address each of these correlates to PA. Although the meta-analysis [11] mentioned previously did include 2 studies within the adolescent age range, both studies incorporated SMS as the only mHealth intervention component. A systematic review of SMS interventions in youth and adolescents has recently been published [19]; thus, this review will focus on native and Web app interventions.

Objectives

The aim of this review was to identify the efficacy or effectiveness of mHealth for facilitating adolescent PA. Scoping reviews permit quick structured mapping of key concepts in a research area, identify gaps in the existing literature, and succinctly summarize emerging research findings [20]. This scoping review is timely as there is no review in the adolescent age group regarding the efficacy or effectiveness of mHealth for PA. In addition, the mHealth market is rapidly growing; therefore, a quicker review process is ideal for dissemination that would be timely to researchers and clinicians looking to improve PA outcomes and provide relevant clinical guidance to patients.

Methods

Search Methods

We followed the scoping review methodology proposed by Arksey and O'Malley [20]. A scoping review involves 5 stages: (1) identifying the research question, (2) identifying relevant studies, (3) selecting the studies, (4) charting the data, and (5) collating the results (writing the manuscript) [20]. Two of the key differences between scoping reviews and systematic reviews lie in the search strategy and the assessment of evidence quality. Scoping reviews have broad research questions and invoke an iterative search process to identify all relevant articles [20]. They also do not seek to assess the quality of evidence for included studies [20]. In stage 1, we identified the following research question: Are mHealth interventions effective for increasing PA among adolescent populations? In stage 2, we identified databases, search terms, and set time constraints. We selected 4 primary databases for searching: PubMed, Google Scholar, PsychINFO, and SportDiscus. Although PubMed serves as the main foundation for publications within medical and public health journals, we added Google Scholar, PsychINFO, and SportDiscus to pull articles that could be in journals more relevant to software engineering, behavior change, and exercise science, respectively. The following terms were identified a priori by AML and FM and then entered into all 4 databases: mHealth or mobile health or apps, "physical activity" or exercise, children or adolescents or teens or "young adults" or kids, and efficacy or effectiveness. Due to differences in search engine functionality, the method by which terms were entered differed per database. See Multimedia Appendix 1 for detailed search methodology. Abstracts outside of the date range (July 2008-October 2018) and abstracts without articles available in English were removed before abstract review. The July 2008 start date marks the launch of the Apple App Store and the presence of smartphone apps in the consumer market. In stage 3, 3 reviewers (AML, SC, and FM) assessed all abstracts against the study criteria. Any uncertainties or differences in opinion were discussed as a group, and if unresolved in discussion, the full article was then reviewed. Inclusion criteria were (1) study participants aged 12 to 18 years, (2) address efficacy or effectiveness of mHealth, (3) mHealth component, and (4) PA-related study outcomes. PA study outcomes were considered to be measures of PA volume and health indicators representative of changes in PA, for example, cardiorespiratory fitness or strength. Exclusion criteria encompassed abstracts without a full article, protocols without results, interventions that only included SMS as the mHealth component, dissertations, or review papers. Reviewers searched all excluded review papers' referenced articles for applicability to scope and added them to the abstract review if deemed plausible to fit. Once the abstract review was complete, the 3 reviewers (AML, SC, and FM) read each of the full articles to examine each article for fit with the criteria. As scoping reviews permit an iterative search process, any articles that were found outside of designated searches were sent to all 3 reviewers as a full article to be evaluated for inclusion into the review. Finally, in stage 4, AML, SC, and VW extracted data for analysis from articles that passed full review. Mendeley citation manager and Microsoft Excel were used to organize the references pulled from searches and to complete our assessment against the inclusion and exclusion criteria, respectively.

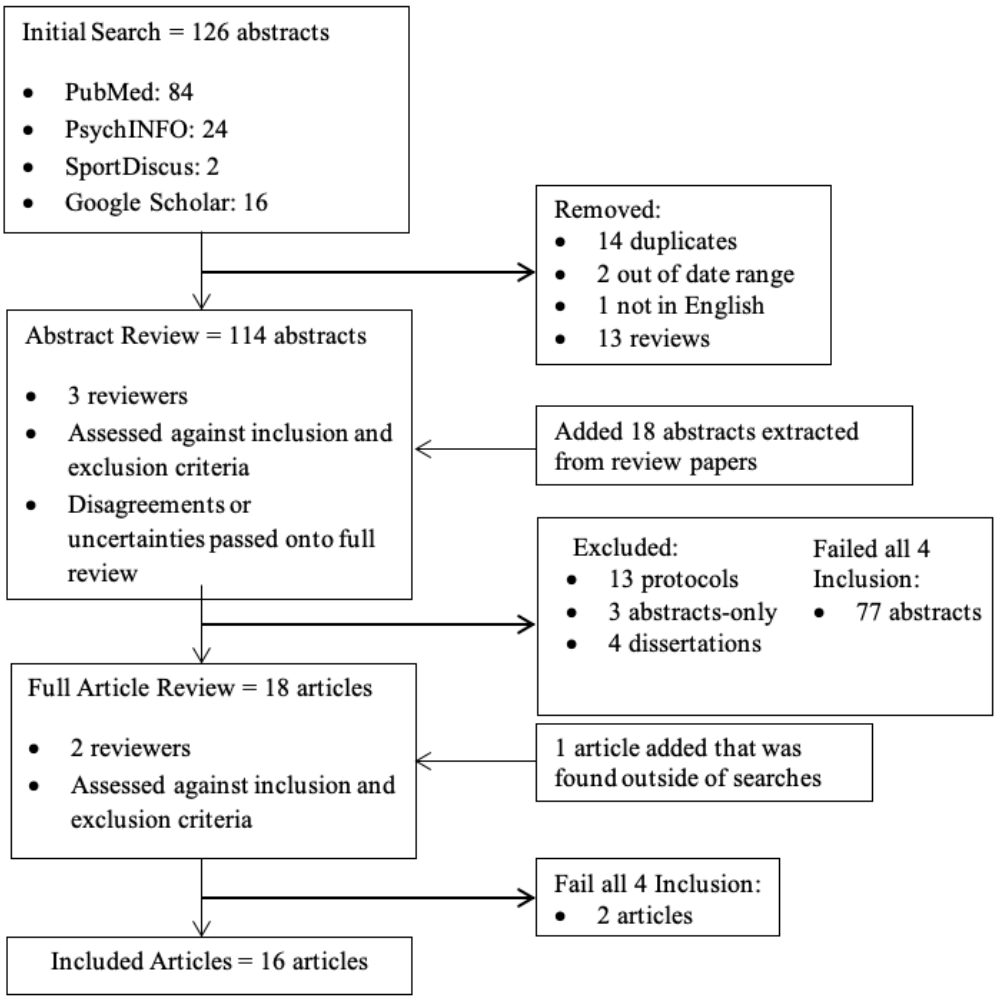
Analysis

We charted the included studies according to key characteristics identified by the authors to delineate the efficacy or effectiveness of mHealth for facilitating PA. The key characteristics identified pre-extraction included year, location, number of participants, age, sex, race and/or ethnicity; study design, setting, duration; mHealth intervention components, additional intervention components, PA outcome measures, and PA outcome results. Setting indicates where recruitment, implementation, and measures of the study occurred. Duration identifies how long the mHealth intervention was used, and subsequent PA outcomes were tracked as previous behavior change research regarding PA has highlighted the difficulties with maintenance [21,22]. Charting the study design provided information on the rigor of the science and whether adequate comparisons were implemented as well as the strength of the conclusions that could be made. Descriptive statistics were calculated regarding the presence or absence of PA characteristics among adolescents in the included studies.

Results

Through our database searches, we initially identified a total of 126 abstracts, of which, 14 were duplicated. Overall, 2 articles were removed as the studies were performed outside the date range of interest, and 1 article that was not in English was excluded as well. In addition, we removed 13 review articles from which we extracted 18 papers of interest for this scoping review, leading to a total of 96 distinct articles. Three reviewers assessed the 88 abstracts for inclusion and exclusion. Abstracts that led to disagreement and/or uncertainty of relevance were included in full review. Overall, 13 articles were protocols, 3 only had an abstract, 4 were dissertations, and 77 failed our inclusion criteria. A total of 17 articles were then considered for full review. Finally, 2 articles were excluded after full review as they did not meet the inclusion criteria. An additional article identified at the end-stage review was sent through full-article review and added to the analysis, totaling 16 studies (see Figure 1). Descriptions of each of the included studies are found in Table 1.

Figure 1. Flowchart of search methods and exclusion process.



Participants

The 16 included studies differed greatly in size (11-607 participants) and location (7 countries). Of them, 8 studies [23-30] took place in the United States and provide varying breakdowns of racial and/or ethnic groups. For each of the main races and/or ethnicities in the United States (white, Hispanic, black, or Asian), there was at least 1 study per group that sampled them in the majority.

Design and Setting

The included studies comprised 9 randomized controlled trials [23-26,30-34], 1 comparative effectiveness trial [28], 3 quasi-experimental studies [35-37], and 3 not randomized or controlled studies [27,29,38]. The most common recruitment setting was at the school (8/16) [28,29,32-36,38], followed by primary care (4/16) [23,26,27,37] and the community (4/16) [24,25,29,31]. Most implementation occurs in a home environment; however, the settings for measurements are split among home, university, school, and primary care.

Study Components

Study components and PA outcomes are described in Table 2. Primary mHealth intervention components for all included studies fell into the following categories: wearables, apps, and websites, and about half of the studies utilized 1 component and the other half used 2 or more mHealth components. The 5 mHealth intervention category combinations utilized were website [24,28,35,37], website+wearable [25,26,29,33,36], app [31,32,38], wearable+app [23,27], and website+wearable+app [30,34]. The most common singular mHealth component is a website. No studies used a wearable alone; however, it was the most common component to be paired with another mHealth component. Overall, 10 out of 16 studies [23-26,28,30,31,33,37,38] isolated mHealth components as the sole intervention method. *Isolate* refers to whether the intervention either solely included mHealth intervention components (versus *additional intervention components*, as described in Table 2) or the design of the study permitted an analysis that could look at the individual effect of the mHealth component on the PA-related outcome variable(s). Additional intervention components ranged from goal-setting guides to educational seminars or group counseling and skills training. In the 12 controlled studies [23-26,30-37], 8 control groups were simply advised to continue usual care or normal PA behaviors [26,30-32,34-37].

Table 1. Description of included studies.

Author (Year), Country	N	Age (years), mean (SD) or range	Study design, duration, and arms	Setting
Chen et al[a] (2017) [23], USA	40	15 (1.56)	RCT[b]; 3 months; Intervention and control	Primary care (recruitment); Home (implementation); Web-based (measures)
Cullen et al (2013) [24], USA	291	12-16	RCT; 8 weeks; Intervention and control	Community (recruitment); Home (implementation); Measurement unclear
Direito et al (2015) [31], New Zealand	51	15.7 (1.2)	RCT; 8 weeks; interventions and 1 control	Community (recruitment); Home (implementation); University (measures)
Gaudet et al (2017) [36], Canada	23	13 (0.3)	Quasi-experimental crossover; 7 weeks; Immediate intervention and delayed intervention (each served as control)	School (recruitment); Home (implementation); Home (measures)
Guthrie et al[a] (2015) [25], USA	182	12.7 (0.9)	RCT; 6 weeks; Active control, passive control, and intervention	Community (recruitment); Home (implementation); Measures unclear
Jimoh et al (2018) [38], United Kingdom	34	16-19	Not randomized or controlled; 8 weeks (4 with paper diary and 4 with app)	School (recruitment, implementation); Home (implementation and measures)
Kennedy et al[a] (2017) [32], Australia	607	14.1 (0.5)	Cluster RCT; 10 weeks; Matched pairs randomization at school level to intervention or control	School (recruitment, implementation and measures); Home (implementation)
Larsen et al (2018) [29], USA	21	14.7 (2.1)	Not randomized or controlled; 12 weeks	School (recruitment); Community (recruitment); Home (implementation)
Lau et al (2012) [35], Hong Kong	78	12-15	Quasi-experimental; 8 weeks; Intervention and control	School (recruitment, measures); Home (implementation)
Mendoza et al (2017) [30], USA	59	16.6 (1.5)	RCT; 10 weeks; Intervention and control	Survivor database and clinic (recruitment); Home (implementation and some measures)
Patrick et al (2013) [26], USA	101	14.3 (1.5)	RCT; 12 months; 3 interventions and 1 control	Primary care (recruitment, some implementation, and measures); Home (implementation)
Schoenfelder et al (2017) [27], USA	11	15.5 (1.4)	Not randomized or controlled; 4 weeks	Primary care (recruitment); Home (implementation and measures)
Slootmaker et al[a] (2010) [33], the Netherlands	87	13-17	RCT; 3 months; Intervention and control	School (recruitment and measures); Home (implementation)
Smith et al[a] (2014) [34], Australia	361	12.7 (0.5)	Cluster RCT; 7 months; School-level matched pairs randomization to intervention or control	School (recruitment, measures, and implementation)
Sousa et al (2015) [37], Portugal	94	12-18	Quasi-experimental; Pre-post design with control; 24 weeks	Clinic (recruitment, implementation, and measures); Home (implementation)
Whittemore et al (2013) [28], USA	384	15.31 (0.69)	Cluster RCT; Class level; 6 months; Comparison of 2 interventions	School (recruitment, implementation, and measures)

[a]Significant between-group results.

[b]RCT: randomized controlled trial.

Table 2. Study components and physical activity outcomes.

Study	Control group	mHealth intervention component	Additional intervention component	Physical activity outcomes	Results
Chen et al[a] (2017) [23]	Given Omron HJ-105 pedometer+blank food diary+online program with 8 modules on general adolescent health	FitBit flex+iStart Smart mobile app	None	Number of days per week with 60 min MVPA[b] (S)	Significant time X group interaction for reported days of PA[c]. Medium effect size
Cullen et al (2013) [24]	Online program with educational materials about nutrition and PA, printable goal sheets	Online program with educational materials about nutrition and PA, role model stories, online self-monitoring, goal review, and problem-solving components	None	60 min of PA on 5 days per week; (S[d]); 60 min of PA 7 days per week; (S)	Post intervention, significantly more adolescents reported being physically active for >60min/day on all 7 days (P<.001); 84% of adolescents reported online program was helpful for increasing PA
Direito et al (2015) [31]	Asked to continue normal PA	Immersive app: Zombies Run! 5K training; Nonimmersive app: Get Running-Couch to 5K	None	Cardiorespiratory fitness (O); Time to run 1 mile (O); Weekly PA (S); Daily MVPA (O)	No difference in VO_2 peak between interventions and control. No intervention effect for self-reported PA or objective MVPA. Group assignment did not have significant effect on mean daily MVPA. For those that used app 3 times/week, statistically significant decline in time to run for nonimmersive app compared with control
Gaudet et al (2017) [36]	Nothing given	FitBit Charge HR and FitBit Web app	Taught SMART and told to set goals	Daily MVPA (O); Step Counts (O)	The immediate intervention group A increased MVPA by 10.9 min (P=.03) during their intervention but Group B did not significantly increase during their intervention. There were no significant differences in MVPA between groups at any time point. Students at baseline that were in the adoption phase experienced a significant increase in MVPA by more than 15 min/day
Guthrie et al[a] (2015) [25]	Active: Received Zamzee activity monitors and Dance Revolution video game; Passive: Received Zamzee activity monitors	Zamzee activity monitors+website with monitor feedback and incentive motivation	None	Daily MVPA (O)	MVPA duration differed significantly across groups (P<.01). Intervention group showed an average 15.26 min/day MVPA, which is 49% greater than passive control (P<.01) and 67% greater than active control (P<.01)
Jimoh et al (2018) [38]	No control	Smartphone app diary for food and exercise recording; SMS[e] with personalized feedback	Paper diaries; In-person meetings with research team to review paper diaries every 2 weeks	Weekly PA (S)	No significant differences between app and paper diary in reported volume of aerobic and strength training during respective intervention period

Study	Control group	mHealth intervention component	Additional intervention component	Physical activity outcomes	Results
Kennedy et al[a] (2017) [32]	Standard school procedures	Web-based smartphone app	Interactive student seminars, structured PA program, lunchtime fitness sessions	Upper body muscular endurance (O[f]), lower body muscular strength (O), Daily MVPA (O), Cardiorespiratory Fitness (O), RT[g] skill competency(O)	Significant intervention effects for upper body muscular endurance at 6 and 12 months; Significant group-by-time interactions for estimated VO_2 max at 6 months; No significant intervention effects for weekday MVPA. Significant group-by-time effects for RT[g] skill competency at 6 and 12 months
Larsen et al (2018) [29]	No control	Website with activity manuals matched to current level of motivational readiness, tailored reports based on regular assessment, tip sheets, local activity resources; Pedometer	One-on-one goal-setting session; Tip sheets for moms of participants	Weekly MVPA (S)	Weekly MVPA increased from 24.7 min at baseline to 79.4 min at follow-up (P<.01). Each participant increased an average of 58.8 min/week (P<.01).
Lau et al (2012) [35]	Nothing given	Internet PA program (behavioral skill training, self-monitoring, tailored feedback, PA planner, chat room; SMS; (virtual friend; message types: behavioral skills, reinforcement of PA benefits, solutions for PA barriers, motivational, informational)	N/A[h]	Weekly PA (S)	Significant increase in intervention group (P=.05) but not in control group (P=.34); however, the intergroup differences were not significant
Mendoza et al (2017) [30]	Usual Care	FitBit Flex and Fitbit app; Facebook group-team and individual motivation badges from staff, and participant discussion of shared experiences; Phone/SMS 1/week to set goal; SMS every other day for reminder and encouragement	N/A	Daily MVPA (O)	Within-group changes in PA from baseline to follow-up not reported. No significant difference in daily MVPA between the intervention and the control group.
Patrick et al (2013) [26]	Usual care	All included: Pedometer; Web only group: program website with Web tutorials+weekly check-in emails; Web+group sessions: program website with Web tutorials; Web+SMS: program website, 3 text messages/week (content, reminders, or questions to counselor)	All 4 conditions included: monthly mailed tip sheets; Web only: also included weekly email check-ins; Web+group also included monthly 90 min group behavioral skills sessions and bimonthly phone calls from a health counselor	Weekly MVPA (S)	No significant differences observed for any PA outcomes

Study	Control group	mHealth intervention component	Additional intervention component	Physical activity outcomes	Results
Schoenfelder et al (2017) [27]	No control	FitBit flex+FitBit mobile app; Facebook group; Daily SMS	None	Step Counts (O)	Daily step counts averaged 8014. From baseline to 4 weeks, step count significantly increased by 3218 steps (P<.01)
Slootmaker et al[a] (2010) [33]	Received brochure with general PA recommendations	PAM accelerometer+PAM coach website	None	Weekly PA (S)	Girls in intervention group significantly increased moderate PA after 3 months (P=.04) compared with control, but effect disappeared at 8 months.
Smith et al[a] (2014) [34]	Standard school procedures	Pedometers; Smartphone app and website	Teacher professional development; Parent newsletters; Researcher-led seminars; Enhanced school sport sessions; Lunchtime PA-mentoring sessions	Weekly PA (O); Weekly MVPA (O); RT Skill competency (O); Upper body maximal strength and endurance (O)	No significant differences in overall PA counts or MVPA; Significant intervention effect on upper body muscular endurance (P=.04) and RT skill competency (P<.01)
Sousa et al (2015) [37]	Standard treatment program of clinical evaluation, medical, psychological, nutritional, and PA counseling	Next Step: e-therapeutic platform (education, self-monitoring, social support, interactive training, and motivational tools)	Standard treatment program of clinical evaluation, medical, psychological, nutritional, and PA counseling	Weekly PA (S)	Significant improvement in PA in intervention group (P<.03)
Whittemore et al (2013) [28]	Comparative effectiveness	HEALTH[e]TEEN website (lessons, goal setting, self-monitoring, health coaching, and social networking)	CST-coping skills training, combined with website for one of the intervention groups	Days per week 60 min MVPA. (S); Days per week of muscle strengthening (S)	In both groups, adolescents significantly increased MVPA over 6 months. No time effects for muscle strengthening.

[a]Significant between-groups results.

[b]MVPA: moderate-to-vigorous physical activity.

[c]PA: physical activity.

[d]S: subjective measure.

[e]SMS: short message system.

[f]O: objective measure.

[g]RT: resistance training.

[h]N/A: not applicable.

Physical Activity Outcomes

The PA outcomes reported can be divided into 5 categories: days per week meeting activity guidelines, cardiorespiratory fitness, MVPA, general activity, muscular strength, and competency. Within each of these categories, individual measures can be isolated by the amount of time they capture and whether they are subjective or objective measures. Parsed out, there are 18 total PA outcomes measured among the 16 studies included (see Table 1, column titled *Physical Activity Outcomes*). A total of 8 measures appear in more than 1 study, the most common being objective daily MVPA (5/13) [25,30-32,36] followed by subjective weekly PA (5/16) [31,33,35,37,38]. Of the 18 PA outcome measures, 44% were subjective, self-reported measures. Significant improvement in a PA outcome over time in the group with the mHealth intervention was observed for all but 2 of the eligible included studies (ineligible: Direito et al and Mendoza et al, which did not present results for this type of analysis); however, this improvement was not always unique to the intervention group. Of the 13 studies that contained a control or comparison group, only 5 [22,24,31,32,33] showed an improvement in a PA outcome that was significantly different from the results of the control or comparison group. Within the 5 studies that showed significant results for between-groups analyses, 3 studies [22,24,32] isolated the mHealth component in the intervention. In addition, 2 of the 3 studies utilized a website and wearable intervention [24,32] and 1 utilized an app and a wearable [23]. The PA outcomes that improved in the 3 studies were objective daily MVPA [25], subjective weekly PA [33], and number of days per week of 60 min of PA [23], respectively.

Table 3. Mobile health intervention components and significant improvement in physical activity outcomes.

Study and modality	Duration ≥12 weeks	Increasing PA[a] outcome over time in experimental group or all groups	Increasing PA outcome between groups or group x time	Isolated the mHealth component
Website				
Lau et al (2012) [35]		✓ Weekly MVPA[b] (S[c])		
Cullen et al (2013) [24]		✓60 min PA 7 days/week (S)		✓
Whittemore et al (2013) [28]	✓	✓ # days/week meeting MVPA rec (S)		✓
Sousa et al (2015) [37]	✓	✓ Weekly PA (S)		✓
App				
Direito et al (2015) [31]		Not reported		✓
Kennedy et al (2017) [32]		✓ Upper body muscular endurance (O[d]), cardiorespiratory fitness (O), RT[e] skill competency (O)	✓ Upper body muscular endurance (O), cardiorespiratory fitness (O), RT skill competency (O)	
Jimoh et al (2018) [38]			Not applicable	✓
Wearable+Website				
Slootmaker et al (2010) [33]	✓	✓ Weekly PA (S)	✓ Weekly PA (S)	✓
Patrick et al (2013) [26]	✓			✓
Guthrie et al (2015) [25]		✓ Daily MVPA (O)	✓ Daily MVPA(O)	✓
Gaudet et al (2017) [36]		✓ Daily MVPA (O)		
Larsen et al (2018) [29]	✓	✓ Weekly MVPA (S)	Not applicable	
Wearable+App				
Schoenfelder et al (2017) [27]		✓ Step count	✓ Not applicable	
Chen et al (2017) [23]	✓	✓ # days/week 60 min PA (S)	✓ # days/week 60 min PA (S)	✓
Wearable+Website+App				
Smith et al (2014) [34]		✓ Upper body muscular endurance (O) and RT skill competency (O)	✓ Upper body muscular endurance (O) and RT skill competency (O)	
Mendoza et al (2017) [30]		Not reported		✓

[a]PA: physical activity.

[b]MVPA: moderate-to-vigorous physical activity.

[c]S: subjective measure.

[d]O: objective measure.

[e]RT: resistance training.

A total of 5 different groupings of mHealth components in the intervention groups (app, wearable+app, website, website+wearable, and website+wearable+app) among 16 studies prevent this review from being able to identify a specific mHealth component as most effective for promoting PA in adolescents. Furthermore, research designs from included interventions did not always isolate the mHealth component for analysis on its sole effect on PA outcomes. The 3 studies that showed significant intervention effects on a PA outcome and isolated the mHealth component in the intervention included 1 study with a wearable+app and 2 studies with a website+wearable. Finally, we considered interventions with a duration longer than 12 weeks as this is frequently observed as the break-up point with respect to adherence. See Table 3 for a visual summary of mHealth components and corresponding improvement in PA outcomes.

Interventions included multiple therapeutic modalities to improve health outcomes, which makes it difficult to disentangle the separate effects for each modality. The extreme diversity in PA outcomes reported across studies also makes comparison among mHealth components challenging. Table 3 expands upon Table 2 and presents a visual representation of PA outcomes that significantly improved compared with the mHealth modality implemented. The main reasons for diversity of PA outcomes

reported were each measure's time unit and whether it was a subjective or objective measure. Roughly half of the measures were subjective, self-reported measures, and a systematic review by Adamo et al showed that self-reported PA data in a pediatric population are often overestimated compared with objective measures of PA, thus caution should be exhibited when interpreting these results [39].

Discussion

Principal Findings

This scoping review synthesizes intervention findings from 16 studies that measured the efficacy or effectiveness of mHealth to improve PA outcomes in adolescents. Our review identified 10 randomized control trials, 3 of which are cluster trials, 3 quasi-experimental interventions with control, and 3 studies without a control group. An initial observation indicates that interventions with an mHealth component could lead to general improvement in PA outcomes over time, as was observed in 12 of the 16 studies. However, when specifically considering studies with a control group, only 5 of 13 showed a significant intervention effect, and an additional 5 of 13 showed improvement among all groups. Of the studies with a control group and objective assessments of PA outcomes, 3 showed statistically significant improvements in the intervention groups compared with the control groups.

It is particularly interesting to see that a majority of the studies included in this scoping review are randomized controlled trials or cluster randomized controlled trials, which are evidently far stronger interventions to demonstrate effect significance. Therefore, despite the small number of studies with a significant intervention effect and objective measurements of PA, we do not think that this scoping review indicates that mHealth interventions are inappropriate to be used to improve PA among adolescents. However, we believe that this forces us to rethink how to use mHealth and how we build mobile apps. Given the growing body of literature showing the impact of mHealth interventions on behavior change [2,6,40,41], we surmise that this is more a reflection of the quality of the mobile apps [12,42-44] and their appropriateness to the specific population we are considering. There is reasonable evidence that apps developed in collaboration with their potential users are more likely to be used and to be effective [45,46]. Therefore, mHealth researchers interested in PA in adolescents should consider building their interventions with their target population from the very beginning. In this review, the study by Chen et al [23] was one of the few studies that found an intervention effect and the only one to create an adolescent stakeholder group for intervention development. This is particularly important given the rapid growth of app development. The company *Flurry* tracks app usage across platforms (iOS and Android). It has observed a 330% growth in health and fitness app usage over the past 3 years [47]. In the meantime, Statista reports that the global mHealth market grew from $21.1 billion in 2016 to $40 billion in 2018 and is expected to reach a staggering $332.7 billion by 2025 [48].

Interestingly, 3 of the 13 controlled studies looked specifically at improvements in muscular strength and conditioning, and 2 of them showed significant improvement. One could conjecture that strength training faring better than aerobic conditioning is simply a reflection of preferences of the specific population under study, particularly boys. Unexpectedly, there was a lack of interventions that leveraged *gamification*, for example, Pokémon GO–type approaches, or competition-type interventions using apps, such as Strava (Strava.com), that allow asynchronous competition in a social media environment. Pokémon GO has been shown to significantly increase the number of young adults who reach 10,000 steps per day [49], and some research shows that gamification could be appealing to adolescents [50].

Recruitment settings mirrored common sites for recruitment in intervention trials that have pediatric or adolescent populations, but the portability of mHealth interventions permitted most studies to use home-based implementation. Home-based implementation reduces participant transportation burden and cost. Measures were spread among 4 different sites, but this was because many of the included studies incorporated biological measures that cannot be measured on the Web or through an app, such as cardiorespiratory fitness and lower body muscular strength. Less than half of the studies were of a duration longer than 12 weeks, thus it is difficult to discern any long-term behavior changes as most studies did not extend far enough to evaluate PA outcome maintenance beyond the usual 12-week end point.

mHealth intervention components utilized within the studies included apps, wearables, and websites. The wearable is the only mHealth component that did not function as a solo intervention component but was always paired with an app or website. Innovations in technology have resulted in most new wearables pairing with an app or website that provides data feedback to the user, with some permitting a social network component, which could explain this observation.

Strengths and Weaknesses

A limitation of scoping reviews is that because of their broad and less targeted approach compared with a traditional systematic review, quantitative results cannot be pooled to understand the effect size of mHealth technologies on PA outcomes in adolescents. However, the purpose of a scoping review methodology is to be broad reaching [20], which is appropriate in the context of mHealth and PA interventions in adolescents. A second limitation arises from the search methodology. There is a possibility that because of different databases requiring modifications of search methods, relevant studies may have been missed. This is primarily a concern with our Google Scholar search as the search engine would not adhere to our Boolean arguments, requiring large amounts of individual title sorting by 1 reviewer (AML). In addition, scoping reviews do not seek to assess the quality of evidence or provide a grading of recommendations assessment, development and evaluation for the included studies; therefore, we cannot comment on the strength or generalizability of the findings [20]. A strength of this review is that it provides a quick snapshot of research conducted in the last decade in the adolescent target population and includes several different kinds of mHealth tools that are employed in intervention work. Even though there have been

reviews pertaining to the use of mHealth and PA among teens, to the authors' knowledge, this paper is the first review to summarize newer tools, such as app and wearable interventions, on PA outcomes for adolescents. This is particularly important given the pace at which technology evolves especially with respect to accuracy of measurements and integration.

Conclusions

In conclusion, the use of mHealth to improve PA in adolescents is an emerging field, complicated by frequent change in the rate at which new technologies develop. There is limited research about the efficacy and effectiveness of each mHealth modality, and future research designs need to provide avenues for analysis of the effect of the mHealth component alone on PA outcomes in adolescents in addition to the combined effect of the total intervention, if applicable. As there is a diversity of mHealth modalities, a reported component analysis of the specific modality utilized in a study, for example, an app, could provide information to compare it with other apps utilized in other studies. For example, apps can have social networking components, notifications, reward systems, gaming features, or education modules, all of which could affect PA outcomes differently; therefore, future work could be improved if these qualities were mapped out for comparison. In this sense, mHealth modalities could be evaluated as efficacious or effective on the basis of component analysis versus referring to their general category such as *website*, *app*, or *wearable*.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Detailed search methods per database.

[DOCX File, 101KB - mhealth_v7i2e11847_app1.docx]

References

1. Ryu S. World Health Organization. 2011. mHealth: New Horizons for Health through Mobile Technologies: Based on the Findings of the Second Global Survey on eHealth (Global Observatory for eHealth Series, Volume 3) URL: https://www.who.int/goe/publications/goe_mhealth_web.pdf [accessed 2019-02-01] [WebCite Cache ID 75rNPpFXE]
2. Hswen Y, Viswanath K. Beyond the Hype: Mobile Technologies and Opportunities to Address Health Disparities. J Mob Technol Med Internet 2015 Jan.
3. Orbis Research. Orbis Research. 2016. The mHealth (Mobile Healthcare) Ecosystem-Opportunities, Challenges, Strategies, and Forecasts URL: https://orbisresearch.com/reports/index/the-mhealth-mobile-healthcare-ecosystem2017-2030-opportunities-challenges-strategies-and-forecasts [accessed 2019-02-01] [WebCite Cache ID 75rOUCeBM]
4. Research 2 Guidance. 2017. mHealth App Economics 2017 URL: https://www.uzelf.org/wp-content/uploads/2017/12/R2G-mHealth-Developer-Economics-2017-Status-And-Trends.pdf [accessed 2019-02-01] [WebCite Cache ID 75rOm3xz9]
5. Anderson M, Jiang J. Pew Internet. 2018. Teens, Social Media & Technology 2018 Internet URL: http://www.pewinternet.org/2018/05/31/teens-social-media-technology-2018/ [accessed 2019-01-30] [WebCite Cache ID 75oeQ4D7H]
6. Handel M. mHealth (Mobile Health)-Using apps for health and wellness. Explor J Sci Heal Internet 2011;7(4):256-261 [FREE Full text] [doi: 10.1016/j.explore.2011.04.011] [Medline: 21724160]
7. 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]
8. Marcolino M, Oliveira J, D'Agostino M, Ribeiro A, Alkmim M, 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]
9. Centers for Disease Control and Prevention. Atlanta; 2015. Physical Activity Basics URL: https://www.cdc.gov/physicalactivity/basics/pa-health/index.htm [accessed 2019-01-30] [WebCite Cache ID 75ofOuV0U]
10. Stephens J, Allen J. Mobile phone interventions to increase physical activity and reduce weight: a systematic review. J Cardiovasc Nurs 2013;28(4):320-329 [FREE Full text] [doi: 10.1097/JCN.0b013e318250a3e7] [Medline: 22635061]
11. Fanning J, Mullen SP, McAuley E. Increasing physical activity with mobile devices: a meta-analysis. J Med Internet Res 2012 Nov 21;14(6):e161 [FREE Full text] [doi: 10.2196/jmir.2171] [Medline: 23171838]
12. Modave F, Bian J, Leavitt T, Bromwell J, Harris III C, Vincent H. Low quality of free coaching apps with respect to the American College of Sports Medicine guidelines: a review of current mobile apps. JMIR Mhealth Uhealth 2015 Jul 24;3(3):e77 [FREE Full text] [doi: 10.2196/mhealth.4669] [Medline: 26209109]
13. Bondaronek P, Alkhaldi G, Slee A, Hamilton F, Murray E. Quality of publicly available physical activity apps: review and content analysis. JMIR Mhealth Uhealth 2018 Mar 21;6(3):e53 [FREE Full text] [doi: 10.2196/mhealth.9069] [Medline: 29563080]

14. Gerald GI, Richard HM, Nissman D, Kathleen LM, Carey T. AHRQ Technical Reviews. 2006. Criteria for distinguishing effectiveness efficacy trials in systematic reviews URL: https://www.donau-uni.ac.at/imperia/md/content/department/evidenzbasierte_medizin/abstracts_publikationen_gerald/criteria_for_distinguishing_effectiveness_from_efficacy_trials_in_systematic_reviews.pdf [accessed 2019-02-01] [WebCite Cache ID 75rPjEgFj]
15. Sawyer S, Afifi R, Bearinger L, Blakemore S, Dick B, Ezeh A, et al. Adolescence: a foundation for future health. Lancet 2012 Apr 28;379(9826):1630-1640. [doi: 10.1016/S0140-6736(12)60072-5] [Medline: 22538178]
16. DiLorenzo T, Stucky-Ropp R, Vander Wal JS, Gotham HJ. Determinants of exercise among children. II. A longitudinal analysis. Prev Med 1998;27(3):470-477. [doi: 10.1006/pmed.1998.0307] [Medline: 9612838]
17. Sallis JF, Prochaska JJ, Taylor WC. A review of correlates of physical activity of children and adolescents. Med Sci Sports Exerc 2000 May;32(5):963-975. [Medline: 10795788]
18. Morrissey JL, Janz KF, Letuchy EM, Francis SL, Levy SM. The effect of family and friend support on physical activity through adolescence: a longitudinal study. Int J Behav Nutr Phys Act 2015 Aug 20;12:103 [FREE Full text] [doi: 10.1186/s12966-015-0265-6] [Medline: 26289232]
19. Ludwig K, Arthur R, Sculthorpe N, Fountain H, Buchan DS. Text messaging interventions for improvement in physical activity and sedentary behavior in youth: systematic review. JMIR Mhealth Uhealth 2018 Sep 17;6(9):e10799 [FREE Full text] [doi: 10.2196/10799] [Medline: 30224335]
20. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodo 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]
21. Nigg CR, Borrelli B, Maddock J, Dishman RK. A theory of physical activity maintenance. Appl Psychol 2008 Jul 08;57(4):544-560. [doi: 10.1111/j.1464-0597.2008.00343.x]
22. Amireault S, Godin G, Vézina-Im LA. Determinants of physical activity maintenance: a systematic review and meta-analyses. Health Psychol Rev 2013 Mar;7(1):55-91. [doi: 10.1080/17437199.2012.701060]
23. Chen JL, Guedes CM, Cooper BA, Lung AE. Short-term efficacy of an innovative mobile phone technology-based intervention for weight management for overweight and obese adolescents: pilot study. Interact J Med Res 2017 Aug 02;6(2):e12 [FREE Full text] [doi: 10.2196/ijmr.7860] [Medline: 28768612]
24. Cullen K, Thompson D, Boushey C, Konzelmann K, Chen TA. Evaluation of a web-based program promoting healthy eating and physical activity for adolescents: teen choice: food and fitness. Health Educ Res 2013 Aug;28(4):704-714. [doi: 10.1093/her/cyt059] [Medline: 23748162]
25. Guthrie N, Bradlyn A, Thompson SK, Yen S, Haritatos J, Dillon F, et al. Development of an accelerometer-linked online intervention system to promote physical activity in adolescents. PLoS One 2015;10(5):e0128639 [FREE Full text] [doi: 10.1371/journal.pone.0128639] [Medline: 26010359]
26. Patrick K, Norman GJ, Davila EP, Calfas KJ, Raab F, Gottschalk M, et al. Outcomes of a 12-month technology-based intervention to promote weight loss in adolescents at risk for type 2 diabetes. J Diabetes Sci Technol 2013 May 01;7(3):759-770 [FREE Full text] [doi: 10.1177/193229681300700322] [Medline: 23759410]
27. Schoenfelder E, Moreno M, Wilner M, Whitlock KB, Mendoza JA. Piloting a mobile health intervention to increase physical activity for adolescents with ADHD. Prev Med Rep 2017 Jun;6:210-213 [FREE Full text] [doi: 10.1016/j.pmedr.2017.03.003] [Medline: 28373931]
28. Whittemore R, Jeon S, Grey M. An internet obesity prevention program for adolescents. J Adolesc Health 2013 Apr;52(4):439-447 [FREE Full text] [doi: 10.1016/j.jadohealth.2012.07.014] [Medline: 23299003]
29. Larsen B, Benitez T, Cano M, Dunsiger SS, Marcus BH, Mendoza-Vasconez A, et al. Web-based physical activity intervention for Latina adolescents: feasibility, acceptability, and potential efficacy of the Niñas Saludables study. J Med Internet Res 2018 May 09;20(5):e170 [FREE Full text] [doi: 10.2196/jmir.9206] [Medline: 29743151]
30. Mendoza JA, Baker KS, Moreno MA, Whitlock K, Abbey-Lambertz M, Waite A, et al. A Fitbit and Facebook mHealth intervention for promoting physical activity among adolescent and young adult childhood cancer survivors: A pilot study. Pediatr Blood Cancer 2017 Dec;64(12):-. [doi: 10.1002/pbc.26660] [Medline: 28618158]
31. Direito A, Jiang Y, Whittaker R, Maddison R. Apps for IMproving FITness and Increasing Physical Activity Among Young People: the AIMFIT pragmatic randomized controlled trial. J Med Internet Res 2015 Aug 27;17(8):e210 [FREE Full text] [doi: 10.2196/jmir.4568] [Medline: 26316499]
32. Kennedy SG, Smith JJ, Morgan PJ, Peralta L, Hilland TA, Eather N, et al. Implementing resistance training in secondary schools: a cluster RCT. Med Sci Sports Exerc 2017 Aug 25 (forthcoming). [doi: 10.1249/MSS.0000000000001410] [Medline: 28846564]
33. Slootmaker SM, Chinapaw MJ, Seidell JC, van Mechelen W, Schuit AJ. Accelerometers and internet for physical activity promotion in youth? Feasibility and effectiveness of a minimal intervention [ISRCTN93896459]. Prev Med 2010 Jul;51(1):31-36. [doi: 10.1016/j.ypmed.2010.03.015] [Medline: 20380847]
34. Smith JJ, Morgan P, Plotnikoff RC, Dally KA, Salmon J, Okely AD, et al. Smart-phone obesity prevention trial for adolescent boys in low-income communities: the ATLAS RCT. Pediatrics 2014 Sep;134(3):e723-e731 [FREE Full text] [doi: 10.1542/peds.2014-1012] [Medline: 25157000]

35. Lau EY, Lau PW, Chung PK, Ransdell LB, Archer E. Evaluation of an Internet-short message service-based intervention for promoting physical activity in Hong Kong Chinese adolescent school children: a pilot study. Cyberpsychol Behav Soc Netw 2012 Aug;15(8):425-434. [doi: 10.1089/cyber.2012.0161] [Medline: 22897473]
36. Gaudet J, Gallant F, Bélanger M. A Bit of Fit: minimalist intervention in adolescents based on a physical activity tracker. JMIR Mhealth Uhealth 2017 Jul 06;5(7):e92 [FREE Full text] [doi: 10.2196/mhealth.7647] [Medline: 28684384]
37. Sousa P, Fonseca H, Gaspar P, Gaspar F. Controlled trial of an Internet-based intervention for overweight teens (Next.Step): effectiveness analysis. Eur J Pediatr 2015 Sep;174(9):1143-1157 [FREE Full text] [doi: 10.1007/s00431-015-2502-z] [Medline: 25772743]
38. Jimoh F, Lund EK, Harvey LJ, Frost C, Lay WJ, Roe MA, et al. Comparing diet and exercise monitoring using smartphone app and paper diary: a two-phase intervention study. JMIR Mhealth Uhealth 2018 Jan 15;6(1):e17 [FREE Full text] [doi: 10.2196/mhealth.7702] [Medline: 29335239]
39. Adamo KB, Prince SA, Tricco AC, Connor-Gorber S, Tremblay M. A comparison of indirect versus direct measures for assessing physical activity in the pediatric population: a systematic review. Int J Pediatr Obes 2009;4(1):2-27. [doi: 10.1080/17477160802315010] [Medline: 18720173]
40. Gold J, Pedrana AE, Stoove MA, Chang S, Howard S, Asselin J, et al. Developing health promotion interventions on social networking sites: recommendations from The FaceSpace Project. J Med Internet Res 2012 Feb 28;14(1):e30 [FREE Full text] [doi: 10.2196/jmir.1875] [Medline: 22374589]
41. Azar KM, Lesser LI, Laing BY, Stephens J, Aurora MS, Burke LE, et al. Mobile applications for weight management: theory-based content analysis. Am J Prev Med 2013 Nov;45(5):583-589. [doi: 10.1016/j.amepre.2013.07.005] [Medline: 24139771]
42. Eng DS, Lee JM. The promise and peril of mobile health applications for diabetes and endocrinology. Pediatr Diabetes 2013 Jun;14(4):231-238 [FREE Full text] [doi: 10.1111/pedi.12034] [Medline: 23627878]
43. Hogan NM, Kerin MJ. Smart phone apps: smart patients, steer clear. Patient Educ Couns 2012 Nov;89(2):360-361. [doi: 10.1016/j.pec.2012.07.016] [Medline: 22897983]
44. Breton ER, Fuemmeler BF, Abroms LC. Weight loss-there is an app for that! But does it adhere to evidence-informed practices? Transl Behav Med 2011 Dec;1(4):523-529 [FREE Full text] [doi: 10.1007/s13142-011-0076-5] [Medline: 24073074]
45. Arnhold M, Quade M, Kirch W. Mobile applications for diabetics: a systematic review and expert-based usability evaluation considering the special requirements of diabetes patients age 50 years or older. J Med Internet Res 2014 Apr 09;16(4):e104 [FREE Full text] [doi: 10.2196/jmir.2968] [Medline: 24718852]
46. 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 01;6(5):1197-1206 [FREE Full text] [doi: 10.1177/193229681200600525] [Medline: 23063047]
47. Kesiraju L, Vogels T. Flurry. 2018. Health & Fitness App Users Are Going the Distance URL: https://flurrymobile.tumblr.com/post/165079311062/health-fitness-app-users-are-going-the-distance [accessed 2019-01-30] [WebCite Cache ID 75ogXz6eC]
48. Statista. 2018. Total mHealth market size worldwide 2025 forecast URL: https://www.statista.com/statistics/938544/mhealth-market-size-forecast-globally/ [accessed 2019-01-30] [WebCite Cache ID 75ogfX0ci]
49. Xian Y, Xu H, Xu H, Liang L, Hernandez A, Wang T, et al. An initial evaluation of the impact of Pokémon GO on physical activity. J Am Heart Assoc 2017 May 16;6(5):- [FREE Full text] [doi: 10.1161/JAHA.116.005341] [Medline: 28512111]
50. Turner T, Spruijt-Metz D, Wen CK, Hingle MD. Prevention and treatment of pediatric obesity using mobile and wireless technologies: a systematic review. Pediatr Obes 2015 Dec;10(6):403-409 [FREE Full text] [doi: 10.1111/ijpo.12002] [Medline: 25641770]

Abbreviations

mHealth: mobile health
MVPA: moderate-to-vigorous physical activity
PA: physical activity
SMS: short message service

Edited by G Eysenbach; submitted 06.08.18; peer-reviewed by L Dugas, N Ridgers, R Newton, R Booth; comments to author 21.09.18; revised version received 07.12.18; accepted 23.01.19; published 12.02.19.

Please cite as:
Lee AM, Chavez S, Bian J, Thompson LA, Gurka MJ, Williamson VG, Modave F
Efficacy and Effectiveness of Mobile Health Technologies for Facilitating Physical Activity in Adolescents: Scoping Review
JMIR Mhealth Uhealth 2019;7(2):e11847
URL: http://mhealth.jmir.org/2019/2/e11847/
doi:10.2196/11847
PMID:30747716

Review

Technology-Supported Self-Guided Nutrition and Physical Activity Interventions for Adults With Cancer: Systematic Review

Nicole Kiss[1,2], BSc, MNutDiet, PhD; Brenton James Baguley[3,4], BHSc, MNutrDiet; Kylie Ball[1], BA, PhD; Robin M Daly[1], PhD; Steve F Fraser[1], PhD; Catherine L Granger[5,6], PhD; Anna Ugalde[7], PhD

[1]Institute for Physical Activity and Nutrition, Deakin University, Geelong, Australia

[2]Department of Cancer Experiences Research, Peter MacCallum Cancer Centre, Melbourne, Australia

[3]School of Psychology, Faculty of Health, Deakin University, Burwood, Australia

[4]School of Human Movement and Nutrition Sciences, The University of Queensland, St Lucia, Australia

[5]School of Physiotherapy, Faculty of Medicine, Dentistry and Health Sciences, University of Melbourne, Melbourne, Australia

[6]Department of Physiotherapy, The Royal Melbourne Hospital, Melbourne, Australia

[7]School of Nursing and Midwifery, Faculty of Health, Deakin University, Burwood, Australia

Corresponding Author:
Nicole Kiss, BSc, MNutDiet, PhD
Institute for Physical Activity and Nutrition
Deakin University
Geelong,
Australia
Phone: 61 3 9246 8858
Email: nicole.kiss@deakin.edu.au

Abstract

Background: Nutrition and physical activity interventions are important components of cancer care. With an increasing demand for services, there is a need to consider flexible, easily accessible, and tailored models of care while maintaining optimal outcomes.

Objective: This systematic review describes and appraises the efficacy of technology-supported self-guided nutrition and physical activity interventions for people with cancer.

Methods: A systematic search of multiple databases from 1973 to July 2018 was conducted for randomized and nonrandomized trials investigating technology-supported self-guided nutrition and physical activity interventions. Risk of bias was assessed using the Cochrane Risk of Bias tool. Outcomes included behavioural, health-related, clinical, health service, or financial measures.

Results: Sixteen randomized controlled trials representing 2684 participants were included. Most studies were web-based interventions (n=9) and had a 12-week follow-up duration (n=8). Seven studies assessed dietary behaviour, of which two reported a significant benefit on diet quality or fruit and vegetable intake. Fifteen studies measured physical activity behaviour, of which eight studies reported a significant improvement in muscle strength and moderate-to-vigorous physical activity. Four of the nine studies assessing the health-related quality of life (HRQoL) reported a significant improvement in global HRQoL or a domain subscale. A significant improvement in fatigue was found in four of six studies. Interpretation of findings was influenced by inadequate reporting of intervention description and compliance.

Conclusions: This review identified short-term benefits of technology-supported self-guided interventions on the physical activity level and fatigue and some benefit on dietary behaviour and HRQoL in people with cancer. However, current literature demonstrates a lack of evidence for long-term benefit.

Trial Registration: PROSPERO CRD42017080346; https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=80346

(JMIR Mhealth Uhealth 2019;7(2):e12281) doi:10.2196/12281

KEYWORDS

cancer; diet; exercise; nutrition; physical activity; self-guided interventions; technology

Introduction

It is estimated that over 32 million people are living with cancer worldwide [1], and the predicted global incidence of cancer is estimated to increase from 14 million new cases in 2012 to more than 17 million in 2020 [1]. Cancer is now considered a chronic disease, and the number of cancer survivors in the United States is estimated to exceed 20 million by 2026 [2]. This rapid increase is adding pressure on health care systems to cope with the growing number of people requiring treatment while maintaining high-quality health care. As a result, there is a need to consider alternative, easily accessible, and flexible models of delivering care, in particular, supportive care interventions, to people with cancer in order to reduce the demand for clinical resources while maintaining optimal clinical and health outcomes.

Nutrition and physical activity interventions are vital components of cancer care [3]. Prevalence studies demonstrate that as many as 40% of cancer patients are affected by malnutrition, which is associated with increased mortality and health care costs and poor health-related quality of life (HRQoL) [4-6]. Sarcopenia, the loss of skeletal muscle mass, strength, or function, is present in 25%-57% of cancer patients and is an independent predictor of survival [7-9]. Nutrition and exercise interventions can improve muscle mass, muscle strength, physical function, nutritional status, fatigue, and HRQoL for people undergoing cancer treatment [4,5,10,11]. Preliminary evidence also shows potential for improved survival from nutrition interventions delivered throughout treatment [12,13]. Conversely, obesity after treatment completion is associated with reduced cancer survival; nonetheless, less than 20% of cancer survivors meet the dietary recommendations and less than 50% meet the physical activity recommendations, demonstrating a clear role for interventions to support healthy eating behaviors and increased physical activity [14-19]. Nutrition and physical activity interventions, particularly in survivors of breast cancer and men with prostate cancer treated with androgen-deprivation therapy, have been shown to produce clinically meaningful, beneficial weight loss or improvements in muscle mass and cardiometabolic health outcomes (eg, reduced insulin resistance and low-density lipoprotein cholesterol levels) [20,21]. Further, exercise training following a cancer diagnosis has a protective effect on cancer-specific mortality, cancer recurrence, and all-cause mortality [10].

Technology-based platforms such as the internet, mobile phone or tablet apps, and telehealth and wearable devices provide a unique opportunity to deliver broad-reaching interventions and health care to people with cancer. In the general population, technology-supported nutrition and physical activity interventions have demonstrated positive, albeit modest, benefits of increasing physical activity levels, reducing dietary fat intake, and increasing fruit and vegetable consumption [22-24]. In cancer populations, a recent systematic review and meta-analysis investigated digital health behavior-change interventions, which targeted diet and physical activity in cancer survivors [25]. The authors reported a mean improvement of 41 minutes per week of moderate-to-vigorous physical activity levels (P=.006) and a pooled reduction in body mass index (BMI)/weight of –0.23 (P=.011) with the use of digital interventions. However, technology-supported interventions may still require considerable facilitation by a health professional and use of clinical resources, inhibiting their practicality within usual care [26]. In the context of the growing demand for supportive health care in the cancer setting, technology-supported nutrition and physical activity interventions that are primarily self-guided have the potential to deliver broad-reaching interventions using minimal clinician resources. “Self-guided” refers to interventions delivered with minimal or no facilitation by a health professional and is a key component for improving sustainable dietary and physical activity behavior change. However, the efficacy of self-guided nutrition and physical activity interventions for people with cancer is yet to be established. This systematic review aims to describe and appraise the literature on the efficacy of technology-supported self-guided nutrition and physical activity interventions for people with cancer.

Methods

Reporting Guidelines

This systematic review was performed in accordance with the reporting requirements of the PRISMA statement [27]. The protocol was registered in the PROSPERO database with the reference number CRD42017080346 [28].

Search Strategy

The following databases were searched for peer-reviewed, English-language papers from 1973 to July 2018: Medline Complete, Scopus, CINAHL, EMBASE, Cochrane Library, and SPORTDiscus. The search terms included (Cancer OR oncology OR tumour OR malignancy OR malignant neoplasm) and (online OR internet OR “web-based” OR website OR ehealth OR app OR apps OR application OR mobile application OR smartphone OR mobile phone OR cell phone OR Mhealth OR telehealth OR telemedicine OR technology) and (nutrition OR diet OR physical activity OR exercise). Reference lists of relevant articles were manually searched to identify additional articles.

Selection Criteria

Studies were eligible for inclusion if they were original research studies on adult participants aged ≥18 years who were diagnosed with any type of cancer at any stage prior to, during, or after cancer treatment, including cancer survivors, and received any treatment modality. Studies were included if they investigated a technology-supported nutrition and physical activity intervention that was largely self-guided and if the technology was accessed primarily outside the clinical setting. An intervention was deemed self-guided when there was minimal or no facilitation by a clinician. Minimal facilitation could encompass activities such as occasional email reminders, an introductory session on navigating the technology platform, or initial exercise prescription. Technology platforms for intervention delivery could be online, mobile phone, or tablet apps or wearable technology. The intervention content needed to focus on nutrition and physical activity. For the purpose of this paper, physical activity also includes exercise interventions. Interventions that included nutrition or physical activity as part

of a broader suite of lifestyle or wellness interventions were excluded, unless nutrition or physical activity was a major component of the intervention, comprising at least a quarter of the content together. Studies were required to have a comparator group that could include active controls such as usual care, waitlist controls, or no treatment. Pilot studies that focused on feasibility and acceptability alone were excluded. Outcomes were any measures focused on health-related (eg, quality of life and fatigue), clinical (eg, weight and body composition), health service (eg, resource utilization, hospital admissions, and patient satisfaction), behavioral (eg, dietary intake and physical activity level), or financial outcomes (eg, cost to patients).

Data Extraction and Quality Assessment

Titles and abstracts were screened for eligibility by two independent reviewers (NK and BB) to exclude articles that were clearly irrelevant. Full-text articles were retrieved and the selection criteria were applied independently by the same two reviewers. Any discrepancies were resolved by discussion.

Each included study was assessed for bias independently by the first author and one of the other authors using the Cochrane Risk of Bias tool [29]. Seven categories were examined to assess selection bias, performance bias, detection bias, attrition bias, reporting bias, and other sources of bias and rated as high risk, low risk, or unclear risk. Ratings were compared and a consensus was reached through discussion. As some of the categories were open to interpretation, all authors agreed on the following: (1) a rating of low risk for performance bias as the nature of these interventions indicated that blinding was not possible and the findings were not likely to be influenced by the lack of blinding, unless another reason was identified to rate otherwise and (2) in the absence of blinding, detection bias was rated as low risk if the outcomes were objectively assessed, and as high risk if outcomes were behavioral or patient reported. All studies were included regardless of the bias rating. The Behaviour Change Taxonomy by Michie et al (2013) was used to describe the type of behavior-change techniques reported within each included study [30]. The heterogeneity of studies and diversity of outcomes indicated that the quantitative synthesis of data was not appropriate.

Results

Study Selection

The literature search identified a total of 3346 articles, and eight more articles were identified from the reference lists of relevant papers. Of the total, 18 articles representing 16 studies were included in the systematic review. Two studies were reported in two separate papers each. One of these studies reported on different outcomes in separate papers [31,32]; the other study reported on outcomes at 6 months [33] and 12 months [34] in separate papers. All articles are included, but the findings are presented as a single study. Figure 1 describes the selection process.

Overview of Included Studies

Tables 1 and 2 present the characteristics of all included studies. All studies were randomized controlled trials presenting data from 2684 participants. The duration of follow-up ranged from 10 weeks to 12 months, and five studies reported follow-up duration of ≥6 months. Comparator groups included waitlist controls in six studies [32,35-38], active controls in five studies [39-43], usual care in four studies [33,44-46], and no treatment controls in one study [47]. Three studies used more than one comparator group, one of which used two active controls [45], one used a no-treatment control and two active controls [46], and one used a waitlist control and an active control [37].

Figure 1. Selection process.

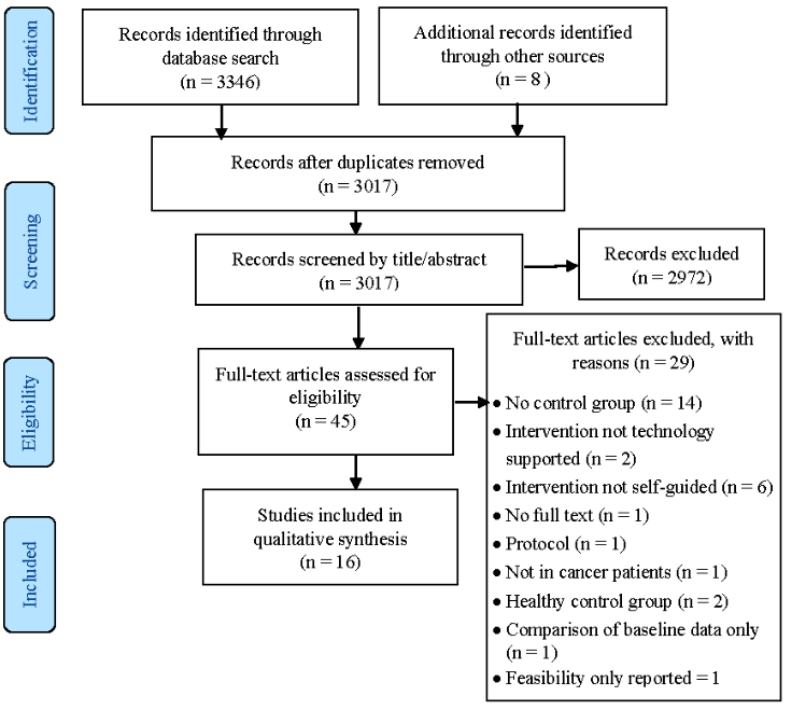


Table 1. Characteristics of the included studies.

Author, year, country	Participants	Study design	Behavior-change technique[a]	Intervention		Type of control group
				Self-guided component	Facilitated component	
Bantum et al, 2014, United States [36]	Cancer survivors between 4 wk and 5 y posttreatment N=352 Consent rate=56% Retention rate=86%	Two-arm RCT[b]	Goal setting (behavior) Feedback on behavior Self-monitoring of outcome(s) of behavior Social support (unspecified) Instructions on how to perform a behavior Social comparison Credible source	Six web-based modules of 22 topics, including healthy eating, exercise, stress management, communication, and fatigue management, accessed over 6 weeks	Baseline training of the website content and weekly goal setting of health behavior change was discussed with a facilitator on the website	Waitlist control
Forbes et al, 2015, Canada [35]	Breast or prostate cancer survivors N=95 Consent rate=32% Retention rate=88%	Two-arm RCT	Goal setting (behavior) Problem solving Feedback on behavior Self-monitoring of behavior Social support (unspecified) Instructions on how to perform a behavior Social comparison Credible source Social reward	Nine web-based exercise behavior-change modules, sequentially published over a 9-week period, including exercise safety, goal setting, benefits, barriers, and strategies on compliance to exercise. Participants logged step counts on an accelerometer website	Weekly email updates about new information and brief summary of previous week PA[c] level	Waitlist control
Galiano-Castillo et al, 2016 & 2017, Spain [31,32]	Breast cancer survivors N=81 Consent rate=89% Retention rate=88%	Two-arm RCT	Feedback on behavior Instructions on how to perform a behavior Demonstration of the behavior Credible source	Web-based 8-week intervention providing access to 24 exercise sessions, 3 per week, tailored to participants, including warm-up, resistance and aerobic exercise, and cool down to meet The American College of Sports Medicine recommendations for cancer survivors	Baseline familiarity with the content to individualize the exercise training. Instant messages and video conference were available if requested by the participant for further exercise support	Waitlist control
Gnagarella et al, 2016, Italy [39]	Any cancer diagnosis during or after treatment N=125 Consent rate=28% Retention rate=59%	Two-arm RCT	Instructions on how to perform a behavior Information about health consequences Credible source	Web-based 6-month intervention with access to weekly forums, blogs, and content on healthy eating to reduce treatment symptoms, control weight loss, or maintaining body mass and guidelines for healthy eating	Option to interact with facilitator upon request	Active control
Kanera et al, 2016 & 2017, Netherlands [33,34]	Cancer survivors between 4 wk and 56 wk posttreatment N=462 Consent rate=36% Retention rate=82.5%	Two-arm RCT	Goal setting (behavior) Problem solving Action planning Discrepancy between current behavior and goals Feedback on behaviors Self-monitoring of behaviors Instructions on how to perform a behavior Information about health consequences Social comparison Credible source Pros and cons	Web-based intervention over 6 months with 8 modules of videos, written content, goal setting, action planning, and problem identification including nutrition, exercise, smoking, fatigue, anxiety, and depression	No facilitation	Usual care

Author, year, country	Participants	Study design	Behavior-change technique[a]	Intervention		Type of control group
				Self-guided component	Facilitated component	
Krebs et al, 2017, United States [44]	Breast or prostate cancer survivors N=86 Consent rate=53% Retention rate=79%	Two-arm RCT	Goal setting (behavior) Problem solving Feedback on behavior Social support (unspecified) Instructions on how to perform a behavior Credible source	DVD[d]-based intervention over 12 weeks of nutrition and exercise advice, derived from the American Cancer Society guidelines for cancer survivors. Components included enhancing knowledge, developing positive expectations, reducing barriers, and supporting self-efficacy	No facilitation	Usual care
Lee et al, 2014, South Korea [40]	Breast cancer survivors N=59 Consent rate=50% Retention rate=97%	Two-arm RCT	Goal setting (behavior) Problem solving Action planning Discrepancy between current behavior and goal Feedback on behavior Self-monitoring of behavior Instructions on how to perform the behavior Information about health consequences Credible source	Web-based 12-week intervention with participants encouraged by SMS[e] to access the tailored Web-based content biweekly, which covered enhancing exercise and diet behaviors, barriers and diet, or exercise guidelines for cancer survivors	30-min training on using the Web-based platform	Active control
Mayer et al, 2014, United States [49]	Colon cancer survivors between 6 wk postoperation and 12 mo postdiagnosis N=284 Consent rate=54% Retention rate=80%	Two-arm RCT	Goal setting (behavior) Self-monitoring of behavior Social support (unspecified) Instructions on how to perform a behavior Credible source	Six-month mobile app physical activity intervention including skill building, information provision, and support services designed to increase daily activity levels. New content added over the period of the intervention	Personal trainer available to answer questions in discussion groups and initiate individually tailored private messages to inactive participants	Control
Ormel et al, 2018, Netherlands [48]	Any cancer diagnosis during active systemic treatment or survivorship N=32 Consent rate=53% Retention rate=96%	Two-arm RCT	Self-monitoring of behavior Prompts/cues Credible source	12-wk mobile app physical activity intervention consisting of physical activity advice and self-monitoring	No facilitation	Usual care
Pope et al, 2018, United States [41]	Breast cancer survivors who completed treatment 3 mo to 10 y prior N=30 Consent rate=77% Retention rate=67%	Two-arm RCT	Problem solving Social support (unspecified) Instructions on how to perform a behavior Demonstration of the behavior Credible source	10-wk smart watch and Facebook intervention including a strength and aerobic training program plus access to a private Facebook page for delivery of health education tips	Tutorial on the use of the smart watch and Facebook page. Participants contacted by researchers every other week to encourage continuation of the intervention	Active control

Author, year, country	Participants	Study design	Behavior-change technique[a]	Intervention		Type of control group
				Self-guided component	Facilitated component	
Rabin et al, 2011, United States [42]	Young adult cancer survivors (18-39 y) N=18 Consent rate=72% Retention rate=94%	Two-arm RCT	Goal setting (behavior) Feedback on behavior Self-monitoring of behavior Social support (unspecified) Instructions on how to perform the behavior Information about health consequences Credible source	Web-based 12-wk intervention with access to purposefully designed website providing information on weekly goal setting and PA resources individually tailored to the stage of change and feedback generated from monthly questionnaires	One-off induction to the website	Active control
Sajid et al, 2016, United States [45]	Men aged ≥65 y with prostate cancer on hormone therapy N=19 Consent rate, not reported Retention rate=68%	Three-arm RCT	Self-monitoring of behavior Instructions on how to perform the behavior Demonstration of the behavior Graded tasks	Intervention arm 1: Wearable device-based 6-wk intervention plus instructions for safe aerobic and tailored resistance training exercise, and total steps per day Intervention arm 2: Wearable device plus an individually tailored Wii-Fit exercise program for 6 wk	Both arms: 45-min exercise introduction with an accredited exercise physiologist and weekly reminders for completion of diary and pedometer assessments	Control
Uhm et al, 2017, South Korea [43]	Breast cancer survivors N=356 Consent rate, not reported Retention rate=95%	Two-arm RCT	Feedback on behavior Self-monitoring Instructions on how to perform the behavior Graded tasks	Mobile app and wearable device-based 12-wk intervention. Content and instructive videos with PA goals were available on the mobile app based on a 2MWT[f] at baseline. Content on the mobile app was additional to the facilitated exercise program	Physiatrists prescribed the amount of 90-150 min of moderate-intensity aerobic exercise and 4-8 resistance training exercises	Active control
Vallance et al, 2007, Canada [46]	Breast cancer survivors N=377 Consent rate=25% Retention rate=89.7%	Four-arm RCT	Self-monitoring of behavior Instructions on how to perform the behavior Credible source	All groups were instructed to perform 30-min MVPA[g] 5 d/wk over a 12-wk intervention. Intervention arm 1: Wearable device and documents to record daily step total Intervention arm 2: Wearable device plus self-perusal of guidelines for exercising among breast cancer survivors and a document to record daily step total Intervention arm 3: Self-perusal of guidelines for exercising among breast cancer survivors	No facilitation	Control

Author, year, country	Participants	Study design	Behavior-change technique[a]	Intervention		Type of control group
				Self-guided component	Facilitated component	
Valle et al, 2017, United States [37]	Breast cancer survivors N=35 Consent rate=70% Retention rate=94%	Three-arm RCT	Goal setting (behavior) Problem solving Goal setting (outcome) Feedback on behavior Self-monitoring of behavior Monitoring of outcome(s) of behavior without feedback Instructions on how to perform a behavior Information about health consequences Graded tasks Credible source Self-talk	Intervention arm 1: Web-based with mobile companion app providing 24-wk intervention on nutrition and PA for body weight and weekly email of standardized content on behavior-change strategies related to weight loss. Web or mobile app accessed to log body weight and PA Intervention arm 2: Same as intervention arm 1 plus a wearable device and additional education on steps per day and meeting the steps-per-day recommendations	Both intervention arms: 1-h face-to-face nutrition consult, 24 weekly emails of behavior change to reach 150-225 min of moderate-intensity PA and for energy reduction by 100 kcal through dietary intake, and tailored feedback on weight were provided	Waitlist control
Yun et al, 2012, Korea [38]	Cancer survivors <24 mo posttreatment N=273 Consent rate=28% Retention rate=89%	Two-arm RCT	Monitoring outcome(s) of behavior by others without feedback Social support (unspecified) Instructions on how to perform a behavior Credible source	Web-based 12-wk intervention via seven education modules with personally tailored information on energy conservation, PA, nutrition, sleep hygiene, pain, distress management, and information on fatigue	No facilitation	Waitlist control

[a]Determined from Michie et al [30].

[b]RCT: randomized controlled trial.

[c]PA: physical activity.

[d]DVD: digital video disk.

[e]SMS: short message service.

[f]2MWT: 2-minute walk test.

[g]MVPA: moderate-to-vigorous physical activity.

Sample sizes ranged from 18 to 462 participants. Thirteen studies included cancer survivors of a variety of cancer types, with a majority of breast or prostate cancer survivors [32,33,35-38,40-44,46,48]; one study included participants receiving active cancer treatment [45]; and two studies included participants who were at any cancer stage prior to, during, or after cancer treatment [39,49].

Nine studies examined a physical activity intervention [32,35,41-43,45,46,48,49], one study examined a nutrition intervention [39], and six studies examined a combined physical activity and nutrition intervention [33,36-38,40,44]. Of the studies that examined a combined intervention, three were part of a broader suite of interventions encompassing areas such as fatigue management, stress management, and pain; however, nutrition and physical activity were major components of these interventions [33,36,38].

Participation rates for the studies ranged from 25% to 89% (median=53%), and four studies reported a consent rate of ≥70%. However, once enrolled in the studies, participant retention was generally high, with 12 of the 16 studies retaining >80% of participants over the study duration. Methods for measuring uptake and adherence to the intervention, including the proportion of patients who accessed the intervention content and the average number of logins to the intervention website, varied considerably, making comparison between studies challenging. Four studies did not report on uptake or intervention adherence [39,43,45,49].

Studies used between 3 and 11 (median=5.5) behavior-change techniques, and provision of instructions to perform a behavior was the most common technique used in all but one of the included studies. Other commonly used techniques included Credible Source to deliver information (14 studies), self-monitor behavior (11 studies), and provide feedback on behavior (9 studies). In some studies, insufficient description of the intervention prevented a full analysis of the behavior-change techniques employed [38,39,46,49]. Insufficient description of the intervention also limited the ability to assess the quality and content of the interventions. However, 14 of the included studies used a behavior change technique—Credible Source—meaning the content was provided by a source that was considered reliable.

Table 2. Outcome measures and findings of the included studies.

Author, year, country	Outcomes	Intervention uptake	Between-group findings
Bantum et al, 2014, United States [36]	Measured at baseline and 6 mo Primary outcomes: fruit and vegetable intake, PA[a], depression, fatigue, and insomnia	67% participants accessed all sessions. Mean number of sessions accessed=5.3 (SD 1.28)	Compared to control group, the intervention group showed a significant improvement in insomnia (9.6 to 10.1 vs 9.6 to 9.2; *P*=.03), strenuous exercise (29 min/wk [no change] vs 32 to 51 min/wk; *P*=.01), stretching (26 to 25 min/wk vs 31 to 46 min/wk; *P*=.01)
Forbes et al, 2015, Canada [35]	Measured at baseline and 10 wk Primary outcome: feasibility of the intervention Secondary outcomes: PA and quality of life	67% viewed the modules at least once. Average number of logins over the 9-wk period=10.3	Improved mental health in the waitlist controls compared to the intervention group (*d*=0.37, *P*=.01)
Galiano-Castillo et al, 2016 & 2017, Spain [31,32]	Measured at baseline, 8 wk, and 6 mo Primary outcomes: strength, quality of life, pain, and fatigue	Adherence rate=93.9% of the scheduled sessions	Compared to the control group, at 8 wk, the intervention group showed improved isometric abdominal (*P*<.001, *d*=1.02), back (*P*<.001, *d*=1.31), lower-body (*P*=.001, *d*=–0.81); and hand grip strength (*P*=.006, *d*=0.66); 6MWT[b] (*P*<.001, *d*=0.92); global quality of life (*P*<.001, *d*=0.89); physical functioning (*P*<.001, *d*=0.9); role functioning (*P*=.001, *d*=0.78); cognitive functioning (*P*=.002, *d*= 0.75); pain severity (*P*=.001, *d*=–0.82); fatigue (*P*<.001, *d*=–0.89); and pain interference (*P*=.05, *d*=–0.47). At 6 months, the intervention effect was maintained for all except role functioning and pain severity
Gnagarella et al, 2016, Italy [39]	Measured at baseline and 6 mo Primary outcomes: nutrition knowledge, food consumption, and quality of life	Not reported	No significant differences between groups for nutrition knowledge or food consumption. Compared to the control group, the intervention group had improved role functioning in the quality of life scale at 6 mo (–6.3 vs 5.1, *P*=.02)
Kanera et al, 2016 & 2017, Netherlands [33,34]	Measured at baseline, 6 mo, and 12 mo Primary outcomes: PA and dietary behavior	Average of 2.23 (SD 1.53) modules followed. PA activity module followed by 25% participants. Diet module followed by 62% participants	Significantly increased vegetable consumption (*P*=.023) and moderate physical activity (*P*=.04) in the intervention group compared to the control group after 6 mo. This effect was sustained at 12 mo for physical activity but not vegetable consumption
Krebs et al, 2017, United States [44]	Measured at baseline and 12 wk Primary outcomes: fruit and vegetable intake and PA	72% viewed the DVD[c], 50% completed the full DVD	No significant between-group differences
Lee et al, 2014, South Korea [40]	Measured at baseline and 12 wk Primary outcomes: diet composition and PA Secondary outcomes: quality of life, anxiety, depression, fatigue, motivational readiness, and self-efficacy	89% of patients consistently participated in the program throughout the intervention	Compared to control group, after 12 wk, the intervention group showed significantly improved proportion of participants meeting the recommendations of 150 min/wk moderate-intensity exercise (35.7% vs 65.5%, *P*<.0001) and five servings of fruit or vegetables/d (32.1% vs 55.2%, *P*=.001) and improved the diet-quality index (9.6 vs 11.1, *P*=.001), physical functioning (75.9 vs 83.6, *P*=.02), fatigue (15.3 vs 13.5, *P*=.03), and self-efficacy for exercise (*P*=.02) and fruit and vegetable intake (*P*=.02)
Mayer et al, 2014, United States [49]	Measured at baseline and 3, 6, and 9 mo Primary outcome: physical activity Secondary outcomes: distress and quality of life	93.8% participants described as users (accessed system at least once). Of the 180 days of possible use, mean use=55.3 days (SD 50.0)	No significant differences between the intervention and control groups
Ormel et al, 2018, Netherlands [48]	Measured at baseline, and 6 and 12 wk Primary outcome: feasibility of the intervention Secondary outcome: PA	Not reported	Compared to the control group, the intervention group had significantly increased total minutes of PA at 6 wk (2348 min/wk vs 3773 min/wk, *P*=.04), but there was no difference in sedentary time between groups. There were no significant between-group differences at 12 weeks

Author, year, country	Outcomes	Intervention uptake	Between-group findings
Pope et al, 2018, United States [41]	Measured at baseline and 10 wk Primary outcomes: PA and energy expenditure Secondary outcomes: anthropometry, body composition, cardiorespiratory fitness, quality of life, and psychosocial constructs	Participants wore the smart watch 6-7 d/wk and accessed the Facebook page 1.2 times/wk on an average	Compared to the intervention, the control group demonstrated improved physical activity-related social support and reduced barriers. No other significant between-group differences were observed
Rabin et al, 2011, United States [42]	Measured at baseline and 12 wk Primary outcomes: feasibility and acceptability Secondary outcomes: PA, mood, and fatigue	Average number of website logins=14.75 (SD 8.46)	Compared to the control group, at 12 weeks, there was a medium effect of the intervention on increasing MVPA[d] (16.5 min/wk vs 102.5 min/wk, *d*=0.64) and a large effect of the intervention on mood (–5.00 vs –25.86, *d*> 0.80) and fatigue (–3.30 vs –11.43, *d*>0.80). However, these differences were not statistically significant
Sajid et al, 2016, United States [45]	Measured at baseline, 6 wk, and 12 wk Primary outcome: physical performance Secondary outcomes: steps per day, lean muscle mass, and chess press repetitions	Not reported	Compared to controls, intervention arm 1 showed greater improvement in physical performance (*P*=.04) and a higher rate of change in steps per day at 12 wk (*P*<.01). There were no additional between-group differences seen either between intervention arm 1 and control or intervention arm 2 and control
Uhm et al, 2017, South Korea [43]	Measured at baseline, 6 wk, and 12 wk Primary outcomes: PA, quality of life, anthropometrical measures, body mass index, blood pressure, functional capacity, and hand grip strength	Not reported	There were no significant between-group differences at 6 and 12 wk
Vallance et al, 2007, Canada [46]	Measured at baseline and 12 wk Primary outcomes: PA Secondary outcomes: quality of life, fatigue, brisk walking, and objective step count	Participants who received the written material reported on the content 2.1 times on an average Participants who received the pedometer recorded their steps on 83.3% of study days	Compared to the control group, a significant improvement was seen in MVPA at 12 wk in intervention arm 1 (30 min/wk vs 59 min/wk, *P*=.02) and intervention arm 2 (30 min/wk vs 87 min/wk, *P*=.02). Compared to the control group, significant improvements were seen in brisk walking at 12 wk in intervention arm 1 (0 min/wk vs 94 min/wk, *P*<.001), intervention arm 2 (0 min/wk vs 58 min/wk, *P*=.03), and intervention arm 3 (0 min/wk vs 72 min/wk, *P*=.006) were observed. No differences in objective step count. Intervention arm 2 showed significant improvement in quality of life (6.9 vs 1.1, *P*=.003) at 12 wk compared to control
Valle et al, 2017, United States [37]	Measured at baseline, 3 mo, and 6 mo Primary outcome: feasibility Secondary outcomes: PA, body mass index, weight, body composition, and metabolic syndrome biomarkers	Intervention arm 1: 100% participants reported reading some/all/most of the email content and email feedback Intervention arm 2: 90% participants reported reading some/all/most of the email content and email feedback	Intervention arm 2 significantly reduced body mass index (–0.4 vs 0.1, *P*=0.046) compared to controls at 6 mo. Intervention arm 1 maintained HBA_{1c}[e] levels as compared to increased HBA_{1c} levels observed in the control group (0.0 vs 0.15, *P*=.02). No other significant between-group differences were observed

Author, year, country	Outcomes	Intervention uptake	Between-group findings
Yun et al, 2012, Korea [38]	Measured at baseline and 12 wk Primary outcome: fatigue Secondary outcomes: nutritional status, quality of life, anxiety, and depression	Intervention completed by 83.1% participants	Compared to the control group, the intervention group had a significantly greater decrease in fatigue (group difference=–0.66, *P*=.001, *d*=0.29) and anxiety (–0.9, *P*=.004, *d*=0.33) and significantly greater increase in nutritional status (0.47, *P*=.04, *d*=0.23), global quality of life (5.22, *P*=.02, *d*=0.26), emotional functioning (4.69, *P*=.02, *d*=0.19), social functioning (4.73, *P*=.03, *d*=0.24), and cognitive functioning (6.09, *P*=.002, *d*=0.24) after 12 wk

[a]PA: physical activity.

[b]6MWT: six minute walk test.

[c]DVD: digital video disk.

[d]MVPA: moderate to vigorous physical activity.

[e]HBA_{1c}: hemoglobin A_{1c}.

Technology Platform and Level of Facilitation

The majority of studies (n=9) were delivered as Web-based modules with content covering nutrition and physical activity recommendations [32,33,35-40,42]. Three studies used a wearable device to deliver a physical activity intervention [41,45,46]. The remaining studies used mobile apps (n=3) or digital video disk (n=1) to deliver content on nutrition and physical activity recommendations [43,44,48,49]. In five studies, there was no facilitation of the self-guided intervention [33,38,44,46,49]. In the remaining studies, facilitation involved baseline training on the intervention or Web platform [32,36,37,40-43,45], emails sent at varying frequencies to provide content updates, reminders or summaries of completed physical activity [35,37,45], and interaction with a facilitator in a discussion group or upon request [39,48].

Risk of Bias Assessment

The outcome of the bias assessment is presented in Table 3. All included studies were randomized controlled trials. However, as evident from Table 3, there was a high degree of variation in the bias of the included studies, in particular, lack of clarity about the randomization process and few studies included blinding of the outcome assessment.

Behavioral Outcomes

Of the five studies that investigated dietary behaviors, two found improvements in the intervention arm. Kanera et al (2016) reported a significant increase in vegetable consumption after 6 months of use of a personalized web-based intervention, although this increase was not sustained at 12 months postintervention [33,34]. Similarly, Lee et al (2014) observed a significant improvement in diet quality and fruit and vegetable intake following a 12-week Web-based intervention, but did not measure longer-term outcomes [40].

One of the studies where no benefit of intervention was observed on dietary behavior did not specify the inclusion criteria for selecting participants who were not meeting the dietary recommendations [36]. The authors subsequently found they had recruited participants who had better-than-average fruit and vegetable intake and little need for health behavior change [36]. The two remaining studies that found no improvement in dietary behavior were not powered to detect between-group differences [39,44].

In total, 14 studies examined interventions to promote physical activity behaviors, of which eight studies reported positive outcomes. Bantum et al (2014) reported significant increases in both strenuous exercise and stretching at 6 months after a 6-week Web-based intervention including a suite of health behavior-change content [36]. Another 8-week Web-based intervention delivering physical activity recommendations tailored to participants found significant improvements in isometric abdominal, back, and lower-body muscle strength at 2 and 6 months following the intervention [32]. A 12-week mobile app intervention providing physical activity advice reported a significant improvement in the total minutes of physical activity at 6 weeks but no difference at 12 weeks [49].

Kanera et al (2016) found a significant increase in moderate physical activity levels after 6 months of use of a personalized Web-based intervention; however, similar to the effect on dietary behavior in this study, the difference was not sustained at 12 months [33,34]. In a 12-week Web-based intervention providing tailored exercise content to participants, significant improvements were observed in the proportion of participants following moderate-intensity exercise recommendations [40]. Sajid et al (2016) found a higher rate of change in the number of steps per day among participants who received instructions for exercising and used a wearable device in combination as compared to the control group. However, the same effect was not observed when participants followed a tailored Wii-Fit exercise program in combination with the wearable device, although the sample size was small and not powered to detect differences [45]. In another study, participants using a wearable device alone or in combination with access to exercise guidelines experienced significant improvements in weekly moderate-to-vigorous intensity physical activity and brisk walking [46]. However, improvements in brisk walking were also observed in participants with access to exercise guidelines alone [46]. Although underpowered to detect a significant difference in moderate-to-vigorous physical activity levels, another study reported a medium effect (102.5 [intervention] vs 16.5 [control] minutes/week) of a Web-based intervention after 12 weeks [42].

Table 3. Risk of bias for included studies.

Author, year	Random sequence generation	Allocation concealment	Blinding of participants/personnel	Blinding of outcome assessment	Incomplete outcome data	Selective reporting	Other bias
Bantum et al, 2014 [36]	Low	Unclear	Low	High	High	High	High
Forbes et al, 2015 [35]	Low	Unclear	Low	Low	Low	Unclear	Low
Galiano-Castillo et al, 2016 [31], 2017 [32]	Low	Unclear	Low	Low	Low	High	Low
Gnagnarella et al, 2016 [39]	Low	Unclear	Low	High	High	High	High
Kanera et al, 2016 [33], 2017 [34]	Low	Low	High	Low	Low	Low	Low
Krebs et al, 2017 [44]	Low	Low	Low	High	Low	Unclear	Low
Lee et al, 2014 [40]	Low	Low	Low	High	Low	Low	Low
Mayer et al, 2018 [49]	Unclear	Unclear	Low	High	Low	Unclear	Low
Ormel et al, 2018 [48]	Low	Unclear	Low	High	Low	Low	Low
Pope et al, 2018 [41]	Low	Unclear	High	Low	High	Unclear	Unclear
Rabin et al, 2011 [42]	Unclear	Low	Low	High	High	Unclear	Unclear
Sajid et al, 2016 [45]	Unclear	Low	Low	Unclear	Unclear	Unclear	High
Uhm et al, 2017 [43]	High	Low	Low	High	High	Unclear	Low
Vallance et al, 2007 [46]	Low	Low	Low	High	Low	High	Low
Valle et al, 2017 [37]	Low	Low	Low	Low	Low	Low	Unclear
Yun et al, 2012 [38]	Low	Low	Low	High	Low	Low	Unclear

Several studies that did not report positive outcomes had limitations. Although Forbes et al (2015) did not specifically select participants who did not meet the physical activity recommendations, a sub-group analysis showed that their Web-based intervention was more effective in participants who were not meeting the recommendations [35]. Similarly, another study failed to select participants with poor physical activity behaviors [41], and two studies were not powered to detect an effect [37,44]. Uhm et al used an active control group that received the same information as the intervention group but without the support of a mobile app [43]. Mayer et al (2018) considered the possibility that the lack of benefit was due to the short duration of the intervention [48]; however, the 6-month intervention was longer than that included in several studies reporting positive outcomes [32,36,40,49].

Overall, technology-supported self-guided interventions appeared to have some benefit on dietary intake; however, the few studies that assessed this outcome had several limitations. A relatively consistent positive benefit was noted for physical activity, although the long-term benefits remain unknown; only one study reported outcomes beyond 6 months and found no effect at that time point. There were no patterns in the type or number of behavior-change techniques supporting the effective interventions in comparison to those for which no effect was observed. Effectiveness of interventions for physical activity level did not depend on whether physical activity was patient reported or objectively measured.

Clinical Outcomes

Clinical outcomes were reported in four of the included studies [37,38,41,43]. One study examined weight change, lean body mass, fat mass, BMI, and metabolic syndrome biomarkers after 6 months of use of a Web- and mobile app-based nutrition and physical activity intervention [37]. The intervention was tested in two groups: one used the intervention alone and other used the intervention along with wearable technology. The results showed an improvement in the BMI in the latter intervention group as compared to the control group. However, only the group that used the intervention alone showed an improvement in metabolic syndrome biomarkers [37]. Another study using a mobile app intervention and wearable device showed no difference in BMI, arm circumference, handgrip strength, and blood pressure between the intervention group and an active control group that received written information [43]. A further study reported a significant improvement in the nutritional status, measured using the Mini Nutrition Assessment tool, at 12 weeks after the use of a combined Web-based nutrition and physical activity intervention [38]. Pope et al (2018) found no differences in weight or body composition using a 10-week smart watch and Facebook physical activity intervention [41].

Overall, changes in clinical outcomes as a result of technology-supported self-guided interventions were inconsistent but were assessed in only a small number of studies, making it difficult to draw any concrete conclusions.

Health-Related Outcomes

HRQoL was assessed in nine studies, five of which used the European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire - Core 30 [32,38-40,43], three used the Functional Assessment of Cancer Therapies questionnaire [35,38,48], and one used the Patient-Reported Outcome Measurement System [41]. Two studies reported

improved global HRQoL [32,38], and four studies reported improvements in one or more of the HRQoL subscales [32,38-40]. Studies demonstrating the most-beneficial HRQoL outcomes tended to be Web-based, used waitlist controls, involved a physical activity intervention, and reported very high adherence to the intervention. A further three-arm randomized trial by Sajid et al (2016) investigating functional capacity alone through use of a short physical performance battery reported improved physical performance in the wearable device intervention group compared to the control group, but this result was not observed for the wearable device plus Wii Fit arm [45].

Fatigue was measured in six studies, three of which used the Brief fatigue Inventory [36,38,40] and the remaining studies used the Piper Fatigue Scale [32], the fatigue scale from the Functional Assessment of Cancer Therapy-Anemia measurement system [46], and the Profile of Mood States scale [42]. Of the six studies, four reported significantly improved fatigue following the intervention [32,38,40,42]. Of the four studies, one study involved a 12-week Web-based intervention with modules covering sleep, hygiene, pain, distress management, and fatigue in addition to nutrition and physical activity. As such, the positive effect on fatigue may have been related to other components of the intervention. Two of the six studies found no significant difference in fatigue [36,46]. No consistent differences were evident between the studies that did and did not observe an effect on fatigue.

Yun et al (2012) observed a significantly greater decrease in anxiety, but not depression, after their 12-week Web-based intervention. However, similar to the effect on fatigue, the effect on anxiety may have been related to other components of the intervention beyond nutrition and physical activity [38]. Four studies on physical activity or combined nutrition and physical activity interventions measured depression [36,40], anxiety [40], distress [48], or mood disturbance [42], and found no significant between-group differences.

One study assessed the effect of a Web-based intervention of modules covering nutrition, exercise, stress management, communication, and fatigue management on insomnia and found a significant improvement following the 6-week intervention [36]. One study reported significant improvement in pain severity after an 8-week Web-based physical activity intervention as compared to the control arm [32].

Overall, HRQoL and fatigue appeared to improve after technology-supported self-guided interventions, and the majority of studies showed a positive benefit. The effects on mental health, pain, and insomnia require further investigation.

No studies investigated health service use or financial outcomes. In addition, no studies examined or compared outcomes according to key potential moderating factors such as gender, socioeconomic position, or ethnicity. Studies including only female patients with breast cancer were as likely to report positive outcomes as those including both male and female participants.

Discussion

The main finding of this review was that technology-supported self-guided interventions appear to improve physical activity behaviors and fatigue in people with cancer in the short term. Although many studies only measured these outcomes in the short term, a few studies that assessed these outcomes in the longer term found that this benefit was not sustained. There was a minor effect of such interventions on dietary behavior and an inconsistent effect on clinical outcomes such as weight, BMI, body composition parameters, and mental health outcomes. This finding is largely consistent with that of previous systematic reviews investigating digital health or eHealth interventions in cancer survivors, without focusing on the self-guided component [25]. For instance, Roberts et al (2017) reported improved physical activity and BMI in cancer survivors through digital health physical activity and diet interventions [25]. Similarly, Haberlin et al (2018) observed improvement in physical activity following the use of eHealth to promote physical activity among cancer survivors [50]. Both these reviews concluded that the effect of these interventions was promising, but studies using objective measures and assessing the impact on long-term outcomes remain a priority. Similar conclusions were drawn in a systematic review of computer-tailored physical activity and dietary behavior-promotion programs in the general population; this review showed improvements in physical activity and dietary behavior, but these improvements were limited to the short or medium term due to the lack of long-term follow-up in studies [23].

Inconsistent reporting of self-guided interventions is an issue that has resulted in a recent recommendation for the development of a standardized reporting framework for these types of interventions in people with cancer [51]. Overall, self-guided interventions are not well described in the literature, which limits definitive extraction of the key components of the intervention associated with improved health outcomes in people with cancer. In particular, this affected the extraction of the behavior-change techniques used within the interventions, and in some studies, sufficient information was not reported to determine the specific strategies used. A further issue was the reporting of intervention compliance, which was measured in different ways across studies and at times, not reported at all. Therefore, in cases where interventions did not report positive outcomes, it was difficult to determine whether the intervention itself was not effective or whether the lack of effect was related to poor compliance or insufficient participant engagement in the intervention. Similar issues were identified in a 2017 systematic review on self-guided interventions for psychosocial distress in people with cancer [26]. Of note, in 2011, a CONSORT checklist was published to improve the reporting of Web-based and mobile health interventions [52]; however, none of the studies in this systematic review reported the use of this framework.

Two of the studies included in this review did not specifically select participants who had poor health behaviors for the study. However, in their dietary and physical activity intervention group, Bantum et al (2014) found that they had recruited participants who had a better-than-average fruit and vegetable

intake, and thus, there was no need to change their dietary behavior, potentially explaining the lack of effect of the intervention [36]. Similarly, Forbes et al (2015) did not target recruitment to participants who were not meeting physical activity recommendations; however, a subsequent subanalysis revealed that the intervention was more effective in participants who did not meet the physical activity recommendations at baseline [35]. Although only two studies in our review discussed this limitation, the eligibility criteria of the included studies revealed that 10 of the 16 studies did not specifically target participants with poor dietary or physical activity behaviors. Failure to specifically recruit participants requiring supportive care interventions has also been identified as an issue in several previous systematic reviews of supportive care interventions in cancer [26,53]. It is imperative that future research in this area targets people who require support but are currently not able to access the care they require.

Technology-supported self-guided interventions require a high level of self-motivation for participants to engage with the intervention. Compliance, or intervention uptake, provides some indication of participant engagement. However, inconsistencies in how the engagement was reported, as previously discussed, lead to difficulties in comparing studies. Although measured post-intervention, and therefore, not a true indicator of engagement with the intervention, patient satisfaction provides some insight into the acceptability of an intervention, which has implications for adherence and engagement [54]. A number of the studies included in this review measured patient satisfaction using a questionnaire or interview. Although patient satisfaction was high, there was no consistent approach or questionnaire used across studies; as such, comparison between studies was not possible.

Of the 16 studies included in this review, only two involved patients who were undergoing active cancer treatment. Personalized nutrition and physical activity interventions delivered individually by a clinician or other health professionals during cancer treatment have demonstrated benefits on body composition, nutritional status, quality of life, fatigue, and functional outcomes [4,5,20,55]. However, considering the increasing incidence of cancer, a personalized approach to nutrition and physical activity interventions may have limited long-term feasibility. Models for stratifying patients by risk categories are proposed for the management of cancer survivors [56] and may need to be considered for interventions delivered during active treatment. The level of facilitation in the studies included in this review was minimal, making this a potential cost-effective approach, although evaluation of the cost effectiveness was not a feature of any of the studies and resource requirements were not reported.

The strengths of this review are reporting according to the PRISMA guidelines; inclusion of risk of bias assessment; and categorization of the behavior change techniques underpinning the interventions, which has not been included in previous systematic reviews. Limitations include substantial heterogeneity among the included studies in terms of sample size, risk of bias, outcome measures, type and duration of interventions, and use of behavior-change techniques, which restricted our ability to distinguish the components of the interventions that were effective and to complete our meta-analyses.

In summary, this systematic review identified a short-term benefit of technology-supported self-guided interventions with regard to physical activity behavior and fatigue and a small benefit with regard to dietary behavior and HRQoL in people with cancer. However, there was considerable heterogeneity in the quality of the included studies and some heterogeneity along with major methodological limitations, which make interpretation of the findings challenging. Despite the potential of technology-supported interventions, there is a lack of evidence for their long-term benefit, which requires further investigation. Furthermore, a high proportion of studies did not actively target people with poor nutrition or physical activity behaviors. Future studies should ensure that interventions are tested in people requiring improvements in nutrition and physical activity who are not currently able to access the care they require.

Acknowledgments

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. AU is supported by a Victorian Cancer Agency Early Career Seed Grant. CG is supported by a Victorian Cancer Agency Clinical Research Fellowship.

Conflicts of Interest

None declared.

References

1. International Agency for Research on Cancer Globocan Database. Estimated Incidence, Mortality and Prevalence Worldwide in 2012 Fact Sheet URL: http://gco.iarc.fr/today/fact-sheets-cancers [accessed 2018-12-05] [WebCite Cache ID 74S7qoR18]
2. Miller KD, Siegel RL, Lin CC, Mariotto AB, Kramer JL, Rowland JH, et al. Cancer treatment and survivorship statistics, 2016. CA Cancer J Clin 2016 Jul;66(4):271-289 [FREE Full text] [doi: 10.3322/caac.21349] [Medline: 27253694]
3. Kushi LH, Doyle C, McCullough M, Rock CL, Demark-Wahnefried W, Bandera EV, American Cancer Society 2010 Nutrition and Physical Activity Guidelines Advisory Committee. American Cancer Society Guidelines on nutrition and physical activity for cancer prevention: reducing the risk of cancer with healthy food choices and physical activity. CA Cancer J Clin 2012;62(1):30-67 [FREE Full text] [doi: 10.3322/caac.20140] [Medline: 22237782]

4. Arends J, Bachmann P, Baracos V, Barthelemy N, Bertz H, Bozzetti F, et al. ESPEN guidelines on nutrition in cancer patients. Clin Nutr 2017 Dec;36(1):11-48. [doi: 10.1016/j.clnu.2016.07.015] [Medline: 27637832]
5. Isenring E, Zabel R, Bannister M, Brown T, Findlay M, Kiss N. Updated evidence-based practice guidelines for the nutritional management of patients receiving radiation therapy and/or chemotherapy. Nutr Diet 2013;70(4):2013-2024. [doi: 10.1111/1747-0080.1]
6. Marshall KM, Loeliger J, Nolte L, Kelaart A, Kiss NK. Prevalence of malnutrition and impact on clinical outcomes in cancer services: A comparison of two time points. Clin Nutr 2018 May 02. [doi: 10.1016/j.clnu.2018.04.007] [Medline: 29789167]
7. Blauwhoff-Buskermolen S, Versteeg KS, de van der Schueren MAE, den Braver NR, Berkhof J, Langius JAE, et al. Loss of Muscle Mass During Chemotherapy Is Predictive for Poor Survival of Patients With Metastatic Colorectal Cancer. J Clin Oncol 2016 Apr 20;34(12):1339-1344. [doi: 10.1200/JCO.2015.63.6043] [Medline: 26903572]
8. Martin L, Birdsell L, Macdonald N, Reiman T, Clandinin MT, McCargar LJ, et al. Cancer cachexia in the age of obesity: skeletal muscle depletion is a powerful prognostic factor, independent of body mass index. J Clin Oncol 2013 Apr 20;31(12):1539-1547. [doi: 10.1200/JCO.2012.45.2722] [Medline: 23530101]
9. Prado CMM, Baracos VE, McCargar LJ, Reiman T, Mourtzakis M, Tonkin K, et al. Sarcopenia as a determinant of chemotherapy toxicity and time to tumor progression in metastatic breast cancer patients receiving capecitabine treatment. Clin Cancer Res 2009 Apr 15;15(8):2920-2926 [FREE Full text] [doi: 10.1158/1078-0432.CCR-08-2242] [Medline: 19351764]
10. Cormie P, Zopf EM, Zhang X, Schmitz KH. The Impact of Exercise on Cancer Mortality, Recurrence, and Treatment-Related Adverse Effects. Epidemiol Rev 2017 Jan 01;39(1):71-92. [doi: 10.1093/epirev/mxx007] [Medline: 28453622]
11. Gardner JR, Livingston PM, Fraser SF. Effects of exercise on treatment-related adverse effects for patients with prostate cancer receiving androgen-deprivation therapy: a systematic review. J Clin Oncol 2014 Feb 01;32(4):335-346. [doi: 10.1200/JCO.2013.49.5523] [Medline: 24344218]
12. Furness K, Silvers MA, Savva J, Huggins CE, Truby H, Haines T. Long-term follow-up of the potential benefits of early nutritional intervention in adults with upper gastrointestinal cancer: a pilot randomised trial. Support Care Cancer 2017 Dec;25(11):3587-3593. [doi: 10.1007/s00520-017-3789-2] [Medline: 28612158]
13. Ravasco P, Monteiro-Grillo I, Camilo M. Individualized nutrition intervention is of major benefit to colorectal cancer patients: long-term follow-up of a randomized controlled trial of nutritional therapy. Am J Clin Nutr 2012 Dec;96(6):1346-1353. [doi: 10.3945/ajcn.111.018838] [Medline: 23134880]
14. Arem H, Chlebowski R, Stefanick ML, Anderson G, Wactawski-Wende J, Sims S, et al. Body mass index, physical activity, and survival after endometrial cancer diagnosis: results from the Women's Health Initiative. Gynecol Oncol 2013 Feb;128(2):181-186 [FREE Full text] [doi: 10.1016/j.ygyno.2012.10.029] [Medline: 23127972]
15. Arem H, Irwin ML. Obesity and endometrial cancer survival: a systematic review. Int J Obes (Lond) 2013 May;37(5):634-639 [FREE Full text] [doi: 10.1038/ijo.2012.94] [Medline: 22710929]
16. Blanchard CM, Courneya KS, Stein K, American Cancer Society's SCS-II. Cancer survivors' adherence to lifestyle behavior recommendations and associations with health-related quality of life: results from the American Cancer Society's SCS-II. J Clin Oncol 2008 May 1;26(13):2198-2204 [FREE Full text] [doi: 10.1200/JCO.2007.14.6217] [Medline: 18445845]
17. Chan DSM, Vieira AR, Aune D, Bandera EV, Greenwood DC, McTiernan A, et al. Body mass index and survival in women with breast cancer-systematic literature review and meta-analysis of 82 follow-up studies. Ann Oncol 2014 Oct;25(10):1901-1914 [FREE Full text] [doi: 10.1093/annonc/mdu042] [Medline: 24769692]
18. Nagle CM, Dixon SC, Jensen A, Kjaer SK, Modugno F, deFazio A, Australian Ovarian Cancer Study Group, Ovarian Cancer Association Consortium. Obesity and survival among women with ovarian cancer: results from the Ovarian Cancer Association Consortium. Br J Cancer 2015 Sep 01;113(5):817-826 [FREE Full text] [doi: 10.1038/bjc.2015.245] [Medline: 26151456]
19. Protani MM, Nagle CM, Webb PM. Obesity and ovarian cancer survival: a systematic review and meta-analysis. Cancer Prev Res (Phila) 2012 Jul;5(7):901-910 [FREE Full text] [doi: 10.1158/1940-6207.CAPR-12-0048] [Medline: 22609763]
20. O'Neill RF, Haseen F, Murray LJ, O'Sullivan JM, Cantwell MM. A randomised controlled trial to evaluate the efficacy of a 6-month dietary and physical activity intervention for patients receiving androgen deprivation therapy for prostate cancer. J Cancer Surviv 2015 Sep;9(3):431-440. [doi: 10.1007/s11764-014-0417-8] [Medline: 25916660]
21. Reeves MM, Terranova CO, Eakin EG, Demark-Wahnefried W. Weight loss intervention trials in women with breast cancer: a systematic review. Obes Rev 2014 Sep;15(9):749-768. [doi: 10.1111/obr.12190] [Medline: 24891269]
22. Bort-Roig J, Gilson ND, Puig-Ribera A, Contreras RS, Trost SG. Measuring and influencing physical activity with smartphone technology: a systematic review. Sports Med 2014 May;44(5):671-686. [doi: 10.1007/s40279-014-0142-5] [Medline: 24497157]
23. Broekhuizen K, Kroeze W, van Poppel MN, Oenema A, Brug J. A systematic review of randomized controlled trials on the effectiveness of computer-tailored physical activity and dietary behavior promotion programs: an update. Ann Behav Med 2012 Oct;44(2):259-286 [FREE Full text] [doi: 10.1007/s12160-012-9384-3] [Medline: 22767052]
24. Fanning J, Mullen SP, McAuley E. Increasing physical activity with mobile devices: a meta-analysis. J Med Internet Res 2012;14(6):e161 [FREE Full text] [doi: 10.2196/jmir.2171] [Medline: 23171838]

25. Roberts AL, Fisher A, Smith L, Heinrich M, Potts HWW. Digital health behaviour change interventions targeting physical activity and diet in cancer survivors: a systematic review and meta-analysis. J Cancer Surviv 2017 Dec;11(6):704-719 [FREE Full text] [doi: 10.1007/s11764-017-0632-1] [Medline: 28779220]
26. Ugalde A, Haynes K, Boltong A, White V, Krishnasamy M, Schofield P, et al. Self-guided interventions for managing psychological distress in people with cancer - A systematic review. Patient Educ Couns 2017 May;100(5):846-857 [FREE Full text] [doi: 10.1016/j.pec.2016.12.009] [Medline: 28081937]
27. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ 2009;339:b2535 [FREE Full text] [Medline: 19622551]
28. PROSPERO International prospective register of systematic reviews. Technology supported self-guided nutrition and physical activity interventions in adults with cancer URL: https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=80346 [accessed 2018-12-12] [WebCite Cache ID 74c9B7Jff]
29. Cochrane Training. Cochrane handbook for systematic reviews of interventions, URL: https://training.cochrane.org/handbook [accessed 2018-12-05] [WebCite Cache ID 74S8TkJuD]
30. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. Ann Behav Med 2013 Aug;46(1):81-95. [doi: 10.1007/s12160-013-9486-6] [Medline: 23512568]
31. Galiano-Castillo N, Arroyo-Morales M, Lozano-Lozano M, Fernández-Lao C, Martín-Martín L, Del-Moral-Ávila R, et al. Effect of an Internet-based telehealth system on functional capacity and cognition in breast cancer survivors: a secondary analysis of a randomized controlled trial. Support Care Cancer 2017 Dec;25(11):3551-3559. [doi: 10.1007/s00520-017-3782-9] [Medline: 28639097]
32. Galiano-Castillo N, Cantarero-Villanueva I, Fernández-Lao C, Ariza-García A, Díaz-Rodríguez L, Del-Moral-Ávila R, et al. Telehealth system: A randomized controlled trial evaluating the impact of an internet-based exercise intervention on quality of life, pain, muscle strength, and fatigue in breast cancer survivors. Cancer 2016 Oct 15;122(20):3166-3174 [FREE Full text] [doi: 10.1002/cncr.30172] [Medline: 27332968]
33. Kanera IM, Bolman CAW, Willems RA, Mesters I, Lechner L. Lifestyle-related effects of the web-based Kanker Nazorg Wijzer (Cancer Aftercare Guide) intervention for cancer survivors: a randomized controlled trial. J Cancer Surviv 2016 Mar 17. [doi: 10.1007/s11764-016-0535-6] [Medline: 26984534]
34. Kanera IM, Willems RA, Bolman CAW, Mesters I, Verboon P, Lechner L. Long-term effects of a web-based cancer aftercare intervention on moderate physical activity and vegetable consumption among early cancer survivors: a randomized controlled trial. Int J Behav Nutr Phys Act 2017 Feb 10;14(1):19 [FREE Full text] [doi: 10.1186/s12966-017-0474-2] [Medline: 28187725]
35. Forbes CC, Blanchard CM, Mummery WK, Courneya KS. Feasibility and Preliminary Efficacy of an Online Intervention to Increase Physical Activity in Nova Scotian Cancer Survivors: A Randomized Controlled Trial. JMIR Cancer 2015 Nov 23;1(2):e12 [FREE Full text] [doi: 10.2196/cancer.4586] [Medline: 28410166]
36. Bantum EO, Albright CL, White KK, Berenberg JL, Layi G, Ritter PL, et al. Surviving and thriving with cancer using a Web-based health behavior change intervention: randomized controlled trial. J Med Internet Res 2014;16(2):e54 [FREE Full text] [doi: 10.2196/jmir.3020] [Medline: 24566820]
37. Valle CG, Deal AM, Tate DF. Preventing weight gain in African American breast cancer survivors using smart scales and activity trackers: a randomized controlled pilot study. J Cancer Surviv 2017 Feb;11(1):133-148. [doi: 10.1007/s11764-016-0571-2] [Medline: 27631874]
38. Yun YH, Lee KS, Kim Y, Park SY, Lee ES, Noh D, et al. Web-based tailored education program for disease-free cancer survivors with cancer-related fatigue: a randomized controlled trial. J Clin Oncol 2012 Apr 20;30(12):1296-1303 [FREE Full text] [doi: 10.1200/JCO.2011.37.2979] [Medline: 22412149]
39. Gnagnarella P, Misotti AM, Santoro L, Akoumianakis D, Del CL, De LF, et al. Nutritional Online Information for Cancer Patients: a Randomized Trial of an Internet Communication Plus Social Media Intervention. J Cancer Educ 2016 Dec;31(3):472-480. [doi: 10.1007/s13187-015-0820-5] [Medline: 25820605]
40. Lee MK, Yun YH, Park H, Lee ES, Jung KH, Noh D. A Web-based self-management exercise and diet intervention for breast cancer survivors: pilot randomized controlled trial. Int J Nurs Stud 2014 Dec;51(12):1557-1567. [doi: 10.1016/j.ijnurstu.2014.04.012] [Medline: 24856854]
41. Pope ZC, Zeng N, Zhang R, Lee HY, Gao Z. Effectiveness of Combined Smartwatch and Social Media Intervention on Breast Cancer Survivor Health Outcomes: A 10-Week Pilot Randomized Trial. J Clin Med 2018 Jun 07;7(6) [FREE Full text] [doi: 10.3390/jcm7060140] [Medline: 29880779]
42. Rabin C, Dunsiger S, Ness KK, Marcus BH. Internet-Based Physical Activity Intervention Targeting Young Adult Cancer Survivors. J Adolesc Young Adult Oncol 2011 Dec;1(4):188-194 [FREE Full text] [doi: 10.1089/jayao.2011.0040] [Medline: 23610737]
43. Uhm KE, Yoo JS, Chung SH, Lee JD, Lee I, Kim JI, et al. Effects of exercise intervention in breast cancer patients: is mobile health (mHealth) with pedometer more effective than conventional program using brochure? Breast Cancer Res Treat 2017 Feb;161(3):443-452. [doi: 10.1007/s10549-016-4065-8] [Medline: 27933450]

44. Krebs P, Shtaynberger J, McCabe M, Iocolano M, Williams K, Shuk E, et al. An eHealth Intervention to Increase Physical Activity and Healthy Eating in Older Adult Cancer Survivors: Summative Evaluation Results. JMIR Cancer 2017 Mar 01;3(1):e4 [FREE Full text] [doi: 10.2196/cancer.6435] [Medline: 28410171]
45. Sajid S, Dale W, Mustian K, Kotwal A, Heckler C, Porto M, et al. Novel physical activity interventions for older patients with prostate cancer on hormone therapy: A pilot randomized study. J Geriatr Oncol 2016 Mar;7(2):71-80 [FREE Full text] [doi: 10.1016/j.jgo.2016.02.002] [Medline: 26916611]
46. Vallance JKH, Courneya KS, Plotnikoff RC, Yasui Y, Mackey JR. Randomized controlled trial of the effects of print materials and step pedometers on physical activity and quality of life in breast cancer survivors. J Clin Oncol 2007 Jun 10;25(17):2352-2359 [FREE Full text] [doi: 10.1200/JCO.2006.07.9988] [Medline: 17557948]
47. Hatchett A, Hallam JS, Ford MA. Evaluation of a social cognitive theory-based email intervention designed to influence the physical activity of survivors of breast cancer. Psychooncology 2013 Apr;22(4):829-836. [doi: 10.1002/pon.3082] [Medline: 22573338]
48. Mayer DK, Landucci G, Awoyinka L, Atwood AK, Carmack CL, Demark-Wahnefried W, et al. SurvivorCHESS to increase physical activity in colon cancer survivors: can we get them moving? J Cancer Surviv 2018 Feb;12(1):82-94. [doi: 10.1007/s11764-017-0647-7] [Medline: 28994035]
49. Ormel HL, van der Schoot GGF, Westerink NL, Sluiter WJ, Gietema JA, Walenkamp AME. Self-monitoring physical activity with a smartphone application in cancer patients: a randomized feasibility study (SMART-trial). Support Care Cancer 2018 May 21. [doi: 10.1007/s00520-018-4263-5] [Medline: 29785635]
50. Haberlin C, O'Dwyer T, Mockler D, Moran J, O'Donnell DM, Broderick J. The use of eHealth to promote physical activity in cancer survivors: a systematic review. Support Care Cancer 2018 Jun 16. [doi: 10.1007/s00520-018-4305-z] [Medline: 29909476]
51. Ugalde A, Kiss N, Livingston P, Aranda S. Towards a framework for reporting self-guided interventions for people with cancer. Current Opinion in Supportive and Palliative Care 2018:1. [doi: 10.1097/spc.0000000000000353]
52. 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]
53. Sanjida S, McPhail SM, Shaw J, Couper J, Kissane D, Price MA, et al. Are psychological interventions effective on anxiety in cancer patients? A systematic review and meta-analyses. Psychooncology 2018 Jun 09. [doi: 10.1002/pon.4794] [Medline: 29885258]
54. Fuertes JN, Boylan LS, Fontanella JA. Behavioral indices in medical care outcome: the working alliance, adherence, and related factors. J Gen Intern Med 2009 Jan;24(1):80-85 [FREE Full text] [doi: 10.1007/s11606-008-0841-4] [Medline: 18972089]
55. Taaffe DR, Newton RU, Spry N, Joseph D, Chambers SK, Gardiner RA, et al. Effects of Different Exercise Modalities on Fatigue in Prostate Cancer Patients Undergoing Androgen Deprivation Therapy: A Year-long Randomised Controlled Trial. Eur Urol 2017 Dec;72(2):293-299 [FREE Full text] [doi: 10.1016/j.eururo.2017.02.019] [Medline: 28249801]
56. Watson EK, Rose PW, Neal RD, Hulbert-Williams N, Donnelly P, Hubbard G, et al. Personalised cancer follow-up: risk stratification, needs assessment or both? Br J Cancer 2012 Jan 03;106(1):1-5 [FREE Full text] [doi: 10.1038/bjc.2011.535] [Medline: 22215103]

Abbreviations

2MWT: two minute walk test
6MWT: six minute walk test
BMI: body mass index
DVD: digital video disk
HBA_{1c}: hemoglobin A_{1c}
HRQoL: health-related quality of life
MVPA: moderate to vigorous physical activity
PA: physical activity
RCT: randomized controlled trial
SMS: short message service

Edited by C Dias; submitted 20.09.18; peer-reviewed by JA Hughson, D Mayer; comments to author 31.10.18; revised version received 08.11.18; accepted 13.11.18; published 12.02.19.

Please cite as:
Kiss N, Baguley BJ, Ball K, Daly RM, Fraser SF, Granger CL, Ugalde A
Technology-Supported Self-Guided Nutrition and Physical Activity Interventions for Adults With Cancer: Systematic Review
JMIR Mhealth Uhealth 2019;7(2):e12281
URL: http://mhealth.jmir.org/2019/2/e12281/
doi:10.2196/12281
PMID:30747720

Review

Evaluation of Self-Management Support Functions in Apps for People With Persistent Pain: Systematic Review

Hemakumar Devan[1], BPT, MPhty, PhD; Devin Farmery[2]; Lucy Peebles[3], BSc; Rebecca Grainger[2], BMedSci, MBChB, FRACP, PhD

[1]Centre for Health, Activity and Rehabilitation Research, School of Physiotherapy, University of Otago, Wellington, New Zealand

[2]Department of Medicine, University of Otago, Wellington, New Zealand

[3]University of Otago, Wellington, New Zealand

Corresponding Author:
Hemakumar Devan, BPT, MPhty, PhD
Centre for Health, Activity and Rehabilitation Research, School of Physiotherapy
University of Otago
PO Box 7343, Newtown
Wellington, 6242
New Zealand
Phone: 64 4 385 5566
Fax: 64 4 385 5427
Email: hemakumar.devan@otago.ac.nz

Abstract

Background: Smartphone apps are a potential mechanism for development of self-management skills in people with persistent pain. However, the inclusion of best-practice content items in available pain management apps fostering core self-management skills for self-management support is not known.

Objective: The aim of the study was to evaluate the contents of smartphone apps providing information on pain management strategies for people with persistent pain facilitating self-management support and to appraise the app quality.

Methods: A systematic search was performed in the New Zealand App Store and Google Play Store. Apps were included if they were designed for people with persistent pain, provided information on pain self-management strategies, and were available in English. App contents were evaluated using an a priori 14-item self-management support (SMS-14) checklist. App quality was assessed using the 23-item Mobile Apps Rating Scale.

Results: Of the 939 apps screened, 19 apps met the inclusion criteria. Meditation and guided relaxation were the most frequently included self-management strategies. Overall, the included apps met a median of 4 (range 1-8) of the SMS-14 checklist. A total of 3 apps (Curable, PainScale-Pain Diary and Coach, and SuperBetter) met the largest number of items (8 out of 14) to foster self-management of pain. Self-monitoring of symptoms (n=11) and self-tailoring of strategies (n=9) were frequently featured functions, whereas a few apps had features facilitating social support and enabling communicating with clinicians. No apps provided information tailored to the cultural needs of the user. The app quality mean scores using Mobile Apps Rating Scale ranged from 2.7 to 4.5 (out of 5.0). Although use of 2 apps (Headspace and SuperBetter) has been shown to improve health outcomes, none of the included apps have been evaluated in people with persistent pain.

Conclusions: Of the 3 apps (Curable, PainScale-Pain Diary and Coach, and SuperBetter) that met the largest number of items to support skills in self-management of pain, 2 apps (PainScale-Pain Diary and Coach and SuperBetter) were free, suggesting the potential for using apps as a scalable, wide-reaching intervention to complement face-to-face care. However, none provided culturally tailored information. Although 2 apps (Headspace and SuperBetter) were validated to show improved health outcomes, none were tested in people with persistent pain. Both users and clinicians should be aware of such limitations and make informed choices in using or recommending apps as a self-management tool. For better integration of apps in clinical practice, concerted efforts are required among app developers, clinicians, and people with persistent pain in developing apps and evaluating for clinical efficacy.

(JMIR Mhealth Uhealth 2019;7(2):e13080) doi:10.2196/13080

KEYWORDS

smartphone; chronic pain; culture; mHealth; self-management; technology

Introduction

Background

Persistent, noncancer pain is the leading cause of disability worldwide, affecting 1 in 5 people [1,2]. Persistent pain includes a wide range of pain conditions such as persistent primary pain (eg, low back pain and neck pain), musculoskeletal pain (eg, osteoarthritis and rheumatoid arthritis), migraine, orofacial pain, neuropathic pain, and pain following trauma and surgery [3]. Recent results from the Global Burden of Diseases suggest persistent low back pain and migraine are the 2 major contributors globally to disability-adjusted life years [4], signifying the large personal, societal, and economic impact of persistent pain.

The current best practice for management of persistent pain involves group-based, multidisciplinary, cognitive behavioral interventions focusing on fostering self-management [5-7]. Self-management support can be achieved by providing individuals with tools and strategies to help them choose healthier behaviors and transforming the patient-clinician relationship into one of collaborative care [8]. Fostering self-management support could be achieved via core self-management skills. Core self-management skills include self-efficacy building, self-tailoring, self-monitoring of symptoms, goal setting and planning, problem solving, and shared decision making. In the context of persistent pain, individuals learn a variety of self-management or active coping strategies (eg, cognitive behavioral therapy, activity pacing, and relaxation techniques) from pain management interventions. In collaboration with health professionals (*shared decision making*), individuals learn to identify meaningful goals (eg, personal, physical, and psychological) to improve pain and functioning (*goal setting*) [8]. Practicing a variety of strategies learned from pain management interventions, individuals develop the capacity to monitor their symptoms (*self-monitoring*) and tailor strategies that work best for them (*self-tailoring*) [5]. With the help and support of peers and health professionals, they can actively problem solve (*problem solving*) and set meaningful goals [9]. Positive reinforcement from learning self-management skills and adhering to practicing those strategies [10] ultimately develops the confidence in one's ability (*self-efficacy building*) to engage in meaningful activities despite pain (*acceptance*), thereby fostering sustained behavioral change [11].

Specialized pain services including multidisciplinary, cognitive behavioral interventions are resource intensive and are often only provided in secondary or tertiary care settings [5]. Therefore, there are many barriers to accessing specialized pain services in Australia and New Zealand, including long-waiting lists after referral from primary care [12], issues related to transport, and physically accessing in-person delivered pain services [13]. Furthermore, there is little capacity to increase services because of a lack of a specialist pain workforce [12]. Therefore, innovative solutions for delivering pain management services are urgently required.

Smartphone apps are a potential mechanism for the development of self-management skills. Interventions delivered via smartphones are more likely to be easily accessible to a wider population and are scalable [14-16], with smartphone user numbers expected to reach 5 billion in 2020 [17]. Evidence suggests smartphone app use can foster self-management skills in other long-term conditions such as diabetes, asthma, and gout [18-20], and it can potentially deliver behavioral interventions to people living with persistent pain [21,22]. For example, pain apps have been generally classified to (1) provide general information on pain, including symptom identification and planning treatments, (2) track daily symptoms such as pain intensity, mood, daily activity, and medications, and (3) provide information on self-management strategies [23].

Previous reviews have examined the quality and content of self-management apps for people with persistent pain [23,24]. However, none of these reviews evaluated the app contents using a comprehensive self-management support checklist [3], including the cultural appropriateness of the contents via addressing cultural beliefs (eg, race, ethnicity, religion, and socioeconomic status). Cultural beliefs have been shown to influence illness perceptions, pain experiences, and attitudes to pain management [25-27], and addressing cultural beliefs of individuals with persistent pain is recommended by international clinical guidelines for pain management [28,29]. Lalloo et al [24] comprehensively reviewed 256 pain management apps in 2014 and classified the contents based on 5 core self-management skills (pain education, self-efficacy building, self-monitoring, social support, and goal setting). The review identified the lack of a comprehensive app, including all the self-management skills and lack of health care provider involvement in over 91.7% (256/279) of the included apps. However, the included apps were not described in detail, and there was no quality evaluation of the apps. Recently, reviews of self-management apps for specific pain conditions, namely persistent low back pain and arthritis, have been performed [30,31]. The included apps in both reviews recommended evidence-based interventions, yet the quality of information in the apps was poor [30,31]. Although Machado et al [30] concluded that there remains a need for higher-quality apps fostering self-management in people with persistent low back pain, Bhattarai et al [31] suggested the need for apps better suited for older adults, the more prevalent population group with arthritis. However, the apps identified from both reviews are not suitable for people with a wide variety of persistent pain conditions.

Objectives

In light of such limitations, this review extends previous work by adopting a comprehensive International Classification of Diseases (ICD-11) definition for persistent pain [3] and evaluating app contents using an evaluation checklist for best-practice content items for self-management support. The

primary purpose of the review was to evaluate the contents of smartphone apps providing information on pain management strategies for people with persistent pain facilitating self-management support and to appraise app quality.

Methods

This review was reported according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines for systematic reviews [32]. The review protocol was submitted for registration at PROSPERO—International Prospective Register of Systematic Review; however, the submission was rejected as PROSPERO does not accept systematic review protocols for apps.

Operational Definition

Persistent pain was defined as recurrent pain for more than 3 months [3], and persistent pain disorders were classified based on the ICD-11 definition [3] as (1) persistent primary pain, (2) persistent posttraumatic and postsurgical pain, (3) persistent neuropathic pain, (4) persistent malignant pain, (5) persistent headache and orofacial pain, (6) persistent visceral pain, and (7) persistent musculoskeletal pain. Culture was defined as a broad construct encompassing ethnicity, religion, socioeconomic status, disability, and sexual orientation [33].

Search Strategy

A systematic search was performed in New Zealand Google Play Store (Android) and the App Store (iOS)—via the website fnd.io [34] —on November 14, 2017. These stores account for 99.22% of the market share of smartphone platforms used in New Zealand [35]. Search terms for both stores included *pain*, *pain management*, and *chronic pain*. An updated search was conducted on December 07, 2018, using the term *pain management* in both Google Play and App stores (via fnd.io [34]) to identify any new apps.

Inclusion Criteria

Apps were included if they were (1) smartphone apps targeted to adults (>16 years) with persistent pain, (2) capable of being used for any persistent pain condition, (3) able to provide information on at least one or more of the following recommended strategies for pain management: pain education, activity pacing, thought and behavioral management, exercises, relaxation or breathing, meditation or mindfulness, distraction techniques, and (4) available in English.

Apps were excluded if (1) a specific pain condition was targeted (eg, low back pain and migraine), (2) only pain monitoring function was provided, and (3) the app was an advertisement for a specific clinic or product.

Data Extraction

App names and descriptions from the app search in Google Play Store and the App Store were screened against a priori selection criteria. Next, eligible apps were downloaded for further screening using either OnePlus 3 smartphone running Android 8.0.0 for Google Play Store apps or Apple iPhone 6S running iOS 11.1.2 for the iOS platform. Apps found on both platforms were only downloaded on the iOS platform [30]. DF conducted the app search and screened apps for inclusion. Unclear app store descriptions were discussed with HD or RG for selection. When a free version (often indicated by *lite* in the title) and fully featured paid version of a same app were available, we assessed only the fully featured paid version so that the full functionality of the app was evaluated. A list of the final included apps was collated into an Excel spreadsheet (Microsoft Corp), with metadata about each app extracted from the relevant app stores. This included information on developer, price, app size (in megabytes), app version, and a brief summary of app contents.

Evaluation of App Contents for Self-Management Support

Currently, there are no widely accepted frameworks or clinical guidelines recommending the key components of a pain self-management program. Therefore, a customized 14-item self-management support (SMS-14) checklist [5,8,24,36] was used for the purpose of this app review to evaluate the contents of the included apps for learning or developing self-management support (Table 1). The SMS-14 checklist included 6 core self-management skills (12 items) and 2 functions (2 items). Core self-management skills included self-efficacy building via recommended pain management strategies (7 items), problem-solving (1 item), goal-setting (1 item), self-tailoring (1 item), self-monitoring (1 item), and partnership between views of the patient and clinician via communication skills (1 item). The functions included providing access to social support (1 item) and culturally tailored information (1 item) by addressing cultural beliefs related to ethnicity, religion, socioeconomic status, disability, and sexual orientation [33]. The items from the SMS-14 checklist were informed from a well-established self-management framework—Stanford Self-management Support model [8,36], an evidence-based persistent pain guideline [28]—and previous reviews on self-management and persistent pain [5,24,31]. Apps scored 1 point for each item featured within the app, with a possible score between 0 and 14. Two reviewers (DF and HD) independently evaluated the SMS-14 checklist of included apps by using all the functions of the downloaded apps for at least 10 min and scored the items based on mutual consensus. RG reviewed and adjudicated any discrepancies.

App Quality Evaluation

The 23-item Mobile Apps Rating Scale (MARS) [37] was used to assess general app quality. The MARS is a reliable tool for assessing app quality using 5 sections: engagement (5 items), functionality (4 items), aesthetics (3 items), information quality (7 items), and a subjective app quality score (4 items). The app quality mean score of an included app was calculated from individual mean scores of engagement, functionality, esthetics, and information quality. Two trained reviewers (LP and DF) independently scored the included apps using the MARS and subsequently compared scores and formed a consensus opinion to give a final reported MARS score [37]. Any difficulties in reaching a consensus were discussed with HD to reach a conclusion. For each MARS section, the interrater reliability of total mean scores between the reviewers was calculated using

intraclass correlation coefficient (ICC)—2-way random-effects model with absolute agreement among single ratings [1,3].

Table 1. Self-management support (SMS-14) checklist for best-practice content items for self-management support of pain.

Skills	Description	Examples (if one or more present, scored *yes*)	Score
Self-efficacy building	Provision of information on self-management or active coping strategies to improve the ability to control one's behavior—Cognitive Behavioral Therapy (CBT) approaches	Pain education: Mechanisms of pain/neurophysiology; Information on pain-stress-depression; Thought management; Fear avoidance; Catastrophizing; Medication use; Sleep management	Yes; No
		Activity pacing: Increments in activity interspersed with short periods of rest or changes in activity not determined by level of pain; (eg, increasing tolerance on activities of preference such as walking and sitting)	Yes; No
		Thought and behavioral management: CBT-based therapy; Acceptance-based therapies: Mindfulness-based stress reduction; Acceptance and commitment therapy	Yes; No
		Exercises (biomechanical or aerobic): Biomechanical—Exercises aimed at changing or improving spinal mechanics (eg, stretching, strengthening, range of motion, Pilates, and McKenzie exercises); Aerobic exercises—Exercises aimed at improving cardiovascular fitness and endurance (eg, walking, running, and swimming)	Yes; No
		Relaxation and breathing: Self-calming methods—Breathing exercises and mental imagery	Yes; No
		Meditation and mindfulness: Interventions primarily focusing on physical, mental, and spiritual focus; Yoga, Tai Chi, and mindfulness.	Yes; No
		Distraction techniques: Shifting attentional focus from pain; Music, art, play, and images	Yes; No
Self-tailoring	Provision of structured information and self-management support based on the individual symptoms/needs	Scope for the individuals to incorporate the self-management strategies learnt to fit their individual needs	Yes; No
Self-monitoring of symptoms	Capacity to help people to monitor their symptoms (eg, mood, thoughts, and pain intensity)	Thought diaries; Daily activity tracking; Pain diaries; Mindfulness; Sleep management	Yes; No
Goal setting and planning	Capacity to identify and log meaningful goals (physical, emotional, social) and track goals	Planning daily activities; Planning a specific activity goal	Yes; No
Problem solving	A systematic approach to be aware of and developing a plan for dealing with stressful or challenging situations	Having a plan for dealing with flare-ups; Ability to deal with stressors (eg, pain, stress, depression, and anxiety)	Yes; No
Partnership between views of patient and clinicians	Opportunity to interact with health care provider and involve people with persistent pain in decision making	Information/training on assertive communication with health professionals for patients	Yes; No
Social support	Access to a community of persons living with persistent pain	Provision of access to other persons in pain through the app for the purpose of providing emotional, informational, and appraisal support	Yes; No
Cultural relevance	Reporting of culturally tailored information applicable for diverse ethnic groups	Tailored information toward different cultures: Ethnicity; Religion; Socioeconomic status; Disability; Sexual orientation	Yes; No

Results

Systematic Search

The initial search identified 600 unique Android apps from the Google Play Store and 339 unique iOS apps from the App Store—Figure 1. Following app name and description screening, 876 apps were excluded. Reasons for app exclusion were not being relevant to people with persistent pain (327/876, 37.3%), not applicable to all pain conditions (256/876, 29.2%), and not containing information related to pain management strategies (114/876, 13.0%)—Figure 1. After full app content screening, 19 apps were included for the review. This included 1 app (Curable) found only in the updated search.

App Characteristics and Content

A summary of key characteristics of included apps is presented in Table 2. Overall, 12 of the 19 included apps were on both platforms, 3 were only on Android, and 4 only on iOS. Apps ranged in size from 3.3 MB to 105.30 MB, with the largest apps being Mindfulness Daily (105.3 MB) and Release Pain with Andrew Johnson (97.3 MB). Moreover, 5 of the apps were free and 11 had a single payment ranging from NZD $2.99 to NZD $14.99. Additionally, 3 apps were available on annual subscription (Aware—Meditation and Mindfulness NZD $49.99, Curable NZD $209.88, and Headspace NZD $139.99), and some of these apps also required additional in-app downloads for specific pain management content. Just more than half of the included apps (n=10) explicitly described the involvement of a health professional in developing app contents. A majority of the included apps were meditation (n=13) and relaxation apps (n=16) in the form of audio-guided imagery and hypnosis. One app provided comprehensive pain neuroscience education and guided relaxation for pain management (Curable). In addition, 2 apps delivered yoga in the form of postures for pain relief (Yoga for Pain therapy) and a specific form of transcendental meditation (Sahaja Kundalini Meditation).

Overall, the 19 apps met a median of 4 (range 1-8) of the SMS-14 checklist (Table 3). Free apps (median=7, range 2-8) scored better than the paid apps (median=1.5, range 1-8), with 2 of the 3 highest scoring apps (Pain Scale—Pain Diary and Coach and SuperBetter both scored 8) being free. Of the paid apps, the Curable app had the top score, including 8 content items from the SMS-14 checklist. Self-tailoring of provided skills and strategies (n=9) and the capacity for self-monitoring of symptoms (diary entries and/or self-reflection) (n=11) were both commonly found self-management skill content. Apps featuring cognitive interventions such as meditation (n=13) scored well for self-management score, as this content often gained 1 point for each of relaxation and breathing, thought and behavioral management, and problem solving. Of the 7 self-efficacy building strategies, pain education, activity pacing, exercises, and distraction techniques were seldom included. Opportunities to communicate with clinicians, social support from peers with persistent pain, and features to actively set goals and problem solve their symptoms were found infrequently, and none of the apps contained tailored information to address cultural beliefs. The mutual agreement between DF and HD's independent evaluations of SMS-14 checklist was substantial; adjusted kappa statistic (κ) was 0.86 (95% CI 0.79-0.93).

Quality Assessment

The app quality mean scores from MARS evaluation ranged from 2.72 to 4.54 (Table 4). Curable (4.54), Headspace (4.48), and PainScale (4.47) were the 3 highest scoring apps. Out of the 5 sections of the MARS, apps scored poorly in engagement and subjective app quality. Overall, 4 apps met criterion for scientific validation (item 19; Relax—Stress and Anxiety Relief, Mindfulness Daily, Headspace, and SuperBetter), of which 2 (Headspace and SuperBetter) had been clinically evaluated to show improved health outcomes [38,39]. No apps scored for scientific validation in improving health outcomes for people with persistent pain. The ICC scores for all sections from our MARS rating were greater than 0.69; in particular, app quality mean score showed a high level of reliability (ICC=0.92), indicating a good consistency across app ratings.

Figure 1. Flowchart for systematic app search from Google Play and App stores.

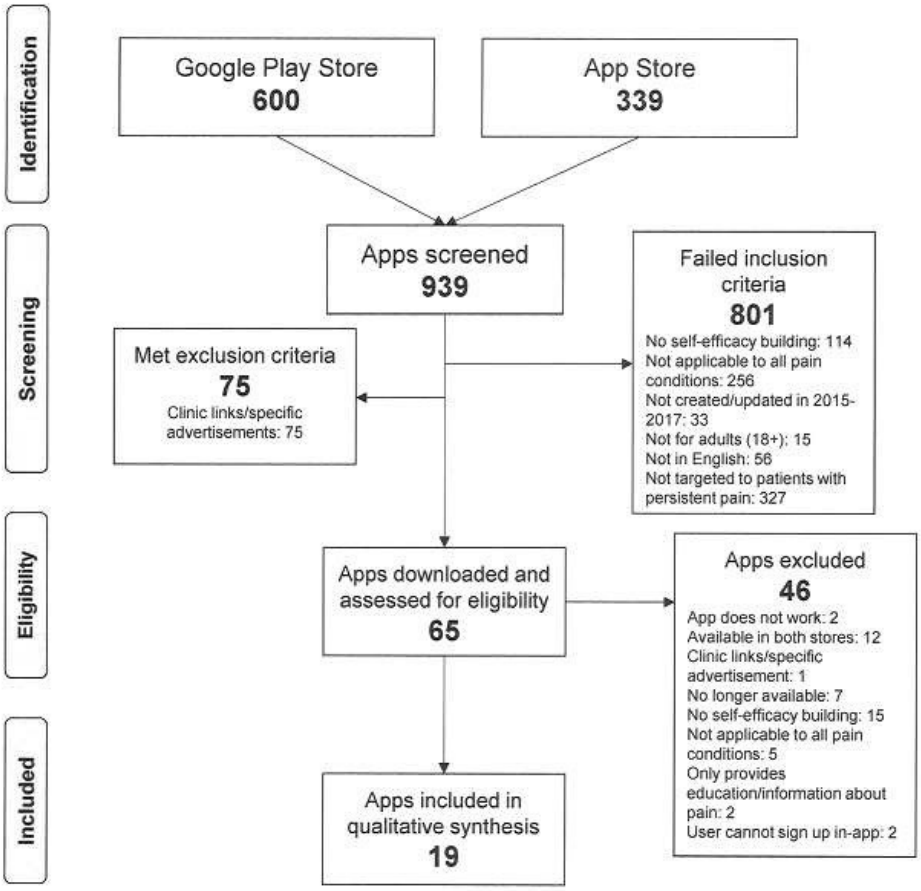


Table 2. Summary of app characteristics.

App Name	Developer	Platform(s)	App Version	Cost (NZD)	App size (MB)	Description
Aware—Meditation and Mindfulness	Uber Health Tech Pvt. Ltd	Android or iOS	2.2	$49.99 per year RRP[a]	8.19	Mindfulness or guided meditation
Counting Down to Relaxation	Healthy Visions	Android or iOS	2.1	$4.49	24.38	Guided relaxation or hypnosis
Curable: Back Pain, Migraine and Chronic Pain Relief	Curable Inc	Android or iOS	4.1.0	$209.88 per year RRP[a]	3.44	Guided mind and body strategies for pain management
Freedom From Pain	Healthy Visions	Android or iOS	2.1	$12.99	24.85	Hypnosis audio track
General Pain Management Guided Imagery	Shelley Spencer-Hellmich	Android or iOS	2	$8.99	48.3	Guided relaxation audio
Headspace: Guided Meditation	Headspace, Inc.	Android or iOS	3.2.4	$139.99 per year RRP[a]	30.28	Guided mindfulness to increase *approach* attention
Meditations for Pain Relief	Highly Meditated	Android	1.0.0	$3.02	37.54	Mindfulness or guided meditation
Mindfulness Coach	US Department of Veterans Affairs (VA)	iOS	1.4	Free	24.8	Mindfulness or guided meditation
Mindfulness Daily	INWARD, INC	iOS	1.4	Free	105.3	Mindfulness or guided meditation
Mindfulness Meditation for Pain Relief—Jon Kabat-Zinn	Sounds True	iOS	101.101.5	$14.99	32.1	Mindfulness or guided meditation
Pain Control Hypnosis by Glenn Harrold	Diviniti Publishing Ltd	Android or iOS	1.3	$5.99	80.54	Hypnosis audio
Pain Relief and Healing Meditation	Drentek	Android or iOS	2	$2.99	51	Mindfulness or guided meditation
Pain Relief Hypnosis PRO	Surf City Apps LLC	Android or iOS	4.5	$5.99	68.3	Hypnosis audio
PainScale-Pain Diary and Coach	Boston Scientific	Android or iOS	1.5	Free	31.7	Pain monitoring and management
Relax—Stress and Anxiety Relief	Saagara	Android or iOS	2.03	$4.49	51	Guided relaxation audio
Release Pain with Andrew Johnson	Michael Schneider	iOS	8.38	$4.49	97.3	Guided relaxation audio
Sahaja Kundalini Meditation	egeApps	Android	1.2.6	$4.49	7.56	Guided Kundalini meditation
SuperBetter	SuperBetter, LLC	Android or iOS	1.1.4	Free	4.91	Gamified tasks and goal-setting
Yoga For Pain Therapy	Harwell Publishing	Android	2	Free	16.02	Yoga postures and meditation

[a]RRP: recommended retail price.

Table 3. Evaluation checklist for self-management skills and functions.

App name	Core self-management skills														Functions		Total (14)
	Self-efficacy building							ST[h]	SM[i]	GS[j]	PS[k]	PV[l]	SS[m]	CR[n]			
	PE[a]	AP[b]	TB[c]	E[d]	R/B[e]	M/M[f]	D[g]										
Aware—Meditation and Mindfulness	__o	—	✓	—	✓	✓	—	✓	✓	—	✓	—	—	—			6
Counting Down to Relaxation	—	—	—	—	✓	—	—	—	—	—	—	—	—	—			1
Curable: Back Pain, Migraine and Chronic Pain Relief	✓	—	✓	—	✓	✓		✓	✓	—	✓	—	✓	—			8
Freedom From Pain	—	—	—	—	✓	—	—	—	—	—	—	—	—	—			1
General Pain Management Guided Imagery	—	—	—	—	✓	—	—	—	—	—	—	—	—	—			1
Headspace: Guided Meditation	—	—	✓	—	✓	✓	—	✓	✓	—	✓	—	—	—			6
Meditations for Pain Relief	—	—	—	—	✓	✓	—	—	✓	—	✓	—	—	—			4
Mindfulness Coach[p]	—	—	✓	—	✓	✓	—	✓	✓	✓	—	—	—	—			6
Mindfulness Daily[p]	—	—	✓	—	✓	✓	—	✓	✓	✓	✓	—	—	—			7
Mindfulness Meditation for Pain Relief—Jon Kabat-Zinn	—	—	✓	—	✓	✓	—	✓	✓	—	✓	—	—	—			6
Pain Control Hypnosis by Glenn Harrold	—	—	—	—	✓	—	—	—	—	—	—	—	—	—			1
Pain Relief & Healing Meditation	—	—	—	—	✓	✓	—	—	—	—	—	—	—	—			2
Pain Relief Hypnosis PRO	—	—	—	—	✓	—	—	—	—	—	—	—	—	—			1
PainScale-Pain Diary and Coach[p]	✓	—	—	✓	✓	✓	—	✓	✓	—	—	✓	✓	—			8
Relax—Stress and Anxiety Relief	—	—	—	—	✓	✓	—	✓	✓	—	—	—	—	—			4
Release Pain with Andrew Johnson	—	—	—	—	✓	—	—	—	—	—	—	—	—				1
Sahaja Kundalini Meditation	—	—	—	—	—	✓	—	—	—	—	—	—	—	—			1
SuperBetter[p]	—	✓	✓	✓	—	✓	—	✓	✓	✓	✓	—	✓	—			8
Yoga For Pain Therapy[p]	—	—	—	—	—	✓	—	—	✓	—	—	—	—	—			2
Count across apps	2	1	7	1	16	13	0	9	11	3	7	1	3	0			—

[a]PE: pain education.

[b]AP: activity pacing.

[c]TB: thoughts and behavioral management.

[d]E: exercises (biomechanical/aerobic).

[e]R/B: relaxation/Breathing.

[f]M/M: meditation/Mindfulness.

[g]D: distraction techniques.

[h]ST: self-tailoring.

[j]SM: self-monitoring of symptoms.

[j]GS: goal setting and planning.

[k]PS: problem solving.

[l]PV: partnership between views of patient and health professionals.

[m]SS: social support.

[n]CR: cultural relevance.

[o]Criteria not met.

[p]Free app.

Table 4. Mean Mobile App Rating Scale scores of included apps.

App Name	MARS[a] Categories					
	Engagement Section A	Functionality Section B	Aesthetics Section C	Information Section D	Subjective Quality Section E	App Quality Mean Score[b] (A+B+C+D)
Aware—Meditation and Mindfulness	3.60	4.25	4.67	4.20	4.50	4.18
Counting Down to Relaxation	1.80	4.25	2.00	2.83	2.00	2.72
Curable: Back Pain, Migraine and Chronic Pain Relief	4.00	5.00	4.66	4.50	4.75	4.54
Freedom From Pain	2.60	4.00	2.67	3.50	1.75	3.19
General Pain Management Guided Imagery	3.00	4.75	2.33	3.50	2.00	3.40
Headspace: Guided Meditation	4.20	4.50	5.00	4.20	5.00	4.48
Meditations for Pain Relief	3.40	5.00	4.33	3.25	3.00	4.00
Mindfulness Coach[c]	2.80	4.75	4.00	4.33	2.75	3.97
Mindfulness Daily[c]	4.20	4.50	5.00	4.14	4.75	4.46
Mindfulness Meditation for Pain Relief-Jon Kabat-Zinn	2.60	3.50	3.00	3.00	1.50	3.03
Pain Control Hypnosis by Glenn Harrold	3.20	5.00	4.00	4.00	3.00	4.05
Pain Relief and Healing Meditation	2.60	4.00	3.00	2.67	1.75	3.07
Pain Relief Hypnosis PRO	3.80	4.75	4.67	3.67	3.25	4.22
PainScale-Pain Diary and Coach[c]	4.20	4.50	4.67	4.50	4.00	4.47
Relax—Stress and Anxiety Relief	3.60	4.75	5.00	3.67	4.00	4.25
Release Pain with Andrew Johnson	3.00	5.00	4.33	3.67	2.75	4.00
Sahaja Kundalini Meditation	2.20	4.50	3.33	3.00	2.00	3.26
SuperBetter[c]	4.00	4.00	4.33	4.33	4.25	4.17
Yoga For Pain Therapy[c]	2.20	4.50	3.33	2.33	1.25	3.09

[a]MARS: Mobile Apps Rating Scale.

[b]App quality mean scores were calculated based on mean scores from engagement, functionality, aesthetics and information quality.

[c]Free app.

Discussion

Principal Findings

Although this app review identified 19 apps that purported to support self-management of persistent pain, the comprehensiveness of app contents fostering ongoing self-management skills and functions was limited. Of interest, 2 of the 3 apps (PainScale-Pain Diary and Coach and SuperBetter) including the largest number of items to support self-management of pain were both free, suggesting the potential for using apps as a scalable, wide-reaching intervention to complement face-to-face care. None of the included apps provided information tailored to cultural needs of the user, and none had been evaluated in people with pain. Furthermore, features facilitating individual goal setting, peer support, and

communicating health information with clinicians were infrequent.

A majority of the included apps (n=11) focused on delivering a single self-management strategy (eg, meditation and guided relaxation) for symptom management but lacked the features to facilitate core self-management skills (eg, goal setting and problem solving). Apps with more features facilitating core self-management skills have the potential for facilitating behavioral change [40], as identified in app-based systematic reviews promoting physical activity [41,42]. However, the self-management skills that are most influential in facilitating long-term behavior change in this clinical population remain unknown. A recent metasynthesis of qualitative studies identified that people with persistent pain valued ongoing support to sustain motivation to incorporate learnt self-management strategies after completing face-to-face treatments [11]. Overall, 3 of the included apps (Curable, Mindfulness Daily, and SuperBetter) did include 5 out of the 6 core self-management skills, suggesting that apps used after or alongside face-to-face care could be a feasible mechanism for providing ongoing self-management support.

Self-management strategies delivered for symptom management were predominantly mindfulness-based cognitive interventions. Mindfulness-based interventions such as stress reduction and cognitive therapy have been shown to improve long-term health outcomes in people with persistent pain [43]. Mindfulness-based apps had sophisticated features to facilitate thought and behavioral management. For example, apps featured weekly modules for the user to complete and had the capacity to self-monitor meditations via logs and setting reminders. By contrast, app facilitating guided relaxation in the form of guided imagery and hypnosis had only audio tracks with limited opportunities for tailoring to user needs. Inclusion of persuasive system design principles by means of task breakdown, providing reminders, and praise for achieving intended behavior and providing peer support with similar users has the potential to be persuasive to facilitate adherence to self-management strategies [44].

Pain education and activity pacing, 2 widely used self-management strategies were infrequently present in the included apps. For the purpose of the review, pain education was defined relative to concepts of pain neuroscience education, information on pain-stress-depression, medication, and sleep management. Pain neuroscience education has been shown to improve long-term improvements in pain-related disability and functioning [45]. Except for the Curable app, which provided comprehensive pain education via an interactive virtual pain coach, it is a significant omission that pain education and support for activity pacing were absent from most apps. For people with pain to successfully pace daily activities, preplanning, establishing user-centered activity goals, scheduling daily activities based on their symptoms, and prioritizing to achieve graded increment of activities are required [46]. These skills need to be supplemented by constant reinforcement, feedback, and monitoring [47]. Recent studies have shown the potential for wearable devices providing real-time feedback during daily activities and encouraging people to improve pacing [48,49]. Given the complexity in defining pacing, theoretical models guiding pacing intervention (operant theory and energy conservation) [46], and individualized nature of pacing [47], the potential to facilitate pacing via apps needs further investigation before widespread inclusion in apps.

Goal setting was found in only 3 of the included apps, a finding similar to previous app-reviews on persistent pain [24,31]. It has been shown to be an effective strategy to foster self-efficacy in people with pain [10]. Only 2 apps (Mindfulness Coach and Mindfulness Daily) used a SMART format (specific, measurable, achievable, realistic, and timed), whereas SuperBetter used gamified terms such as *quests*, *power ups*, and *bad guys* to foster goal setting. Despite the differences, further research is required to identify best-practice delivery of the goal-setting features in pain apps that would address the needs of users [31].

Social support can foster self-management [50], but it was featured in only 3 included apps. Although most of the apps allowed users to post activities to social media, which could be considered a means of facilitating peer support, this criterion was defined as providing access to peers with a similar health condition by providing informational, appraisal, and emotional support from their common disease experiences [51]. An app meeting this criterion would provide peer support via patients' stories of living well, with pain or discussion forums embedded in the app. Online support communities to connect and share information have been shown to improve pain-related outcomes such as improved self-efficacy and self-reported global health [52,53]. Although issues related to confidentiality and privacy need to be considered before incorporation of these features into apps, people with persistent pain value the importance of peer support and peer validation as an important skill for long-term self-management [11].

None of the apps provided culturally tailored information, a finding similar to our previous review on pain self-management websites [54]. Although meeting all these cultural aspects in an app may not be possible and would be country or region specific, there is a need for providing information tailored to cultural beliefs, considering the existing health disparities across cultures (eg, ethnic and racial minorities) in pain prevalence and access to services [55]. From the New Zealand context, providing culturally appropriate information is important to meet the needs of New Zealand Māori—Indigenous people of New Zealand—who are at high risk of reporting persistent pain [56] and experience greater health inequities than any other ethnic group in New Zealand [13,57]. It is interesting to note that none of the previous app reviews [23,24,30,31,58] have investigated the provision of culturally tailored information despite the evidence to support that cultural beliefs and practices affect individual pain experiences [59] and adherence to active self-management strategies [27]. Currently, there is no guideline available for cultural tailoring of apps. Codesigning resources with specific cultural groups by involving significant others (eg, family members), using visual aids such as pictures, cartoons, and videos, and simplified text with less jargon have been shown to positively influence health beliefs and self-management strategies [60,61], suggesting the potential for apps as a medium for delivering culturally relevant information.

The included apps had high rating on the MARS. Apps in general scored highly in esthetics and often earned higher scores in engagement, resulting in apps that are easy to use, interactive, and promoted repeat use. Although 2 of our included apps (Headspace and SuperBetter) scored on the MARS item 19 for *evidence base* for improved health outcomes in people with anxiety and depression [38,39], no apps were validated in people with persistent pain, a finding similar to a previous app review on persistent back pain [30]. Furthermore, the mean MARS total score (3.82) for the included apps was higher than in the previous pain app review (2.36) [30]. Differences could be because of the large number of apps reviewed (n=61) compared with the present review (n=19).

The majority of included apps required payment (n=14,) ranging from a single payment of NZD $2.99 to NZD $25.99 per month, raising the question if pain apps can be an equitable intervention. This is particularly important when considering existing inequities among people with persistent pain. As those in lower socioeconomic areas are overrepresented in this clinical group [56], they are also therefore disproportionately impacted by financial barriers to accessing self-management interventions [62,63]. Interestingly, the 2 of the 3 highest scored apps from our self-management support evaluation and MARS rating (PainScale-Pain Diary and Coach and SuperBetter) were both free, suggesting that providing high-quality self-management support via apps can be feasible and equitable.

Limitations

Although this is the first review to comprehensively assess the app contents using best-practice items to support self-management skills, including cultural tailoring in people with persistent pain, the following limitations need to be acknowledged in interpreting the results from this review. First, during our app screening process, we identified a number of popular apps for symptom monitoring (eg, Manage My Pain and My Pain Diary), which failed to meet our inclusion criteria because the present review had an inclusion criterion that required apps to provide information on pain management strategies. Although pain intensity assessment is an important aspect of overall self-management, pain intensity alone is a poor indicator of suffering and disability associated with persistent pain [64]. Furthermore, multidisciplinary cognitive behavioral therapy–based interventions focus on improving pain-related function instead of pain intensity [65]. Therefore, a comprehensive evaluation of symptom monitoring apps is beyond the scope of this review and would require another systematic review of apps. Although we were able to identify 2 apps (Brigham and Women's Hospital (BWH) Pain app and WebMD Pain Coach) from previous studies [31,66], these apps were not available at the New Zealand App Store and Google Play Store. Two popular apps (Pain Tricks and Pain Squad) were not included as they focused on children with persistent pain.

Secondly, our customized SMS-14 checklist, although based on evidence-based recommendations, has not been tested for validity. However, evidence suggests that technological interventions with more features facilitating core self-management skills (self-efficacy building, self-monitoring, goal setting, and problem solving) and support functions (peer support and cultural appropriateness) have the potential to foster long-term adaptive behavioral change [40-42]. Therefore, clinicians can be reasonably confident in the comprehensiveness of features facilitating self-management support from the apps included in the review. We acknowledge this is an area for further exploration as the most effective component of the multidisciplinary behavioral interventions remains unknown [5]. Next, only apps available in the New Zealand stores were included. This could have missed comprehensive pain self-management apps that are currently available in other countries (eg, BWH Pain app and Pain Toolkit). Finally, only app descriptions were used in the first screening stage, a standard practice in app-based reviews. Therefore, apps containing suitable self-management skills may have been missed if these skills were not described in the app store description.

Conclusions

This app review evaluated the contents of apps fostering self-management support for people with persistent pain. Of the 3 apps (Curable, PainScale-Pain Diary and Coach and SuperBetter) that met the largest number of items to support skills in self-management of pain, 2 apps (PainScale-Pain Diary and Coach and SuperBetter) were free, suggesting the potential for using apps as a scalable, wide-reaching intervention to complement face-to-face care. However, none provided culturally tailored information. Although 2 apps (Headspace and SuperBetter) were validated to show improved health outcomes, none were tested in people with persistent pain. Both users and clinicians have to be aware of such limitations and make informed choices in using and recommending apps as a self-management tool. Concerted efforts are required among app developers, clinicians, and people with persistent pain in developing and testing apps for better integration of such technologies in clinical practice.

Acknowledgments

The authors would like to thank Arthritis New Zealand and University of Otago Summer Research Scholarship for their financial support for this research. The authors extend thanks to Associate Professor Jean Hay-Smith, PhD, University of Otago, Dunedin, New Zealand for her guidance in completing an app review while undertaking a similar project concurrently. Furthermore, thanks are given to Anya Dai, BSc (Jilin), MSc (Rennin), for her assistance with statistical analysis.

Conflicts of Interest

None declared.

References

1. Hoy DG, Smith E, Cross M, Sanchez-Riera L, Blyth FM, Buchbinder R, et al. Reflecting on the global burden of musculoskeletal conditions: lessons learnt from the global burden of disease 2010 study and the next steps forward. Ann Rheum Dis 2015 Jan;74(1):4-7. [doi: 10.1136/annrheumdis-2014-205393] [Medline: 24914071]
2. GBD 2015 Disease Injury Incidence Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. Lancet 2016 Dec 08;388(10053):1545-1602 [FREE Full text] [doi: 10.1016/S0140-6736(16)31678-6] [Medline: 27733282]
3. Treede RD, Rief W, Barke A, Aziz Q, Bennett MI, Benoliel R, et al. A classification of chronic pain for ICD-11. Pain 2015 Jun;156(6):1003-1007 [FREE Full text] [doi: 10.1097/j.pain.0000000000000160] [Medline: 25844555]
4. GBD 2016 Disease Injury Incidence Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet 2017 Sep 16;390(10100):1211-1259 [FREE Full text] [doi: 10.1016/S0140-6736(17)32154-2] [Medline: 28919117]
5. Nicholas MK, Blyth FM. Are self-management strategies effective in chronic pain treatment? Pain Manag 2016;6(1):75-88. [doi: 10.2217/pmt.15.57] [Medline: 26678703]
6. Vranceanu AM, Stone M, Wallace T, Kulich R. Cognitive Behavioral Therapy for chronic pain. In: Vranceanu A, Greer J, Safren S, editors. The Massachusetts General Hospital Handbook Of Behavioral Medicine. New York: Humana Press; 2017:93-114.
7. Flor H, Turk DC. Chronic Pain: An Integrated Biobehavioral Approach. Seattle: IASP Press; 2011.
8. de Silva D. What is self-management support? In: Evidence: Helping people help themselves. London: Health Foundation; 2011:1-3.
9. Okifuji A, Turk D. Engagement of patients in the self-management of pain. In: O'Donohue W, James L, Snipes C, editors. Practical Strategies and Tools to Promote Treatment Engagement. Cham: Springer; 2017:229-251.
10. Nicholas MK, Asghari A, Corbett M, Smeets RJ, Wood BM, Overton S, et al. Is adherence to pain self-management strategies associated with improved pain, depression and disability in those with disabling chronic pain? Eur J Pain 2012 Jan;16(1):93-104. [doi: 10.1016/j.ejpain.2011.06.005] [Medline: 21705246]
11. Devan H, Hale L, Hempel D, Saipe B, Perry MA. What works and does not work in a self-management intervention for people with chronic pain? Qualitative systematic review and meta-synthesis. Phys Ther 2018 May 01;98(5):381-397. [doi: 10.1093/ptj/pzy029] [Medline: 29669089]
12. Hogg MN, Gibson S, Helou A, DeGabriele J, Farrell MJ. Waiting in pain: a systematic investigation into the provision of persistent pain services in Australia. Med J Aust 2012 Apr 2;196(6):386-390. [doi: 10.5694/mja12.10140] [Medline: 22471539]
13. McGavock Z. Massey University. Wellington; 2011. A life transformed: the lived experiences of Māori with chronic pain URL: https://mro.massey.ac.nz/bitstream/handle/10179/3430/02_whole.pdf?sequence=1&isAllowed=y
14. Househ M, Borycki E, Kushniruk A. Social Media Mobile Technologies for Healthcare. Hershey: IGI Global; 2014.
15. Sundararaman LV, Edwards RR, Ross EL, Jamison RN. Integration of mobile health technology in the treatment of chronic pain: a critical review. Reg Anesth Pain Med 2017;42(4):488-498. [doi: 10.1097/AAP.0000000000000621] [Medline: 28570436]
16. Becker S, Miron-Shatz T, Schumacher N, Krocza J, Diamantidis C, Albrecht U. mHealth 2.0: experiences, possibilities, and perspectives. JMIR Mhealth Uhealth 2014;2(2):e24 [FREE Full text] [doi: 10.2196/mhealth.3328] [Medline: 25099752]
17. Statistica. Number of smartphone users worldwide from 2014 to 2020 (in billions) URL: https://www.statista.com/statistics/330695/number-of-smartphone-users-worldwide/ [WebCite Cache ID 74ZOi3aKh]
18. Arnhold M, Quade M, Kirch W. Mobile applications for diabetics: a systematic review and expert-based usability evaluation considering the special requirements of diabetes patients age 50 years or older. J Med Internet Res 2014;16(4):e104 [FREE Full text] [doi: 10.2196/jmir.2968] [Medline: 24718852]
19. Huckvale K, Car M, Morrison C, Car J. Apps for asthma self-management: a systematic assessment of content and tools. BMC Med 2012;10:144 [FREE Full text] [doi: 10.1186/1741-7015-10-144] [Medline: 23171675]
20. Nguyen AD, Baysari MT, Kannangara DR, Tariq A, Lau AY, Westbrook JI, et al. Mobile applications to enhance self-management of gout. Int J Med Inform 2016 Oct;94:67-74. [doi: 10.1016/j.ijmedinf.2016.06.021] [Medline: 27573313]
21. Slater H, Dear BF, Merolli MA, Li LC, Briggs AM. Use of eHealth technologies to enable the implementation of musculoskeletal Models of Care: Evidence and practice. Best Pract Res Clin Rheumatol 2016 Jun;30(3):483-502 [FREE Full text] [doi: 10.1016/j.berh.2016.08.006] [Medline: 27886943]
22. Christiansen S, Gupta A. Can mobile technology improve treatment of chronic pain? Pain Med 2014 Aug;15(8):1434-1435. [doi: 10.1111/pme.12479] [Medline: 25040764]
23. Rosser BA, Eccleston C. Smartphone applications for pain management. J Telemed Telecare 2011;17(6):308-312. [doi: 10.1258/jtt.2011.101102] [Medline: 21844177]

24. Lalloo C, Jibb LA, Rivera J, Agarwal A, Stinson JN. "There's a Pain App for That": review of patient-targeted smartphone applications for pain management. Clin J Pain 2015 Jun;31(6):557-563. [doi: 10.1097/AJP.0000000000000171] [Medline: 25370138]
25. Gatchel RJ, Peng YB, Peters ML, Fuchs PN, Turk DC. The biopsychosocial approach to chronic pain: scientific advances and future directions. Psychol Bull 2007 Jul;133(4):581-524. [doi: 10.1037/0033-2909.133.4.581] [Medline: 17592957]
26. Anderson KO, Green CR, Payne R. Racial and ethnic disparities in pain: causes and consequences of unequal care. J Pain 2009 Dec;10(12):1187-1204. [doi: 10.1016/j.jpain.2009.10.002] [Medline: 19944378]
27. Brady B, Veljanova I, Chipchase L. Are multidisciplinary interventions multicultural? A topical review of the pain literature as it relates to culturally diverse patient groups. Pain 2016 Feb;157(2):321-328. [doi: 10.1097/j.pain.0000000000000412] [Medline: 26588691]
28. Savigny P, Kuntze S, Watson P, Underwood M, Ritchie G, Cotterell M. National Institute for Health and Clinical Excellence. 2009. Low back pain: early management of persistent non-specific low back pain URL: http://www.sif-fisioterapia.it/wp-content/uploads/2014/12/Management-of-Low-Back-Pain-UK-2009.pdf [accessed 2019-02-05] [WebCite Cache ID 75xYMrWeG]
29. Pain Australia. Pyrmont; 2011. National Pain Strategy URL: https://www.painaustralia.org.au/static/uploads/files/national-pain-strategy-2011-exec-summary-with-stakeholders-wftfjtpbdaij.pdf [accessed 2019-02-02] [WebCite Cache ID 75rtCYC7F]
30. Machado GC, Pinheiro MB, Lee H, Ahmed OH, Hendrick P, Williams C, et al. Smartphone apps for the self-management of low back pain: a systematic review. Best Pract Res Clin Rheumatol 2016 Dec;30(6):1098-1109. [doi: 10.1016/j.berh.2017.04.002] [Medline: 29103552]
31. Bhattarai P, Newton-John TR, Phillips JL. Quality and usability of arthritic pain self-management apps for older adults: a systematic review. Pain Med 2018 Mar 01;19(3):471-484. [doi: 10.1093/pm/pnx090] [Medline: 28541464]
32. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009 Jul 21;6(7):e1000097 [FREE Full text] [doi: 10.1371/journal.pmed.1000097] [Medline: 19621072]
33. Anderson LM, Scrimshaw S, Fullilove MT, Fielding JE, Normand J. Culturally competent healthcare systems. American Journal of Preventive Medicine 2003 Apr;24(3):68-79. [doi: 10.1016/S0749-3797(02)00657-8] [Medline: 12668199]
34. Fnd. Fnd: Experience the App Store and iTunes Anywhere URL: https://fnd.io/ [accessed 2019-02-05] [WebCite Cache ID 75xdg0E0x]
35. StatCounter. Mobile Operating System Market Share New Zealand | StatCounter Global Stats 2017 URL: http://gs.statcounter.com/os-market-share/mobile/new-zealand/#monthly-201611-201710 [WebCite Cache ID 74ZOHrY08]
36. Lorig KR, Holman HR. Self-management education: history, definition, outcomes, and mechanisms. Ann Behav Med 2003 Aug;26(1):1-7. [doi: 10.1207/S15324796ABM2601_01] [Medline: 12867348]
37. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. JMIR Mhealth Uhealth 2015;3(1):e27 [FREE Full text] [doi: 10.2196/mhealth.3422] [Medline: 25760773]
38. Economides M, Martman J, Bell M, Sanderson B. Improvements in stress, affect, and irritability following brief use of a mindfulness-based smartphone app: a randomized controlled trial. Mindfulness (N Y) 2018;9(5):1584-1593 [FREE Full text] [doi: 10.1007/s12671-018-0905-4] [Medline: 30294390]
39. Roepke AM, Jaffee SR, Riffle OM, McGonigal J, Broome R, Maxwell B. Randomized controlled trial of SuperBetter, a smartphone-based/internet-based self-help tool to reduce depressive symptoms. Games Health J 2015 Jun;4(3):235-246. [doi: 10.1089/g4h.2014.0046] [Medline: 26182069]
40. Webb TL, Joseph J, Yardley L, Michie S. Using the internet to promote health behavior change: a systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. J Med Internet Res 2010;12(1):e4 [FREE Full text] [doi: 10.2196/jmir.1376] [Medline: 20164043]
41. Middelweerd A, Mollee JS, van der Wal CN, Brug J, te Velde SJ. Apps to promote physical activity among adults: a review and content analysis. Int J Behav Nutr Phys Act 2014;11:97 [FREE Full text] [doi: 10.1186/s12966-014-0097-9] [Medline: 25059981]
42. Direito A, Dale LP, Shields E, Dobson R, Whittaker R, Maddison R. Do physical activity and dietary smartphone applications incorporate evidence-based behaviour change techniques? BMC Public Health 2014;14:646 [FREE Full text] [doi: 10.1186/1471-2458-14-646] [Medline: 24965805]
43. Veehof MM, Trompetter HR, Bohlmeijer ET, Schreurs KM. Acceptance- and mindfulness-based interventions for the treatment of chronic pain: a meta-analytic review. Cogn Behav Ther 2016;45(1):5-31. [doi: 10.1080/16506073.2015.1098724] [Medline: 26818413]
44. Oinas-Kukkonen H, Harjumaa M. Persuasive systems design: key issues, process model, and system features. Commun Assoc Inf Sys 2009;24:28 [FREE Full text] [doi: 10.17705/1CAIS.02428]
45. Moseley GL, Butler DS. Fifteen years of explaining pain: the past, present, and future. J Pain 2015 Sep;16(9):807-813. [doi: 10.1016/j.jpain.2015.05.005] [Medline: 26051220]

46. Nielson WR, Jensen MP, Karsdorp PA, Vlaeyen JW. Activity pacing in chronic pain: concepts, evidence, and future directions. Clin J Pain 2013 May;29(5):461-468. [doi: 10.1097/AJP.0b013e3182608561] [Medline: 23247005]
47. Murphy SL, Kratz AL. Activity pacing in daily life: a within-day analysis. Pain 2014 Dec;155(12):2630-2637 [FREE Full text] [doi: 10.1016/j.pain.2014.09.028] [Medline: 25267209]
48. Felipe S, Singh A, Bradley C, Williams AC, Bianchi-Berthouze N. Roles for personal informatics in chronic pain. In: Proceedings of the 9th International Conference on Pervasive Computing Technologies for Healthcare. 2015 Presented at: 9th International Conference on Pervasive Computing Technologies for Healthcare; May 20-23, 2015; Istanbul, Turkey p. 161-168.
49. Singh A, Piana S, Pollarolo D, Volpe G, Varni G, Tajadura-Jiménez A, et al. Go-with-the-flow: tracking, analysis and sonification of movement and breathing to build confidence in activity despite chronic pain. Hum Comput Interact 2015 Nov 25;31(3-4):335-383 [FREE Full text] [doi: 10.1080/07370024.2015.1085310]
50. López-Martínez AE, Esteve-Zarazaga R, Ramírez-Maestre C. Perceived social support and coping responses are independent variables explaining pain adjustment among chronic pain patients. J Pain 2008 Apr;9(4):373-379. [doi: 10.1016/j.jpain.2007.12.002] [Medline: 18203665]
51. Dennis CL. Peer support within a health care context: a concept analysis. Int J Nurs Stud 2003 Mar;40(3):321-332. [Medline: 12605954]
52. Ruehlman LS, Karoly P, Enders C. A randomized controlled evaluation of an online chronic pain self management program. Pain 2012 Feb;153(2):319-330 [FREE Full text] [doi: 10.1016/j.pain.2011.10.025] [Medline: 22133450]
53. Lorig KR, Ritter PL, Laurent DD, Plant K. The internet-based arthritis self-management program: a one-year randomized trial for patients with arthritis or fibromyalgia. Arthritis Rheum 2008 Jul 15;59(7):1009-1017 [FREE Full text] [doi: 10.1002/art.23817] [Medline: 18576310]
54. Devan H, Perry M, Shipton E, Dowell A, Jones B, Ingham T, et al. Health Informatics New Zealand (HiNZ) Conference. 2017. Co-creating a self-management eHealth resource by the people and for the people with persistent pain URL: https://cdn.ymaws.com/hinz.site-ym.com/resource/resmgr/conference_2017/HiNZ17_proceedings.pdf [accessed 2018-12-10] [WebCite Cache ID 74ZP3Cxae]
55. Henschke N, Lorenz E, Pokora R, Michaleff ZA, Quartey JN, Oliveira VC. Understanding cultural influences on back pain and back pain research. Best Pract Res Clin Rheumatol 2016 Dec;30(6):1037-1049. [doi: 10.1016/j.berh.2017.08.004] [Medline: 29103548]
56. Ministry Of Health. Wellington: Ministry of Health. 2017. Annual Data Explorer URL: https://minhealthnz.shinyapps.io/nz-health-survey-2016-17-annual-data-explorer/_w_d6435d9e/_w_6c9f494d/ [accessed 2018-12-10] [WebCite Cache ID 74ZPAup0E]
57. Mcgavock ZC, Barnes HM, Mccreanor T. Māori and pain: a literature review. AlterNative 2012 Jun;8(2):163-175 [FREE Full text] [doi: 10.1177/117718011200800205]
58. Portelli P, Eldred C. A quality review of smartphone applications for the management of pain. Br J Pain 2016 Aug;10(3):135-140 [FREE Full text] [doi: 10.1177/2049463716638700] [Medline: 27583140]
59. Brady B, Veljanova I, Chipchase L. An exploration of the experience of pain among culturally diverse migrant communities. Rheumatol Adv Pract 2017;1(1):1-14 [FREE Full text] [doi: 10.1093/rap/rkx002]
60. Lin IB, Ryder K, Coffin J, Green C, Dalgety E, Scott B, et al. Addressing disparities in low back pain care by developing culturally appropriate information for Aboriginal Australians: "My Back on Track, My Future". Pain Med 2017 Nov 01;18(11):2070-2080. [doi: 10.1093/pm/pnw314] [Medline: 28087847]
61. Jones L, Jacklin K, O'Connell ME. Development and use of health-related technologies in indigenous communities: critical review. J Med Internet Res 2017 Dec 20;19(7):e256 [FREE Full text] [doi: 10.2196/jmir.7520] [Medline: 28729237]
62. Pincus T, Esther R, DeWalt DA, Callahan LF. Social conditions and self-management are more powerful determinants of health than access to care. Ann Intern Med 1998 Sep 01;129(5):406-411. [doi: 10.7326/0003-4819-129-5-199809010-00011] [Medline: 9735069]
63. Jerant AF, von Friederichs-Fitzwater MM, Moore M. Patients' perceived barriers to active self-management of chronic conditions. Patient Educ Couns 2005 Jun;57(3):300-307. [doi: 10.1016/j.pec.2004.08.004] [Medline: 15893212]
64. Sullivan MD, Ballantyne JC. Must we reduce pain intensity to treat chronic pain? Pain 2016 Jan;157(1):65-69. [doi: 10.1097/j.pain.0000000000000336] [Medline: 26307855]
65. Kamper SJ, Apeldoorn AT, Chiarotto A, Smeets RJ, Ostelo RW, Guzman J, et al. Multidisciplinary biopsychosocial rehabilitation for chronic low back pain. Cochrane Database Syst Rev 2014 Sep 02(9):CD000963. [doi: 10.1002/14651858.CD000963.pub3] [Medline: 25180773]
66. Jamison RN, Jurcik DC, Edwards RR, Huang C, Ross EL. A pilot comparison of a smartphone app with or without 2-way messaging among chronic pain patients: who benefits from a pain app? Clin J Pain 2017 Aug;33(8):676-686 [FREE Full text] [doi: 10.1097/AJP.0000000000000455] [Medline: 27898460]

Abbreviations

BWH: Brigham and Women's Hospital

ICC: intraclass correlation coefficient
ICD: International Classification of Diseases
MARS: Mobile App Rating Scale
SMS-14: 14-item self-management support checklist

Edited by G Eysenbach; submitted 10.12.18; peer-reviewed by B Thompson, T Janmohamed; comments to author 02.01.19; revised version received 10.01.19; accepted 13.01.19; published 12.02.19.

Please cite as:
Devan H, Farmery D, Peebles L, Grainger R
Evaluation of Self-Management Support Functions in Apps for People With Persistent Pain: Systematic Review
JMIR Mhealth Uhealth 2019;7(2):e13080
URL: http://mhealth.jmir.org/2019/2/e13080/
doi:10.2196/13080
PMID:30747715

Review

The Current State of Mobile Phone Apps for Monitoring Heart Rate, Heart Rate Variability, and Atrial Fibrillation: Narrative Review

Ka Hou Christien Li[1,2,3*]; Francesca Anne White[3*]; Timothy Tipoe[1,2]; Tong Liu[4], MD, PhD; Martin CS Wong[5], FFPH, FESC, FACC, MD; Aaron Jesuthasan[3], BSc; Adrian Baranchuk[6], FACC, FRCPC, FCCS, MD; Gary Tse[2,7*], FESC, FACC, FHRS, FRCP, PhD; Bryan P Yan[7,8,9], MBBS, FRACP, FRCP, FESC, FACC, MD

[1]Department of Medicine and Therapeutics, Faculty of Medicine, The Chinese University of Hong Kong, Hong Kong, China (Hong Kong)

[2]Li Ka Shing Institute of Health Sciences, Faculty of Medicine, The Chinese University of Hong Kong, Hong Kong, China (Hong Kong)

[3]Faculty of Medicine, Newcastle University, Newcastle upon Tyne, United Kingdom

[4]Tianjin Key Laboratory of Ionic-Molecular Function of Cardiovascular Disease, Department of Cardiology, Tianjin Institute of Cardiology, Second Hospital of Tianjin Medical University, Tianjin, China

[5]Division of Family Medicine and Primary Health Care, JC School of Public Health and Primary Care, Faculty of Medicine, The Chinese University of Hong Kong, Hong Kong, China (Hong Kong)

[6]Division of Cardiology, Kingston General Hospital, Queen's University, Kington, ON, Canada

[7]Division of Cardiology, Department of Medicine and Therapeutics, Faculty of Medicine, The Chinese University of Hong Kong, Hong Kong, China (Hong Kong)

[8]Division of Cardiology, Department of Medicine and Therapeutics, Prince of Wales Hospital, Hong Kong, China (Hong Kong)

[9]Institute of Vascular Medicine, The Chinese University of Hong Kong, Hong Kong, China (Hong Kong)

[*]these authors contributed equally

Corresponding Author:
Bryan P Yan, MBBS, FRACP, FRCP, FESC, FACC, MD
Division of Cardiology
Department of Medicine and Therapeutics, Faculty of Medicine
The Chinese University of Hong Kong
9/F, Lui Che Woo Clinical Sciences Building
Prince of Wales Hospital, Shatin
Hong Kong,
China (Hong Kong)
Phone: 852 35051750
Email: bryan.yan@cuhk.edu.hk

Abstract

Background: Mobile phone apps capable of monitoring arrhythmias and heart rate (HR) are increasingly used for screening, diagnosis, and monitoring of HR and rhythm disorders such as atrial fibrillation (AF). These apps involve either the use of (1) photoplethysmographic recording or (2) a handheld external electrocardiographic recording device attached to the mobile phone or wristband.

Objective: This review seeks to explore the current state of mobile phone apps in cardiac rhythmology while highlighting shortcomings for further research.

Methods: We conducted a narrative review of the use of mobile phone devices by searching PubMed and EMBASE from their inception to October 2018. Potentially relevant papers were then compared against a checklist for relevance and reviewed independently for inclusion, with focus on 4 allocated topics of (1) mobile phone monitoring, (2) AF, (3) HR, and (4) HR variability (HRV).

Results: The findings of this narrative review suggest that there is a role for mobile phone apps in the diagnosis, monitoring, and screening for arrhythmias and HR. Photoplethysmography and handheld electrocardiograph recorders are the 2 main techniques adopted in monitoring HR, HRV, and AF.

Conclusions: A number of studies have demonstrated high accuracy of a number of different mobile devices for the detection of AF. However, further studies are warranted to validate their use for large scale AF screening.

(JMIR Mhealth Uhealth 2019;7(2):e11606) doi:10.2196/11606

KEYWORDS

mobile phone apps; atrial fibrillation; heart rate; arrhythmia; photoplethysmography; electrocardiography; mobile health

Introduction

Over the past years, there has been a significant increase in the number of mobile phone apps focusing on cardiovascular diseases. These apps are designed to monitor cardiovascular risk factors such as obesity, smoking, sedentary lifestyle, diabetes, and hypertension as well as to prevent and manage chronic conditions such as atrial fibrillation (AF) [1-5]. There are currently more than 100,000 available mobile health apps on iTunes and Google Play as well as more than 400 wearable activity monitors [6]. Approximately 64% of adults possess a mobile phone in the United States, and 62% of mobile phone owners use their phones to access health information and obtain education about diseases and health conditions [6]. Previous research has reported the benefits of using mobile health in modifying behavior to improve cardiovascular health through dietary management, physical activity promotion, smoking cessation, weight control, blood pressure control, cholesterol management, and blood sugar measurement [7-12]. However, there remains some ambiguity when it comes to the use of such apps in heart rate (HR) and HR variability (HRV) because of the lack of reliable statistical tests used in studies and the lack of standardized recommendations from professional bodies in terms of screening and monitoring of these parameters, as highlighted by Pecchia et al recently [13].

Currently, mobile phone–based photoplethysmography (PPG) and handheld electrocardiograph (ECG) recorders are the 2 main concepts adopted in monitoring HR, HRV, and AF [14,15]. PPG is an optical technique that detects heartbeats by analyzing changes in skin color and light absorption. Using a mobile app, the PPG sensor detects variations in light intensity via transmission through or reflection from the tissue (the reflectance and transmittance model). The variations in the light intensity are related to changes in the blood perfusion of the tissue, and based on these changes, heart-related information can be retrieved [14,16]. Similarly, the handheld ECG recorder is also based primarily on a mobile phone and mobile phone app interface. However, an additional external component, an ECG sensor unit, is required for the app to function adequately. This will then allow for a standard lead I ECG to be recorded [15]. For the purpose of this review, the term *handheld ECG recorder* will refer specifically to the mobile phone app involving an external ECG sensor component, which is not required by the PPG approach.

Studies have been conducted using mobile phone apps, such as AliveCor, to detect arrhythmias such as AF [17-19], atrial flutter, atrial and ventricular premature beats, bundle branch blocks, and ST-segment abnormalities [20]. As the use of mobile phone apps in cardiovascular care is a rapidly evolving field, this literature review seeks to act as a checkpoint, encapsulating the current state of mobile phone–based PPG and ECG recorders in monitoring HR, HRV, and AF. This will put into perspective what has been effective and applicable along with the relevant challenges encountered to better navigate future research projects into this niche.

Methods

Overview

A PubMed and EMBASE search that primarily concerned the use of mobile phone apps for monitoring HR, HRV, and AF was conducted. The following search string was employed: (“mobile applications” OR “smartphone” OR “digital health”) AND (“atrial fibrillation” OR “heart rate”). The search period was determined as beginning from the earliest available publications in the databases to October 14, 2018. No language restrictions were included.

The search results were subsequently transferred to Microsoft Excel (Microsoft Corp, Redmond WA), and all potentially relevant reports were retrieved as complete manuscripts and abstracts and assessed for compliance with relevance to the 4 core principles of this narrative review: (1) mobile phone apps, (2) HR, (3) HRV, and (4) AF. Overall, 3 authors (CL, FW, and TT) independently reviewed each publication and report, and findings were agreed upon by consensus with input from 2 senior researchers (BY and GT). A total of 1006 papers were found, and 822 publications were excluded after screening due to irrelevance and noncompliance with principles explored. Thus, 124 papers were used for this narrative review (Multimedia Appendix 1).

Results

Atrial Fibrillation and Other Arrhythmias

Atrial Fibrillation—KardiaMobile App

The KardiaMobile app, developed by AliveCor [21], is a commonly used device in studies investigating the monitoring of AF. The equipment comprises a finger pad sensor for the left and right hands, which is attached to the mobile phone, and an app that receives the transmitted information. A 2017 study by Tu et al gives a detailed account of the technology used by the device [22]. When fingers are placed on the sensors, recording begins. Using a 19,000 Hz center frequency and a modulation index of 200 Hz/mV, electrical activity is transmitted to the mobile phone via ultrasound signal frequency modulation to obtain the ECG trace. The frequency and modulation index signal that is transmitted to the microphone of the mobile phone are digitized at 44.1 kHz and 24-bit resolution, respectively. The app is then able to produce an ECG trace from the received signal (300 samples/second, 16-bit resolution), which can either

be viewed as it is produced or be stored as a PDF file for later viewing. On the basis of the absence of P waves and irregularity of the RR (R wave to R wave) interval, the associated algorithm within the app produces a potential AF diagnosis [22].

The Lau et al study [23] took a group of 109 patients, 39 of which were in AF, and used the AliveCor algorithm to analyze the rhythms, obtaining a sensitivity of 87% and specificity of 97% when compared with interpretation by a cardiologist. The algorithm was then altered by increasing the weighting of absent P waves in the diagnostic process, and recordings were conducted on this same set of patients. The optimized algorithm was found to have a sensitivity of 100% and specificity of 96%. A second set of 204 patients, with 48 having known AF, were then analyzed by the optimized algorithm, which obtained a sensitivity of 98% and specificity of 97% in this second dataset. Other studies have largely found the sensitivity of the AliveCor device to be in the region of 90% to 98%, with specificity values ranging from 29.2% to 99% in various studies [24-29]. However, a study on the use of AliveCor in inpatients in cardiology and geriatric wards only found a sensitivity of 54.5% and 78.9%, respectively, for these groups [26], whereas another study published in 2017 used the AliveCor algorithm as a control and found a sensitivity of 71.4% for AliveCor [30]. A study of 13,222 patients identified 101 participants to have previously undiagnosed AF, with 65.3% of these patients reporting no symptoms before this diagnosis [24]. Within the field of pediatrics, accurate single-lead ECG tracings were also obtained in both healthy children and children with arrhythmic abnormalities [31]. The device is generally well tolerated, and patients report that it is easy to use [27]. Implementation of the AliveCor in community health care settings has also been investigated. Patients generally responded positively to the suggestion of an ECG being conducted by a pharmacist while they were attending the pharmacy for other purposes [32]. However, some patients were unwilling to have this conducted for the fear of discovering something was wrong. In a primary care setting, the use of AliveCor was well received by patients, but the staff had some reservations about being involved because of a self-perceived lack of knowledge [33]. Patients also report that they felt more confident in managing their AF by using the app [27,34]. The AliveCor device [18,35-37] has also been incorporated with patient education to increase the knowledge of the patients about AF [38]. Mobile apps have been extended to monitor paroxysmal AF for intermittent anticoagulation [39]. This evidently raises public health issues of systematic and opportunistic screening. In a recent evaluation of the European Heart Rhythm Association Consensus on screening AF, opportunistic screening was recommended for individuals aged above 65 years and younger patients at high risk of stroke [40]. As such, it is important to understand local screening guidelines before systematic or opportunistic screenings commence as large number of false positives yields unnecessary, expensive, and potentially harmful follow-ups that might affect patients' mental health.

Atrial Fibrillation—CardiioRhythm App

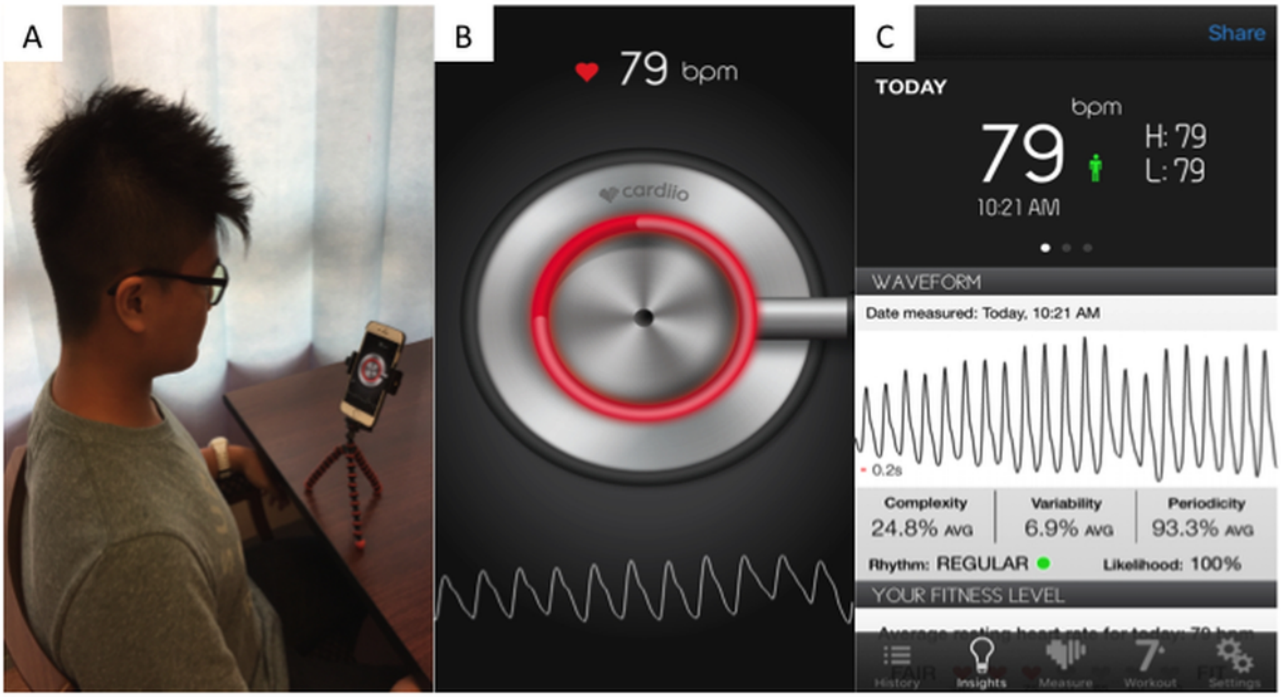
The detection of AF by analysis of finger and facial PPG signals using the CardiioRhythm mobile phone app has been tested in 2 recent studies. In the study by Chan et al [30], finger PPG waveforms of 20-seconds duration were captured 3 times using an iPhone 4S equipped with the CardiioRhythm app and compared with the AliveCor single-lead ECG as referenced. Finger PPG waveforms were captured using the light emitting diode flash of the iPhone and the camera, which detected reflected light to measure the arterial pulsation. These waveforms were recorded at 30 Hz, with the duration of each recording being 17.1 seconds, and were filtered with a bandpass filter using a range of 0.7 Hz to 4.0 Hz. AF was diagnosed from this if 2 of the 3 PPG recordings were irregular. This was defined as a lack of repeating pattern, using a Support Vector Machine to classify the waveforms as similar or nonsimilar to other waveforms. Analysis of the diagnoses put forward by the CardiioRhythm app in this study showed a sensitivity of 92.9% and specificity of 97.7% in detecting AF compared with a sensitivity of 71.4% for the AliveCor device that was obtained by this study. This study found the specificities for the 2 devices to be comparable. The first prospective, international, 2-center, clinical validation study (DETECT AF PRO) based on similar finger PPG technology was also conducted recently to demonstrate the feasibility of such apps alone with a 5-min sensitivity and specificity of 91.5% and 99.6%, respectively [41].

In another study by Yan et al, the CardiioRhythm app was used to analyze facial PPG signals in 217 patients recorded using the front camera of an iPhone 6S without physical contact (Figure 1) [42]. Overall, 3 successive 20-second recordings were acquired per patient. Pulse irregularity in 1 or more PPG readings or 3 uninterpretable PPG readings were considered a positive AF screening result. The CardiioRhythm facial PPG app demonstrated a sensitivity of 94.7% and specificity of 95.8%. The positive and negative predictive values were 92.2% and 97.1%, respectively.

Atrial Fibrillation—FibriCheck App

A recent study conducted by Mortelmans et al [43] used the PPG app FibriCheck. The patient places the left index finger over the flashlight and camera, holding their finger horizontally and keeping the finger in place for 1 min. To ensure an accurate reading, the screen becomes red when appropriate contact has been made, and the practitioner manually commences the measurement. The amount of light that is reflected onto the camera is then captured and used to calculate the variation in local arteriole blood volume pulse variation. The rhythm of the pulse is then identified based on the RR interval. The app in the mobile phone then judges the quality of the signal based on the detection of the pulse. If the detection was masked with noise, or beats were absent, these recordings were not included in the app's analysis. When measurements were detected as a good signal, these data were interpreted by the AF algorithm.

Figure 1. (A) Patient setup; (B) cardiioRhythm app facial photoplethysmography analysis interface; and (C) complete facial photoplethysmography signals report.



Concurrently with the PPG measurement, a single-lead ECG was taken and analyzed by the FibriCheck app. Selection of the data took place based on quality, and good quality measurements were analyzed by the AF algorithm using variability between RR intervals to detect irregularities. The FibriCheck app was used on an iPhone 5S. With 242 participants, this study found a sensitivity of 98%, specificity of 88%, and accuracy of 93% for the FibriCheck algorithm when the data were obtained via PPG. Interpretation by the app of the single-lead ECG performed better, yielding a sensitivity of 98%, specificity of 90%, and accuracy of 94% when compared with standard ECG. When comparing the 2 methods by false-positive results, when analyzed by FibriCheck, 8 false positives were produced from a trace obtained by PPG measurement and 11 false positives were produced from a trace obtained by a single-lead ECG.

Atrial Fibrillation—Other Methods of Detection

Some studies have looked at the use of the inertial measurement unit, using modern microelectromechanical (MEMS) sensors already present in mobile phones, to detect AF [44]. The patient is advised to lie supine and place the mobile phone on their chest. The accelerometer of the mobile phone, which detects the orientation of the phone, can detect cardiogenic movements of the chest. The movements detected are those caused by the opening of the aortic valve, and so, the interval between each successive valve opening is recorded and considered to determine the presence of AF. This yielded a sensitivity of 93.8% and specificity of 100% in the detection of AF. A similar study used the MEMS sensors of mobile phones to obtain measurements, which when analyzed by machine learning methods found a sensitivity of 98.5% and specificity of 95.2% in its best performing method [45].

The feasibility of producing a wearable sensor that records HR and transmits information to a mobile phone app was examined [46]. The way this app identifies AF is similar to the FibriCheck app. As an example, a 42-year-old male admitted with new-onset AF of undetermined duration was deemed appropriate for electrocardioversion because of his FitBit Charge HR device recording the time at which increased HR was observed, allowing the onset of AF to be identified [47]. In a study by Yan et al, the diagnostic accuracy of the CardiioDeepRhythm app, a deep convolutional neural network for detecting AF from the PPG signal acquired using an off-the-shelf wrist-worn device (Empatica E4, Milan, Italy), was tested in 51 in-hospital patients reporting a sensitivity and specificity of 93% and 94%, respectively [48]. These findings demonstrate the promising value of PPG sensors for ambulatory AF monitoring. Krivoshei et al [49] conducted a study investigating the use of PPG in AF, with a view to implementing the technology in a smartwatch; the protocol for this trial was published as an abstract during a conference [50]. Table 1 summarizes the sensitivities and specificities of the apps in AF detection. Despite some studies conducting direct comparisons between different technological modalities [30], it is still done with reference to the gold standard of a conventional 12-lead ECG within a short period. This is especially important because none of the new app technologies have been properly integrated into the wider health care network as compared with the 12-lead ECG, which is a validated instrument commonly available within primary and secondary care settings.

Table 1. Sensitivities and specificities for the 5 main technologies applied in atrial fibrillation detection.

Modalities	Sensitivity (%)	Specificity (%)
Original AliveCor algorithm (finger pad sensor)	54.5-98.0	29.2-99.0
Optimized AliveCor algorithm (finger pad sensor)	98-100	96-97
CardiioRhythm (finger PPG[a])	92.6-92.9	94.8-97.7
CardiioRhythm (facial PPG)	94.7	95.8
CardiioRhythm (wristband PPG)	93	94
FibriCheck (finger PPG)	98	88-90
Inertial measurement unit (chest accelerometer)	93.8-98.5	95.2-100

[a]PPG: photoplethysmography.

Heart Rate and Heart Rate Variability

Heart Rate

There has been considerable innovation in the use of mobile phone apps for the purpose of HR monitoring in adults. In particular, numerous studies have investigated the potential of 2 specific methods—seismocardiogram (SCG) and PPG—in producing accurate HR measurements. Landreani et al, for instance, described the application of SCG and ballistocardiogram (BCG) signals for HR monitoring apps in mobile phone–embedded accelerometers [51]. This is achieved by determining the RR interval of sufficient amplitude and subsequently detecting the fiducial peak from the SCG and BCG signals. In PPG, the accuracy of this method was also demonstrated in a study by Alaleef et al, who found a 99.7% accuracy and maximum absolute error of 0.4 beats/min [52]. The accuracy and feasibility of PPG signals from mobile phone apps also varied between different types of PPG apps. In Parpinel et al's study, higher feasibilities and accuracies were found for contact PPG–based apps compared with the noncontact PPG–based ones [53]. Koenig et al [54] found that their algorithm using PPG to assess HR had a consistent accuracy when compared with ECG (correlation index R>.99).

Gold Standards and Validation of Mobile Phone Heart Rate Apps

Although there is currently no consensus on the gold standard for the validation of HR apps, Vanderberk et al suggested the comparison of the HR on mobile phone apps with an ECG system via RR intervals. With regard to the actual accuracy of these mobile phone apps, there was no significant difference (P=.92) found between the interval measurements of the HR app and the ECG system, suggesting that the HR apps have a comparable accuracy with ECGs [55]. This is further supported by a recent comparison of 3 HR apps against simultaneous standard ECG monitoring, which found an acceptable correlation. However, accuracy became questionable in specific cases of irregular rhythms such as AF [56].

Heart Rate Apps in Different Fields

The use of HR apps has been tested in different clinical settings. In a specific app reviewed by Chaudhry et al, named *Unique Heart Rate Monitor*, PPG was used for the measurement and categorization of workout intensity. It also allows subjects to check their HR in response to medical therapy [57]. The clinical use of HR apps during exercise was also assessed by numerous studies. In particular, some studies noted that the measurement accuracy tended to differ at different exercise intensities. In 1 study, the app had higher measurement errors at increased exercise intensities when compared with Holter echocardiography monitors; however, it had similar accuracy to the Holter monitor at resting and recovery stages [58]. Yan et al presented similar findings, although the correlations between the facial PPG–estimated HR and ECG HR had smaller variations (r=.997 on resting compared with r=.982 with postmoderate intensity exercise) [59]. Furthermore, it was also found that both iOS and Android operating systems, with the same HR app, had concurrent validity with an FT7 Polar HR monitor at rest and at postexercise time points [60]. Although many studies reported the use of mobile phones as the media for measuring and interpreting HRs, some studies used mobile phone apps as a feedback system rather than the actual HR sensor. By using Bluetooth communication, subjects had their HR measured by a specially devised HR sensor at different exercise intensities, and this information was transferred via Bluetooth back to the mobile phone for analysis based on existing information during the pre-exercise period [61]. In addition, some cardiac HR monitors have a built-in alert system with real-time bradycardia and tachycardia arrhythmia detection. In the study by Golzar et al, users reported almost zero delays with data transmission and a 91.62% performance accuracy in comparison with regular ECG monitoring [62].

Apart from using HR apps for exercise HR monitoring, other studies also investigated the potential use of HR monitoring to reduce the exercise-induced risk of hypoglycemia. This is done within the context of control-to-range closed-loop artificial pancreas systems. By integrating HR as part of the feedback system in the Android app, it was found that the inclusion of HR monitoring better controlled blood glucose decline during exercise (P=.02) and resulted in fewer hypoglycemic events during exercise (P=.16 for none vs 2 events) [63].

The use of camera-based PPG has also been suggested for use in other age groups, including infants and newborns. Although the method has reported inaccuracies because of subject movements throughout the day, Kevat et al found that this method of HR measurement had increased clarity and precision in neonates receiving phototherapy [64]. Although this specific application of PPG may have potential use in pediatric patients, the inaccuracies of this method remain. Thus, when PPG is

compared with normal ECG monitoring, different phone apps had lower accuracies for HR measurements using the finger or toe but higher accuracies at the earlobes. HR measurements were inaccurate above 120 bpm [65]. These results suggest that the technique is still evolving and requires considerable research before implementing it as a standard alternative to regular ECG HR monitoring for pediatric populations.

Apart from HR measurements in neonates and children, Android-based HR apps have also been developed to track and increase adherence to certain activities such as breathing awareness meditation. In the study by Gregoski et al, a tension tamer HR app used PPG via phone camera lens to transmit time-stamped HRs back to the hospital server for real-time adherence collection [66]. However, the very small sample size prevents definite conclusions to be drawn. The inaccuracies of these mobile phone apps were also found in the context of fetal HR monitoring by Soffer et al, who reviewed 30 unique apps and found that over 33% of the apps did not put disclaimers and/or provided false medical information [67].

Although PPG has been widely investigated, the use of mobile phones to generate phonocardiograms (PCGs) has been reviewed by many investigators such as Chen et al, who used iPhone 4S to record heart sounds from subjects at rest or postexercise and used these sounds for HR calculation in PCGs via the peak-based detection method [68]. Although the use of this Web-based PCG-template extraction and matching was found to be accurate, further research needs to be done to validate this finding.

Heart Rate Variability

Apart from HR, HRV measurements in mobile phone apps have also been extensively researched. Sometimes coined as the *RR interval*, HRV is the measure of the variation in time intervals between heartbeats, usually with reference to ventricular contraction. Classically, HRV is measured using the ECG using mainly the time-domain and frequency-domain approaches introduced in 1996 [69,70].

With the use of PPG, Bolkovsky et al used both Android and iPhone mobile phones to obtain RR intervals and subsequently deduce HRV via complex HRV algorithms. Although the results were statistically same as the ECG gold standard, the main issue was the insufficient sampling rate in both phones (20 Hz in Android and 30 Hz in iPhone), which was below the suggested rate of 250 Hz [71]. mobile phone PPG has also been advocated by Plews et al, who reported almost perfect correlations of PPG with the ECGs (R=.99), with an acceptable technical error of estimates and trivial differences in standardized differences [72]. The accuracy and reliability of mobile phone PPGs have also been validated in other studies; in some cases, mobile phone PPGs were found to have comparable accuracies with commonly used HRV computer software programs (R=.92) [73]. Other methods of HRV measurement have also been studied, such as seismocardiography. Although sampling frequency from mobile phone devices accounted for a significant source of error, RR series measurements differed by less than 10 ms in some studies, suggesting the comparable accuracy of this measuring technique [74].

Although many studies support mobile phone PPG, common errors in using this low-cost technology include frequent noise and artifacts in measurement. To reduce the impact of noise, however, Huang et al proposed the use of a continuous wavelet transform denoising technique to extract the pulse signal and subsequently deduce the RR intervals in the denoised signal. The experiments showed low mean absolute errors of only 3.53 ms, highlighting the efficacy of this proposed method in reducing mobile phone PPG errors. The use of this technique for error reduction should thus be investigated further [75].

Although methods have been proposed to reduce the impact of noise and artifacts in HRV measurement, other studies have investigated the effect of mobile phone models on the SD of beat-to-beat error measurement (SDE) of HRV indices. In 1 study, 2 different mobile phone models (Samsung S5 and Motorola) showed significant device influence in the supine posture measurement, with the Motorola model having a higher SDE than the Samsung S5 [76].

Application of Heart Rate Variability

In 1 case report, Lai et al reported the potential use of HRV measurement analysis in monitoring concussed athletes as well as assessing their capacity to *return-to-play* [77]. Via an *AliveCor mobile phone ECG* app, HRV parameters were statistically significant in symptomatic and recovered concussed athletes. However, the study noted that this difference was affected by the severity of traumatic brain injury, whereas other similar studies reported no difference in HRV parameters [78,79]. Other studies also proposed the potential use of mobile phone–derived HRV in athletic training programs; for instance, collection of daily HRV data on mobile phones using ultrashort HRV measures provides trainers with numerical indicators on athlete coping and adaptation [80]. Similar studies by Flatt et al also found that measuring the log-transformed root mean square of successive RR intervals (RMSSD) multiplied by 20 (lnRMSSDx20) obtained by mobile phone apps were sensitive markers to the changes in training load in soccer team's training program [81].

HRV by mobile phone apps have also been used as an accurate predictor of acute mountain sickness (AMS) at high altitudes. In 1 study, Mellor et al reported lower HRV scores in those with mild and severe AMS compared with those without AMS (P=.007) and found that a reduction in HRV greater than 5 had an 83% sensitivity and 60% specificity of identifying severe AMS. As this is the first study of its kind, further studies should be conducted to confirm these findings [82].

HRV measurement by mobile phone apps was also investigated with regard to the autonomic nervous system and mental health. For instance, Heathers et al found that pulse rate variability (PRV) could be measured by mobile phone substitutes with accuracies ranging from 2% to 5% for low- and high-frequency spectral power, respectively [83]. Mobile phone apps, thus, may play a future role in psychophysiological research. In addition, mobile phones have also been used in stress classification via the use of night-time HRV data and as a monitoring tool for mental stress in different psychological settings [84,85]. The drawbacks of these studies, however, are the fairly low accuracies (59%) and that the HRV measurements were confined

to ultra-short periods. The potential for long-term measurement should thus be investigated [85].

Other Arrhythmias

Although a large proportion of the research has investigated the use of mobile apps to detect AF, other arrhythmias have been investigated. A study presented in abstract in 2016 by Dimarco et al [86] used the AliveCor device to investigate 148 patients who had reported palpitations for arrhythmias. The device was used as an alternative to ambulatory monitoring.

An ECG trace was obtained using the device, and the trace was then interpreted by a cardiologist. The traces obtained were of suitable quality to diagnose arrhythmias, with diagnoses of supraventricular or ventricular ectopy in 27.4% of patients, sinus tachycardia in 18.6%, AF in 7.1%, and supraventricular tachycardia in 5.3%, which could be managed appropriately. In the remaining 41.6%, they could be reassured that they remained in sinus rhythm.

When the AliveCor device was used to measure corrected QT (Q wave to T wave) interval in those with sinus rhythm and antiarrhythmic drugs, results suggested that it was comparable with 12-lead ECG [87]. With a sensitivity of 64% and specificity of 97%, this suggested that the device could be used to monitor those with arrhythmias. A study comparing the AliveCor device with a 14-day event monitor found that the device was equivalent to the event monitor in detecting AF and premature atrial contractions (PAC) but was slightly less effective in identifying episodes of premature ventricular contractions (PVC) [88].

The Smartphone Pediatric ElectrocARdiogram trial [89] investigated the use of the AliveCor device in pediatric patients with known paroxysmal arrhythmias. The device produced a trace, which was then reviewed by a cardiologist. Of the 240 traces recorded by 20 patients, 231 were suitable for diagnostic purposes. A total of 35 patients were initially included in the trial; 15 of these patients did not transmit ECGs. Of the suitable traces, sinus rhythm was detected in 43% of the traces. Moreover, 29% displayed sinus tachycardia, supraventricular tachycardia was present in 16%, and AF was present in 8%. Furthermore, 98% of parents of patients involved in the trial reported greater comfort in managing their child's arrhythmia, with 93% expressing a desire to continue using the device.

A study by Mc et al [90] assessed whether the Pulse Waveform Analysis app could distinguish sinus rhythm and AF from PAC and PVC. The patient covered the camera and lamp of an iPhone 4S with their finger for 2 min, allowing a pulse waveform to be recorded and a video of blood flow to be recorded. A total of 3 statistical techniques were then used in combination—RMSSD, Poincare plot, and Shannon Entropy (ShE)—to distinguish AF from ectopic beats. Results for each of the arrhythmias were as follows: for AF, a sensitivity of 97%, specificity of 93.5%, and accuracy of 95.1%; for PAC, a sensitivity of 66.7%, specificity of 98%, and accuracy of 95.5%; and for PVC, a sensitivity of 73.3%, specificity of 97.6%, and accuracy of 96%. A similar study assessed the use of a mobile phone app identifying PVC using RR intervals, finding a sensitivity of 90.13% and specificity of 82.52% [91]. A study using similar methods found that the best results were obtained when RMSSD and ShE were combined to analyze the trace. When these 2 techniques were combined in a study of 76 participants, the results were 100% accurate for detecting AF and 96.05% accurate for detecting sinus rhythm [92]. Using simulated ECG signals from the Massachusetts Institute of Technology–Beth Israel Hospital arrhythmia database, a study tested the ability of an Android program to detect arrhythmias from ECG traces [93]. The analog ECG traces were converted to a digital format and then made available for transmission via Bluetooth. The Pan-Tompkins algorithm [94] was then used to identify RR intervals to determine the heartbeat. The algorithm produced an accuracy of 98.98% at detecting sinus rhythm and an average accuracy of 98.34% at detecting 7 different possible heart rhythms.

Discussion

Limitations

The main limitations of this narrative review are the lack of focus on more specific arrhythmias beyond normal sinus rhythm and AF. However, normal sinus rhythm and AF are epidemiologically the most common types of heart rhythm; hence, it is important to establish a strong foundation for these apps with these rhythms first. Second, it is common for narrative reviews to circumvent important criteria to mitigate bias because of the lack of a rigorous evidence-based methodology. This inevitably leads to selection bias based on expert opinions. This bias was addressed in this narrative review by adapting the Preferred Reporting Items for Systematic Reviews and Meta-Analyses criteria to select relevant publications. In doing so, all evidence pertaining to the issues discussed in this review has been considered to provide a balanced understanding and discussion.

Photoplethysmography

Despite being effective in accessing HR and HRV, the applications of PPG monitoring are limited by multiple confounders such as finger pressure, skin tone, light intensities, and user movement leading to artefactual measurements [95,96]. As such, this will have an impact on the feasibility and reliability of mobile phone–based PPG within clinical practice. A minimum sampling rate is necessary for clinically accurate measurements—30 Hz for HR and 200 Hz for HRV measurements [97,98]. However, the frame rate of mobile phones usually operates around 30 Hz, which is a major limitation identified by Bolkhovsky et al [71]. It was proposed that using cubic interpolation, the second derivative, and the zero-crossing algorithm instead of minima detection would overcome this limitation and allow for better HR detection [99-102]. A filter is required to remove artifacts without compromising the original signal when conducting a time-domain analysis to evaluate small variations occurring in normal-to-normal intervals [103]. Examples of filters are independent component analysis using accelerometer data to remove artifacts [104] or employing a fourth-order bandpass filter [99,101]. A heating problem was also noticed by Garcia-Agundez et al when testing their algorithm, which resulted in complaints about the discomfort of holding the mobile phone [102]. Furthermore, although the SCG

accelerometers could be used without supporting ECG signals, HR detection was only possible if patients were motionless and supine, leading to considerable variability in data collection because of different measurement positions. There were also difficulties encountered in setting appropriate R-peak thresholds and in comparing the data with simultaneous ECG signals (the gold standard technique) [105,106]. Another area to consider is the use of PP (P wave to P wave) intervals and PRV instead of the RR interval and HRV. Although most studies use RR and HRV approaches with supporting literature that time domain, frequency domain, and Poincare plot HRV parameters computed using RR interval and PP interval methods showed no significant differences, the PP variability was found to be accurate (0.1 ms) compared with RR variability [107]. Furthermore, most studies usually compare between normal sinus rhythm and AF. However, a simple ectopic will be able to derange specificity to unacceptable levels. In this respect, more research needs to done to evaluate the impact of short-term arrhythmic abnormalities on PPG-based and single-lead ECG–based AF detection modalities.

Handheld Electrocardiograph Recorders

Handheld single-lead ECG recorders, which correspond to the standard lead I, do not cater to potential positional variability. For example, in cases where the heart is positioned more vertically, the amplitude of the QRS complexes may be reduced and comparable with the amplitude of artifacts [15]. Furthermore, a single recorded lead I ECG also does not allow for differentiation between types of narrow and wide complex QRS tachycardia [15]. Although small studies have been conducted to show that it is plausible to use a single-lead ECG to diagnose the ST-segment elevation acute coronary syndrome [108], it is not sufficient to warrant routine clinical use. In some situations, new materials and sensors (eg, biopatches) have allowed these ECGs to be recorded from atypical places (eg, mastoid area) [109] and under different environmental conditions such as after immersion in water or in areas with magnetic fields between 1.5 and 3 T (during magnetic resonance imaging) [110-113]. Novel biopatches have also allowed other parameters such as respiratory rate, body position, temperature, and quality of sleep or physical activity to be monitored, which aids in excluding ST depression caused by increased physical activity or changes in body position [19,114,115]. However, when this is coupled with fewer leads, not only will there be suboptimal signal-to-noise ratio but signals generated might also be different from those produced by standard ECG leads [15]. Furthermore, the utility of biopatches comes with a drawback of having a short intersensor distance, which could potentially compromise the quality of the P wave recorded [15].

Health Care Infrastructural Implementation

Despite being a promising concept, there are currently clear limitations to the use of mobile phone apps for HR, HRV, and AF monitoring. Looking beyond the technological drawbacks, there is little evidence as to how mobile phone monitoring will be matched with health care infrastructural changes to allow for such data to sync with the electronic medical record [116,117]. Furthermore, not all patients will have the same device or mobile phone, which may result in different recorded results. Special clinical attention also has to be given to incidental findings of ECG abnormalities such as ventricular and supraventricular tachycardia in asymptomatic low-risk patients [15]. The impact of app use on health care services is also not to be underestimated. The necessity for a health care professional to confirm arrhythmias detected by these apps will result in a significant burden on services. Given the transient nature of some arrhythmias, an app with a high specificity is likely to be more beneficial to avoid over investigation and treatment of patients, reducing the impact on existing services.

A recent large-scale community AF screening program on more than 10,000 patients aged above of 65 years was conducted, proving the feasibility of integrating this technological within-health care services to identify with new AF [118]. The main drawback illustrated by this study is the lack of downstream management pathways. This raises the medico-legal aspect of introducing mobile phone apps. First, there must be an understanding as to how these apps will *support* decision making or purport to intervene in clinical decisions. This, in turn, will lay the foundation for a robust government framework to evaluate its effect on clinical outcomes and potential unintended consequences [119,120]. A deeper engagement in this respect was raised in January 2016 when 2 class action lawsuits were filed against such manufacturers, calling into question the reliability and accuracy of these devices and apps [121]. Perhaps the issue is in classifying these devices and apps as certified *medical devices*. Despite the intuitive need for medical devices to be of high quality, local regulations for *medical devices* have to be followed. Such regulations range from the Medical Devices Framework concept of *intended purpose* to the risk-based case-by-case approach employed by the US Food and Drug Administration [122].

Looking specifically at the European regulations, the Medical Device Directive was replaced by the new Medical Device Regulation. Although the new regulatory framework still revolves around *intended purpose*, the bar for *medical device* classification has been lowered because of a broader definition assigned to the term *medical purpose*. In short, despite the undisputed potential of such apps in the detection of arrhythmias, adhering to appropriate laws and regulations remains a significant hurdle to address [123].

Conclusions and Future Directions

In this study, we conducted a narrative review of the literature surrounding the usage of mobile phone apps in the monitoring of HR and rhythm. The findings of this narrative review suggest that there is a role for mobile phone app in the diagnosis, monitoring, and screening of arrhythmias and HR. The usage of apps in specific situations, such as during and following exercise or to measure corrected QT interval following the administration of medications, has also shown a role for the apps in more specific scenarios. Although the majority of literature reviewed focused on adult patients, the use of PPG apps in pediatric and neonatal patient populations requires further studies.

Some problems identified with the use of these devices have included patient movement resulting in artifact and in positional variability in patient usage affecting results. This has been

demonstrated in studies investigating both adult and pediatric patients [89]. There is also an issue regarding consistency and availability of mobile phones, with variation in the devices used. Within the context of HR monitoring and AF detection, given the impressive degree of sensitivity (>90%) and specificity (>90%) in most cases or apps, neither sensitivity nor specificity is more important than the other. Instead, it is important for both sensitivity and specificity to be maintained at a high level or be on par with the standard ECG, especially in patients with pacemakers or implantable cardioverter defibrillators [124]. Therefore, further studies are required to address these issues before they can be routinely implemented. As such, there are little implications for current practice at this point.

Moving forward, further large-scale validation studies looking beyond normal sinus rhythm and AF are required. For example, it is important for these apps to not mistake malignant ventricular ectopics as sinus rhythm or AF as it will be potentially harmful to patients. With regard to AF detection, multiple studies have been conducted to validate its feasibility. Perhaps it would be best to start conducting large-scale AF screening programs. Doing so will help address the following points: (1) the medico-legal aspect of implementing these apps as valid *medical devices* for systematic and opportunistic screening nationally, (2) the cost-effectiveness of identifying new AF patients through such a screening program, and (3) refinement of a suitable management pathway for new AF patients identified through this process. Such an initiative has the potential to reduce AF burden internationally and increase the efficiency of picking up AF in the community. Within the clinical setting, apps used by patients can prompt clinicians to perform further investigations such as a confirmatory ECG.

Acknowledgments

GT would like to thank the Croucher Foundation of Hong Kong for their support. BY would like to acknowledge the Hong Kong Research Grants Council–General Research Fund for their support (reference number 14118314). The International Research Mentorship Programme organized by the Newcastle University Cardiovascular Society supported FW and TT.

Authors' Contributions

KL, TL, MW, GT, and BY designed and conceptualized the study. FW and TT were responsible for the data collection and literature screening. KL, FW, and TT drafted the manuscript. FW, MW, AJ, and AB revised the manuscript. BY contributed to critical revision of this manuscript. All authors reviewed the manuscript and approved for publication.

Conflicts of Interest

None declared.

Multimedia Appendix 1

A description outlining the search process of papers.

[PDF File (Adobe PDF File), 22KB - mhealth_v7i2e11606_app1.pdf]

References

1. Artinian NT, Fletcher GF, Mozaffarian D, Kris-Etherton P, Van Horn L, Lichtenstein AH, American Heart Association Prevention Committee of the Council on Cardiovascular Nursing. Interventions to promote physical activity and dietary lifestyle changes for cardiovascular risk factor reduction in adults: a scientific statement from the American Heart Association. Circulation 2010 Jul 27;122(4):406-441 [FREE Full text] [doi: 10.1161/CIR.0b013e3181e8edf1] [Medline: 20625115]
2. Moerenhout T, Devisch I, Cornelis GC. E-health beyond technology: analyzing the paradigm shift that lies beneath. Med Health Care Philos 2018 Mar;21(1):31-41. [doi: 10.1007/s11019-017-9780-3] [Medline: 28551772]
3. Appelboom G, LoPresti M, Reginster JY, Sander Connolly E, Dumont EP. The quantified patient: a patient participatory culture. Curr Med Res Opin 2014 Dec;30(12):2585-2587. [doi: 10.1185/03007995.2014.954032] [Medline: 25118077]
4. Martin SS, Feldman DI, Blumenthal RS, Jones SR, Post WS, McKibben RA, et al. mActive: a randomized clinical trial of an automated mHealth intervention for physical activity promotion. J Am Heart Assoc 2015 Nov 9;4(11) [FREE Full text] [doi: 10.1161/JAHA.115.002239] [Medline: 26553211]
5. Dower GE, Yakush A, Nazzal SB, Jutzy RV, Ruiz CE. Deriving the 12-lead electrocardiogram from four (EASI) electrodes. J Electrocardiol 1988;21 Suppl:S182-S187. [Medline: 3216172]
6. Kelli HM, Witbrodt B, Shah A. The future of mobile health applications and devices in cardiovascular health. Euro Med J Innov 2017 Jan 10;2017:92-97 [FREE Full text] [Medline: 28191545]
7. Wing RR, Tate DF, Gorin AA, Raynor HA, Fava JL. A self-regulation program for maintenance of weight loss. N Engl J Med 2006 Oct 12;355(15):1563-1571. [doi: 10.1056/NEJMoa061883] [Medline: 17035649]
8. Franz MJ, VanWormer JJ, Crain AL, Boucher JL, Histon T, Caplan W, et al. Weight-loss outcomes: a systematic review and meta-analysis of weight-loss clinical trials with a minimum 1-year follow-up. J Am Diet Assoc 2007 Oct;107(10):1755-1767. [doi: 10.1016/j.jada.2007.07.017] [Medline: 17904936]

9. Burke LE, Wang J, Sevick MA. Self-monitoring in weight loss: a systematic review of the literature. J Am Diet Assoc 2011 Jan;111(1):92-102 [FREE Full text] [doi: 10.1016/j.jada.2010.10.008] [Medline: 21185970]
10. Jospe MR, Fairbairn KA, Green P, Perry TL. Diet app use by sports dietitians: a survey in five countries. JMIR Mhealth Uhealth 2015 Jan 22;3(1):e7 [FREE Full text] [doi: 10.2196/mhealth.3345] [Medline: 25616274]
11. Turner-McGrievy GM, Beets MW, Moore JB, Kaczynski AT, Barr-Anderson DJ, Tate DF. Comparison of traditional versus mobile app self-monitoring of physical activity and dietary intake among overweight adults participating in an mHealth weight loss program. J Am Med Inform Assoc 2013 May 1;20(3):513-518 [FREE Full text] [doi: 10.1136/amiajnl-2012-001510] [Medline: 23429637]
12. Ghorai K, Akter S, Khatun F, Ray P. mHealth for smoking cessation programs: a systematic review. J Pers Med 2014 Jul 18;4(3):412-423 [FREE Full text] [doi: 10.3390/jpm4030412] [Medline: 25563359]
13. Pecchia L, Castaldo R, Montesinos L, Melillo P. Are ultra-short heart rate variability features good surrogates of short-term ones? State-of-the-art review and recommendations. Healthc Technol Lett 2018 Mar 14;5(3):94-100 [FREE Full text] [doi: 10.1049/htl.2017.0090] [Medline: 29922478]
14. Allen J. Photoplethysmography and its application in clinical physiological measurement. Physiol Meas 2007 Mar;28(3):R1-39. [doi: 10.1088/0967-3334/28/3/R01] [Medline: 17322588]
15. Guzik P, Malik M. ECG by mobile technologies. J Electrocardiol 2016;49(6):894-901. [doi: 10.1016/j.jelectrocard.2016.07.030] [Medline: 27524474]
16. Kamal AA, Harness JB, Irving G, Mearns AJ. Skin photoplethysmography--a review. Comput Methods Programs Biomed 1989 Apr;28(4):257-269. [Medline: 2649304]
17. Lowres N, Neubeck L, Salkeld G, Krass I, McLachlan AJ, Redfern J, et al. Feasibility and cost-effectiveness of stroke prevention through community screening for atrial fibrillation using iPhone ECG in pharmacies. The SEARCH-AF study. Thromb Haemost 2014 Jun;111(6):1167-1176. [doi: 10.1160/TH14-03-0231] [Medline: 24687081]
18. Hickey KT, Hauser NR, Valente LE, Riga TC, Frulla AP, Masterson Creber R, et al. A single-center randomized, controlled trial investigating the efficacy of a mHealth ECG technology intervention to improve the detection of atrial fibrillation: the iHEART study protocol. BMC Cardiovasc Disord 2016 Jul 16;16:152 [FREE Full text] [doi: 10.1186/s12872-016-0327-y] [Medline: 27422639]
19. Shih CM, Lin CW, Clinciu DL, Jian WS, Kuo TB, Nguyen PA, et al. Managing mass events and competitions with difficult-to-access locations using mobile electrocardiac monitoring. Comput Methods Programs Biomed 2015 Sep;121(2):109-115. [doi: 10.1016/j.cmpb.2015.04.014] [Medline: 26027939]
20. Le Page P, MacLachlan H, Anderson L, Penn LA, Moss A, Mitchell AR. The efficacy of a smartphone ECG application for cardiac screening in an unselected island population. Br J Cardiol 2015;22:31-33. [doi: 10.5837/bjc.2015.009]
21. AliveCor. KardiaMobile URL: https://shop.gb.alivecor.com/products/kardiamobile [accessed 2018-12-12] [WebCite Cache ID 74bX3N04K]
22. Tu HT, Chen Z, Swift C, Churilov L, Guo R, Liu X, et al. Smartphone electrographic monitoring for atrial fibrillation in acute ischemic stroke and transient ischemic attack. Int J Stroke 2017 Oct;12(7):786-789. [doi: 10.1177/1747493017696097] [Medline: 28884653]
23. Lau JK, Lowres N, Neubeck L, Brieger DB, Sy RW, Galloway CD, et al. iPhone ECG application for community screening to detect silent atrial fibrillation: a novel technology to prevent stroke. Int J Cardiol 2013 Apr 30;165(1):193-194. [doi: 10.1016/j.ijcard.2013.01.220] [Medline: 23465249]
24. Chan NY, Choy CC. Screening for atrial fibrillation in 13 122 Hong Kong citizens with smartphone electrocardiogram. Heart 2017 Dec 1;103(1):24-31. [doi: 10.1136/heartjnl-2016-309993] [Medline: 27733533]
25. Chan PH, Wong CK, Pun L, Wong YF, Wong MM, Chu DW, et al. Head-to-head comparison of the AliveCor heart monitor and Microlife WatchBP Office AFIB for atrial fibrillation screening in a primary care setting. Circulation 2017 Jan 3;135(1):110-112 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.116.024439] [Medline: 28028066]
26. Desteghe L, Raymaekers Z, Lutin M, Vijgen J, Dilling-Boer D, Koopman P, et al. Performance of handheld electrocardiogram devices to detect atrial fibrillation in a cardiology and geriatric ward setting. Europace 2017 Jan;19(1):29-39. [doi: 10.1093/europace/euw025] [Medline: 26893496]
27. Lowres N, Mulcahy G, Gallagher R, Ben Freedman S, Marshman D, Kirkness A, et al. Self-monitoring for atrial fibrillation recurrence in the discharge period post-cardiac surgery using an iPhone electrocardiogram. Eur J Cardiothorac Surg 2016 Jul;50(1):44-51. [doi: 10.1093/ejcts/ezv486] [Medline: 26850266]
28. Orchard J, Lowres N, Freedman SB, Ladak L, Lee W, Zwar N, et al. Screening for atrial fibrillation during influenza vaccinations by primary care nurses using a smartphone electrocardiograph (iECG): a feasibility study. Eur J Prev Cardiol 2016 Oct;23(2 suppl):13-20. [doi: 10.1177/2047487316670255] [Medline: 27892421]
29. Koltowski L, Balsam P, Glowczynska R, Peller M, Maksym J, Blicharz L, et al. Comparison of Kardia Mobile (one lead ECGs records) with 12-lead ECGs in 100 consecutive patients with various cardiovascular disorders. EP Europace 2017 Jun;19(3):iii353. [doi: 10.1093/ehjci/eux158.263]
30. Chan PH, Wong CK, Poh YC, Pun L, Leung WW, Wong YF, et al. Diagnostic performance of a smartphone-based photoplethysmographic application for atrial fibrillation screening in a primary care setting. J Am Heart Assoc 2016 Jul 21;5(7) [FREE Full text] [doi: 10.1161/JAHA.116.003428] [Medline: 27444506]

31. Gropler MRF, Dalal AS, Van Hare GF, Silva JNA. Can smartphone wireless ECGs be used to accurately assess ECG intervals in pediatrics? A comparison of mobile health monitoring to standard 12-lead ECG. PLoS One 2018 Sep 27;13(9):e0204403 [FREE Full text] [doi: 10.1371/journal.pone.0204403] [Medline: 30260996]
32. Lowres N, Krass I, Neubeck L, Redfern J, McLachlan AJ, Bennett AA, et al. Atrial fibrillation screening in pharmacies using an iPhone ECG: a qualitative review of implementation. Int J Clin Pharm 2015 Dec;37(6):1111-1120. [doi: 10.1007/s11096-015-0169-1] [Medline: 26202627]
33. Orchard J, Freedman SB, Lowres N, Peiris D, Neubeck L. iPhone ECG screening by practice nurses and receptionists for atrial fibrillation in general practice: the GP-SEARCH qualitative pilot study. Aust Fam Physician 2014 May;43(5):315-319 [FREE Full text] [Medline: 24791776]
34. Kropp C, Ellis J, Nekkanti R, Sears S. Monitoring patients with implantable cardioverter defibrillators using mobile phone electrocardiogram: case study. JMIR Cardio 2018;2(1):e5. [doi: 10.2196/cardio.8710]
35. Treskes RW, Gielen W, Wermer MJ, Grauss RW, van Alem AP, Dehnavi RA, et al. Mobile phones in cryptogenic strOke patients Bringing sIngle Lead ECGs for Atrial Fibrillation detection (MOBILE-AF): study protocol for a randomised controlled trial. Trials 2017 Aug 29;18(1):402 [FREE Full text] [doi: 10.1186/s13063-017-2131-0] [Medline: 28851409]
36. Gwynne K, Flaskas Y, O'Brien C, Jeffries TL, McCowen D, Finlayson H, et al. Opportunistic screening to detect atrial fibrillation in Aboriginal adults in Australia. BMJ Open 2016 Nov 15;6(11):e013576 [FREE Full text] [doi: 10.1136/bmjopen-2016-013576] [Medline: 27852724]
37. Soni A, Karna S, Patel H, Fahey N, Raithatha S, Handorf A, et al. Study protocol for Smartphone Monitoring for Atrial fibrillation in Real-Time in India (SMART-India): a community-based screening and referral programme. BMJ Open 2017 Dec 14;7(12):e017668 [FREE Full text] [doi: 10.1136/bmjopen-2017-017668] [Medline: 29247089]
38. Guhl EN, Schlusser CL, Henault LE, Bickmore TW, Kimani E, Paasche-Orlow MK, et al. Rationale and design of the Atrial Fibrillation health Literacy Information Technology Trial: (AF-LITT). Contemp Clin Trials 2017 Nov;62:153-158 [FREE Full text] [doi: 10.1016/j.cct.2017.09.005] [Medline: 28923492]
39. Stavrakis S, Stoner JA, Kardokus J, Garabelli PJ, Po SS, Lazzara R. Intermittent vs. Continuous Anticoagulation theRapy in patiEnts with Atrial Fibrillation (iCARE-AF): a randomized pilot study. J Interv Card Electrophysiol 2017 Jan;48(1):51-60 [FREE Full text] [doi: 10.1007/s10840-016-0192-8] [Medline: 27696012]
40. Özcan EE, Görenek B. Lessons from the current European Heart Rhythm Association consensus document on screening for atrial fibrillation. Anatol J Cardiol 2018 Mar;19(3):222-224 [FREE Full text] [doi: 10.14744/AnatolJCardiol.2018.37043] [Medline: 29521317]
41. Brasier N, Raichle CJ, Dörr M, Becke A, Nohturfft V, Weber S, et al. Detection of atrial fibrillation with a smartphone camera: first prospective, international, two-centre, clinical validation study (DETECT AF PRO). Europace 2018 Jul 31. [doi: 10.1093/europace/euy176] [Medline: 30085018]
42. Yan BP, Lai WHS, Chan CKY, Chan SC, Chan L, Lam KM, et al. Contact-free screening of atrial fibrillation by a smartphone using facial pulsatile photoplethysmographic signals. J Am Heart Assoc 2018 Apr 5;7(8) [FREE Full text] [doi: 10.1161/JAHA.118.008585] [Medline: 29622592]
43. Mortelmans C, Van Haelst R, Van Der Auwera J, Grieten L, Vandervoort P, Vaes B. Validation of a new smartphone application for the diagnosis of atrial fibrillation in primary care. EP Europace 2017 Jun 1;19(3):iii16. [doi: 10.1093/ehjci/eux136.008]
44. Lahdenoja O, Hurnanen T, Iftikhar Z, Nieminen S, Knuutila T, Saraste A, et al. Atrial fibrillation detection via accelerometer and gyroscope of a smartphone. IEEE J Biomed Health Inform 2018 Jan;22(1):108-118. [doi: 10.1109/JBHI.2017.2688473] [Medline: 28391210]
45. Koivisto T, Lahdenoja O, Hurnanen T, Knuutila T, Vasankari T, Kiviniemi T, et al. Detecting atrial fibrillation via existing smartphones without any add-ons. 2016 Presented at: ESC Congress; August 27-31, 2016; Rome.
46. Rincon F, Grassi PR, Khaled N, Atienza D, Sciuto D. Automated real-time atrial fibrillation detection on a wearable wireless sensor platform. 2012 Presented at: 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society; August 28-September 1, 2012; San Diego, CA, USA. [doi: 10.1109/EMBC.2012.6346465]
47. Rudner J, McDougall C, Sailam V, Smith M, Sacchetti A. Interrogation of patient smartphone activity tracker to assist arrhythmia management. Ann Emerg Med 2016 Sep;68(3):292-294. [doi: 10.1016/j.annemergmed.2016.02.039] [Medline: 27045694]
48. Yan BP, Lai HS, Fong AH, Lai PS, Cheng OW, Chan CK, et al. Validation of a deep convolutional network for detecting atrial fibrillation with a wrist-worn wearable device. 2018 Presented at: Heart Rhythm Society's Annual Scientific Sessions; May 9-12, 2018; Boston, MA URL: https://static1.squarespace.com/static/58334ce08419c209d6bf1882/t/5b9abc3fcd8366c8440de450/1536867391780/HRS-Boston-2018-Cardiio.pdf
49. Krivoshei L, Weber S, Burkard T, Maseli A, Brasier N, Kühne M, et al. Smart detection of atrial fibrillation†. Europace 2017 May 1;19(5):753-757 [FREE Full text] [doi: 10.1093/europace/euw125] [Medline: 27371660]
50. Doerr M, Koenig N, Seeck A, Stoeckli R, Rhinisperger M, Djurdjevic A, et al. Trial design of the WATCH AF trial - SmartWATCHes for detection of atrial fibrillation. EP Europace 2017 Jun 1;19(3):iii41. [doi: 10.1093/ehjci/eux141.014]
51. Landreani F, Caiani EG. Smartphone accelerometers for the detection of heart rate. Expert Rev Med Devices 2017 Nov 26;14(12):935-948. [doi: 10.1080/17434440.2017.1407647] [Medline: 29161916]

52. Alafeef M. Smartphone-based photoplethysmographic imaging for heart rate monitoring. J Med Eng Technol 2017 Jul;41(5):387-395. [doi: 10.1080/03091902.2017.1299233] [Medline: 28300460]
53. Parpinel M, Scherling L, Lazzer S, Della Mea V. Reliability of heart rate mobile apps in young healthy adults: exploratory study and research directions. J Innov Health Inform 2017 Jun 30;24(2):921 [FREE Full text] [Medline: 28749320]
54. Koenig N, Seeck A, Eckstein J, Mainka A, Huebner T, Voss A, et al. Validation of a new heart rate measurement algorithm for fingertip recording of video signals with smartphones. Telemed J E Health 2016 Aug;22(8):631-636. [doi: 10.1089/tmj.2015.0212] [Medline: 26938673]
55. Vandenberk T, Stans J, Mortelmans C, Van Haelst R, Van Schelvergem G, Pelckmans C, et al. Clinical validation of heart rate apps: mixed-methods evaluation study. JMIR Mhealth Uhealth 2017 Aug 25;5(8):e129 [FREE Full text] [doi: 10.2196/mhealth.7254] [Medline: 28842392]
56. Pipitprapat W, Harnchoowong S, Suchonwanit P, Sriphrapradang C. The validation of smartphone applications for heart rate measurement. Ann Med 2018 Nov 17:1-7. [doi: 10.1080/07853890.2018.1531144] [Medline: 30269602]
57. Chaudhry BM. Heart rate monitoring mobile applications. Mhealth 2016 Apr 22;2:17 [FREE Full text] [doi: 10.21037/mhealth.2016.04.01] [Medline: 28293594]
58. Lee ES, Lee JS, Joo MC, Kim JH, Noh SE. Accuracy of heart rate measurement using smartphones during treadmill exercise in male patients with ischemic heart disease. Ann Rehabil Med 2017 Feb;41(1):129-137 [FREE Full text] [doi: 10.5535/arm.2017.41.1.129] [Medline: 28289645]
59. Yan BP, Chan CK, Li CK, To OT, Lai WH, Tse G, et al. Resting and postexercise heart rate detection from fingertip and facial photoplethysmography using a smartphone camera: a validation study. JMIR Mhealth Uhealth 2017 Mar 13;5(3):e33 [FREE Full text] [doi: 10.2196/mhealth.7275] [Medline: 28288955]
60. Mitchell K, Graff M, Hedt C, Simmons J. Reliability and validity of a smartphone pulse rate application for the assessment of resting and elevated pulse rate. Physiother Theory Pract 2016 Aug;32(6):494-499. [doi: 10.1080/09593985.2016.1203046] [Medline: 27459148]
61. Lee H, Chung H, Ko H, Jeong C, Noh SE, Kim C, et al. Dedicated cardiac rehabilitation wearable sensor and its clinical potential. PLoS One 2017 Oct 31;12(10):e0187108 [FREE Full text] [doi: 10.1371/journal.pone.0187108] [Medline: 29088260]
62. Golzar M, Fotouhi-Ghazvini F, Rabbani H, Zakeri FS. Mobile cardiac health-care monitoring and notification with real time tachycardia and bradycardia arrhythmia detection. J Med Signals Sens 2017;7(4):193-202 [FREE Full text] [Medline: 29204376]
63. Breton MD, Brown SA, Karvetski CH, Kollar L, Topchyan KA, Anderson SM, et al. Adding heart rate signal to a control-to-range artificial pancreas system improves the protection against hypoglycemia during exercise in type 1 diabetes. Diabetes Technol Ther 2014 Aug;16(8):506-511 [FREE Full text] [doi: 10.1089/dia.2013.0333] [Medline: 24702135]
64. Kevat AC, Bullen DV, Davis PG, Kamlin CO. A systematic review of novel technology for monitoring infant and newborn heart rate. Acta Paediatr 2017 May;106(5):710-720. [doi: 10.1111/apa.13786] [Medline: 28199732]
65. Ho CL, Fu YC, Lin MC, Chan SC, Hwang B, Jan SL. Smartphone applications (apps) for heart rate measurement in children: comparison with electrocardiography monitor. Pediatr Cardiol 2014 Apr;35(4):726-731. [doi: 10.1007/s00246-013-0844-8] [Medline: 24259012]
66. Gregoski MJ, Vertegel A, Shaporev A, Treiber FA. Tension Tamer: delivering meditation with objective heart rate acquisition for adherence monitoring using a smart phone platform. J Altern Complement Med 2013 Jan;19(1):17-19 [FREE Full text] [doi: 10.1089/acm.2011.0772] [Medline: 22967280]
67. Soffer M, Chen K. Identification and evaluation of fetal heart rate monitoring mobile applications [33I]. Obstet Gynecol 2017;129(5):101S. [doi: 10.1097/01.AOG.0000514972.96123.fb]
68. Chen YH, Chen HH, Chen TC, Chen LG. Robust heart rate measurement with phonocardiogram by on-line template extraction and matching. Conf Proc IEEE Eng Med Biol Soc 2011;2011:1957-1960. [doi: 10.1109/IEMBS.2011.6090552] [Medline: 22254716]
69. Guzik P, Piekos C, Pierog O, Fenech N, Krauze T, Piskorski J, et al. Classic electrocardiogram-based and mobile technology derived approaches to heart rate variability are not equivalent. Int J Cardiol 2018 May 1;258:154-156. [doi: 10.1016/j.ijcard.2018.01.056] [Medline: 29544922]
70. -. Heart rate variability: standards of measurement, physiological interpretation and clinical use. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. Circulation 1996 Mar 1;93(5):1043-1065 [FREE Full text] [Medline: 8598068]
71. Bolkhovsky JB, Scully CG, Chon KH. Statistical analysis of heart rate and heart rate variability monitoring through the use of smart phone cameras. Conf Proc IEEE Eng Med Biol Soc 2012;2012:1610-1613. [doi: 10.1109/EMBC.2012.6346253] [Medline: 23366214]
72. Plews DJ, Scott B, Altini M, Wood M, Kilding AE, Laursen PB. Comparison of heart-rate-variability recording with smartphone photoplethysmography, Polar H7 chest strap, and electrocardiography. Int J Sports Physiol Perform 2017 Nov 1;12(10):1324-1328. [doi: 10.1123/ijspp.2016-0668] [Medline: 28290720]

73. Perrotta AS, Jeklin AT, Hives BA, Meanwell LE, Warburton DER. Validity of the Elite HRV smartphone application for examining heart rate variability in a field-based setting. J Strength Cond Res 2017 Aug;31(8):2296-2302. [doi: 10.1519/JSC.0000000000001841] [Medline: 28195974]
74. Ramos-Castro J, Moreno J, Miranda-Vidal H, García-González MA, Fernández-Chimeno M, Rodas G, et al. Heart rate variability analysis using a seismocardiogram signal. Conf Proc IEEE Eng Med Biol Soc 2012;2012:5642-5645. [doi: 10.1109/EMBC.2012.6347274] [Medline: 23367209]
75. Huang RY, Dung LR. Measurement of heart rate variability using off-the-shelf smart phones. Biomed Eng Online 2016 Jan 29;15:11 [FREE Full text] [doi: 10.1186/s12938-016-0127-8] [Medline: 26822804]
76. Guede-Fernandez F, Ferrer-Mileo V, Ramos-Castro J, Fernandez-Chimeno M, Garcia-Gonzalez MA. Real time heart rate variability assessment from Android smartphone camera photoplethysmography: postural and device influences. Conf Proc IEEE Eng Med Biol Soc 2015;2015:7332-7335. [doi: 10.1109/EMBC.2015.7320085] [Medline: 26737985]
77. Lai E, Boyd K, Albert D, Ciocca M, Chung EH. Heart rate variability in concussed athletes: a case report using the smartphone electrocardiogram. HeartRhythm Case Rep 2017 Sep 6;3(11):523-526 [FREE Full text] [doi: 10.1016/j.hrcr.2017.08.009] [Medline: 29204350]
78. Gall B, Parkhouse W, Goodman D. Heart rate variability of recently concussed athletes at rest and exercise. Med Sci Sports Exerc 2004 Aug;36(8):1269-1274. [Medline: 15292731]
79. La Fountaine MF, Heffernan KS, Gossett JD, Bauman WA, De Meersman RE. Transient suppression of heart rate complexity in concussed athletes. Auton Neurosci 2009 Jun 15;148(1-2):101-103. [doi: 10.1016/j.autneu.2009.03.001] [Medline: 19303821]
80. Flatt AA, Esco MR. Evaluating individual training adaptation with smartphone-derived heart rate variability in a collegiate female soccer team. J Strength Cond Res 2016 Feb;30(2):378-385. [doi: 10.1519/JSC.0000000000001095] [Medline: 26200192]
81. Flatt AA, Esco MR. Smartphone-derived heart-rate variability and training load in a women's soccer team. Int J Sports Physiol Perform 2015 Nov;10(8):994-1000. [doi: 10.1123/ijspp.2014-0556] [Medline: 25756657]
82. Mellor A, Bakker-Dyos J, OÊ¼Hara J, Woods DR, Holdsworth DA, Boos CJ. Smartphone-enabled heart rate variability and acute mountain sickness. Clin J Sport Med 2018 Jan;28(1):76-81. [doi: 10.1097/JSM.0000000000000427] [Medline: 28407652]
83. Heathers JA. Smartphone-enabled pulse rate variability: an alternative methodology for the collection of heart rate variability in psychophysiological research. Int J Psychophysiol 2013 Sep;89(3):297-304. [doi: 10.1016/j.ijpsycho.2013.05.017] [Medline: 23751411]
84. Muaremi A, Arnrich B, Tröster G. Towards measuring stress with smartphones and wearable devices during workday and sleep. Bionanoscience 2013;3:172-183 [FREE Full text] [doi: 10.1007/s12668-013-0089-2] [Medline: 25530929]
85. Salahuddin L, Cho J, Jeong MG, Kim D. Ultra short term analysis of heart rate variability for monitoring mental stress in mobile settings. Conf Proc IEEE Eng Med Biol Soc 2007;2007:4656-4659. [doi: 10.1109/IEMBS.2007.4353378] [Medline: 18003044]
86. Dimarco AD, Onwordi E, Walters E, Mullan N, Willis L, Tanner M. Diagnostic utility of smartphone ECG technology in the initial investigation of palpitations. EP Europace 2017 Jun 1;19(3):iii102. [doi: 10.1093/ehjci/eux141.184]
87. Garabelli P, Stavrakis S, Albert M, Koomson E, Parwani P, Chohan J, et al. Comparison of QT interval readings in normal sinus rhythm between a smartphone heart monitor and a 12-lead ECG for healthy volunteers and inpatients receiving sotalol or dofetilide. J Cardiovasc Electrophysiol 2016 Jul;27(7):827-832. [doi: 10.1111/jce.12976] [Medline: 27027653]
88. Narasimha D, Curtis A. Smartphone based ECG recorder versus an event monitor for the evaluation of palpitations. 2016 Presented at: Heart Rhythm Society's 37th Annual Scientific Sessions; May 4-7, 2016; San Francisco.
89. Nguyen HH, Van Hare GF, Rudokas M, Bowman T, Silva JN. SPEAR Trial: Smartphone Pediatric ElectrocARdiogram Trial. PLoS One 2015 Aug 21;10(8):e0136256 [FREE Full text] [doi: 10.1371/journal.pone.0136256] [Medline: 26295569]
90. McManus DD, Chong JW, Soni A, Saczynski JS, Esa N, Napolitano C, et al. PULSE-SMART: pulse-based arrhythmia discrimination using a novel smartphone application. J Cardiovasc Electrophysiol 2016 Jan;27(1):51-57. [doi: 10.1111/jce.12842] [Medline: 26391728]
91. Cuesta P, Lado MJ, Vila XA, Alonso R. Detection of premature ventricular contractions using the RR-interval signal: a simple algorithm for mobile devices. Technol Health Care 2014;22(4):651-656. [doi: 10.3233/THC-140818] [Medline: 24898863]
92. McManus DD, Lee J, Maitas O, Esa N, Pidikiti R, Carlucci A, et al. A novel application for the detection of an irregular pulse using an iPhone 4S in patients with atrial fibrillation. Heart Rhythm 2013 Mar;10(3):315-319 [FREE Full text] [doi: 10.1016/j.hrthm.2012.12.001] [Medline: 23220686]
93. Yen TH, Chang CY, Yu SN. A portable real-time ECG recognition system based on smartphone. 2013 Presented at: 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC); July 3-7, 2013; Osaka, Japan. [doi: 10.1109/EMBC.2013.6611234]
94. Pan J, Tompkins WJ. A real-time QRS detection algorithm. IEEE Trans Biomed Eng 1985 Mar;32(3):230-236. [doi: 10.1109/TBME.1985.325532] [Medline: 3997178]

95. Gregoski MJ, Mueller M, Vertegel A, Shaporev A, Jackson BB, Frenzel RM, et al. Development and validation of a smartphone heart rate acquisition application for health promotion and wellness telehealth applications. Int J Telemed Appl 2012;2012:696324 [FREE Full text] [doi: 10.1155/2012/696324] [Medline: 22272197]
96. Hassan MA, Malik AS, Saad N, Fofi D, Meriaudeau F. Effect of motion artifact on digital camera based heart rate measurement. Conf Proc IEEE Eng Med Biol Soc 2017 Jul;2017:2851-2854. [doi: 10.1109/EMBC.2017.8037451] [Medline: 29060492]
97. Polimeni G, Scarpino A, Barbé K, Lamonaca F, Grimaldi D. Evaluation of the number of PPG harmonics to assess smartphone effectiveness. 2014 Presented at: 2014 IEEE International Symposium on Medical Measurements and Applications (MeMeA); June 11-12, 2014; Lisboa, Portugal. [doi: 10.1109/MeMeA.2014.6860101]
98. Voss A, Wessel N, Sander A, Malberg H, Dietz R. Influence of low sampling rate on heart rate variability analysis based on non-linear dynamics. 1995 Presented at: Computers in Cardiology 1995; September 10-13, 1995; Vienna, Austria. [doi: 10.1109/CIC.1995.482758]
99. Altini M, Amft O. HRV4Training: large-scale longitudinal training load analysis in unconstrained free-living settings using a smartphone application. Conf Proc IEEE Eng Med Biol Soc 2016 Aug;2016:2610-2613. [doi: 10.1109/EMBC.2016.7591265] [Medline: 28268857]
100. Elgendi M. Heart Rate Variability Measurement Using the Second Derivative Photoplethysmogram. In: Proceedings of the Third International Conference on Bio-inspired Systems and Signal Processing. Setúbal, Portugal: SciTePress; 2010:82-87.
101. Ferrer-Mileo V, Guede-Fernandez F, Fernandez-Chimeno M, Ramos-Castro J, Garcia-Gonzalez MA. Accuracy of heart rate variability estimation by photoplethysmography using an smartphone: processing optimization and fiducial point selection. Conf Proc IEEE Eng Med Biol Soc 2015 Aug;2015:5700-5703. [doi: 10.1109/EMBC.2015.7319686] [Medline: 26737586]
102. Garcia-Agundez A, Dutz T, Goebel S. Adapting smartphone-based photoplethysmograpy to suboptimal scenarios. Physiol Meas 2017 Feb;38(2):219-232. [doi: 10.1088/1361-6579/aa51db] [Medline: 28099163]
103. Peng F, Zhang Z, Gou X, Liu H, Wang W. Motion artifact removal from photoplethysmographic signals by combining temporally constrained independent component analysis and adaptive filter. Biomed Eng Online 2014 Apr 24;13:50 [FREE Full text] [doi: 10.1186/1475-925X-13-50] [Medline: 24761769]
104. Fukushima H, Kawanaka H, Bhuiyan MS, Oguri K. Estimating heart rate using wrist-type photoplethysmography and acceleration sensor while running. Conf Proc IEEE Eng Med Biol Soc 2012;2012:2901-2904. [doi: 10.1109/EMBC.2012.6346570] [Medline: 23366531]
105. Friedrich D, Aubert XL, Fuhr H, Brauers A. Heart rate estimation on a beat-to-beat basis via ballistocardiography - a hybrid approach. Conf Proc IEEE Eng Med Biol Soc 2010;2010:4048-4051. [doi: 10.1109/IEMBS.2010.5627626] [Medline: 21097094]
106. Di Rienzo M, Vaini E, Castiglioni P, Merati G, Meriggi P, Parati G, et al. Wearable seismocardiography: towards a beat-by-beat assessment of cardiac mechanics in ambulant subjects. Auton Neurosci 2013 Nov;178(1-2):50-59. [doi: 10.1016/j.autneu.2013.04.005] [Medline: 23664242]
107. Selvaraj N, Jaryal A, Santhosh J, Deepak KK, Anand S. Assessment of heart rate variability derived from finger-tip photoplethysmography as compared to electrocardiography. J Med Eng Technol 2008;32(6):479-484. [doi: 10.1080/03091900701781317] [Medline: 18663635]
108. Muhlestein JB, Le V, Albert D, Moreno FL, Anderson JL, Yanowitz F, et al. Smartphone ECG for evaluation of STEMI: results of the ST LEUIS Pilot Study. J Electrocardiol 2015;48(2):249-259. [doi: 10.1016/j.jelectrocard.2014.11.005] [Medline: 25601407]
109. Da He D, Winokur ES, Sodini CG. A continuous, wearable, and wireless heart monitor using head ballistocardiogram (BCG) and head electrocardiogram (ECG). Conf Proc IEEE Eng Med Biol Soc 2011;2011:4729-4732. [doi: 10.1109/IEMBS.2011.6091171] [Medline: 22255394]
110. Batchelor JC, Yeates SG, Casson AJ. Conformal electronics for longitudinal bio-sensing in at-home assistive and rehabilitative devices. Conf Proc IEEE Eng Med Biol Soc 2016 Aug;2016:3159-3162. [doi: 10.1109/EMBC.2016.7591399] [Medline: 28268978]
111. Kim DH, Lu N, Ma R, Kim YS, Kim RH, Wang S, et al. Epidermal electronics. Science 2011 Aug 12;333(6044):838-843 [FREE Full text] [doi: 10.1126/science.1206157] [Medline: 21836009]
112. Reyes BA, Posada-Quintero HF, Bales JR, Clement AL, Pins GD, Swiston A, et al. Novel electrodes for underwater ECG monitoring. IEEE Trans Biomed Eng 2014 Jun;61(6):1863-1876. [doi: 10.1109/TBME.2014.2309293] [Medline: 24845297]
113. Lobodzinski SS, Laks MM. New devices for very long-term ECG monitoring. Cardiol J 2012;19(2):210-214 [FREE Full text] [Medline: 22461060]
114. Engel JM, Chakravarthy BL, Rothwell D, Chavan A. SEEQ™ MCT wearable sensor performance correlated to skin irritation and temperature. Conf Proc IEEE Eng Med Biol Soc 2015;2015:2030-2033. [doi: 10.1109/EMBC.2015.7318785] [Medline: 26736685]
115. Jovanov E. Preliminary analysis of the use of smartwatches for longitudinal health monitoring. Conf Proc IEEE Eng Med Biol Soc 2015 Aug;2015:865-868. [doi: 10.1109/EMBC.2015.7318499] [Medline: 26736399]

116. Goodloe JM, Arthur A, Rhodes T, Holder P, Winham J, Thomas S. Emergency medical technician-basic acquisition and transmission of 12-lead ECG using a novel device (Readylink™). J Am Coll Cardiol 2013 Mar;61(10). [doi: 10.1016/S0735-1097(13)60113-8]
117. de Ruvo E, Sciarra L, Martino AM, Rebecchi M, Iulianella RV, Sebastiani F, et al. A prospective comparison of remote monitoring systems in implantable cardiac defibrillators: potential effects of frequency of transmissions. J Interv Card Electrophysiol 2016 Jan;45(1):81-90. [doi: 10.1007/s10840-015-0067-4] [Medline: 26467152]
118. Chan NY, Choy CC, Chan CK, Siu CW. Effectiveness of a nongovernmental organization-led large-scale community atrial fibrillation screening program using the smartphone electrocardiogram: an observational cohort study. Heart Rhythm 2018 Sep;15(9):1306-1311. [doi: 10.1016/j.hrthm.2018.06.006] [Medline: 29890321]
119. Charani E, Castro-Sánchez E, Moore LS, Holmes A. Do smartphone applications in healthcare require a governance and legal framework? It depends on the application!. BMC Med 2014 Feb 14;12:29 [FREE Full text] [doi: 10.1186/1741-7015-12-29] [Medline: 24524344]
120. Nicolas PT, Lindsay FW. Liability for mobile health and wearable technologies. Ann Health Law 2016;25:62 [FREE Full text]
121. Lang M. Beyond Fitbit: a critical appraisal of optical heart rate monitoring wearables and apps, their current limitations and legal implications. ALB L J Sci & Tech 2017:39 [FREE Full text]
122. Quinn P. The EU commission's risky choice for a non-risk based strategy on assessment of medical devices. CLSR 2017;33(3):361-370 [FREE Full text] [doi: 10.1016/j.clsr.2017.03.019]
123. Lang M. Heart rate monitoring apps: information for engineers and researchers about the new European Medical Devices Regulation 2017/745. JMIR Biomed Eng 2017 Aug 18;2(1):e2. [doi: 10.2196/biomedeng.8179]
124. White RD, Flaker G. Smartphone-based arrhythmia detection: should we encourage patients to use the ECG in their pocket? J Atr Fibrillation 2017 Apr 30;9(6):1605 [FREE Full text] [doi: 10.4022/jafib.1605] [Medline: 29250298]

Abbreviations

AF: atrial fibrillation
AMS: acute mountain sickness
BCG: ballistocardiogram
ECG: electrocardiograph
HR: heart rate
HRV: heart rate variability
MEMS: microelectromechanical sensors
PAC: premature atrial contractions
PCG: phonocardiogram
PP: P wave to P wave
PPG: photoplethysmography
PRV: pulse rate variability
PVC: premature ventricular contractions
QT: Q wave to T wave
RMSSD: Root Mean Square of Successive Difference
RR: R wave to R wave
SCG: seismocardiogram
SDE: SD of beat-to-beat error measurement
ShE: Shannon Entropy

Edited by G Eysenbach; submitted 18.07.18; peer-reviewed by L Borges, KC Wong, M Lang, A Mohammadabadi; comments to author 13.09.18; revised version received 06.11.18; accepted 25.11.18; published 15.02.19.

Please cite as:
Li KHC, White FA, Tipoe T, Liu T, Wong MCS, Jesuthasan A, Baranchuk A, Tse G, Yan BP
The Current State of Mobile Phone Apps for Monitoring Heart Rate, Heart Rate Variability, and Atrial Fibrillation: Narrative Review
JMIR Mhealth Uhealth 2019;7(2):e11606
URL: http://mhealth.jmir.org/2019/2/e11606/
doi:10.2196/11606
PMID:30767904

Review

Toward Developing a Standardized Core Set of Outcome Measures in Mobile Health Interventions for Tuberculosis Management: Systematic Review

Seohyun Lee[1], PhD; Youngji Lee[2], MA; Sangmi Lee[2], MPH; Sheikh Mohammed Shariful Islam[3], PhD; Sun-Young Kim[2], PhD

[1]Institute of Health & Environment, Center for Global Health Research, Graduate School of Public Health, Seoul National University, Seoul, Republic of Korea

[2]Department of Public Health Sciences, Graduate School of Public Health, Seoul National University, Seoul, Republic of Korea

[3]Institute for Physical Activity and Nutrition, School of Exercise and Nutrition Sciences, Faculty of Health, Deakin University, Melbourne, Australia

Corresponding Author:
Sun-Young Kim, PhD
Department of Public Health Sciences
Graduate School of Public Health
Seoul National University
1 Gwanak-ro, Gwanak-gu
Seoul, 08826
Republic of Korea
Phone: 82 880 2768
Fax: 82 2 762 2888
Email: sykim22@snu.ac.kr

Abstract

Background: Tuberculosis (TB) management can be challenging in low- and middle-income countries (LMICs) not only because of its high burden but also the prolonged treatment period involving multiple drugs. With rapid development in mobile technology, mobile health (mHealth) interventions or using a mobile device for TB management has gained popularity. Despite the potential usefulness of mHealth interventions for TB, few studies have quantitatively synthesized evidence on its effectiveness, presumably because of variability in outcome measures reported in the literature.

Objective: The aim of this systematic review was to evaluate the outcome measures reported in TB mHealth literature in LMICs.

Methods: MEDLINE, EMBASE, and the Cochrane Database of Systematic Reviews were searched to identify mHealth intervention studies for TB (published up to May 2018) that reported any type of outcome measures. The extracted information included the study setting, types of mHealth technology used, target population, study design, and categories of outcome measures. Outcomes were classified into 13 categories including treatment outcome, adherence, process measure, perception, technical outcome, and so on. The qualitative synthesis of evidence focused on the categories of outcome measures reported by the type of mHealth interventions.

Results: A total of 27 studies were included for the qualitative synthesis of evidence. The study designs varied widely, ranging from randomized controlled trials to economic evaluations. A total of 12 studies adopted short message service (SMS), whereas 5 studies used SMS in combination with additional technologies or mobile apps. The study populations were also diverse, including patients with TB, patients with TB/HIV, health care workers, and general patients attending a clinic. There was a wide range of variations in the definition of outcome measures across the studies. Among the diverse categories of outcome measures, treatment outcomes have been reported in 14 studies, but only 6 of them measured the outcome according to the standard TB treatment definitions by the World Health Organization.

Conclusions: This critical evaluation of outcomes reported in mHealth studies for TB management suggests that substantial variability exists in reporting outcome measures. To overcome the challenges in evidence synthesis for mHealth interventions, this study can provide insights into the development of a core set of outcome measures by intervention type and study design.

(JMIR Mhealth Uhealth 2019;7(2):e12385) doi:10.2196/12385

KEYWORDS

mHealth; tuberculosis; outcome measures; evidence synthesis; low-and middle-income countries

Introduction

Tuberculosis (TB) is one of the deadly infectious diseases that have claimed millions of lives worldwide. According to the World Health Organization (WHO), globally, there were 10.4 million new TB cases causing approximately 1.2 million deaths in 2016 [1,2]. The mortality rate of TB is disproportionately higher in low- and middle-income countries (LMICs). Over 95% of TB deaths occurred in these countries, and 7 LMICs (India, Indonesia, China, Philippines, Pakistan, Nigeria, and South Africa) accounted for 64% of the total burden [3]. In fact, previous studies have shown that there is empirical evidence of positive associations between poverty indicators and TB incidence both at the macro and individual levels [4]. Considering the vicious cycle of poverty and TB, alleviating the burden of TB is more challenging for the LMICs because it requires adequate resources for "prolonged treatment with multiple drugs [5]." The 6-month course of first-line therapy can be burdensome with the possibility of adverse reactions and the treatment of multidrug-resistant (MDR)-TB requires more toxic and expensive drugs [6]. For example, the cost of bedaquiline, a second-line medication to treat MDR-TB, was US $3000 per treatment in middle-income countries and US $900 in low-income countries [7]. In fact, premature discontinuation of the treatment, which can lead to MDR-TB, is common among TB patients not only for its toxicity but also for socioeconomic costs associated with it [8]. Therefore, the management of TB is notoriously difficult especially in LMICs.

In this context, using mobile devices for TB treatment has been recognized as an innovative approach for LMICs where mobile subscription rates have dramatically increased over the past decade. Mobile health (mHealth) interventions involving mobile devices in the management of TB have the potential for reducing costs of information delivery and improving the quality of communication [9]. mHealth can be useful for TB treatment adherence support such as short message service (SMS) for medication reminders or mobile apps for remote directly observed treatment (DOT) strategy [10,11].

Despite the potential of mHealth interventions for improving TB management, the empirical evidence on its effectiveness is mixed. Some studies have demonstrated the effectiveness and feasibility of the mHealth interventions for TB [12,13], whereas others have shown no significant impact [14,15]. Moreover, no study has attempted to synthesize the results quantitatively to rigorously evaluate the effectiveness of mHealth in TB management. Presumably, one of the reasons for such difficulty in synthesizing and evaluating the findings comes from wide variations in the outcomes reported from the mHealth studies for TB.

To respond to this knowledge gap, this study aimed to systematically review previous mHealth studies for TB management and critically evaluate and categorize the outcome measures for different mobile technologies and study designs. The goal of this study was to provide researchers insights into the development of a core set of outcome measures for mHealth interventions intended to improve TB treatment adherence. In doing so, the study can facilitate the evidence synthesis of mHealth interventions for TB.

Methods

Search Strategy and Review Process

Electronic databases (MEDLINE, EMBASE, and Cochrane Database of Systematic Reviews) were searched to identify peer-reviewed studies of mHealth interventions for TB. The systematic search was supplemented by reviewing relevant review papers identified from the initial search. The search strategy for the study population includes key terms describing LMICs such as "resource poor" or "developing country." The search strategy for the mHealth intervention combined multiple keywords such as "mHealth" and "text-messaging." As for the target disease, "tuberculosis," "TB," "multi-drug resistant tuberculosis," and "MDR-TB" were used as search terms. No restrictions were applied to the publication type or publication date, but the language filter was applied to identify studies published in English. The search included articles published up to May 2018. The full search strategy is available in Multimedia Appendix 1.

Furthermore, 3 authors (ShL, YL, and SmL) independently reviewed the retrieved studies throughout the selection process. Each study identified from the databases was screened by 2 reviewers and then a full-text review was conducted for the potentially eligible studies. The disagreement on the selection process was resolved by the other authors who were not involved in the review of the specific study under discussion.

Eligibility for Review

The inclusion criteria for this systematic review were as follows: First, studies conducted in the context of LMICs, as defined by the World Bank's income cutoffs [16]; second, studies involving an intervention using mobile devices (ie, mHealth intervention); third, the target disease of the study should be TB or MDR-TB; fourth, studies designed to evaluate the effectiveness or benefits of mHealth interventions for TB (eg, observational study, mixed-methods study or implementation project, randomized controlled trial [RCT]); fifth, studies reporting more than one type of outcome; finally, only full-text studies published in English were considered eligible. In addition, the authors attempted to identify individual studies from reviews or systematic reviews, which were included in this study.

Data Extraction and Analysis

The qualitative synthesis of evidence focused on the outcome measures reported in each type of mHealth intervention. Information about the study setting, mHealth technologies used, target populations, and types of outcome measures was extracted. To classify the diverse types of detailed outcome measures, the following categories were used: (1) treatment outcome; (2) treatment outcome as defined by WHO; (3) adherence; (4) process measure; (5) perception; (6) technical

outcome; (7) health outcome; (8) quality of life; (9) knowledge; (10) cost-effectiveness; (11) cost; (12) psychosocial outcome; and (13) mortality. Some explanations about these measures are provided in the next few paragraphs.

The treatment outcome includes any outcome measure that deals with the result of TB treatment, such as sputum smear conversion or microscopy test result. The treatment outcome following the WHO definition was separately categorized [17]. The WHO definition was developed to make a distinction for treatment outcomes between the drug-susceptible TB and drug-resistant TB, which are mutually exclusive groups. According to the WHO definition, any patient with TB should belong to either group and then 1 of the 7 treatment outcome cohorts: (1) cured; (2) treatment completed; (3) treatment failed; (4) died; (5) lost to follow-up; (6) not evaluated; and (7) treatment success. The WHO definitions for each of these 7 categories differ between the drug-susceptible TB and drug-resistant TB as described in Multimedia Appendix 2.

The adherence outcome includes medication adherence or treatment adherence. The process measure is any outcome measure related to treatment process, including the receipt of diagnostic test, attendance to appointments, or reporting of adverse events. The perception indicates any outcome measure that captures the user's thoughts on mHealth for TB management. The technical outcome relates to the outcome measures that investigate the technical feasibility such as processing times or system installation. Health outcome, quality of life, knowledge (eg, patients' understanding of the disease or the technology), cost-effectiveness (ie, the extent to which an alternative provides value for money), cost (ie, costs associated with an intervention from different perspectives), psychosocial outcome, and mortality outcomes are additional categories which are self-explanatory.

Risk of Bias: Quality Assessment

To evaluate the quality of individual studies included for our review, risk of bias was assessed with the existing tools. As this systematic review includes various types of studies, it is important to have a coherent set of quality assessment tools for different study designs. Therefore, we used the modified version of the Critical Appraisal Skills Program (CASP) that provides the checklists specific to various types of studies ranging from RCTs and qualitative studies to economic evaluation [18]. In case of mixed-methods studies whose CASP tool has not been developed yet, the quality assessment criteria for mixed-methods studies from the previous study were employed [19]. The quality assessment of studies is presented in Multimedia Appendix 3.

Results

Overview of Included Studies

Among the 312 studies identified after removing duplicates, 260 articles were excluded during the screening process based on the titles and abstracts. Therefore, 52 articles were assessed for eligibility through a full-text review. Of those, 27 studies were included for the qualitative synthesis of evidence. A flow diagram for the selection process, based on the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines, is provided in Figure 1 [20].

Qualitative Synthesis of Evidence

Table 1 presents the results of qualitative synthesis of the mHealth studies for TB management. Approximately, half (15 out of 27) of the studies were conducted in African countries. The study designs were diverse, including 6 RCTs, 5 mixed-methods studies, 1 cohort study, 3 qualitative studies, 4 observational studies, 6 implementation projects, and 2 economic evaluation studies. The types of mHealth technologies utilized were diverse as well: 12 studies employed SMS, 6 studies used mobile app, 5 studies used SMS plus other technology, 3 studies utilized phone calls, and only 1 study applied mHealth for mobile data collection. With regard to the study population, a majority (20 out of 27) of the studies were targeted for TB patients or TB/HIV patients, but there were a few studies that examined the experience of health care workers or general patients at the clinic for TB test results notification.

In terms of the outcome measures, there was a wide range of variants in their definitions even within each category of the outcomes. For instance, both Mohammed et al and Bediang et al defined treatment success as the primary outcome measure of SMS intervention for TB medication adherence in their RCT, but they defined treatment success differently [15,22]. Mohammed et al defined it as "the sum of patients clinically reported as cured (ie, a patient whose sputum smear or culture was positive at the beginning of treatment but who was smear- or culture-negative in the last month of treatment and on at least one previous occasion) or treatment completed (ie, a patient who completed treatment but who does not have a negative sputum smear or culture result in the last month of treatment and on at least one previous occasion)" [22]. On the contrary, Bediang et al defined treatment success as "having completed 6 months treatment and having negative sputum smears at 5 months" [15].

Summary of Outcome Measures

Table 2 summarizes the types of outcome measures reported in the included studies, using the categories defined in this review. The most frequently reported outcome type was treatment success. Approximately half of the studies reviewed (14 out of 27) included treatment outcome but only 6 studies among them followed the WHO definition for the treatment outcome (Multimedia Appendix 2). The second most frequently reported category was perception on the mHealth intervention (13 out of 27). Other categories reported were diverse and included technical outcome, medication or treatment adherence, process measure, etc. However, there was substantial variability within each category of outcome, as shown in Table 1. For example, acceptability and satisfaction within the perception category were defined differently from one study to another [15,28-30,32,35,36,40,44]. Also, SMS-only intervention studies did not focus on technical outcome [12,15,22,29,30,32,35,36,42,43,45,46] whereas studies involving other mHealth technologies such as an app or mobile data collection did so [21,28,33,37,39-41]. On the contrary, outcomes related to cost or cost-effectiveness were reported only via studies involving SMS [29,45,46].

Figure 1. Flow diagram for selection process following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guideline. mHealth: mobile health; TB: tuberculosis.

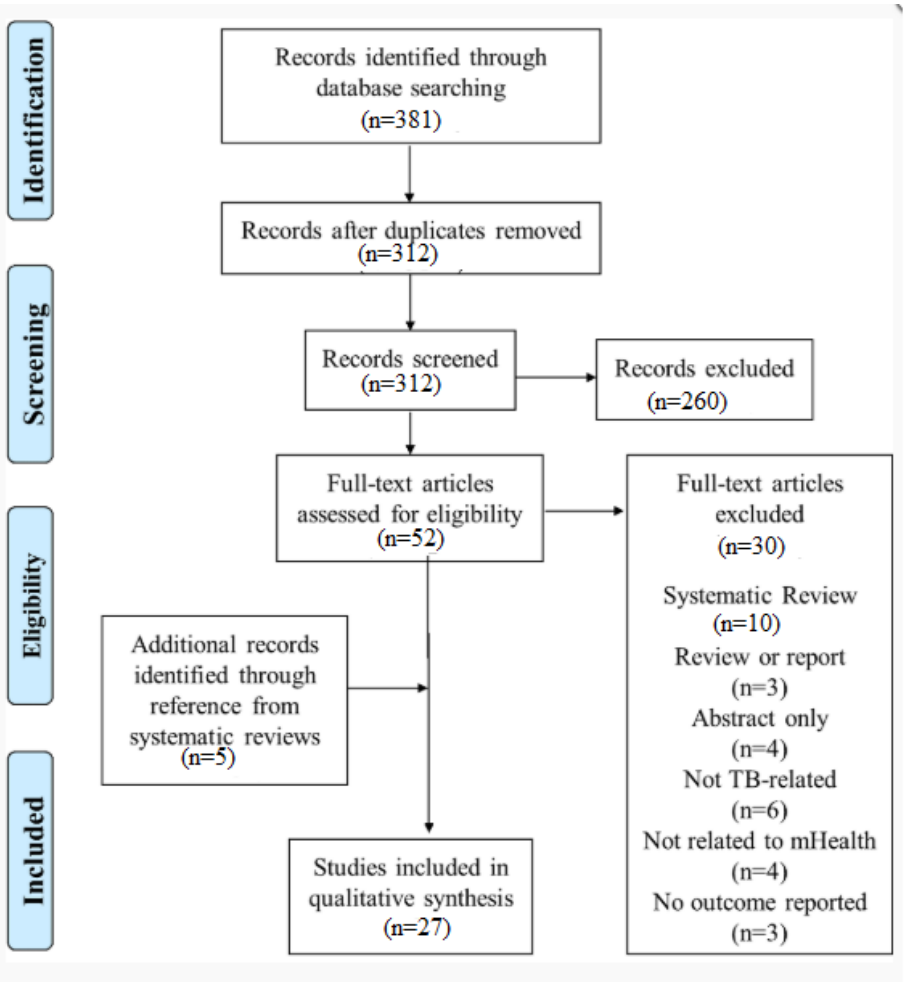


Table 1. Summary of included studies.

Source	Country	Study design	Mobile health technology	Population	Outcome measures	Outcome category	Purpose
Blaya, 2009 [21]	Peru	RCT[a]	Mobile data collection	Health centers	Processing times, frequency of errors, the number of work-hours expended by data collectors	Technical outcome	For laboratory data collection
Mohammed, 2016 [22]	Pakistan	RCT	SMS[b]	TB[c] patients	Primary: clinically recorded treatment success based upon intention-to-treat; Secondary: treatment outcomes (WHO definitions[d]), self-reported medication adherence, self-reported psychological and physical health measures	Treatment outcome, adherence, health outcome	For medication adherence
Bassett, 2013 [23]; Bassett, 2016 [24]	South Africa	RCT protocol (2013); RCT (2016)	SMS and phone calls	Patients at clinic	Primary: treatment completion; Secondary: mortality, receipt of CD4 count and TB test results, and repeat CD4 counts for those not antiretroviral therapy (ART)–eligible at baseline	Treatment outcome, mortality, process measure	For appointment and test result reminder and psychosocial support
Huang, 2017 [25]	China	Cluster RCT protocol	Mobile app	TB patients	Primary: TB treatment result (WHO definitions[d]); Secondary: treatment adherence (the percentage of patients receiving TB treatment who missed fewer than 5% of doses), self-reported adherence, knowledge about TB, quality of life (QoL)	Treatment outcome, adherence, knowledge, QoL	For Bracelet- and self-directed observational therapy
Bediang, 2014 [26]; Bediang, 2018 [15]	Cameroon	RCT protocol (2014); RCT (2018)	SMS	TB patients	Primary: cure rate (absence of Koch's bacilli in the sputum), treatment success (having completed 6 months' treatment and having negative sputum smears at 5 months); Secondary: treatment adherence (drug prescriptions collected and doses taken), attendance to appointments, punctuality of appointments, treatment outcome (WHO definitions[d]), the number of participants who develop resistance, satisfaction	Treatment outcome, adherence, process measure, perception	For medication adherence
Khachadourian, 2015 [27]	Armenia	RCT protocol	SMS and phone calls	TB patients	Primary: physician-reported treatment outcome (WHO definitions[d]); Secondary: patients' knowledge, depression, QoL, within-family TB-related stigma, family social support, self-reported treatment adherence	Treatment outcome, knowledge, psychosocial outcome, QoL, adherence	For medication adherence
Chaiyachati, 2013 [28]	South Africa	Mixed-methods study	Mobile app	Health care workers	Primary: proportion of weekly adverse events forms submitted vs expected by mobile health care workers; Secondary: acceptability (perceived comfort levels with using mobile phone technology), quality of adverse events monitoring, proportion of reportable adverse events being captured; Technical outcomes: phone usage patterns, technical problems experienced	Process measure, perception, technical outcome	For adverse events reporting
Howard, 2016 [29]	Lesotho	Mixed-methods, cluster-randomized trial protocol	SMS	TB/HIV patients, health care workers	Primary: ART initiation, retention, and TB treatment success; Secondary: time to ART initiation, adherence, change in cluster of differentiation 4 (CD4) count, sputum smear conversion, cost-effectiveness, acceptability	Process measure, treatment outcome, adherence, cost-effectiveness, perception	For treatment adherence

Source	Country	Study design	Mobile health technology	Population	Outcome measures	Outcome category	Purpose
Iribarren, 2013 [30]	Argentina	Mixed-methods study (including RCT)	SMS	TB patients	Primary: feasibility (access to mobile phones, familiarity with texting, rate of participant refusal, suboptimal TB understanding), and acceptability (feeling cared for patient's treatment, self-reporting adherence); Secondary: initial efficacy (microscopy test result from positive to negative, treatment outcome)	Process measure, perception, adherence, treatment outcome	For treatment adherence
Hirsch-Moverman, 2017a [31]	Lesotho, Ethiopia	Mixed-methods; implementation science study	Phone calls	TB/HIV patients, TB patients	The number of call attempts per participant for each month, completeness of monthly calls, success rates, challenges	Process measure	For medication adherence
Hirsch-Moverman, 2017b [32]	Lesotho	Mixed-methods implementation science study, cluster-randomized trial protocol	SMS	TB patients, health care workers, caregivers	Primary: the number of child contacts Isoniazid preventive therapy (IPT) initiation, IPT completion; Secondary: HIV testing, yield of active prevalent TB among child contacts, acceptability, and utilization of community-based intervention components	Treatment outcome, perception, process measure	For medication adherence and appointment reminders
Nguyen, 2017 [33]	Vietnam	Cohort study	SMS and mobile app	TB patients	Primary: proportion of patients completing all doses of self-administered treatment; Secondary: proportion of videos uploaded as scheduled, proportion of patients discontinuing using Video DOT (VDOT)	Treatment outcome, technical outcome, process measure	For medication adherence
Daftary, 2017 [34]	Ethiopia	Qualitative study	Interactive voice response (IVR)	HIV patients	Perceptions and attitude, perceived benefits and challenges	Perception	For preventive therapy adherence
Albino, 2014 [35]	Peru	Qualitative study	SMS	TB patients	Perceptions and acceptability	Perception	For treatment adherence
Nhavoto, 2017 [36]	Mozambique	Qualitative study	SMS	TB patients, Health care workers	Usefulness, perceived benefits, ease of use, satisfaction, risks of the SMS system	Perception	For treatment adherence
Hoffman, 2010 [37]	Kenya	Observational study	Mobile app	TB patients	Primary: technical feasibility (patient and health provider receptivity to remote directly observed treatment [DOT]); Secondary: patient preferences and receptivity to receiving TB health message on a mobile phone	Technical outcome, perception	For Mobile Direct Observation of Treatment
de Sumaride Boer, 2016 [12]	Tanzania	Observational pilot study	SMS	HIV patients, TB patients	Quantitative: percentage of doses taken on time, percentage of sent reminders (divided by total intake prescription), percentage of correct reminders (after missed doses), percentage of incorrect reminders (after opening the pillbox but the signal was not sent), percentage of extra openings, percentage of missed doses, percentage of adherence with the exclusion of doses that were taken after a reminder; Qualitative: general experience with using the device	Process measure, adherence, perception	For medication adherence
Garfein, 2015 [38]	Mexico, USA	Observational pilot study	Mobile app	TB patients	Primary: adherence rate (the number of medication doses observed in videos divided by the number of doses expected during the treatment period); Secondary: perceptions of VDOT	Adherence, perception	For VDOT

Source	Country	Study design	Mobile health technology	Population	Outcome measures	Outcome category	Purpose
Dwolatzky, 2006 [39]	South Africa	Observational pilot study	Mobile app	Patients at clinic	Time taken to locate the households	Technical outcome	For locating patients' homes by global positioning system and personal digital assistant
Ha, 2016 [40]	Botswana	Implementation project	Mobile app	TB patients	Cases screened for contact tracing, time required to complete TB contact tracing per contact, quality of data collected, user satisfaction with usability, operational considerations	Technical outcome, perception	For contact tracing
Cowan, 2016 [41]	Mozambique	Implementation project	SMS and mobile app	Health facility	System installation on computers, development of Web-based interface and automated SMS and email messages, test results uploaded to the system, SMS notifications sent to key personnel, the number of users	Technical outcome	For remote monitoring solution
Kunawararak, 2011 [14]	Thailand	Implementation project	Phone calls	TB patients	Cure rates, completion rates, failure rates and success rates, conversion rates	Treatment outcome	For medication adherence
Lorent, 2014 [42]	Cambodia	Implementation project	SMS	General population	TB case detection- smear-positivity, clinical TB treatment uptake–time to treatment initiation outcome–treatment outcomes (WHO definitions[d]), delay in linkage to care	Process measure, treatment outcome	For test result notification
Mahmud, 2010 [43]	Malawi	Implementation project	SMS	Health care workers	Operational net savings, worker time gained, patient enrollment	Process measure	For continuity of care
Narasimhan, 2014 [44]	India	Implementation project	SMS and phone calls	TB patients	Treatment completion and cure rates (WHO definitions[d]), treatment adherence rates, adverse drug reaction rates, stigma associated with TB, patient satisfaction, usage of the mHealth initiative	Treatment outcome, adherence, perception	For medication adherence
Broomhead, 2012 [45]	South Africa	Cost minimization analysis	SMS	TB patients	Smear conversion rate, TB cure rate, reduced average cost per patient	Treatment outcome, cost	For treatment adherence
Hunchangsith, 2012 [46]	Thailand	Cost-effectiveness analysis	SMS	TB patients	Disability-adjusted life years (DALYs) averted, costs (health care perspective), effects of interventions, success rate, failure rate, transfer out rate, death rate	Health outcome, treatment outcome, cost, cost-effectiveness	For medication adherence

[a]RCT: randomized controlled trial.

[b]SMS: short message service.

[c]TB: tuberculosis.

[d]World Health Organization definitions: presented in Multimedia Appendix 2.

Table 2. Reported outcomes by mHealth intervention type.

Intervention type (number of studies) and reference	Categories of outcome measure												
	Treatment outcome	Treatment outcome by WHO[a] definition	Adherence	Process measure	Perception	Technical outcome	Health outcome	QoL[b]	Knowledge	Cost-effectiveness	Cost	Psychosocial outcome	Mortality
Short message service (SMS; 12 studies)													
[22]	✓[c]	✓	✓	—[d]	—	—	✓	—	—	—	—	—	—
[15]	✓	✓	✓	✓	✓	—	—	—	—	—	—	—	—
[29]	✓	—	✓	✓	✓	—	—	—	—	✓	—	—	—
[30]	✓	—	✓	✓	✓	—	—	—	—	—	—	—	—
[32]	✓	—	—	✓	✓	—	—	—	—	—	—	—	—
[35]	—	—	—	—	✓	—	—	—	—	—	—	—	—
[36]	—	—	—	—	✓	—	—	—	—	—	—	—	—
[12]	—	—	✓	✓	✓	—	—	—	—	—	—	—	—
[42]	✓	✓	—	✓	—	—	—	—	—	—	—	—	—
[43]	—	—	—	✓	—	—	—	—	—	—	—	—	—
[45]	✓	—	—	—	—	—	—	—	—	—	✓	—	—
[46]	✓	—	—	—	—	—	✓	—	—	✓	✓	—	—
SMS plus others (5 studies)													
[24]	✓	—	—	✓	—	—	—	—	—	—	—	—	✓
[27]	✓	✓	✓	—	—	—	—	✓	✓	—	—	✓	—
[33]	✓	—	—	✓	—	✓	—	—	—	—	—	—	—
[41]	—	—	—	—	—	✓	—	—	—	—	—	—	—
[44]	✓	✓	✓	—	✓	—	—	—	—	—	—	—	—
Mobile app (6 studies)													
[25]	✓	✓	✓	—	—	—	—	✓	✓	—	—	—	—
[28]	—	—	—	✓	✓	✓	—	—	—	—	—	—	—
[37]	—	—	—	—	✓	✓	—	—	—	—	—	—	—
[38]	—	—	✓	—	✓	—	—	—	—	—	—	—	—
[39]	—	—	—	—	—	✓	—	—	—	—	—	—	—
[40]	—	—	—	—	✓	✓	—	—	—	—	—	—	—
Phone calls or interactive voice response (3 studies)													
[31]	—	—	—	✓	—	—	—	—	—	—	—	—	—
[34]	—	—	—	—	✓	—	—	—	—	—	—	—	—
[14]	✓	—	—	—	—	—	—	—	—	—	—	—	—
Mobile data collection (1 study)													
[21]	—	—	—	—	—	✓	—	—	—	—	—	—	—

[a]WHO: World Health Organization.

[b]QoL: quality of life.

[c]Tick marks indicate that the specific category of outcome measure was reported.

[d]Outcome measure was not reported.

Discussion

Principal Findings

This systematic review critically evaluated the outcomes reported in mHealth studies for TB management in LMICs. The reason why rigorous evidence synthesis is warranted is that recent literature for TB reports mixed results despite the rapid implementation of mHealth technology for TB management. The fragmented pieces of evidence on the effectiveness partly resulted from the wide variations in the definitions of outcome measures in TB mHealth interventions. Even though treatment outcome has been reported by many studies, they often did not adopt the standard definition recommended by the WHO [17].

The WHO definition of TB treatment outcome is part of an effort to standardize outcome measures for TB at the global level. To promote the use of standardized sets of outcome measures for TB, WHO provided the standard definitions and classifications of TB in terms of diagnosis or treatment outcomes [17]. The intention for this standardization effort was to coordinate international comparison of TB treatment outcomes through health information systems. However, the findings of our review revealed that mHealth studies for TB have not comprehensively adopted this standardized approach for TB treatment. Only 6 out of 27 interventions chose to report the treatment outcome according to the WHO definition. Interventions involving phone calls, interactive voice response, or mobile data collection did not consider the WHO definition.

Our findings also suggest that, to rigorously evaluate the effectiveness of mHealth interventions for TB, future studies should be carefully designed with regard to the selection of outcome measures. Indeed, using standard definitions for outcome measures within some commonly reported categories can improve comparability across different studies. As assessed in this study, examples of such categories include treatment outcome (preferably using the WHO definition), perception, process measure, adherence, and technical outcome.

The value of this systematic review can be found in its potential to motivate and facilitate consensus on standard definitions of outcome measures used in mHealth interventions for TB so that such effort can guide more effective mHealth intervention designs for improving TB management. Although current literature shows considerable variability in the definition of outcome measures, discussion and coordination among researchers can promote standardized methods in measuring outcomes. Specifically, the outcomes should be comparable, promote transparent communication, and maintain consistency in terminology. Coordination at the global level is necessary to develop a core set of outcome measures for TB mHealth interventions by study design and technology type utilized.

When choosing a core set of outcome measures with standard definitions, there are additional issues to consider. First, the time point for reporting outcome measures should be clinically meaningful and feasible [47]. Second, a detailed description of the measure should be provided, such as calculation method or definitions. Third, a clear explanation on the target population for each outcome measure can be useful. For example, some outcome measures may be more appropriate for MDR-TB patients rather than TB patients on their first-line therapy course. Finally, long-term outcome measures should be considered to establish fundamental evidence for TB mHealth interventions. The long-term outcomes can be related to physical, psychosocial, or mental health. Despite its importance, our review showed that only 2 out of 27 studies reported long-term health outcomes; Mohammed et al, reported self-reported psychological and physical health measures [22] and Hunchangsith et al, reported DALYs averted [46].

Another issue to consider is related to evidence for cost. As this study suggested, insufficient evidence exists in terms of cost-effectiveness or cost of mHealth interventions for TB management. Those previous studies that have evaluated the cost-effectiveness or cost of mHealth interventions only considered SMS as mHealth channels and did not consider or evaluate other mHealth channels and technologies. However, other mHealth channels and technologies such as mobile apps or global positioning system are now available for TB patients [48]. Therefore, future studies need to assess the cost-effectiveness of such channels and technologies for improving TB management in LMICs.

This study has some limitations. First, the database used for identifying relevant studies is limited to the 3 most frequently cited sources, namely MEDLINE, EMBASE, and Cochrane Database of Systematic Reviews. Grey literature or other sources of information can supplement our findings. To complement this limitation, we attempted to identify additional related studies from relevant systematic reviews searched from our study. Second, the effectiveness of the mHealth interventions for TB was not quantitatively evaluated because of the heterogeneity of the outcomes reported.

Despite these limitations, this study provides an overview of the currently reported outcome measures for mHealth interventions intended to improve TB management in the context of LMICs. The results from this review can be used as a starting point for discussion to adopt standardized definitions within different categories of outcome measures for future mHealth interventions for TB management in LMICs.

Conclusions

This systematic review of mHealth studies for TB suggests that substantial variability exists with regard to the definitions of outcome measures across studies. Our review highlights that a standardized method for measuring the different outcomes is warranted to improve comparability of outcome measures across studies for a more rigorous and reliable evaluation of the effectiveness of mHealth interventions for TB. In doing so, the coordination among researchers and the development of a core set of outcome measures based on standardized methods would be necessary. Our study provides useful information for researchers to better assess the effectiveness of mHealth interventions for TB. In addition, the study provides insights into the possibility of developing a core set of outcome measures by intervention type and study design based on a standardized or coordinated set of methods.

Acknowledgments

ShL was involved in the conceptualization of the study and drafted the manuscript. ShL, YL, and SmL developed the search strategy, collected data, and reviewed the identified studies. SMSI critically reviewed the manuscript. SYK critically reviewed the manuscript and interpreted the findings. All authors reviewed and approved the final version of the manuscript. This work was supported by the National Research Foundation of Korea Grant funded by the Korean Government (#21B20151213037).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Search strategy.

[PDF File (Adobe PDF File), 129KB - mhealth_v7i2e12385_app1.pdf]

Multimedia Appendix 2

Definitions of TB treatment outcome by the WHO [17].

[PDF File (Adobe PDF File), 34KB - mhealth_v7i2e12385_app2.pdf]

Multimedia Appendix 3

Risk of bias assessment.

[PDF File (Adobe PDF File), 794KB - mhealth_v7i2e12385_app3.pdf]

References

1. World Health Organization. Global Health Observatory (GHO) Data URL: http://www.who.int/gho/tb/en/ [accessed 2018-11-20] [WebCite Cache ID 743tX88Tq]
2. Institute for Health Metrics and Evaluation. GBD Results Tool URL: http://ghdx.healthdata.org/gbd-results-tool [accessed 2018-11-20] [WebCite Cache ID 743u1Lk8z]
3. World Health Organization. Tuberculosis URL: http://www.who.int/news-room/fact-sheets/detail/tuberculosis [accessed 2018-11-20] [WebCite Cache ID 743u9swGr]
4. Oxlade O, Murray M. Tuberculosis and poverty: why are the poor at greater risk in India? PLoS One 2012 Nov;7(11):e47533 [FREE Full text] [doi: 10.1371/journal.pone.0047533] [Medline: 23185241]
5. Roberts D, Rylands J, Sinclair D. Interventions using mobile devices (phones, smart phones, or tablets) to improve adherence to treatment for HIV or tuberculosis. Cochrane Database Syst Rev 2016 Sep 15;9. [doi: 10.1002/14651858.CD012353]
6. Nagarajan S, Whitaker P. Management of adverse reactions to first-line tuberculosis antibiotics. Curr Opin Allergy Clin Immunol 2018 Aug;18(4):333-341. [doi: 10.1097/ACI.0000000000000462] [Medline: 29901473]
7. Wirtz VJ, Hogerzeil HV, Gray AL, Bigdeli M, de Joncheere CP, Ewen MA, et al. Essential medicines for universal health coverage. Lancet 2017 Jan 28;389(10067):403-476. [doi: 10.1016/S0140-6736(16)31599-9] [Medline: 27832874]
8. Babiarz KS, Suen SC, Goldhaber-Fiebert JD. Tuberculosis treatment discontinuation and symptom persistence: an observational study of Bihar, India's public care system covering >100,000,000 inhabitants. BMC Public Health 2014 May 01;14:418 [FREE Full text] [doi: 10.1186/1471-2458-14-418] [Medline: 24886314]
9. Aranda-Jan CB, Mohutsiwa-Dibe N, Loukanova S. Systematic review on what works, what does not work and why of implementation of mobile health (mHealth) projects in Africa. BMC Public Health 2014 Feb 21;14:188 [FREE Full text] [doi: 10.1186/1471-2458-14-188] [Medline: 24555733]
10. Story A, Garfein RS, Hayward A, Rusovich V, Dadu A, Soltan V, et al. Monitoring therapy compliance of tuberculosis patients by using video-enabled electronic devices. Emerg Infect Dis 2016 Mar;22(3):538-540 [FREE Full text] [doi: 10.3201/eid2203.151620] [Medline: 26891363]
11. Ngwatu BK, Nsengiyumva NP, Oxlade O, Mappin-Kasirer B, Nguyen NL, Jaramillo E, Collaborative group on the impact of digital technologies on TB. The impact of digital health technologies on tuberculosis treatment: a systematic review. Eur Respir J 2018 Jan;51(1) [FREE Full text] [doi: 10.1183/13993003.01596-2017] [Medline: 29326332]
12. de Sumari-de Boer IM, van den Boogaard J, Ngowi KM, Semvua HH, Kiwango KW, Aarnoutse RE, et al. Feasibility of real time medication monitoring among HIV infected and TB patients in a resource-limited setting. AIDS Behav 2016 May;20(5):1097-1107. [doi: 10.1007/s10461-015-1254-0] [Medline: 26604004]

13. Devi BR, Syed-Abdul S, Kumar A, Iqbal U, Nguyen PA, Li YC, et al. mHealth: an updated systematic review with a focus on HIV/AIDS and tuberculosis long term management using mobile phones. Comput Methods Programs Biomed 2015 Nov;122(2):257-265. [doi: 10.1016/j.cmpb.2015.08.003] [Medline: 26304621]
14. Kunawararak P, Pongpanich S, Chantawong S, Pokaew P, Traisathit P, Srithanaviboonchai K, et al. Tuberculosis treatment with mobile-phone medication reminders in northern Thailand. Southeast Asian J Trop Med Public Health 2011 Nov;42(6):1444-1451. [Medline: 22299414]
15. Bediang G, Stoll B, Elia N, Abena JL, Geissbuhler A. SMS reminders to improve adherence and cure of tuberculosis patients in Cameroon (TB-SMS Cameroon): a randomised controlled trial. BMC Public Health 2018 May 2;18(1):583 [FREE Full text] [doi: 10.1186/s12889-018-5502-x] [Medline: 29720146]
16. World Bank. World Bank Country and Lending Groups URL: https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups [accessed 2018-11-20] [WebCite Cache ID 743uHlrLZ]
17. World Health Organization. Definitions and reporting framework for tuberculosis–2013 revision URL: https://www.who.int/tb/publications/definitions/en/ [accessed 2019-01-14] [WebCite Cache ID 75Qsg5qy4]
18. Critical Appraisal Skills Programme. CASP Checklist URL: https://casp-uk.net/casp-tools-checklists/ [accessed 2018-12-03] [WebCite Cache ID 74NiKvDGi]
19. Atkins S, Launiala A, Kagaha A, Smith H. Including mixed methods research in systematic reviews: examples from qualitative syntheses in TB and malaria control. BMC Med Res Methodol 2012 Apr 30;12:62 [FREE Full text] [doi: 10.1186/1471-2288-12-62] [Medline: 22545681]
20. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009 Jul 21;6(7):e1000097 [FREE Full text] [doi: 10.1371/journal.pmed.1000097] [Medline: 19621072]
21. Blaya JA, Cohen T, Rodríguez P, Kim J, Fraser HS. Personal digital assistants to collect tuberculosis bacteriology data in Peru reduce delays, errors, and workload, and are acceptable to users: cluster randomized controlled trial. Int J Infect Dis 2009 May;13(3):410-418 [FREE Full text] [doi: 10.1016/j.ijid.2008.09.015] [Medline: 19097925]
22. Mohammed S, Glennerster R, Khan AJ. Impact of a daily SMS medication reminder system on Tuberculosis treatment outcomes: a randomized controlled trial. PLoS One 2016 Nov 1;11(11):e0162944 [FREE Full text] [doi: 10.1371/journal.pone.0162944] [Medline: 27802283]
23. Bassett IV, Giddy J, Chaisson CE, Ross D, Bogart LM, Coleman SM, et al. A randomized trial to optimize HIV/TB care in South Africa: design of the Sizanani trial. BMC Infect Dis 2013 Aug 26;13:390 [FREE Full text] [doi: 10.1186/1471-2334-13-390] [Medline: 23972276]
24. Bassett IV, Coleman SM, Giddy J, Bogart LM, Chaisson CE, Ross D, et al. Sizanani: a randomized trial of health system navigators to improve linkage to HIV and TB care in South Africa. J Acquir Immune Defic Syndr 2016 Oct 01;73(2):154-160 [FREE Full text] [doi: 10.1097/QAI.0000000000001025] [Medline: 27632145]
25. Huang R, Ren G, Hu J. Bracelet- and self-directed observational therapy for control of tuberculosis: study protocol for a cluster randomized controlled trial. Trials 2017 Jul 4;18(1):286 [FREE Full text] [doi: 10.1186/s13063-017-1996-2] [Medline: 28673323]
26. Bediang G, Stoll B, Elia N, Abena JL, Nolna D, Chastonay P, et al. SMS reminders to improve the tuberculosis cure rate in developing countries (TB-SMS Cameroon): a protocol of a randomised control study. Trials 2014 Jan 24;15:35 [FREE Full text] [doi: 10.1186/1745-6215-15-35] [Medline: 24460827]
27. Khachadourian V, Truzyan N, Harutyunyan A, Thompson ME, Harutyunyan T, Petrosyan V. People-centered tuberculosis care versus standard directly observed therapy: study protocol for a cluster randomized controlled trial. Trials 2015 Jun 22;16:281 [FREE Full text] [doi: 10.1186/s13063-015-0802-2] [Medline: 26093675]
28. Chaiyachati KH, Loveday M, Lorenz S, Lesh N, Larkan LM, Cinti S, et al. A pilot study of an mHealth application for healthcare workers: poor uptake despite high reported acceptability at a rural South African community-based MDR-TB treatment program. PLoS One 2013 May 28;8(5):e64662 [FREE Full text] [doi: 10.1371/journal.pone.0064662] [Medline: 23724075]
29. Howard AA, Hirsch-Moverman Y, Frederix K, Daftary A, Saito S, Gross T, et al. The START Study to evaluate the effectiveness of a combination intervention package to enhance antiretroviral therapy uptake and retention during TB treatment among TB/HIV patients in Lesotho: rationale and design of a mixed-methods, cluster-randomized trial. Glob Health Action 2016;9:31543 [FREE Full text] [doi: 10.3402/gha.v9.31543] [Medline: 27357074]
30. Iribarren S, Beck S, Pearce PF, Chirico C, Etchevarria M, Cardinale D, et al. TextTB: a mixed method pilot study evaluating acceptance, feasibility, and exploring initial efficacy of a text messaging intervention to support TB treatment adherence. Tuberc Res Treat 2013;2013:349394 [FREE Full text] [doi: 10.1155/2013/349394] [Medline: 24455238]
31. Hirsch-Moverman Y, Burkot C, Saito S, Frederix K, Pitt B, Melaku Z, et al. Reaching the end of the line: operational issues with implementing phone-based unannounced pill counts in resource-limited settings. PLoS One 2017 Oct 19;12(10):e0185549 [FREE Full text] [doi: 10.1371/journal.pone.0185549] [Medline: 29049382]
32. Hirsch-Moverman Y, Howard AA, Frederix K, Lebelo L, Hesseling A, Nachman S, et al. The PREVENT study to evaluate the effectiveness and acceptability of a community-based intervention to prevent childhood tuberculosis in Lesotho: study

protocol for a cluster randomized controlled trial. Trials 2017 Nov 21;18(1):552 [FREE Full text] [doi: 10.1186/s13063-017-2184-0] [Medline: 29157275]
33. Nguyen TA, Pham MT, Nguyen TL, Nguyen VN, Pham DC, Nguyen BH, et al. Video directly observed therapy to support adherence with treatment for tuberculosis in Vietnam: a prospective cohort study. Int J Infect Dis 2017 Dec;65:85-89 [FREE Full text] [doi: 10.1016/j.ijid.2017.09.029] [Medline: 29030137]
34. Daftary A, Hirsch-Moverman Y, Kassie GM, Melaku Z, Gadisa T, Saito S, et al. A qualitative evaluation of the acceptability of an interactive voice response system to enhance adherence to Isoniazid preventive therapy among people living with HIV in Ethiopia. AIDS Behav 2017 Nov. [doi: 10.1007/s10461-016-1432-8] [Medline: 27221743]
35. Albino S, Tabb KM, Requena D, Egoavil M, Pineros-Leano MF, Zunt JR, et al. Perceptions and acceptability of short message services technology to improve treatment adherence amongst tuberculosis patients in Peru: a focus group study. PLoS One 2014 May 14;9(5):e95770 [FREE Full text] [doi: 10.1371/journal.pone.0095770] [Medline: 24828031]
36. Nhavoto JA, Grönlund Å, Klein GO. Mobile health treatment support intervention for HIV and tuberculosis in Mozambique: perspectives of patients and healthcare workers. PLoS One 2017 Apr 18;12(4):e0176051 [FREE Full text] [doi: 10.1371/journal.pone.0176051] [Medline: 28419149]
37. Hoffman JA, Cunningham JR, Suleh AJ, Sundsmo A, Dekker D, Vago F, et al. Mobile direct observation treatment for tuberculosis patients: a technical feasibility pilot using mobile phones in Nairobi, Kenya. Am J Prev Med 2010 Jul;39(1):78-80. [doi: 10.1016/j.amepre.2010.02.018] [Medline: 20537846]
38. Garfein RS, Collins K, Muñoz F, Moser K, Cerecer-Callu P, Raab F, et al. Feasibility of tuberculosis treatment monitoring by video directly observed therapy: a binational pilot study. Int J Tuberc Lung Dis 2015 Sep;19(9):1057-1064 [FREE Full text] [doi: 10.5588/ijtld.14.0923] [Medline: 26260824]
39. Dwolatzky B, Trengove E, Struthers H, McIntyre JA, Martinson NA. Linking the global positioning system (GPS) to a personal digital assistant (PDA) to support tuberculosis control in South Africa: a pilot study. Int J Health Geogr 2006 Aug 16;5:34 [FREE Full text] [doi: 10.1186/1476-072X-5-34] [Medline: 16911806]
40. Ha YP, Tesfalul MA, Littman-Quinn R, Antwi C, Green R, Mapila TO, et al. Evaluation of a mobile health approach to Tuberculosis contact tracing in Botswana. J Health Commun 2016 Oct;21(10):1115-1121 [FREE Full text] [doi: 10.1080/10810730.2016.1222035] [Medline: 27668973]
41. Cowan J, Michel C, Manhiça I, Mutaquiha C, Monivo C, Saize D, et al. Remote monitoring of Xpert® MTB/RIF testing in Mozambique: results of programmatic implementation of GxAlert. Int J Tuberc Lung Dis 2016 Mar;20(3):335-341. [doi: 10.5588/ijtld.15.0535] [Medline: 27046714]
42. Lorent N, Choun K, Thai S, Kim T, Huy S, Pe R, et al. Community-based active tuberculosis case finding in poor urban settlements of Phnom Penh, Cambodia: a feasible and effective strategy. PLoS One 2014 Mar;9(3):e92754 [FREE Full text] [doi: 10.1371/journal.pone.0092754] [Medline: 24675985]
43. Mahmud N, Rodriguez J, Nesbit J. A text message-based intervention to bridge the healthcare communication gap in the rural developing world. Technol Health Care 2010;18(2):137-144. [doi: 10.3233/THC-2010-0576] [Medline: 20495253]
44. Narasimhan P, Bakshi A, Kittusami S, Prashant S, Mathai D, Bakshi K, et al. A customized m-Health system for improving Tuberculosis treatment adherence and follow-up in south India. Health Technol 2013 Dec 18;4(1):1-10. [doi: 10.1007/s12553-013-0067-2]
45. Broomhead S, Mars M. Retrospective return on investment analysis of an electronic treatment adherence device piloted in the Northern Cape Province. Telemed J E Health 2012;18(1):24-31. [doi: 10.1089/tmj.2011.0143] [Medline: 22150713]
46. Hunchangsith P, Barendregt JJ, Vos T, Bertram M. Cost-effectiveness of various tuberculosis control strategies in Thailand. Value Health 2012;15(1 Suppl):S50-S55 [FREE Full text] [doi: 10.1016/j.jval.2011.11.006] [Medline: 22265067]
47. Needham DM. Understanding and improving clinical trial outcome measures in acute respiratory failure. Am J Respir Crit Care Med 2014 Apr 15;189(8):875-877 [FREE Full text] [doi: 10.1164/rccm.201402-0362ED] [Medline: 24735026]
48. Bakibinga P, Kamande E, Omuya M, Ziraba AK, Kyobutungi C. The role of a decision-support smartphone application in enhancing community health volunteers' effectiveness to improve maternal and newborn outcomes in Nairobi, Kenya: quasi-experimental research protocol. BMJ Open 2017 Jul 20;7(7):e014896 [FREE Full text] [doi: 10.1136/bmjopen-2016-014896] [Medline: 28729309]

Abbreviations

DALY: disability-adjusted life years
DOT: directly observed treatment
IPT: isoniazid preventive therapy
IVR: interactive voice response
LMICs: low- and middle-income countries
MDR-TB: multidrug resistant-TB
mHealth: mobile health
RCT: randomized controlled trial
SMS: short message service

TB: tuberculosis
WHO: World Health Organization

Edited by G Eysenbach; submitted 02.10.18; peer-reviewed by A Martinez-Millana, A Gray; comments to author 11.11.18; revised version received 03.12.18; accepted 09.12.18; published 19.02.19.

Please cite as:
Lee S, Lee Y, Lee S, Islam SMS, Kim SY
Toward Developing a Standardized Core Set of Outcome Measures in Mobile Health Interventions for Tuberculosis Management: Systematic Review
JMIR Mhealth Uhealth 2019;7(2):e12385
URL: http://mhealth.jmir.org/2019/2/e12385/
doi:10.2196/12385
PMID:30777847

Viewpoint

An Ecological Approach to Smart Homes for Health Care Services: Conceptual Framework of a Smart Servicescape Wheel

Hyo-Jin Kang[1], BSc, MBA, PhD; Jieun Han[2], BSc, PhD; Gyu Hyun Kwon[2], BSc, MSc, PhD

[1]Department of Service Design Engineering, Sungshin Women's University, Seoul, Republic of Korea

[2]Graduate School of Technology and Innovation Management, Hanyang University, Seoul, Republic of Korea

Corresponding Author:
Gyu Hyun Kwon, BSc, MSc, PhD
Graduate School of Technology and Innovation Management
Hanyang University
222 Wangsimni-ro
Seongdong-gu
Seoul, 04763
Republic of Korea
Phone: 82 2 2220 2414
Email: ghkwon@hanyang.ac.kr

Abstract

Background: Smart homes are considered effective solutions for home health care for the elderly, as smart home technologies can reduce care costs and improve elderly residents' independence. To develop a greater understanding of smart homes for health care services (SHHSs), this study accentuated the necessity of ecological approaches with an emphasis on environmental constraints. This study was based on 2 rationales: (1) users are inclined to perceive the service quality and service experience from environments (ie, servicescape) owing to the intangibility of health care and the pervasiveness of smart home technologies, and (2) both service domains are complex adaptive systems in which diversified and undefined service experiences—not only a few intended service flows—can be generated by complex combinations of servicescape elements.

Objective: This study proposed the conceptual framework of a Smart Servicescape Wheel (SSW) as an ecological approach delineating the extensive spectrum of environmental constraints in SHHSs.

Methods: The SSW framework was established based on a literature review.

Results: Generally divided by perceptible and imperceptible servicescapes, the SSW consists of the perceptible Physical scape (ie, hardware components, environmental cues, and human states) and Social scape (ie, service relationships and social relationships) as well as the imperceptible Datascape (ie, computing intelligence, databases, and communication networks). Following the ecological approach, each category of the SSW is subdivided and defined at the level of components or functions.

Conclusions: The SSW's strengths lie in the various application opportunities for SHHSs. In terms of service planning and development, the SSW can be utilized to (1) establish the requirements for SHHS development, (2) associate with work domain analysis by defining component layers, and (3) understand the real contexts of SHHSs for the enhanced prediction of diverse service experiences. Regarding service management, it can be applied to develop measurement items for the operation and evaluation of SHHSs.

(JMIR Mhealth Uhealth 2019;7(2):e12425) doi:10.2196/12425

KEYWORDS

health care information management; system analysis, smart homes for health care services; ecological approach; conceptual framework; smart servicescape wheel

Introduction

Growing Needs for Informal Care for the Elderly

Increased life expectancy and declining birthrates have caused the rapid expansion of an aging population. Moreover, because of common chronic diseases (eg, heart disease, stroke, cancer, dementia, and diabetes), the demand and costs required for the care of elderly people are immensely growing [1,2]. To diminish these care costs in a relatively effective way, the solution focus is shifting away from formal care in hospitals and care centers

toward informal care in personal home environments [3] because of the cost-effectiveness of informal care and less intrusiveness in one's personal life [4,5]. Elderly people often require frequent and immediate medical interventions to prevent emergencies through continuous monitoring of their physiological parameters and activities [6]. This continuous monitoring, which might cause a significant financial burden in the case of formal inpatient care, can be realized more comfortably through informal care using a smart home platform [7]. Highly developed information and communication technologies can make home environments intelligent and provide remote, nonintrusive health care monitoring [6].

Smart Homes as Effective Solutions for Home Health Care Services for the Elderly

Smart homes are being realized using ubiquitous computing and internet of things (IoT) technology, which connects devices and systems [8,9]. Previous studies have characterized the smart home as integrating technologies such as (1) home automation that allows devices and systems to be controlled automatically, (2) communication networks that connect the key electrical appliances and services, (3) remote access and control that allows the system to be operated at a distance, and (4) home intelligence that is aware of users' individual contexts [10-13]. With those characteristics, scholars have defined the smart home's purpose as providing a better home life experience by promoting safety, security, comfort, communication, and entertainment through technical management in the home environment [14,15]. Meanwhile, other scholars have pointed out the concerns over ethical and legal issues of smart home technologies regarding privacy, security, and confidentiality because of the highly identifiable nature of data [16-21]. Nevertheless, with a judicious approach to their ethical and legal risks, smart homes could be effective for continuous and remote monitoring of elderly health and for disease prevention, and they can reduce the costs of care for the aging generation while improving their independence and quality of life [2,3,22,23]. Various terms have been utilized to describe this technology, such as *health smart home* [24,25], *U-health smart home* [26], *ubiquitous health care* [6,27,28], and *smart homes in or for (elderly) health care* [2,3].

Meanwhile, the focus of previous research on smart homes has primarily been on the development and application of smart home technologies within the boundaries of computer science and engineering research [11,12,15,29-32] because smart home production is inherently technology-intensive [33]. Likewise, according to several literature reviews about health care in smart homes [2,3,22,34], smart homes for health care have largely been investigated from the perspective of technology application. Such studies have analyzed and classified literature on smart homes for health care according to the types of sensors, network or communication technologies, and the algorithm models of data processing [2,3,12].

However, health care services are cocreated by elderly residents' experiences, and their experiences directly influence the perceived quality of the health care service [35,36]. Moreover, health care service experience is shaped by the interactions of elderly users with numerous touchpoints in the context of service [37]. In this sense, the viewpoints of service experience and service context are substantial in health care service, but they have not actively been engaged in research on smart homes for health care. Therefore, we accentuated the service experience and context perspectives of health care in smart homes in this study. To be consistent in our use of terminology, we adopted the term *smart homes for health care services* (SHHSs) and established an operationalized definition encompassing both technical and experiential viewpoints as follows: SHHSs are residences integrated with ubiquitous computing and IoT technology, which have the characteristics of home automation, communication network, home intelligence, and remote control and access by authorized health care personnel. They provide informal health care services such as real-time and long-term health monitoring, disease prevention by detecting anomalies, and unobtrusive activity support that does not interfere with individuals' quotidian activities; thus, they reduce care costs, enable a satisfactory service experience in a private and comfortable home environment, and improve the independence of elderly residents.

Why Do We Focus on Service Environments Regarding the Health Care Service Experiences in Smart Homes?

The academic perspective of service experience originates from service management research. In this field, the significance of service environments and service context engaged with service experience has been emphasized, as the service provision and service interactions occur in such environments and contexts [38]. First, *service interactions*, widely called *service encounters* [39] in service management research, are direct or indirect interactions between customers and service providers. Service interactions have been underscored because they can be diversified according to individual customers' varied past experiences and preferences as well as probable differences between service providers [40]. Second, *service environment*, which Bitner defined as the *servicescape* [41], is the physical and social environment where the service interactions take place, and the servicescape can influence the different ways in which customers perceive and experience a service [41,42]. Finally, *service context* refers to the social and cultural structures, such as social norms and institutions, which could influence the way a service's stakeholders interact with each other [43,44]. Therefore, service experiences are created by the service provisions and service interactions in association with diverse service environments (ie, servicescapes) and are concurrently profoundly influenced by service context [38]. In other words, service experience could be cocreated by multidimensional combinations of those 3 layers—service interactions, servicescapes, and service contexts.

The main focus of this study—the *servicescape* is a fundamental and componential level in the creation of service experiences as it can serve as an inducing or restricting factor on usage behaviors. Consequently, it can influence the service user's emotional response, perception of the intangible service quality, satisfaction with the whole service experience, and intention of continuous usage [42,45-48].

In line with the significance of servicescapes in service experience, the servicescape perspective is also required for

SHHSs, which has the characteristics of both health care services and smart services. Generally, health care is a credence-based service and patients tend to have difficulty assessing the technical quality of service (ie, the professional credibility of medical care) [36]. Instead, functional quality (ie, the service quality as it relates to the physical environment—the servicescape) is the primary determinant that affects patients' perception of the health care service process [35,37]. Regarding smart services, intelligent devices and environments communicate and collect data in real time based on advanced networks and pervasive computing technology [49,50], but data transfer and technological elements are invisible or intangible to service users. Furthermore, direct interactions between users and service providers (ie, service encounters) are largely supplemented or replaced by interactions with smart devices and environmental elements. Consequently, users are inclined to perceive the service quality and service experience from the environments of smart services—namely, the *smart servicescape* [51]—thus, from the direct, visible, and tangible interactions with the smart servicescape. In this way, health care services and smart services commonly have a particular significance of servicescape from the viewpoint of service experience; for this reason, this study used the concept of smart servicescape for SHHSs.

The importance of the smart servicescape in SHHSs can be supported from another perspective; the ecological approach for a complex adaptive system. Rouse defined a *complex adaptive system* as a nonlinear and dynamic system that is composed of independent and intelligent agents whose behavior patterns emerge rather than being designed into or controlled by the system [52]. He also stated that one cannot force or command such systems to abide by behavioral and performance directions. The health care service falls under this system because of its diversified stakeholders and interests [53]. Meanwhile, the smart home can also be regarded as a complex adaptive system for 2 reasons. First, it is a dynamic and therefore complex system consisting of diversified intelligent agents (ie, smart devices and environments) based on real-time communication. Second, the smart home is an adaptive system that is continuously customized in accordance with the emerging behavioral patterns of its resident users because it is a private space where usage behaviors cannot be easily controlled to comply with service providers' intended directions. The SHHS can, therefore, be considered a complex adaptive system.

Ecological Approach to Smart Homes for Health Care Services

Then how can service experiences in a complex adaptive system such as an SHHS be understood? In the research field of service management or service design, service experiences have been analyzed and profiled, adopting *service blueprinting* or *customer journey maps*. Service blueprinting, pioneered by Shostack [54], is a diagrammatic approach in which the key activities and their linkages involved in service delivery are plotted. It clarifies the series of customer actions, physical evidence, and frontstage and backstage interactions to emphasize the perspective of service users [55,56]. Customer journey maps are also a diagrammatic method in which customers' steps in engaging with multiple touchpoints in a service are illustrated [57]. These maps conceptualize service experiences as a chronological process of a customer's journey with a service provider [58]. These approaches can provide an understanding of the sequential flow of tasks in a service itinerary (ie, the customer's journey) that a service provider has intentionally designed. However, they may not be suitable for describing a complex adaptive system such as SHHSs, because diversified—hence unintended and undefined—behavioral patterns can be generated given the specificities of SHHSs (ie, multiple stakeholders, private home circumstances, and perpetual service duration for continuous health monitoring).

For this reason, the SHHS requires the *ecological approach*. This approach, which originated from Barker and Gibson's ecological psychology theory [59,60], considers environmental and social factors as parts of a system. According to this approach, individuals are surrounded by a complex combination of physical and social variables that operate in both direct and indirect ways to influence human activities [61,62]. Vicente advocated for the ecological approach in the field of human-computer interaction (HCI) and claimed that a work analysis for complex sociotechnical systems needs to start with and prioritize environmental constraints [63]. Vicente classified the demands (or constraints) of human works into (1) *cognitive constraints* relevant to human cognitive systems such as mental models and (2) *environment constraints* that constitute the context in which humans are situated, such as their physical and social realities [63]. The dominant viewpoint in psychology and HCI has been the cognitivist approach, which emphasizes cognitive mental models such as the sequential flow of an instructional service itinerary. Alternatively, the ecological approach leads to the analysis of environmental constraints, and it enables individuals to understand the real context that may shape actual behaviors (ie, *behavior-shaping constraints*) [64] and allows them to deal with unexpected and variable situations [63]. Given that SHHSs are complex adaptive systems in which diversified and undefined behavioral patterns can be generated by complex combinations of contextual and environmental factors, the ecological approach is appropriate for understanding service experiences in SHHSs by identifying the smart servicescape. Considering the necessity of the ecological approach to SHHSs, this study proposed a conceptual framework of smart servicescape to delineate complex environmental elements of SHHSs.

Fundamental Concepts for the Framework: Smart Servicescape

Ahead of investigating the smart servicescape, the concept of *servicescape* needs to be clarified. As briefly mentioned in the Introduction, Bitner [41] defined the servicescape as a man-made physical and social environment in which service encounters are framed. Bitner emphasized the effects of physical environments and classified them into 3 dimensions: (1) ambient conditions (circumstantial attributes, such as temperature, air quality, noise, music, and odor); (2) space and function (the arrangement and layout of the machinery, equipment, and furnishings); and (3) signs, symbols, and artifacts (visible communicators on the exterior and interior) [41]. Since then, many scholars have explored the categorization of servicescapes in diverse service sectors (eg, restaurants, leisure, hospitality,

and health care facilities) [45,47,65-67]. Furthermore, the *social servicescape* has also been underlined including social or noncommercial relationships, such as direct interactions with service providers, indirect interactions with other customers, social density, and connectedness [42,68].

Founded on the notion of servicescape embracing physical and social aspects, we had proposed the concept of *smart servicescape* regarding smart services in a prior study [51]. Analyzing the cases of smart home service experiences, new dimensions, including *datascape*, were introduced to reflect the key characteristics of smart services (ie, real-time data collection and the continuous data exchange of intelligent objects) [51].

Meanwhile, several studies have been conducted on the architecture of smart homes in association with health care services [2,3,6,26]. Researchers have commonly proposed a four-layer architecture for SHHSs even though their detailed elements or labels differ. Generally, they have agreed on the layers of (1) sensors and actuators, (2) communication or network, (3) computing or data processing, and (4) services. The layer of sensors and actuators is relevant to physical devices, such as the units of home automation and home control, biosensors, and environment sensors. The communication layer refers to the wired or wireless home networking required for information gathering, service discovery, and appliance discovery. The computing and processing layer pertains to the information analysis and machine learning technology engaged in knowledge management and decision making. Finally, the service layer varies from the types of health care services to those of health care service providers and organizations. Although the layered architecture represents the overall structure of the SHHS well, the criteria for classification of the 4 layers are not clearly defined. Moreover, the aforementioned studies have been conducted in computer science and engineering fields; therefore, the research focus lies more on a technical perspective investigating which types of technologies are applied in each layer. The ecological approach has been absent from previous studies; particularly, the perspective of service experience and that of smart servicescape have not yet been investigated in SHHSs.

The Smart Servicescape Wheel

The Focus of the Smart Servicescape Wheel

To infuse the ecological approach with the understanding of service experiences in SHHSs—in other words, to delineate the extensive spectrum of environmental constraints in the complex adaptive system of SHHSs—we introduced a conceptual framework of the smart servicescape, the *Smart Servicescape Wheel* (SSW), as illustrated in Figure 1.

The Smart Servicescape Wheel for SHHSs globally classifies conventional servicescapes, such as the *Physical scape* and *Social scape* as *perceptible servicescapes*. The *Datascape*, which is a distinctive characteristic of smart services, is categorized as an *imperceptible servicescape* because its elements are hardly perceived by elderly residents unless they are intentionally visualized with certain service objectives. Moreover, the interactions that occur between the service user (ie, the elderly resident) and the perceptible servicescape can usually be recognized as *direct interactions*, such as the interactions of the elderly with smart devices (eg, physiological monitoring devices) or caregivers (eg, clinicians at health care centers and medical staff at hospitals). However, a large part of the interactions in the smart servicescape can hardly be perceived by the resident. These *ambient interactions* [69] are among the Datascape, Physical scape, and Social scape without the user's direct intervention. Subsequently, the whole scope of direct and ambient interactions constitutes the *cyber-socio-physical interactions* [70].

Nevertheless, the focus of the SSW from the ecological perspective is not on the interactions diversely shaped in the spectrum of smart servicescape but on the detailed servicescape elements, namely the environmental constraints themselves. Moreover, the environmental constraints should be defined independently of any particular device, event, task, or interface because the ecological approach's implication lies in designing a diverse spectrum of future service practices combining the environmental constraints rather than in designing a single flow of a service [63]. This implication implies that the smart servicescape elements need to be defined at the level of components or functions. Founded on this rationale, the Physical scape consists of (1) hardware components, (2) environmental cues, and (3) human states; the Social scape is composed of (1) service relationships and (2) social relationships; and the Datascape comprises (1) computing intelligence, (2) databases, and (3) communication networks. Accordingly, this section will discuss in detail each component of the SSW which is categorized in Table 1.

Physical Scape

Hardware Components

The SHHS depends on data-collecting equipment and devices to monitor the state of residents and their environments [12]. At the level of components rather than devices, the hardware components consist of sensors and actuators [2,6,26]. Sensors are used to detect states and changes in the residents and their environments by measuring environmental or physiological parameters, and they are often seamlessly integrated into the living space and equipment [3,71]. Adopting the taxonomy of sensors by Amiribesheli et al [3], sensors are classified into (1) physiological sensors, (2) environmental sensors, (3) multimedia sensors, and (4) binary sensors.

First, physiological sensors are required to monitor a resident's health condition. They acquire diverse biometric data (eg, body temperature, weight, blood pressure, pulse rate, blood glucose, and respiration—namely *physiological cues*, to be discussed in a later section) through various forms of medical equipment (eg, electrocardiography, electromyography, and electroencephalography) [3,12,72,73]. Second, environmental sensors are utilized to detect environmental data in smart homes, such as light, noise, temperature, and humidity—namely the *environmental cues* to be explicated in the following section. Light sensors that assess the illumination intensity and temperature sensors that measure data for heating or cooling air are commonly used [12,74]. Third, multimedia sensors include cameras and microphones to collect audiovisual data

[3]. Audiovisual data enable the behavior monitoring and activity recognition of residents with high accuracy [75,76], but it involves privacy concerns [77]. Finally, binary sensors' output data are in a discrete state of 0 or 1. These sensors are commonly used to detect the state of objects or residents because of the simple form of data and the unobtrusiveness in residents' daily lives [3]. Common binary sensors include passive infrared sensors that collect data about residents' movements or stillness [78], contact switch sensors that detect the state of objects (eg, opening and closing of door), and radio-frequency identification (RFID) tags that identify objects and people to track their locations [79].

As another hardware component, actuators are required to respond to commands or feedback from residents or to perform curated service actions decided from computing intelligence (to be explicated in the Datascape section), such as the control of ambience or home appliances [2].

Figure 1. Smart servicescape wheel for smart homes for health care services.

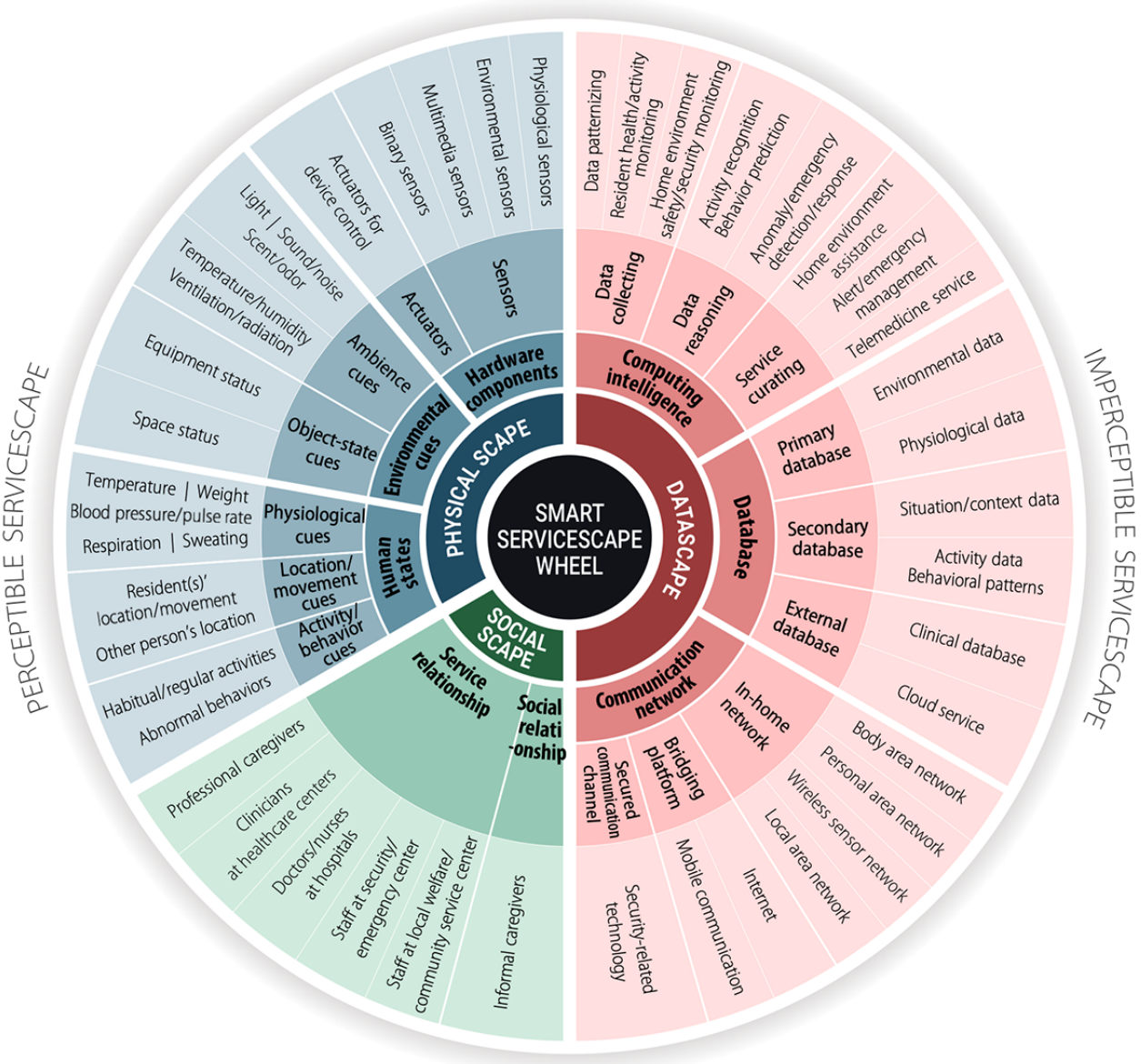


Table 1. Smart servicescape elements in smart homes for health care services.

Smart servicescape categories	Smart servicescape elements in SHHSs[a]	Literature source
Perceptible servicescape		
Physical scape		
Hardware components		
Sensors	Physiological sensors; environmental sensors; multimedia sensors; binary sensors	[2,3,12]
Actuators	Actuators for device control	[2]
Environmental cues		
Ambience cues	Light; sound or noise; scent or odor; temperature and humidity; ventilation and radiation	[2,12]
Object-state cues	Equipment status (eg, home appliances); space status (eg, home security and energy management)	[16]
Human states		
Physiological cues	Body temperature; weight; blood pressure and pulse rate; blood glucose; respiration and sweating	[12,16]
Location or movement cues	Resident's location and movement; other person's location	[3]
Activity or behavior cues	Habitual or regular activities; abnormal behaviors	[3]
Social scape		
Service relationship	Professional caregivers; clinicians at health care centers[b]; doctors and nurses at hospitals; staff at security or emergency centers[b]; staff at local welfare or community service centers[b]	[3]
Social relationship	Informal caregivers (eg, family members and neighbors)	[3]
Imperceptible servicescape		
Datascape		
Computing intelligence		
Data collecting	Resident health or activity monitoring; home environment (safety and security) monitoring; data patternizing[b]	[12,22]
Data reasoning	Activity recognition and behavior prediction; anomaly and emergency detection	[2,3,22]
Service curating	Home environment assistance; alert and emergency management; telemedicine service	[12,22]
Database		
Primary database	Environmental data; physiological data	[2,3,12]
Secondary database	Situation or context data[b]; activity data and behavioral patterns	[3]
External database	Clinical database[b]; cloud service[b]	N/A[c]
Communication network		
In-home network	Body area network; personal area network; wireless sensor network; local area network	[2]
Bridging platform	Internet; mobile communication	[2]
Secured communication channel	Security-related technology	[2]

[a]SHHSs: smart homes for health care services.

[b]These elements were newly added. Most of the paper sources for the elements of smart servicescape were literature review papers. The terms of elements were attuned to maintain consistent terminology, and the category labels were newly established in this study.

[c]N/A: not available.

Environmental Cues

The second category of the physical scape, environmental cues, is composed of *ambience cues* and *object-state cues*. Ambience cues include light, sound/noise, scent/odor, and air quality (ie, temperature, humidity, ventilation, or radiation), and they are usually detected by environmental sensors. Object-state cues involve (1) the equipment status (eg, the opening/closure or on/off of home appliances and the location of personal items) [3,12] and (2) space status (eg, home security or energy consumption status) [22], which are normally detected by binary sensors. Ambience cues and object-state cues captured by

pertinent sensors are collected as input data for home environment monitoring, and they can also appear as output results of *service curation by computing intelligence* (to be explained in the section of Datascape). For instance, an increase in a bathroom's humidity level, the opening and subsequent closure of the bathroom door, and an increase in water consumption can all be interpreted as the resident's shower (or bath)-taking activity. Consequently, monitoring service functions can be activated to prepare for possible emergencies, such as slips or falls. Although the data collected from environmental cues may be straightforward, the interpretation to capture the full activities or behavior patterns requires additional knowledge of the environment and context.

Human States

Human states are essential elements as a target of health care services with real-time long-term health monitoring. Physiological cues, location/movement cues, and activity/behavior cues constitute this category. First, physiological cues are captured by diverse physiological sensors and collected as biometric data. Body temperature, heart rate, blood pressure, respiration rate, and body weight are the frequently monitored basic parameters [72,73]. Blood glucose can be regularly monitored for diabetic patients; oxygen saturation of blood or sweating can be measured, for instance, to monitor an elderly resident's physical status during remedial exercises. Second, location or movement cues can also be detected by multimedia sensors (eg, video cameras) or binary sensors (eg, presence or motion sensors and pressure sensors). For example, when the resident is in a bedroom, his or her location can be detected by a closed-circuit television camera or a motion sensor installed in the room's ceiling or more simply by pressure sensors equipped in a bed or foot mat. Location cues can be tracked for the individual that is being cared for or multiple residents, including other family members or cohabitants. The location of a third person (eg, visitors) can also be considered depending on service functions. Finally, activity/behavior cues can be considered advanced data derived from the interpretation of the former physiological cues and/or location or movement cues combined with environmental cues. For instance, a cooking activity can be detected combining the resident's location in the kitchen by pressure sensors on a kitchen mat, his or her movement by a motion sensor in the kitchen, the opening/closing of a refrigerator door, the on/off of a gas stove switch, and so on [3]. In this way, habitual/regular activities, such as cooking, dining, sleeping, resting, or bathing, can be detected. At the same time, abnormal behaviors, such as slips or falls, can be captured based on irregular events, such as a sudden impact on a foot mat or an unusual duration of standing in locations [6,26].

Social Scape

Service Relationship

In the circumstances of SHHSs, the residents themselves are the very target of health monitoring service. Therefore, they are included in the Physical scape as the source of human states. Meanwhile, the resident can interact with diverse stakeholders of health care services in service relationships. Professional caregivers visit the resident's home regularly or irregularly to provide professional care services. Clinicians at health care centers are those who directly monitor the residents' health status and provide consultations to maintain the resident's health. Doctors and nurses can also be contacted through a telemedicine system; they may be interested in receiving updates about the progress of the resident's disease based on health-monitoring data [3]. Meanwhile, other service relationships exist that are not directly relevant to health or medical services. When emergencies or security problems occur, the staff at security or emergency centers can be contacted by the residents or family members, or they can be automatically informed by an intelligent SHHS system detecting abnormal behavior. Staff at local welfare centers or community service centers may also be a part of nonmedical service relationships, and their visits to a resident's home are important to check the living conditions of elderly people who live alone.

Social Relationship

In addition to service relationships, social relationships can be included in the social scape for those who work as *informal caregivers*. For instance, family members who live apart from the elderly resident, friends, neighbors, or any acquaintances can build informal relationships by routinely or irregularly contacting the elderly residents or by visiting them. Such relationships could be meaningful because the informal caregivers can help elderly residents engage in affective relations with emotional support for their daily lives.

Datascape

Computing Intelligence

As previously noted, the Datascape is a distinctive characteristic of smart services as an imperceptible servicescape, and *computing intelligence* is a foremost part of Datascape that can make health care services *smart* by intelligent processing. Computing intelligence is composed of 3 elements, data collecting, data reasoning, and service curating, and the computing process in SHHSs also follows this order. Meanwhile, it seems that those elements of computing intelligence might be considered more as functions rather than servicescape elements. However, Datascape is considered an imperceptible environment in the SSW, so the collecting and processing of data that are inevitable parts to constitute the Datascape will also be regarded as environmental elements in this conceptual framework.

First, *data collecting* is an initial phase in which primary data related to human states (ie, physiological cues and location/movement cues) or environmental cues (ie, ambience cues and object-state cues) are gathered. The collected data are patternized and stored in a database to be analyzed for the next reasoning phase. For instance, data from sensors that detect environmental cues can be patternized using the following format: *Detection [Date] [Hour] [Sensor number] [Object code] [Location code]* [3]. Subsequently, the collected primary data are used to (1) monitor the residents' health status and activity and (2) monitor the safety and security of the home environment.

Second, *data reasoning* uses a variety of knowledge engineering and data processing algorithms [3,12], such as artificial neural

networks [80], fuzzy logic [81], hidden Markov models [82], and context-based reasoning [83] to analyze the primary data. The aforementioned algorithms are used to learn and develop models for the resident's behavioral and physiological patterns as well as the home environmental patterns [2]. Consequently, data reasoning aims to (1) recognize the resident's daily activities, (2) predict the resident's behaviors based on activity recognition, and (3) detect anomalies or emergencies as the activities are carried out [3,22,84,85].

Finally, *service curating* is a substantial element of computing intelligence because it implies the final service solutions in SHHSs that have been generated from extensive interactions among the smart servicescape elements. Moreover, the service curating elements are directly experienced and perceived by the resident as *health care services*. According to the process of data reasoning and the decision-making process [2,26], computer intelligence can curate final solution services, such as (1) home environment assistance (eg, temperature control, gas/power control, or door opening/locking), (2) alert/emergency management (eg, a reminder to take pills, a warning when the gas is on, or an emergency call to family members/care centers), and (3) telemedicine service (eg, a progress check of chronic diseases, a consultation of health training status, or a prescription for medications).

Database

As the second category of the Datascape, the database is composed of a primary database, secondary database, or external database. First, the primary database stores environmental data and physiological data directly collected from relevant sensors or devices. It is usually a local database, and it provides those data to central servers—the computing intelligence—as inputs for further processing. Second, the secondary database accumulates the analyzed output data from the computing intelligence, such as situation/context data, activity data, and behavioral patterns. As the secondary database becomes strengthened by data accumulation cycles, the recognition accuracy of activities or behavioral patterns can be augmented by machine learning through computing intelligence [3,86,87]. Finally, the external database can be accessed by the database of SHHSs with authorized personnel's permission, for instance, to access a clinical database such as personal electronic medical records (EMRs) from health care centers. Moreover, a generic database such as cloud services can be accessed and utilized by SHHSs.

Communication Network

The sensors and actuators of the hardware components in SHHSs are connected with computing intelligence and the database through a communication network [3,12]. Consequently, the environmental cues and the human state cues captured by the sensors are transmitted to the central computing server over a wired and/or wireless communication medium [2].

In-home network refers to the communication platform that provides data communication inside the smart home. Body area network (BAN), personal area network (PAN), wireless sensor network (WSN), and local area network (LAN) can be included in this category, in order of range [2,12]. Physiological sensors integrated in wearable devices usually worn by the resident form the BAN. In addition to the wearable on-body sensors, diverse sensors on personal devices can be connected through RFID or Bluetooth, forming the PAN. The BAN or PAN can be connected with other environmental sensors and actuators through the WSN (eg, ZigBee) [88]. The central platform of computing intelligence can communicate with any sensors and actuators in the smart home using the WSN to collect data or send feedback for pertinent actions [2]. Consequently, those short-range and low-powered communication platforms of in-home networks form the LAN using a technology such as Wi-Fi.

Moreover, bridging platforms, such as the internet and mobile communication, are cost-effective and readily available solutions for remote communication to access the external database [12] and transmit various types of data, such as text, image, voice, and video [3]. Furthermore, secured communication channels [12] are required particularly for health care services to transmit data to/from the clinical database, such as personal EMRs in health care centers, because of privacy and authentication issues [89].

Discussion

How to Apply the Smart Servicescape Wheel

Implications for Service Planning and Development

As previously noted, the value of adopting the ecological approach by proposing the SSW lies in delineating the extensive spectrum of environmental constraints in the complex adaptive system of SHHSs. Diversified and undefined service experiences can be generated by multifaceted combinations of smart servicescape elements that could influence and shape the domain of users' behavior. Therefore, the SSW defined the smart servicescape elements at the level of components or functions. Founded on this conceptual framework, it would be worthwhile to suggest how to apply the SSW in research or practice for the development of SHHSs.

First, the SSW can be utilized to establish requirements for developing SHHSs. When service planners and developers contemplate elements to combine to provide appropriate service functions and service contexts, the SSW can serve as a *map* or a *list of candidates* to demonstrate various and possible options for smart servicescapes. It can help service planners and developers to consider the various types of perceptible servicescape elements—namely, the types of sensors or actuators and the types of information (eg, information from environmental cues, human states, or social scape relationships)—and their diverse combinations. Associated with their types and combinations, the types of imperceptible Datascape elements would be determined, such as the types of input or output data in terms of database, the types of data modeling and decision-making algorithms regarding computing intelligence, and the types of service curation. Moreover, depending on the service curation types, the pertinent elements of a Physical scape or Social scape can be selected for activation in response to planned actions.

Second, the SSW can be applied in association with work domain analysis (WDA). Founded on the ecological approach, WDA is a part of cognitive work analysis that identifies the functional structure of a system (ie, a work domain) independent of activities [63,90,91]. WDA's worth lies in the fact that the separation of structure from activities helps bring an important source of order to the analysis of complex adaptive systems such as SHHSs [92]. The WDA requires the determination of physical resources, technical functions, domain functions, domain values, and system purpose [92]. Therefore, the SSW can provide candidate elements for the layer of physical resources—and partially includes the layer of technical or domain functions in case of the elements of computing intelligence. In accordance with the specific properties of service concepts in SHHSs, service planners/developers can select relevant smart servicescape elements and connect them to pertinent functions and values to realize an intended service purpose effectively. Conversely, based on the various combinations of smart servicescape elements, innovative functions or values of service domains can be newly defined, and a novel service system can be proposed.

Finally, the SSW can contribute to the understanding of real contexts in SHHSs and the anticipation of diversified service experiences. The extensive spectrum of environmental elements illustrated by the SSW can support the anticipation of possible behavioral patterns in the resident's service experiences. Diverse combinations of the smart servicescape elements and their subsequent patterns of user behaviors can be accumulated in the database and learned by the computing intelligence employing machine learning technologies. Consequently, this process can improve the accuracy of behavior prediction and enhance the appropriateness of service curation in SHHSs.

Implications for Service Management

From the perspective of service management, the SSW can be utilized to develop measurement items for the operation and evaluation of SHHSs. To determine the items to evaluate services, the SSW can serve as a comprehensive list of elements to be considered. If the detailed evaluation measurement scales are developed, it could be applied in various ways such as to (1) diagnose the current state of service operation, (2) determine improvement points by identifying weaknesses in terms of components, functions, and services, (3) enhance the pertinence of actions against unexpected events or service failures, and (4) strengthen the thoroughness of SHHS management. Moreover, this kind of evaluation can find an opportunity to extend its application domains toward other services based on SHHSs.

Limitations and Future Research

Despite the value and opportunities presented by the SSW, several limitations of the framework stem from its focus and conceptual nature. As noted in the Introduction, service experiences could be cocreated by multidimensional association among 3 layers: servicescape, service interactions, and service contexts. However, the focus of the SSW is limited in describing detailed servicescape elements rather than in determining the interactions that occur between those elements or defining probable influences from/on social and cultural structures (ie, service context). For the sake of service planning and development, the other 2 layers of service interactions and service contexts need to be profoundly investigated.

First, in terms of the service interaction layer, the next phase after defining the environmental components and functions could be the establishment of detailed interactions and activities shaped by the combination of smart servicescape elements. Direct interactions such as those between the service user (ie, the elderly resident) and the perceptible servicescape need to be identified. Furthermore, the *ambient interactions* among the imperceptible and perceptible servicescapes (ie, the Datascape, Physical scape, and Social scape) without the user's direct intervention require a more comprehensive exploration considering the ambient characteristic of smart services. The layer of service interactions might allow us to consider innovative ways of direct and ambient interactions for more natural and user-centered experiences of SHHSs.

Second, regarding the service context layer, social and cultural structures such as social norms and institutions could influence the ways of service interaction and experience formation. Particularly, social structures and institutions pertinent to SHHSs need to be investigated in terms of ethical and legal issues because personal data collected through smart home technologies require a great degree of protection [17,20]. Therefore, norms, regulations, policy, and legislation systems in association with privacy, security, and confidentiality issues in SHHSs would require high attention in future research [93].

Third, as the SSW is a theoretically established conceptual framework, further empirical studies to evaluate the validity of the framework are needed; it should be applied in the development of a new system of SHHSs or in the analysis of real use cases of SHHSs. Moreover, because the SSW is defined at the conceptual component or function level, it is limited in its elucidation of real products or systems that could be the complete combination of smart servicescape elements. Therefore, further research on the association of the SSW components to describe the cases of actual products or systems could be possible.

Finally, although the aspect of Social scape was covered less extensively than other scape categories in this study, it does not mean that it is less significant. The Social scape has been investigated largely from the perspective of stakeholders in service research, and the stakeholders in SHHSs would be complicated considering their intricate relationships. Therefore, it could be required to explore the interplays of Social scape parties in association with the elements of Physical scape and Datascape.

Conclusions

This study asserted the value of an ecological approach with emphasis on environmental constraints for understanding SHHSs. This study is based on 2 rationales: (1) users tend to perceive service quality and service experiences through the servicescape because of the intangibility of health care and the pervasiveness of smart home services and (2) both service domains are complex adaptive systems in which diversified and undefined service experiences—not only a few intended service

flows—can be generated by complex combinations of servicescape elements.

Accordingly, the conceptual framework of the SSW was proposed as an ecological approach delineating the extensive spectrum of environmental constraints in SHHSs. Generally divided into perceptible and imperceptible servicescapes, the SSW consists of the perceptible Physical scape (ie, hardware components, environmental cues, and human states), the Social scape (ie, service relationships and social relationships), and the imperceptible Datascape (ie, computing intelligence, database, and communication networks).

The strengths of the SSW lie in its various application opportunities. The SSW can be utilized in service planning and development to (1) establish the requirements for SHHS development, (2) associate with WDA by defining component layers, and (3) understand the real contexts of SHHSs to enhance the prediction of diverse service experiences, as well as in service management to develop measurement items for the operation and evaluation of SHHSs.

Acknowledgments

This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea under grant NRF-2017S1A5A8019275. This fund has no specific role or influence in the research process.

Conflicts of Interest

None declared.

References

1. Wimo A, Jönsson L, Bond J, Prince M, Winblad B, Alzheimer Disease International. The worldwide economic impact of dementia 2010. Alzheimers Dement 2013 Jan;9(1):1-11.e3. [doi: 10.1016/j.jalz.2012.11.006] [Medline: 23305821]
2. Majumder S, Aghayi E, Noferesti M, Memarzadeh-Tehran H, Mondal T, Pang Z, et al. Smart homes for elderly healthcare-recent advances and research challenges. Sensors (Basel) 2017 Oct 31;17(11):E2496 [FREE Full text] [doi: 10.3390/s17112496] [Medline: 29088123]
3. Amiribesheli M, Benmansour A, Bouchachia A. A review of smart homes in healthcare. J Ambient Intell Human Comput 2015 Mar 14;6(4):495-517. [doi: 10.1007/s12652-015-0270-2]
4. Wiles JL, Jayasinha R. Care for place: the contributions older people make to their communities. J Aging Stud 2013 Apr;27(2):93-101. [doi: 10.1016/j.jaging.2012.12.001] [Medline: 23561274]
5. Majumder S, Mondal T, Deen MJ. Wearable sensors for remote health monitoring. Sensors (Basel) 2017 Jan 12;17(1):130 [FREE Full text] [doi: 10.3390/s17010130] [Medline: 28085085]
6. Deen MJ. Information and communications technologies for elderly ubiquitous healthcare in a smart home. Pers Ubiquit Comput 2015 Jun 26;19(3-4):573-599. [doi: 10.1007/s00779-015-0856-x]
7. van Hoof J, Demiris G, Wouters EJ. Handbook of Smart Homes, Health Care and Well-being. Basel, Switzerland: Springer; 2017.
8. Atzori L, Iera A, Morabito G. The Internet of Things: a survey. Comput Networks: Int J Comput Telecommunicat Networking 2010 Oct;54(15):2787-2805. [doi: 10.1016/j.comnet.2010.05.010]
9. Bassi A, Horn G. European Commission: Information Society and Media. 2008 May 27. Internet of Things in 2020: a roadmap for the future URL: https://docbox.etsi.org/erm/Open/CERP%2020080609-10/Internet-of-Things_in_2020_EC-EPoSS_Workshop_Report_2008_v1-1.pdf [accessed 2019-02-02] [WebCite Cache ID 75seJXAmf]
10. van Berlo A, Allen B. Design Guidelines on Smart Homes: A COST 219bis Guidebook. Brussels, Belgium: COST, European Co-Operation in the Field of Scientific and Technical Research; 1999:40.
11. Park SH, Won SH, Lee JB, Kim SW. Smart home: digitally engineered domestic life. Pers Ubiquitous Comput 2003 Jul 1;7(3-4):189-196. [doi: 10.1007/s00779-003-0228-9]
12. Alam MR, Reaz MB, Ali MA. A review of smart homes—past, present, and future. IEEE Trans Syst Man Cybern C 2012 Nov;42(6):1190-1203. [doi: 10.1109/TSMCC.2012.2189204]
13. Balta-Ozkan N, Davidson R, Bicket M, Whitmarsh L. The development of smart homes market in the UK. Energy 2013 Oct 1;60:361-372. [doi: 10.1016/j.energy.2013.08.004]
14. Poulson D, Nicolle CA, Galley M. Loughborough University Institutional Repository. Loughborough, UK: Loughborough University; 2002. Review of the current status of research on smart homes and other domestic assistive technologies in support of the TAHI trials URL: https://dspace.lboro.ac.uk/dspace-jspui/handle/2134/1030 [accessed 2019-02-02] [WebCite Cache ID 75sf3J7vA]
15. Aldrich FK. Smart homes: past, present and future. In: Harper R, editor. Inside The Smart Home. London: Springer; 2003:17-39.

16. Nebeker C, Lagare T, Takemoto M, Lewars B, Crist K, Bloss CS, et al. Engaging research participants to inform the ethical conduct of mobile imaging, pervasive sensing, and location tracking research. Transl Behav Med 2016 Dec;6(4):577-586. [doi: 10.1007/s13142-016-0426-4] [Medline: 27688250]
17. Nelson BW, Allen NB. Extending the passive-sensing toolbox: using smart-home technology in psychological science. Perspect Psychol Sci 2018 Nov;13(6):718-733. [doi: 10.1177/1745691618776008] [Medline: 30217132]
18. Carter A, Liddle J, Hall W, Chenery H. Mobile phones in research and treatment: ethical guidelines and future directions. JMIR Mhealth Uhealth 2015;3(4):e95 [FREE Full text] [doi: 10.2196/mhealth.4538] [Medline: 26474545]
19. Pisani AR, Wyman PA, Mohr DC, Perrino T, Gallo C, Villamar J, et al. Human subjects protection and technology in prevention science: selected opportunities and shallenges. Prev Sci 2016 Aug;17(6):765-778 [FREE Full text] [doi: 10.1007/s11121-016-0664-1] [Medline: 27220838]
20. Torous J, Nebeker C. Navigating ethics in the digital age: introducing connected and open research ethics (CORE), a tool for researchers and institutional review boards. J Med Internet Res 2017 Feb 8;19(2):e38 [FREE Full text] [doi: 10.2196/jmir.6793] [Medline: 28179216]
21. Kelly P, Marshall SJ, Badland H, Kerr J, Oliver M, Doherty AR, et al. An ethical framework for automated, wearable cameras in health behavior research. Am J Prev Med 2013 Mar;44(3):314-319. [doi: 10.1016/j.amepre.2012.11.006] [Medline: 23415131]
22. Demiris G, Hensel BK. Technologies for an aging society: a systematic review of "smart home" applications. Yearb Med Inform 2008:33-40. [Medline: 18660873]
23. Fisk MJ. The implications of smart home technologies. In: Peace SM, Holland C, editors. Inclusive Housing in an Ageing Society: Innovative Approaches. Bristol, UK: Policy Press Scholarship; Mar 2012.
24. Rialle V, Duchene F, Noury N, Bajolle L, Demongeot J. Health "Smart" home: information technology for patients at home. Telemed J E Health 2002;8(4):395-409. [doi: 10.1089/15305620260507530] [Medline: 12626109]
25. Bartolomeu P, Fonseca J, Vasques F. Challenges in health smart homes. 2008 Presented at: Second International Conference on Pervasive Computing Technologies for Healthcare; Jan 30-Feb 1, 2008; Tampere, Finland p. 19-22. [doi: 10.1109/PCTHEALTH.2008.4571016]
26. Agoulmine N, Deen MJ, Lee JS, Meyyappan M. U-Health smart home. IEEE Nanotechnology Mag 2011 Sep;5(3):6-11. [doi: 10.1109/MNANO.2011.941951]
27. Kang DO, Kang K, Lee HJ, Ko EJ, Lee J. A Systematic design tool of context aware system for ubiquitous healthcare service in a smart home. : IEEE; 2007 Presented at: Future Generation Communication and Networking (FGCN 2007); December 6-8, 2008; Jeju, South Korea p. 49-54. [doi: 10.1109/FGCN.2007.45]
28. Rahmani AM, Thanigaivelan NK, Tuan NG, Granados J, Negash B, Liljeberg P, et al. Smart e-health gateway: bringing intelligence to Internet-of-Things based ubiquitous healthcare systems. 2015 Presented at: 12th Annual IEEE Consumer Communications and Networking Conference (CCNC); January 9-12, 2015; Las Vegas, NV, USA p. 826-834. [doi: 10.1109/CCNC.2015.7158084]
29. Friedewald M, Da Costa O, Punie Y, Alahuhta P, Heinonen S. Perspectives of ambient intelligence in the home environment. Telemat Inform 2005 Aug;22(3):221-238. [doi: 10.1016/j.tele.2004.11.001]
30. Davidoff S, Lee MK, Yiu C, Zimmerman J, Dey AK. Principles of smart home control. Berlin: Springer Berlin Heidelberg; 2006 Presented at: UbiComp 2006: Ubiquitous Computing 8th International Conference; September 17-21, 2006; CA, USA p. 19-34. [doi: 10.1007/11853565_2]
31. Xiaojing Y, Junwei H. A framework for cloud-based smart home. 2011 Presented at: International Conference on Computer Science and Network Technology; December 24-26, 2011; Harbin, China p. 894-897. [doi: 10.1109/ICCSNT.2011.6182105]
32. Saad al-sumaiti A, Ahmed MH, Salama MM. Smart home activities: a literature review. Electr Pow Compo Sys 2014 Feb 5;42(3-4):294-305. [doi: 10.1080/15325008.2013.832439]
33. Wang G, Song D. Smart home services using the Internet of Things. In: Geng H, editor. Internet Of Things and Data Analytics Handbook. Hoboken, NJ: John Wiley & Sons; 2017:613-630.
34. Chan M, Estève D, Escriba C, Campo E. A review of smart homes- present state and future challenges. Comput Methods Programs Biomed 2008 Jul;91(1):55-81. [doi: 10.1016/j.cmpb.2008.02.001] [Medline: 18367286]
35. Butler D, Oswald SL, Turner DE. The effects of demographics on determinants of perceived health-care service quality. The case of users and observers. J Manag Med 1996;10(5):8-20. [Medline: 10166035]
36. Naidu A. Factors affecting patient satisfaction and healthcare quality. Int J Health Care Qual Assur 2009;22(4):366-381. [doi: 10.1108/09526860910964834] [Medline: 19725209]
37. Han J, Kang HJ, Kwon GH. Understanding the servicescape of nurse assistive robot: the perspective of healthcare service experience. 2017 Presented at: 14th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI); June 28-July 1, 2017; Jeju, South Korean p. 644-649. [doi: 10.1109/URAI.2017.7992693]
38. Akaka MA, Vargo SL. Extending the context of service: from encounters to ecosystems. J Serv Market 2015 Sep 14;29(6/7):453-462. [doi: 10.1108/jsm-03-2015-0126]
39. Surprenant CF, Solomon MR. Predictability and personalization in the service encounter. J Mark 1987 Apr;51(2):86. [doi: 10.2307/1251131]

40. Bitner MJ, Booms BH, Tetreault MS. The service encounter: diagnosing favorable and unfavorable incidents. J Mark 1990 Jan;54(1):71-84. [doi: 10.2307/1252174]
41. Bitner MJ. Servicescapes: the impact of physical surroundings on customers and employees. J Mark 1992 Apr;56(2):57-71. [doi: 10.2307/1252042]
42. Rosenbaum MS, Massiah C. An expanded servicescape perspective. J Serv Manag 2011 Aug 9;22(4):471-490. [doi: 10.1108/09564231111155088]
43. Akaka MA, Vargo SL, Lusch RF. The complexity of context: a service ecosystems approach for international marketing. J Int Mark 2013 Dec;21(4):1-20. [doi: 10.1509/jim.13.0032]
44. Akaka MA, Vargo SL, Schau HJ. The context of experience. J Serv Manag 2015 Apr 20;26(2):206-223. [doi: 10.1108/josm-10-2014-0270]
45. Wakefield KL, Blodgett JG. The effect of the servicescape on customers' behavioral intentions in leisure service settings. J Serv Mark 1996 Dec;10(6):45-61. [doi: 10.1108/08876049610148594]
46. Wakefield KL, Blodgett JG. Customer response to intangible and tangible service factors. Psychol Mark 1999 Jan;16(1):51-68. [doi: 10.1002/(SICI)1520-6793(199901)16:1<51::AID-MAR4>3.0.CO;2-0]
47. Lin IY. Evaluating a servicescape: the effect of cognition and emotion. Int J Hosp Manage 2004 Jun;23(2):163-178. [doi: 10.1016/j.ijhm.2003.01.001]
48. Reimer A, Kuehn R. The impact of servicescape on quality perception. Eur J Mark 2005 Jul;39(7/8):785-808. [doi: 10.1108/03090560510601761]
49. Allmendinger G, Lombreglia R. Four strategies for the age of smart services. Harv Bus Rev 2005 Oct;83(10):131-142. [Medline: 16250631]
50. Satyanarayanan M. Pervasive computing: vision and challenges. IEEE Pers Commun 2001 Aug;8(4):10-17. [doi: 10.1109/98.943998]
51. Kang HJ, Kwon GH, Kim B, Park E. A Framework for smart servicescape: a case of smart home service experience. 2017 Presented at: IASDR 2017 (International Association of Societies of Design Research); Oct 31-Nov 3, 2017; Cincinnati, OH, USA p. 1665-1677. [doi: 10.7945/C2RH6N]
52. Rouse WB. Managing complexity: disease control as a complex adaptive system. Inf Knowledge Syst Manag 2000 Apr;2(2):143-165 [FREE Full text]
53. Rouse WB. Health care as a complex adaptive system: implications for design and management. Bridge-Nat Acad Eng 2008 Dec 3;38(1):17 [FREE Full text]
54. Shostack GL. Designing services that deliver. Harv Bus Rev 1984;62(1):133-139 [FREE Full text]
55. Patrício L, Fisk RP, Falcão e Cunha JF, Constantine L. Multilevel service design: from customer value constellation to service experience blueprinting. J Serv Res 2011 Mar 29;14(2):180-200. [doi: 10.1177/1094670511401901]
56. Bitner MJ, Ostrom AL, Morgan FN. Service blueprinting: a practical technique for service innovation. Calif Manage Rev 2008 Apr;50(3):66-94. [doi: 10.2307/41166446]
57. Richardson A. Using customer journey maps to improve customer experience. Harv Bus Rev 2010;15(1):2-5 [FREE Full text]
58. Lemon KN, Verhoef PC. Understanding Customer Experience Throughout the Customer Journey. J Mark 2016 Nov;80(6):69-96. [doi: 10.1509/jm.15.0420]
59. Barker RG. Ecological Psychology: Concepts And Methods For Studying The Environment Of Human Behavior. Palo Alto, CA: Stanford University Press; Jun 1, 1969.
60. Gibson JJ. The Ecological Approach to Visual Perception. New York: Psychology Press; Nov 22, 2014:346.
61. Bronfenbrenner U. The Ecology Of Human Development: Experiments By Nature And Design. Cambridge, MA: Harvard University Press; Sep 15, 1981:352.
62. Bonnes M, Secchiaroli G. Environmental Psychology: A Psycho-Social Introduction. London, UK: SAGE Publications; 1995:240.
63. Vicente KJ. Cognitive Work Analysis: Toward Safe, Productive, And Healthy Computer-Based Work. Mahwah, NJ: Taylor & Francis, Inc; Apr 3, 1999:416.
64. Rasmussen J, Pejtersen AM, Schmidt K. Taxonomy for cognitive work analysis. Roskilde, Denmark: Risø National Laboratory; 1990:156.
65. Han J, Kang HJ, Kwon GH. A systematic underpinning and framing of the servicescape: reflections on future challenges in healthcare services. Int J Environ Res Public Health 2018 Mar 13;15(3):509 [FREE Full text] [doi: 10.3390/ijerph15030509] [Medline: 29534035]
66. Han J, Kang HJ, Kwon GH. A measurement for evaluating the environmental quality of advanced healthcare facilities: Intelligent healthscape quality for medical staff. Build Environ 2018 Oct 15;144:532-541. [doi: 10.1016/j.buildenv.2018.08.023]
67. Ryu K, Jang SS. The effect of environmental perceptions on behavioral intentions through emotions: the case of upscale restaurants. J Hosp Tour Res 2016 Jul 26;31(1):56-72. [doi: 10.1177/1096348006295506]
68. Tombs AG, McColl-Kennedy JR. Social-servicescape conceptual model. Mark Theory 2016 Aug 11;3(4):447-475. [doi: 10.1177/1470593103040785]

69. Hoffman DL, Novak TP. Emergent experience and the connected consumer in the smart home assemblage and the Internet of Things. SSRN J 2015 Aug 22 [FREE Full text] [doi: 10.2139/ssrn.2648786]
70. Ng IC, Wakenshaw SY. The Internet-of-Things: review and research directions. Int J Res Mark 2017 Mar;34(1):3-21. [doi: 10.1016/j.ijresmar.2016.11.003]
71. Orwat C, Graefe A, Faulwasser T. Towards pervasive computing in health care - a literature review. BMC Med Inform Decis Mak 2008 Jun 19;8:26 [FREE Full text] [doi: 10.1186/1472-6947-8-26] [Medline: 18565221]
72. Virone G, Noury N, Demongeot J. A system for automatic measurement of circadian activity deviations in telemedicine. IEEE Trans Biomed Eng 2002 Dec;49(12):1463-1469. [doi: 10.1109/TBME.2002.805452] [Medline: 12542242]
73. Williams G, Doughty K, Bradley D. A systems approach to achieving CarerNet-an integrated and intelligent telecare system. IEEE Trans Inf Technol Biomed 1998 Mar;2(1):1-9. [doi: 10.1109/4233.678527]
74. Helal S, Mann W, El-Zabadani H, King J, Kaddoura Y, Jansen E. The gator tech smart house: a programmable pervasive space. Comput J 2005 Mar;38(3):50-60. [doi: 10.1109/MC.2005.107]
75. Yamazaki T. Beyond the smart home. : IEEE; 2006 Presented at: 2006 International Conference on Hybrid Information Technology; Nov 9-11, 2006; Jeju, South Korea p. 350-355. [doi: 10.1109/ICHIT.2006.253633]
76. Yamazaki T. The ubiquitous home. Int J Smart Home 2007;1(1):17-22 [FREE Full text]
77. Demiris G, Oliver DP, Giger J, Skubic M, Rantz M. Older adults' privacy considerations for vision based recognition methods of eldercare applications. Technol Health Care 2009;17(1):41-48. [doi: 10.3233/THC-2009-0530] [Medline: 19478404]
78. Reeder B, Meyer E, Lazar A, Chaudhuri S, Thompson HJ, Demiris G. Framing the evidence for health smart homes and home-based consumer health technologies as a public health intervention for independent aging: a systematic review. Int J Med Inform 2013 Jul;82(7):565-579 [FREE Full text] [doi: 10.1016/j.ijmedinf.2013.03.007] [Medline: 23639263]
79. Juels A. RFID security and privacy: a research survey. IEEE J Sel Areas Commun 2006 Feb;24(2):381-394. [doi: 10.1109/JSAC.2005.861395]
80. Mozer MC. The neural network house: an environment that adapts to its inhabitants. 1998 Presented at: 15th National Conference of American Association for Artificial Intelligence; July 26–30, 1998; Madison, WI, USA p. 110-114 URL: http://www.aaai.org/Papers/Symposia/Spring/1998/SS-98-02/SS98-02-017.pdf
81. Vainio AM, Valtonen M, Vanhala J. Proactive fuzzy control and adaptation methods for smart homes. IEEE Intell Syst 2008 Mar;23(2):42-49. [doi: 10.1109/MIS.2008.33]
82. Brdiczka O, Langet M, Maisonnasse J, Crowley JL. Detecting human behavior models from multimodal observation in a smart home. IEEE Trans Autom Sci Eng 2009 Oct;6(4):588-597. [doi: 10.1109/TASE.2008.2004965]
83. Tinghuai M, Yong-Deak K, Qiang M, Meili T, Weican Z. Context-aware implementation based on CBR for smart home. : IEEE; 2005 Presented at: IEEE International Conference on Wireless And Mobile Computing, Networking And Communications; August 22-24, 2005; Montreal, Que, Canada. [doi: 10.1109/WIMOB.2005.1512957]
84. Xu K, Wang X, Wei W, Song H, Mao B. Toward software defined smart home. IEEE Commun Mag 2016 May;54(5):116-122. [doi: 10.1109/MCOM.2016.7470945]
85. Das SK, Cook DJ, Battacharya A, Heierman EO, Tze-Yun L. The role of prediction algorithms in the MavHome smart home architecture. IEEE Wireless Commun 2002 Dec;9(6):77-84. [doi: 10.1109/MWC.2002.1160085]
86. Ramos C, Augusto JC, Shapiro D. Ambient Intelligence—the next step for artificial intelligence. IEEE Intell Syst 2008 Mar;23(2):15-18. [doi: 10.1109/MIS.2008.19]
87. Rashidi P, Cook DJ. Keeping the resident in the loop: adapting the smart home to the User. IEEE Trans Syst Man Cybern A 2009 Sep;39(5):949-959. [doi: 10.1109/TSMCA.2009.2025137]
88. Yongping J, Zehao F, Du X. Design and application of wirless sensor network Web server based on S3C2410 and Zigbee Protocol. : IEEE; 2009 Presented at: International Conference on Networks Security, Wireless Communications and Trusted Computing; April 25-26, 2009; Wuhan, Hubei, China p. 28-31. [doi: 10.1109/NSWCTC.2009.404]
89. Yoo S, Rho D, Cheon G, Choi J. A central repository for biosignal data. : IEEE; 2008 Presented at: International Conference on Information Technology and Applications in Biomedicine; May 30-31, 2008; Shenzhen, China p. 275-277. [doi: 10.1109/ITAB.2008.4570556]
90. Naikar N, Hopcroft R, Moylan A. Defence Science and Technology Organisation Victoria (Australia) Air Operations Div. 2005 Feb. Work domain analysis: theoretical concepts and methodology URL: https://apps.dtic.mil/docs/citations/ADA449707 [accessed 2019-02-02] [WebCite Cache ID 75srwKjhR]
91. Rasmussen J, Pejtersen AM, Goodstein LP. Cognitive Systems Engineering. New York: Wiley-Interscience; Aug 1994.
92. Lintern G. Cognitive Systems Design. 2016. Tutorial: work domain analysis URL: http://www.cognitivesystemsdesign.net/Tutorials/Work%20Domain%20Analysis%20Tutorial.pdf [accessed 2018-01-18] [WebCite Cache ID 75iL2KuOk]
93. US Department Of Health And Human Services. A Treatment Improvement Protocol - Using Technology-Based Therapeutic Tools In Behavioral Health Services - TIP 60. Morrisville, North Carolina: Lulu; Nov 18, 2016:206.

Abbreviations

BAN: body area network

EMR: electronic medical record
HCI: human-computer interaction
IoT: internet of things
LAN: local area network
PAN: personal area network
RFID: radio-frequency identification
SHHS: smart homes for health care service
SSW: Smart Servicescape Wheel
WDA: work domain analysis
WSN: wireless sensor network

Edited by G Eysenbach; submitted 05.10.18; peer-reviewed by B Nelson; comments to author 22.11.18; revised version received 13.01.19; accepted 20.01.19; published 08.02.19.

Please cite as:
Kang HJ, Han J, Kwon GH
An Ecological Approach to Smart Homes for Health Care Services: Conceptual Framework of a Smart Servicescape Wheel
JMIR Mhealth Uhealth 2019;7(2):e12425
URL: http://mhealth.jmir.org/2019/2/e12425/
doi:10.2196/12425
PMID:30735145

Original Paper

EVIDENT Smartphone App, a New Method for the Dietary Record: Comparison With a Food Frequency Questionnaire

Jose I Recio-Rodriguez[1,2,3,4], PhD; Carmela Rodriguez-Martin[1,2,3], BSc; Jesus Gonzalez-Sanchez[1,2,3,5], PhD; Emiliano Rodriguez-Sanchez[1,2,3], MD, PhD; Carme Martin-Borras[6,7], PhD; Vicente Martínez-Vizcaino[8], MD, PhD; Maria Soledad Arietaleanizbeaskoa[9], PhD; Olga Magdalena-Gonzalez[10], BSc; Carmen Fernandez-Alonso[11], MD, PhD; Jose A Maderuelo-Fernandez[1,2,3], MD, PhD; Manuel A Gomez-Marcos[1,2,3,12*], MD, PhD; Luis Garcia-Ortiz[1,2,3,13*], MD, PhD; EVIDENT Investigators[14*]

[1]Primary Health Care Research Unit, Institute of Biomedical Research of Salamanca, La Alamedilla Health Center, Salamanca, Spain

[2]Health Service of Castilla y León, Valladolid, Spain

[3]Primary Care Prevention and Health Promotion Research Network, Barcelona, Spain

[4]Faculty of Health Sciences, University of Burgos, Burgos, Spain

[5]Department of Nursing, University of Extremadura, Plasencia Campus, Salamanca, Spain

[6]Department of Physical Activity and Sport Sciences, Faculty of Psychology, Education and Sport Sciences Blanquerna, Ramon Llull University, Barcelona, Spain

[7]Department of Physical Therapy, Faculty of Health Sciences Blanquerna, Ramon Llull University, Barcelona, Spain

[8]Center for Health and Social Research, University of Castilla-La Mancha, Cuenca, Spain

[9]Primary Health Care Research Unit of Bizkaia, Basque Health Service-Osakidetza, Bilbao, Spain

[10]Torre Ramona Health Center, Aragon Health Service, Zaragoza, Spain

[11]Casa del Barco Health Center, Castilla and León Health Service, Valladolid, Spain

[12]Department of Medicine, University of Salamanca, Salamanca, Spain

[13]Department of Biomedical and Diagnostic Sciences, University of Salamanca, Salamanca, Spain

[14]redIAPP: Spanish Research Network for Preventive Activities and Health Promotion in Primary Care, Salamanca, Spain

[*]these authors contributed equally

Corresponding Author:
Jesus Gonzalez-Sanchez, PhD
Department of Nursing
University of Extremadura, Plasencia Campus
Avda Comuneros 27-31
Salamanca,
Spain
Phone: 34 923231859
Email: jesusgonzsan@gmail.com

Abstract

Background: More alternatives are needed for recording people's normal diet in different populations, especially adults or the elderly, as part of the investigation into the effects of nutrition on health.

Objective: The aim of this study was to compare the estimated values of energy intake, macro- and micronutrient, and alcohol consumption gathered using the EVIDENT II smartphone app against the data estimated with a food frequency questionnaire (FFQ) in an adult population aged 18 to 70 years.

Methods: We included 362 individuals (mean age 52 years, SD 12; 214/362, 59.1% women) who were part of the EVIDENT II study. The participants registered their food intake using the EVIDENT app during a period of 3 months and through an FFQ. Both methods estimate the average nutritional composition, including energy intake, macro- and micronutrients, and alcohol. Through the app, the values of the first week of food recording, the first month, and the entire 3-month period were estimated. The FFQ gathers data regarding the food intake of the year before the moment of interview.

Results: The intraclass correlation for the estimation of energy intake with the FFQ and the app shows significant results, with the highest values returned when analyzing the app's data for the full 3-month period (.304, 95% CI 0.144-0.434; $P<.001$). For

this period, the correlation coefficient for energy intake is .233 (P<.001). The highest value corresponds to alcohol consumption and the lowest to the intake of polyunsaturated fatty acids (r=.676 and r=.155; P<.001), respectively. The estimation of daily intake of energy, macronutrients, and alcohol presents higher values in the FFQ compared with the EVIDENT app data. Considering the values recorded during the 3-month period, the FFQ for energy intake estimation (Kcal) was higher than that of the app (a difference of 408.7, 95% CI 322.7-494.8; P<.001). The same is true for the other macronutrients, with the exception g/day of saturated fatty acids (.4, 95% CI −1.2 to 2.0; P=.62).

Conclusions: The EVIDENT app is significantly correlated to FFQ in the estimation of energy intake, macro- and micronutrients, and alcohol consumption. This correlation increases with longer app recording periods. The EVIDENT app can be a good alternative for recording food intake in the context of longitudinal or intervention studies.

Trial Registration: ClinicalTrials.gov NCT02016014; http://clinicaltrials.gov/ct2/show/NCT02016014 (Archived by WebCite at http://www.webcitation.org/760i8EL8Q)

(JMIR Mhealth Uhealth 2019;7(2):e11463) doi:10.2196/11463

KEYWORDS

technology assessment, biomedical; telemedicine; energy intake; diet records; surveys and questionnaires

Introduction

Background

Within the framework of intervention studies, the analysis of current or recent dietary intake can be approached with a wide variety of methods. In general, these methods are based on either a record of actual food intake over several days or an estimation using tools that assess the frequency of consumption of different foods [1]. Food frequency questionnaires (FFQs) allow us to estimate normal food consumption habits over a relatively recent period (months or years) by recording the frequency and amount with which the foods included in a set list are eaten. This type of questionnaire is most frequently used in epidemiological studies because they are cheap, simple, fast, and allow individual intake to be estimated with just 1 administration. The main drawback is that they usually do not have complete data, their list of foods is finite, and they require specific information from the population of interest, including the main sources of nutrients in the normal diet of the zone in which the study is conducted [2,3].

The review by Illner et al [2] suggests that food data collected with cell phones could improve collection compared with using conventional records because of the possibility of recording in real time. In addition, higher intraobserver reliability has been observed in comparison with written formats. Estimates of dietary intake based on records made in periods of 3 to 7 days could reduce costs. However, validity still appears to be limited when estimating an individual's intake [4], given the necessary prior training and increased effort involved in data processing.

Different apps have proven their validity against a conventional method such as 24-hour recall. The apps Easy diet diary, My meal mate, and Food now obtained good correlations and a nonsignificant difference for energy intake [5-7] and small differences of means for some food groups [8]. In addition, other apps such as My fitness PAL or e-DIA correlate positively regarding the consumption of certain food groups [8,9], although they may overestimate that of carbohydrates and lipids, especially in individuals with a higher energy intake [9]. All these apps, however, have been validated in a relatively young audience and little is known about the behavior of older smartphone users.

Research into the effect of food on health requires new tools and technologies to quantify food intake in a valid, flexible, and simple way [3]. However, more evidence is needed before recommending the app of these alternatives for recording the normal diet of different populations, especially among adults or the elderly [2]. In the EVIDENT II study [10], an app for the recording of dietary patterns was developed, which, in addition to promoting the adoption of healthy lifestyles by the user, can be useful for gathering data regarding food intake and their subsequent use in nutritional intervention studies.

Objectives

The objective of this study was to compare the estimated values of energy consumption, macro- and micronutrients, and alcohol collected with the EVIDENT II smartphone app in a week, a month, and 3 months, with the data estimated by means of a frequency questionnaire of food consumption in an adult population aged between 18 and 70 years.

Methods

Study Design and Population

Participants were selected from among those included in the EVIDENT II clinical trial intervention group [10]. The objective of this study was to assess the effectiveness of adding the use of a smartphone app to brief lifestyle improvement counseling (physical activity and eating habits) in a sample of the adult population. Overall, 6 groups from the Red Española de Investigación para Actividades Preventivas y Promoción de la Salud en Atención Primaria (REDIAPP; Spanish Research Network for Preventive Activities and Health Promotion in Primary Care) participated in the study. Subjects aged between 18 and 70 years were included. Subjects were excluded if they were unable to do exercise or follow the Mediterranean diet, or if they met any of the exclusion criteria. These criteria were the known presence of coronary or cerebrovascular atherosclerotic disease; heart failure; moderate or severe chronic obstructive pulmonary disease; musculoskeletal disease involving limited walking; advanced respiratory, renal, or hepatic disease; severe

mental disease; or a treated oncological disease diagnosed in the last 5 years [10]. Of the 415 subjects in the intervention group (brief counseling + smartphone app) of the EVIDENT II study [10], 53 did not use the app to record their diet; hence, for the study presented in this manuscript, 362 subjects were included. This sample size allows the detection of a correlation coefficient of .15 or higher between the energy intake estimated by the FFQ and that estimated with the app, accepting an alpha risk of .05 and a beta risk of .20 in a bilateral contrast.

Variables and Measurement Instruments

Estimation of Average Daily Nutritional Composition With the EVIDENT App

The EVIDENT app is a smartphone app developed by the computer company CGB and the Grupo de Investigaciónen Atención Primaria de Castilla y León (Castilla y León Primary Care Research Group) of the REDIAPP (intellectual property registration number 00/2014/2207).

The app allows the user to enter information regarding lifestyle habits such as diet. Within the app, foods are organized by groups that include dairy products, eggs, meat, fish, vegetables, fresh fruits, pulses, cereals, oils and fats, pastries, cakes and cookies, processed foods, snacks, beverages, bread, nuts, pasta, rice, salads, seafood, sauces, soups, and creams. The app was configured with the characteristics of each participant (age, sex, weight, and height). The subjects were required to enter their food intake daily (breakfast, midmorning snack, lunch, afternoon snack, and dinner) and select dishes and foods from the app menu (Figure 1). In the case of breakfast, midmorning, and afternoon snacks, the user selects each of the foods and their quantity listed on the app in the standard measures used by the reference population (Figure 2). For the lunch and dinner record, the type of dish and the size are selected (Figure 1). Other foods consumed at different times are recorded with the meal closest to them in time. On completing the daily record, the user is provided with 2 types of information by the app (Figure 3). The first covers the average nutritional composition of the corresponding day in terms of energy intake and macronutrients. This information was estimated using food composition tables and serving sizes provided by the user. The second concerns the nutritional composition of each ingested food corresponding to the portion size selected. In addition, each individual receives general weekly notifications with feedback messages about how well the general recommendations regarding certain aspects of the diet such as the consumption of fruits and vegetables or olive oil are being applied. This study presents the daily estimates of energy intake, consumption of macro- and micronutrients, and alcohol during the first week of recording, the first month, and the total time that individuals used the app (3 months). The average daily energy intake is expressed in Kcal, whereas the intake of macronutrients and alcohol is expressed in g/day. The micronutrients are expressed in terms of their respective units of consumption.

Figure 1. EVIDENT app main screen and selection of dishes.

Figure 2. EVIDENT app food menu.

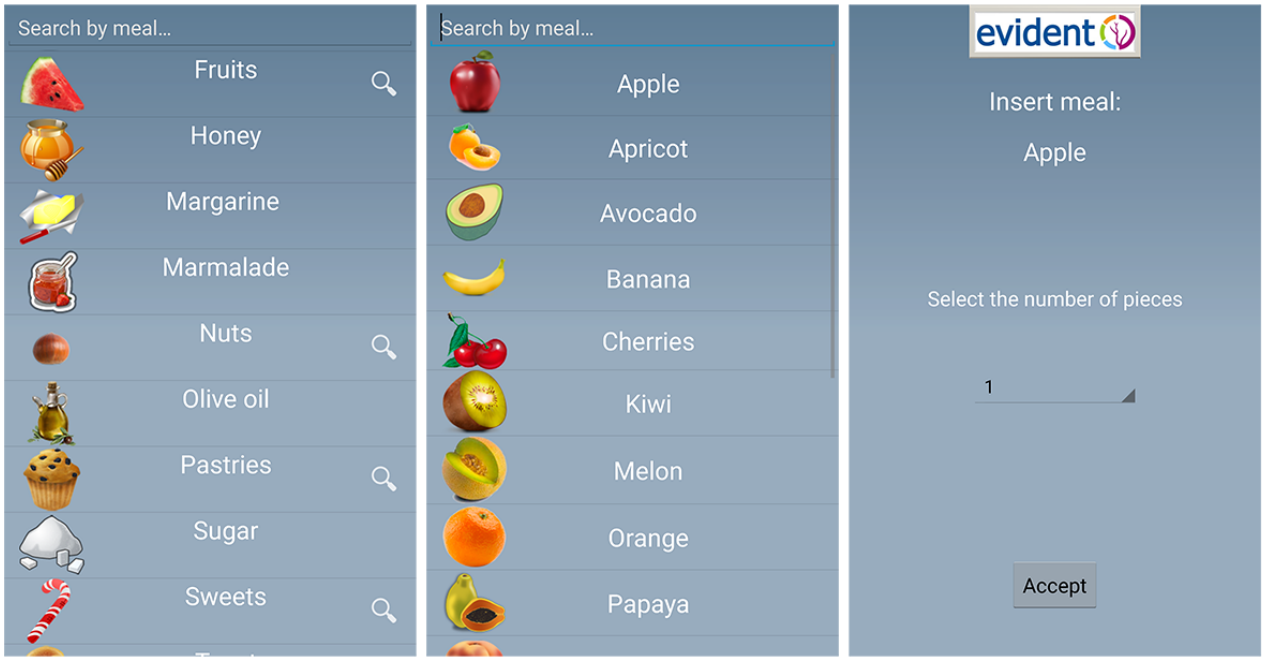
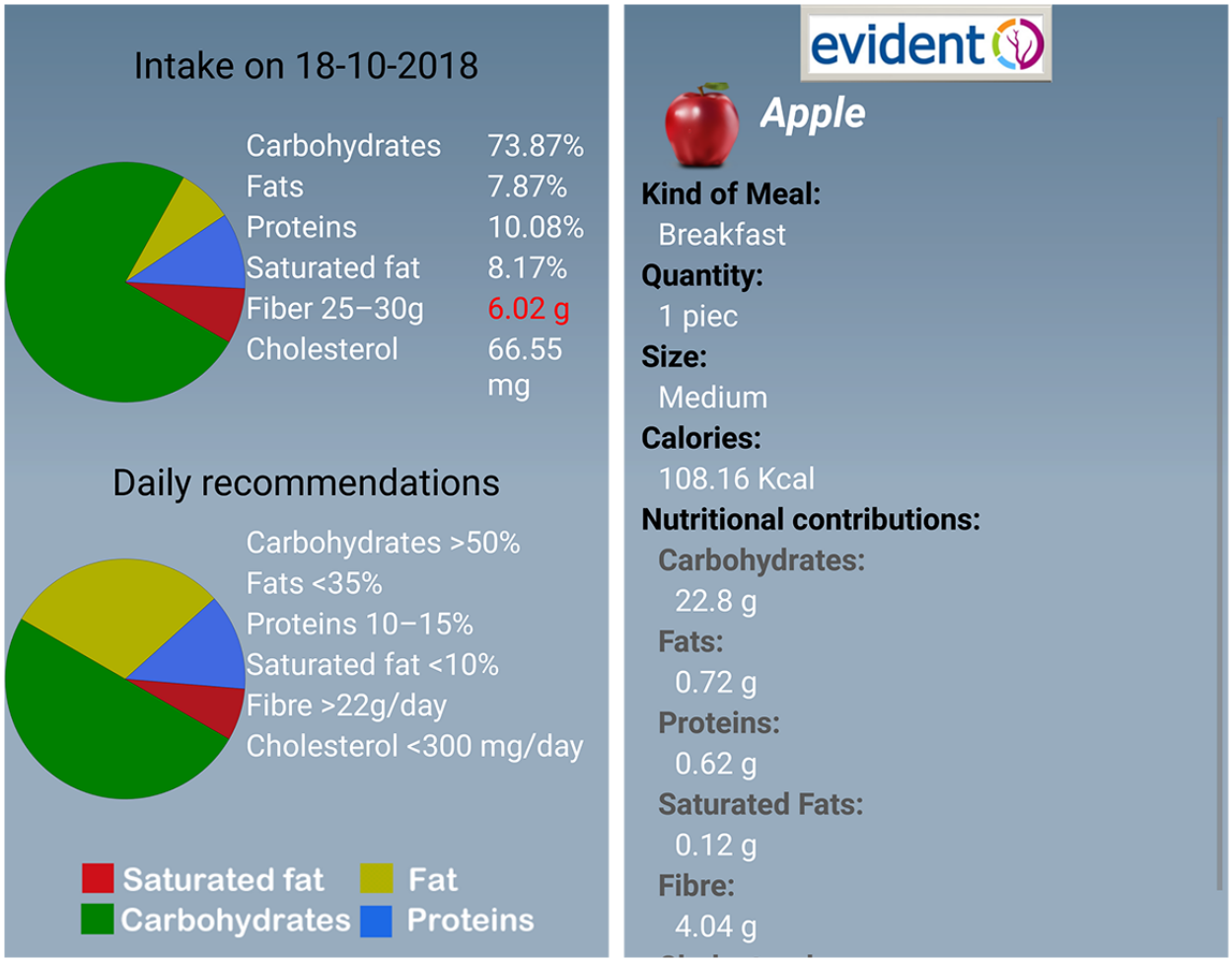


Figure 3. EVIDENT app feedback information.



Estimation of Average Daily Nutritional Composition With the Food Frequency Questionnaire (FFQ)

A self-administered FFQ validated for the Spanish population was used [11]. This questionnaire covers 137 foods frequently used among the reference population, along with their standard portion sizes. These foods are organized into groups that include dairy products, eggs, meat, fish, vegetables, fresh fruits, pulses, cereals, oils and fats, pastries and cakes, processed foods, snacks, and beverages. Within each of these categories, a certain number of products are included. After receiving instructions from researchers, participants indicated the frequency with which each item was consumed during the last year on a scale of 9 response options (never or almost never, 1-3 times a month, once a week, 2-4 times a week, 5-6 times a week, once a day, 2-3 times a day, 4-6 times a day, or more than 6 times a day). The questionnaire data were used to estimate daily energy consumption expressed in Kcal and intake of macronutrients and alcohol expressed in g/day. Micronutrients were quantified in terms of their respective units of consumption. For this study, we used data obtained from the FFQ conducted for 3 months.

Other Measurements

Data on the sociodemographic characteristics of the population (age and sex); educational level and occupation; smoking and personal history of hypertension, dyslipidemia, and diabetes mellitus were collected.

Hypertension was considered present with figures of ≥140 mm Hg systolic blood pressure and/or ≥90 mm Hg diastolic blood pressure or if the subject was being treated with antihypertensive agents [12]. Dyslipidemia was determined with total cholesterol of ≥240 mg/dL or triglycerides of ≥200 mg/dL or lipid-lowering drug treatment [13]. For diabetes mellitus 2, we applied the American Diabetes Association criteria (glycated hemoglobin ≥6.5%, fasting plasma glucose ≥126 mg/dL, plasma glucose ≥200 mg/dL after 2 hours during an oral tolerance test of glucose, random plasma glucose ≥200 mg/dL, or the use of antidiabetic treatment) [14].

Smoking history was assessed through questions about the participant's smoking status (smoker or nonsmoker). We considered smokers to be subjects who currently smoke or who stopped smoking less than 1 year ago.

Ethical Considerations

The study was approved by the clinical research ethics committee of the Salamanca health care area (June 21, 2013), and all participants gave written informed consent in accordance with the general recommendations of the Declaration of Helsinki [15].

Statistical Analysis

Descriptive statistics regarding clinical and sociodemographic characteristics of the studied population were expressed by means and SDs for the continuous variables and frequency distribution for the qualitative variables.

Concordance of the energy intake estimations produced by the app and the FFQ per week, month, and 3 months was analyzed using the intraclass correlation coefficient (ICC). In addition, the Bland-Altman method was used to calculate the limits of agreement between the 2 measurement tools. The comparison of means regarding energy consumption, macronutrients, and alcohol between the 2 measuring instruments was conducted with the *Studentt* test. In addition, the Pearson correlation coefficient was used to analyze the relationship between the estimations of macro- and micronutrients and alcohol intake from the 2 measurement tools and between the 3 estimates of the app.

All analyses were performed with SPSS version 23.0 (IBM Corporation, Armonk, NY, USA) and an alpha risk of .05 was set as the limit of statistical significance.

Results

Characteristics of the Study Population

Our sample comprised 362 individuals with a mean age of 52 (SD 12) years (214/362, 59.1% women). Regarding cardiovascular risk factors, 35.9% (130/362) had a diagnosis of arterial hypertension, 28.5% (103/362) had dyslipidemia, 20.4% (74/362) were smokers, and 7.5% (27/362) had diabetes mellitus type 2 (Table 1). Among the 362 participants, 235 (235/362, 64.9%) registered their food intake more than 61 days (2 months) on the app, 62 (62/362, 17.1%) between 31 and 60 days (between 1 and 2 months), and 65 (65/362, 18.0%) less than 30 days (<1 month).

Comparision Between Food Frequency Questionnaire and EVIDENT Smartphone App

ICC between the FFQ and app for the estimation of energy intake at 1 week, 1 month, and 3 months shows a significant association in all 3 cases, with the highest ICC yielded by the 3-month record (.304, 95% CI 0.144-0.434; Table 2). The Bland-Altman plot was used to analyze the concordance of energy intake estimated by the FFQ at 3 months and the EVIDENT app at 1 week, 1 month, and 3 months (Figure 4), with a limit of agreement at 3 months of 408 Kcal (95% CI −1223 to 2040).

Figures for the estimation of daily energy, macronutrient, and alcohol intakes are higher with the FFQ than those provided by the EVIDENT app (Table 3). The lowest values for energy, macronutrients, and alcohol were found after 3 months of data recording. As can be seen, the value estimated through the FFQ for energy (Kcal) was higher than that yielded by the app, with a difference of 408.7 (95% CI 322.7-494.8; P<.001). This is repeated with the remaining macronutrients, with the exception of saturated fatty acids (g/day), in which a difference of 0.4 (95% CI −1.2 to 2.0; P=.62) was observed. The average values of energy, macronutrients, and alcohol estimates recorded with the app can be seen in Table 4.

Table 1. Baseline characteristics of the study population (N=362).

Baseline characteristics	Statistics[a]
Age (years), mean (SD)	52 (12)
Females, n (%)	214 (59.1)
Work situation, n (%)	
Works outside of home	196 (54.1)
Homemaker	48 (13.3)
Retired	70 (19.3)
Student	7 (1.9)
Unemployed	41 (11.3)
Educational level, n (%)	
University studies	102 (28.2)
Middle or high school	184 (50.8)
Elementary school	76 (21.0)
Smoking, n (%)	
Nonsmoker	168 (46.4)
Smoker	74 (20.4)
Former smoker	120 (33.1)
BMI[b] categories, n (%)	
BMI <25	97 (26.8)
BMI 25 to 30	152 (42.0)
BMI >30	113 (31.2)
Hypertension, n (%)	130 (35.9)
Dyslipidemia, n (%)	103 (28.5)
Type 2 diabetes mellitus, n (%)	27 (7.5)

[a]Categorical variables are expressed as n (%) and continuous variables as mean (SD).

[b]BMI: body mass index.

Table 2. Intraclass correlation coefficient for energy intake.

Comparisons	Intraclass correlation (95% CI)	*P* value
FFQ[a] and EVIDENT app: 1 week	.203 (0.021-0.352)	.02
FFQ and EVIDENT app: 1 month	.267 (0.099-0.404)	.01
FFQ and EVIDENT app: 3 months	.304 (0.144-0.434)	<.001
EVIDENT app (1 week, 1 month, and 3 months)	.941 (0.929-0.951)	<.001

[a]FFQ: food frequency questionnaire.

Figure 4. Bland-Altman Plots with the differences in energy intake estimation (Kcal) between the food frequency questionnaire (FFQ; 3 months) and data records of the application (1 week, 1 month and 3 months). (a) Limit of agreement for the estimation of the energy intake between FFQ and EVIDENT application (1 week): 374 Kcal (-1373 to 2121); (b) Limit of agreement for the estimation of the energy intake between FFQ and EVIDENT application (1 month): 386 Kcal (-1291 to 2063); (c) Limit of agreement for the estimation of the energy intake between FFQ and EVIDENT application (3 months): 408 Kcal (-1223 to 2040).

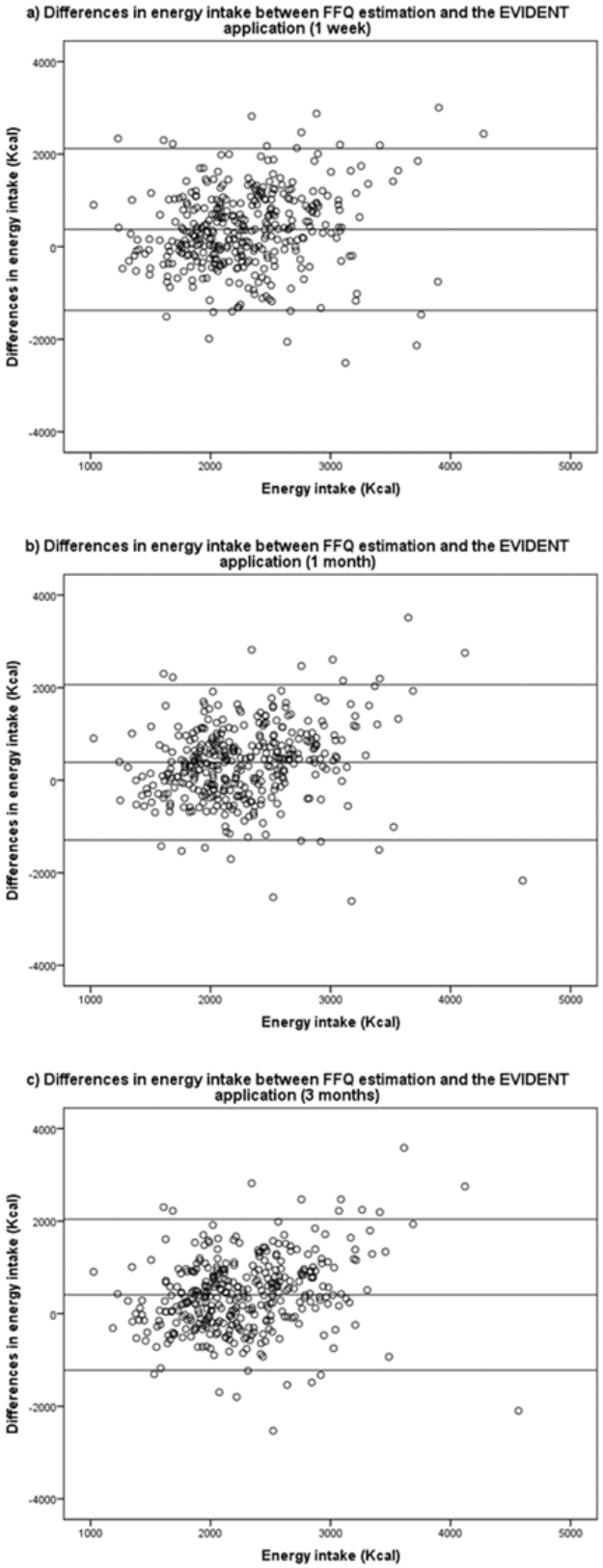


Table 3. Comparison of consumption of energy, macronutrients, and alcohol between the food frequency questionnaire and EVIDENT app.

Nutritional composition	FFQ[a], mean (SD)	Mean difference of FFQ-EVIDENT app at 1 week (95% CI)	Mean difference of FFQ-EVIDENT app at 1 month (95% CI)	Mean difference of FFQ-EVIDENT app at 3 months (95% CI)
Energy intake (Kcal/day)	2467.3 (729.8)	374.0 (281.9 to 466.1)	386.2 (297.7 to 474.6)	408.7 (322.7 to 494.8)
Proteins (g/day)	106.9 (29.5)	20.2 (16.5 to 23.9)	20.7 (17.1 to 24.3)	21.5 (18.0 to 25.1)
Carbohydrates (g/day)	260.4 (88.1)	59.8 (49.6 to 70.0)	60.8 (50.9 to 70.7)	63.8 (54.1 to 73.4)
Fats (g/day)	103.8 (37.9)	9.0 (4.1 to 14.0)	9.8 (4.9 to 14.6)	10.5 (5.8 to 15.2)
Saturated fats (g/day)	30.2 (12.3)	0.1 (−1.5 to 1.8)[b]	0.1 (−1.5 to 1.8)[b]	0.4 (−1.2 to 2.0)[b]
Monounsaturated (g/day)	45.9 (17.6)	1.7 (−0.6 to 4.1)[b]	2.2 (0.0 to 4.3)[b]	2.5 (0.4 to 4.6)
Polyunsaturated (g/day)	16.9 (8.6)	4.4 (3.4 to 5.4)	4.4 (3.4 to 5.4)	4.5 (3.5 to 5.5)
Cholesterol (g/day)	458.8 (175.9)	125.5 (105.2 to 145.8)	127.2 (107.5 to 146.9)	130.2 (110.9 to 149.4)
Fiber (g/day)	27.5 (10.5)	1.9 (0.7 to 3.2)	1.9 (0.7 to 3.2)	2.5 (1.3 to 3.7)
Alcohol (g/day)	9.2 (13.3)	4.0 (2.9 to 5.2)	3.9 (2.8 to 5.0)	3.8 (2.7 to 5.0)

[a]FFQ: food frequency questionnaire.

[b]*P*>.05.

Table 4. Estimation of the average daily consumption of energy, macronutrients, and alcohol with data collected through the EVIDENT app during the first week, first month, and throughout the 3 months.

Nutritional composition	EVIDENT app 1 week, mean (SD)	EVIDENT app 1 month, mean (SD)	EVIDENT app 3 months, mean (SD)
Energy intake (Kcal/day)	2093.2 (602.5)	2081.1 (577.0)	2058.5 (557.9)
Carbohydrates (g/day)	86.7 (24.6)	86.2 (22.3)	85.4 (21.9)
Proteins (g/day)	200.6 (60.5)	199.6 (58.6)	196.6 (57.6)
Fats (g/day)	94.7 (34.6)	94.0 (33.1)	93.2 (31.7)
Saturated fats (g/day)	30.0 (13.9)	30.0 (13.9)	29.8 (13.2)
Monounsaturated fats (g/day)	44.2 (15.7)	43.7 (14.1)	43.4 (13.5)
Polyunsaturated fats (g/day)	12.5 (4.9)	12.5 (4.9)	12.4 (4.8)
Cholesterol (g/day)	333.3 (125.1)	331.6 (113.0)	328.6 (106.9)
Fiber (g/day)	25.5 (8.3)	25.5 (8.3)	25.0 (8.2)
Alcohol (g/day)	5.2 (7.7)	5.3 (7.6)	5.3 (7.6)

Table 5 shows the correlations between the EVIDENT app's 1-week estimates for daily energy intake and the consumption of macronutrients and alcohol on the one hand, and the 1-month and 3-month records on the other. The values for energy correlate highly (r=.823 and r=.774; P<.001 all), although the highest correlations correspond to alcohol consumption (r=.894 and r=.865; P<.001), whereas the lowest correlations are those regarding cholesterol intake (r=.757 and r=.669; P<.001).

Table 6 shows the correlations between energy, macronutrient, and alcohol intakes recorded by the FFQ and the EVIDENT app for the first week, the first month, and at 3 months. The coefficients are higher when the FFQ is correlated with the 3-month data of the app. Energy intake yields a correlation coefficient of .233 (P<.001). The highest value corresponds to alcohol consumption and the lowest to the intake of polyunsaturated fatty acids (r=.676 and r=.155; P<.001), respectively.

Among the micronutrients (vitamins and minerals), we found significant correlations between FFQ and app estimates (Table 7). The highest concordance values were found for calcium and vitamin C intake (r=.316 and r=.319; P<.001), respectively, whereas the lowest correlations corresponded to phosphorus and vitamin D (r=.106 and r=.110; P<.05), respectively.

Table 5. Correlation of the consumption of nutrients between the intake estimated by the EVIDENT app in the first week, registration of the first month, and the record of the 3 months.

Nutritional composition	Correlation of the consumption between first week and first month	Correlation of the consumption between first week and three months
Energy intake (Kcal/day)	.823[a]	.774[a]
Carbohydrates (g/day)	.845[a]	.800[a]
Proteins (g/day)	.789[a]	.722[a]
Fats (g/day)	.800[a]	.731[a]
Saturated fats (g/day)	.825[a]	.763[a]
Monounsaturated fats (g/day)	.775[a]	.715[a]
Polyunsaturated fats (g/day)	.802[a]	.733[a]
Cholesterol (g/day)	.757[a]	.669[a]
Fiber (g/day)	.835[a]	.777[a]
Alcohol (g/day)	.894[a]	.865[a]

[a]*P*<. 01.

Table 6. Correlation of the nutrient consumption between the estimated values with the food frequency questionnaire and the registration through the EVIDENT app of the first week, first month, and 3 months.

Nutritional composition	Correlation between FFQ[a] with first week records by the app	Correlation between FFQ with first month records by the app	Correlation between FFQ with 3 months records by the app
Energy intake (Kcal/day)	.135[b]	.202[c]	.233[c]
Carbohydrates (g/day)	.147[c]	.224[c]	.268[c]
Proteins (g/day)	.172[c]	.207[c]	.204[c]
Fats (g/day)	.165[c]	.220[c]	.232[c]
Saturated fats (g/day)	.276[c]	.291[c]	.311[c]
Monounsaturated fats (g/day)	.118[b]	.220[c]	.236[c]
Polyunsaturated fats (g/day)	.111[b]	.130[b]	.155[c]
Cholesterol (g/day)	.154[c]	.193[c]	.204[c]
Fiber (g/day)	.204[c]	.280[c]	.314[c]
Alcohol (g/day)	.647[c]	.675[c]	.676[c]

[a]FFQ: food frequency questionnaire.

[b]*P*<.05.

[c]*P*<.01.

Table 7. Correlation of mineral and vitamin consumption between the values estimated by the food frequency questionnaire (FFQ) and data collected through the registration of the EVIDENT app of the first week, first month, and 3 months.

Micronutrients	Correlation between FFQ[a] with first week records by the app	Correlation between FFQ with first month records by the app	Correlation between FFQ with 3 months records by the app
Minerals			
Calcium (mg/day)	.259[b]	.295[b]	.316[b]
Iron (mg/day)	.190[b]	.254[b]	.259[b]
Iodo (µg/day)	.155[b]	.178[b]	.198[b]
Magnesium (mg/day)	.175[b]	.277[b]	.299[b]
Zinc (mg/day)	.175[b]	.210[b]	.221[b]
Selenium (µg/day)	.189[b]	.244[b]	.232[b]
Phosphorus (mg/day)	.093	.109[c]	.106[c]
Sodium (mg/day)	.124[c]	.157[b]	.151[b]
Potassium (mg/day)	.197[b]	.267[b]	.278[b]
Vitamins			
Vitamin A (mcg/day)	.141[b]	.167[b]	.166[b]
Vitamin D (mcg/day)	.094	.092	.110[c]
Vitamin C (mg/day)	.262[b]	.308[b]	.319[b]
Vitamin B1 (mg/day)	.155[b]	.189[b]	.224[b]
Vitamin B2 (mg/day)	.205[b]	.244[b]	.248[b]
Niacin (mg/day)	.130[c]	.181[b]	.171[b]
Vitamin B6 (mg/day)	.210[b]	.236[b]	.225[b]
Folic acid (mcg/day)	.231[b]	.276[b]	.298[b]
Vitamin B12 (mcg/day)	.139[b]	.196[b]	.196[b]
Retinol (mcg/day)	.163[b]	.151[b]	.162[b]
Carotenoids (mcg/day)	.230[b]	.275[b]	.274[b]

[a]FFQ: food frequency questionnaire.

[b]$P<.01$.

[c]$P<.05$.

Discussion

Principal Findings

Dietary records with the EVIDENT app yielded significant intraclass correlation with the FFQ in terms of energy intake. Similarly, the correlation between both instruments regarding the estimation of daily macro- and micronutrients intake and alcohol consumption was statistically significant. Furthermore, the correlations between the values obtained with the FFQ and with the EVIDENT app increase with longer app recording periods. The results of this study allow us to hypothesize that the use of electronic devices can be an alternative to FFQ for food recording in the context of longitudinal or intervention studies.

The increasing use of smartphones has produced an increase in the number of apps aimed at recording food intake. Many of these apps [6,7,16,17] have been compared against a traditional method of collecting information such as 24-hour recall with a young population. These apps have in common a high correlation for the estimation of energy intake, with small and insignificant differences between the app and 24-hour recall methods and with slightly higher values recorded by 24-hour recall for all measurements. The EVIDENT app's estimates of energy intake are significantly lower than those estimated with the FFQ. The correlation between both instruments of measurement is significant and rises as the food recording period using the app increases. However, this correlation is low compared with other works [6,7,16,17]. There are several explanations for this finding. First of all, the EVIDENT app was compared with an FFQ rather than a 24-hour recall. This

aspect is key to understanding the results. The 24-hour recall consists of collecting the most detailed information possible about the foods and beverages consumed in the preceding 24 hours. It is a method widely used in cross-sectional studies, but the great variety of available foods and ways of preparation makes it more difficult to estimate the rations consumed. In addition, one of the main limitations is that a single day of analysis does not reflect the usual pattern of consumption. When a comparison is made with a measuring instrument that explores food consumption over a short period (24 hours), a high correlation is expected as the food intake in that period is limited and the time gap between the records of both tools is closer, so there is less risk of memory bias. FFQs, meanwhile, estimate the nutritional intake over a longer in this case, 1 year. During this period, variability in foods is much higher and may be influenced by aspects such as seasonality or festivities that could alter food consumption. The FFQ seem to be a good instrument to estimate the nutritional composition over long periods (1 year or more) and the 24-hour recall in shorter periods (between 1 day and a week). However, for the estimation of energy and nutrient intake in the context of clinical investigations, it is necessary to use instruments capable of reporting beyond recent intake. The results of this study suggest that the use of the EVIDENT app can be a good measuring instrument for the estimation of nutritional composition in the context of longitudinal or intervention studies, although the implementation in clinical practice of this type of food registration tools could be difficult if there is not enough motivation to register them for long periods.

Comparison With Prior Work

Among the macronutrients, the highest correlation is found for carbohydrates and fiber and the lowest, although significant, for polyunsaturated fatty acids. However, the correlation coefficients maintain similar values for all dietary components, which lend the app a high degree of cohesion. Again, the higher the number of days recorded, the greater the correlation. The estimation for energy intake and consumption of macronutrients show higher values when FFQ is used, which reinforces the conclusions of many other studies that indicate that FFQs may overestimate food consumption compared with other assessment instruments [18,19]. There are also notable differences in the process of data collection between the 2 tools for estimating the nutritional composition. Both the app and the FFQ collect the same types of food. However, the amount of products available for selection, within each category, is greater in the app. In this way, for example, in the selection of fish, the app has a greater range of products to select, distinguishing between each of them in a particular way. Although the FFQ tends to group more products into subcategories such as blue fish, white fish, or salted fish, not all fishes within the same subcategory have the same nutritional composition. In this sense, the app could be much finer in the estimation of macro- and micronutrients.

In the study conducted by Ashman et al [16], which also analyzed the results obtained with an app and a questionnaire for the estimation of micronutrient intake, moderate correlations (r=.47-.94; P<.05 in all of them) were found. However, among other information, the study subjects contributed characteristics regarding the preparation of the food consumed, which may have been influenced by subjective aspects linked to particular preferences in cooking as well as to certain customs. The population consisted of pregnant women, possibly under specific care and dietary care regimes. Moreover, previous research has shown that the 3 days, during which the app was used, would not be a sufficient period to accurately estimate micronutrient intake [20]. Regarding alcohol consumption, Ambrosini et al [6], in line with our results, also found a moderate correlation using an app and the 24-hour recall questionnaire. However, the study concluded that the app may have underestimated the consumption of alcohol compared with the 24-hour recall questionnaire. In addition, subjects only used the app for 4 days. A possible explanation of the difference between the different correlations found in the studies of Ashman and Ambrosini [6,16], for the consumption of macro- and micronutrients, and our study is that in the EVIDENT study, throughout the registration period, participants received daily notifications with recommendations to improve their dietary habits. These recommendations were part of the intervention of the EVIDENT trial and were aimed at achieving greater adherence to the Mediterranean diet and better adaptation to the recommendations for consuming macronutrients in the context of the usual diet. Another aspect that can explain the results are the different characteristics of the population studied in terms of age and health. The samples studied by Ashman [16] and Ambrosini [6] are relatively small (N=25 and N=50), with a young age average (29 and 31 years, respectively), and with a high percentage of women. However, the participants of the EVIDENT study are between 18 and 70 years of age, with an average of 52 years. In addition, the percentage of cardiovascular risk factors (hypertension, diabetes, dyslipidemia, and smoking) is slightly higher than in the general population. These circumstances highlight the heterogeneity of this sample and, therefore, the difficulty of finding more accurate correlations of nutrients between the different estimation instruments.

Limitations

One of the main limitations of this study is the traditional instrument to which the app's records were compared, an FFQ. This questionnaire was validated for the reference population (Spanish population) and uses the nutritional composition data of 137 normally consumed foods in Spain for the daily estimation. The data involved are those of the year before the interview. Other apps have used shorter periods (24 or 48 hours) and have compared their results against 24-hour recall. These studies have obtained better correlation results but these results were to be expected, given the very limited period during which intake data were collected.

Conclusions

The EVIDENT app correlates significantly with FFQ in the estimation of energy intake, macro- and micronutrients, and alcohol consumption. This correlation grows as the app's food recording period increases. The EVIDENT app can be a good alternative for gathering information on energy intake and the consumption of macronutrients, in the context of longitudinal or intervention studies.

Acknowledgments

This study was funded by the Spanish Ministry of Science and Innovation and Carlos III Health Institute/European Regional Development Fund (FIS: PI13/00618, PI13/01526, PI13/00058, PI13/01635, PI13/02528, PI12/01474; RETICS: RD12/0005, RD16/0007), Regional Health Management of Castilla and León (GRS 1191/B/15, GRS 909/B/14, GRS 770/B/13), and the Infosalud Foundation. None of the funders were involved in the design, implementation, analysis, or interpretation of the data.

Authors' Contributions

JIRR, MAGM, JGS, CRM, and LGO designed the research; JIRR, CRM, ERS, JGS, CMB, VMV, MSA, OMG, and CFA conducted the research; LGO, JAMF, and JIRR analyzed the data; JIRR, JGS, MAGM, CRM, and LGO wrote the paper; and JIRR had the primary responsibility for final content. All authors read and approved the final manuscript.

The study was approved by the clinical research ethics committee (CEIC) of the health care area of Salamanca (CEIC of Area de salud de Salamanca, June 21, 2013) as a coordinating center. It was also approved by the ethics committees of the 5 collaborating centers (CEIC of Aragón, CEIC of IDIAP Jordi Gol, CEIC of Euskadi, CEIC of the Castilla la Mancha, and CEIC of the Area de salud de Valladolid Oeste). Subjects signed informed consent forms before inclusion in the study, in accordance with the Declaration of Helsinki.

Conflicts of Interest

None declared.

References

1. Nelson M, Bingham SA. Assessment of food consumption nutrient intake. In: Margetts MB, Nelson M, editors. Design concepts in nutritional epidemiology (2nd edition). New York: Oxford University Press; 1997:123-169.
2. Illner AK, Freisling H, Boeing H, Huybrechts I, Crispim SP, Slimani N. Review and evaluation of innovative technologies for measuring diet in nutritional epidemiology. Int J Epidemiol 2012 Aug;41(4):1187-1203 [FREE Full text] [doi: 10.1093/ije/dys105] [Medline: 22933652]
3. Stumbo PJ. New technology in dietary assessment: a review of digital methods in improving food record accuracy. Proc Nutr Soc 2013 Feb;72(1):70-76. [doi: 10.1017/S0029665112002911] [Medline: 23336561]
4. Sharp DB, Allman-Farinelli M. Feasibility and validity of mobile phones to assess dietary intake. Nutrition 2014;30(11-12):1257-1266. [doi: 10.1016/j.nut.2014.02.020] [Medline: 24976425]
5. Pendergast FJ, Ridgers ND, Worsley A, McNaughton SA. Evaluation of a smartphone food diary application using objectively measured energy expenditure. Int J Behav Nutr Phys Act 2017 Dec 14;14(1):30 [FREE Full text] [doi: 10.1186/s12966-017-0488-9] [Medline: 28288657]
6. Ambrosini GL, Hurworth M, Giglia R, Trapp G, Strauss P. Feasibility of a commercial smartphone application for dietary assessment in epidemiological research and comparison with 24-h dietary recalls. Nutr J 2018 Jan 09;17(1):5 [FREE Full text] [doi: 10.1186/s12937-018-0315-4] [Medline: 29316930]
7. Carter MC, Burley VJ, Nykjaer C, Cade JE. 'My Meal Mate' (MMM): validation of the diet measures captured on a smartphone application to facilitate weight loss. Br J Nutr 2013 Feb 14;109(3):539-546. [doi: 10.1017/S0007114512001353] [Medline: 22717334]
8. Rangan AM, Tieleman L, Louie JC, Tang LM, Hebden L, Roy R, et al. Electronic Dietary Intake Assessment (e-DIA): relative validity of a mobile phone application to measure intake of food groups. Br J Nutr 2016 Jun;115(12):2219-2226. [doi: 10.1017/S0007114516001525] [Medline: 27121045]
9. Teixeira V, Voci SM, Mendes-Netto RS, da Silva DG. The relative validity of a food record using the smartphone application MyFitnessPal. Nutr Diet 2018 Apr;75(2):219-225. [doi: 10.1111/1747-0080.12401] [Medline: 29280547]
10. Recio-Rodríguez JI, Martín-Cantera C, González-Viejo N, Gómez-Arranz A, Arietaleanizbeascoa MS, Schmolling-Guinovart Y, et al. Effectiveness of a smartphone application for improving healthy lifestyles, a randomized clinical trial (EVIDENT II): study protocol. BMC Public Health 2014;14:254 [FREE Full text] [doi: 10.1186/1471-2458-14-254] [Medline: 24628961]
11. Fernández-Ballart JD, Piñol JL, Zazpe I, Corella D, Carrasco P, Toledo E, et al. Relative validity of a semi-quantitative food-frequency questionnaire in an elderly Mediterranean population of Spain. Br J Nutr 2010 Jun;103(12):1808-1816. [doi: 10.1017/S0007114509993837] [Medline: 20102675]
12. ESH/ESC Task Force for the Management of Arterial Hypertension. 2013 Practice guidelines for the management of arterial hypertension of the European Society of Hypertension (ESH) and the European Society of Cardiology (ESC): ESH/ESC Task Force for the Management of Arterial Hypertension. J Hypertens 2013 Oct;31(10):1925-1938. [doi: 10.1097/HJH.0b013e328364ca4c] [Medline: 24107724]
13. Jellinger PS, Smith DA, Mehta AE, Ganda O, Handelsman Y, Rodbard HW, AACE Task Force for Management of Dyslipidemia and Prevention of Atherosclerosis. American Association of Clinical Endocrinologists' Guidelines for Management of Dyslipidemia and Prevention of Atherosclerosis. Endocr Pract 2012;18(Suppl 1):1-78. [Medline: 22522068]

14. Chamberlain JJ, Rhinehart AS, Shaefer Jr CF, Neuman A. Diagnosis and management of diabetes: synopsis of the 2016 American Diabetes Association Standards of Medical Care in Diabetes. Ann Intern Med 2016 Apr 19;164(8):542-552. [doi: 10.7326/M15-3016] [Medline: 26928912]
15. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. J Am Med Assoc 2013 Nov 27;310(20):2191-2194. [doi: 10.1001/jama.2013.281053] [Medline: 24141714]
16. Ashman AM, Collins CE, Brown LJ, Rae KM, Rollo ME. Validation of a smartphone image-based dietary assessment method for pregnant women. Nutrients 2017 Jan 18;9(1):E73 [FREE Full text] [doi: 10.3390/nu9010073] [Medline: 28106758]
17. Bucher DT, Carrard I, Farina E, Danuser B, Kruseman M. Development and evaluation of e-CA, an electronic mobile-based food record. Nutrients 2017 Jan 18;9(1):E76 [FREE Full text] [doi: 10.3390/nu9010076] [Medline: 28106767]
18. Kowalkowska J, Slowinska MA, Slowinski D, Dlugosz A, Niedzwiedzka E, Wadolowska L. Comparison of a full food-frequency questionnaire with the three-day unweighted food records in young Polish adult women: implications for dietary assessment. Nutrients 2013 Jul 19;5(7):2747-2776 [FREE Full text] [doi: 10.3390/nu5072747] [Medline: 23877089]
19. Thompson FE, Byers T. Dietary assessment resource manual. J Nutr 1994 Nov;124(11 Suppl):2245S-2317S. [Medline: 7965210]
20. Basiotis PP, Welsh SO, Cronin FJ, Kelsay JL, Mertz W. Number of days of food intake records required to estimate individual and group nutrient intakes with defined confidence. J Nutr 1987 Sep;117(9):1638-1641. [doi: 10.1093/jn/117.9.1638] [Medline: 3655942]

Abbreviations

BMI: body mass index
FFQ: food frequency questionnaire
ICC: intraclass correlation coefficients
REDIAPP: Red Española de Investigación para Actividades Preventivas y Promoción de la Salud en Atención Primaria [Spanish Research Network for Preventive Activities and Health Promotion in Primary Care]

Edited by G Eysenbach; submitted 09.07.18; peer-reviewed by J Kay; comments to author 12.10.18; revised version received 30.10.18; accepted 22.11.18; published 08.02.19.

Please cite as:
Recio-Rodriguez JI, Rodriguez-Martin C, Gonzalez-Sanchez J, Rodriguez-Sanchez E, Martin-Borras C, Martínez-Vizcaino V, Arietaleanizbeaskoa MS, Magdalena-Gonzalez O, Fernandez-Alonso C, Maderuelo-Fernandez JA, Gomez-Marcos MA, Garcia-Ortiz L, EVIDENT Investigators
EVIDENT Smartphone App, a New Method for the Dietary Record: Comparison With a Food Frequency Questionnaire
JMIR Mhealth Uhealth 2019;7(2):e11463
URL: http://mhealth.jmir.org/2019/2/e11463/
doi:10.2196/11463
PMID:30735141

Original Paper

Evaluation of Electronic and Paper-Pen Data Capturing Tools for Data Quality in a Public Health Survey in a Health and Demographic Surveillance Site, Ethiopia: Randomized Controlled Crossover Health Care Information Technology Evaluation

Atinkut Alamirrew Zeleke[1,2], MPH; Abebaw Gebeyehu Worku[3], PhD; Adina Demissie[2], MPH; Fabian Otto-Sobotka[4], PhD; Marc Wilken[1], MSc; Myriam Lipprandt[1], PhD; Binyam Tilahun[2], PhD; Rainer Röhrig[1], MD

[1]Division of Medical Informatics, Department of Health Services Research, Carl von Ossietzky University of Oldenburg, Oldenburg, Germany

[2]Department of Health Informatics, Institute of Public Health, University of Gondar, Gondar, Ethiopia

[3]Department of Reproductive Health, Institute of Public Health, University of Gondar, Gondar, Ethiopia

[4]Division of Epidemiology and Biometry, Department of Health Services Research, Carl von Ossietzky University of Oldenburg, Oldenburg, Germany

Corresponding Author:
Atinkut Alamirrew Zeleke, MPH
Division of Medical Informatics
Department of Health Services Research
Carl von Ossietzky University of Oldenburg
Building V04-1-133
Ammerländer Heerstraße 140
Oldenburg, 26129
Germany
Phone: 49 1732587251
Email: atinkut.alamirrew.zeleke@uni-oldenburg.de

Abstract

Background: Periodic demographic health surveillance and surveys are the main sources of health information in developing countries. Conducting a survey requires extensive use of paper-pen and manual work and lengthy processes to generate the required information. Despite the rise of popularity in using electronic data collection systems to alleviate the problems, sufficient evidence is not available to support the use of electronic data capture (EDC) tools in interviewer-administered data collection processes.

Objective: This study aimed to compare data quality parameters in the data collected using mobile electronic and standard paper-based data capture tools in one of the health and demographic surveillance sites in northwest Ethiopia.

Methods: A randomized controlled crossover health care information technology evaluation was conducted from May 10, 2016, to June 3, 2016, in a demographic and surveillance site. A total of 12 interviewers, as 2 individuals (one of them with a tablet computer and the other with a paper-based questionnaire) in 6 groups were assigned in the 6 towns of the surveillance premises. Data collectors switched the data collection method based on computer-generated random order. Data were cleaned using a MySQL program and transferred to SPSS (IBM SPSS Statistics for Windows, Version 24.0) and R statistical software (R version 3.4.3, the R Foundation for Statistical Computing Platform) for analysis. Descriptive and mixed ordinal logistic analyses were employed. The qualitative interview audio record from the system users was transcribed, coded, categorized, and linked to the International Organization for Standardization 9241-part 10 dialogue principles for system usability. The usability of this open data kit–based system was assessed using quantitative System Usability Scale (SUS) and matching of qualitative data with the isometric dialogue principles.

Results: From the submitted 1246 complete records of questionnaires in each tool, 41.89% (522/1246) of the paper and pen data capture (PPDC) and 30.89% (385/1246) of the EDC tool questionnaires had one or more types of data quality errors. The overall error rates were 1.67% and 0.60% for PPDC and EDC, respectively. The chances of more errors on the PPDC tool were multiplied by 1.015 for each additional question in the interview compared with EDC. The SUS score of the data collectors was 85.6. In the qualitative data response mapping, EDC had more positive suitability of task responses with few error tolerance characteristics.

Conclusions: EDC possessed significantly better data quality and efficiency compared with PPDC, explained with fewer errors, instant data submission, and easy handling. The EDC proved to be a usable data collection tool in the rural study setting. Implementation organization needs to consider consistent power source, decent internet connection, standby technical support, and security assurance for the mobile device users for planning full-fledged implementation and integration of the system in the surveillance site.

(JMIR Mhealth Uhealth 2019;7(2):e10995) doi:10.2196/10995

KEYWORDS

public health; maternal health; surveillance; survey; data collection; data quality; tablet computer; mHealth; Ethiopia

Introduction

Scientific Background

In most of the low- and middle-income countries, millions of people born and die without registration in any legal and statistical records. Health and social policy planning and evaluations are performed with statistical assumptions for unrecorded lives [1,2]. The lack of a fully functional civil registration system in those countries force them to rely on other interim measure data sources such as censuses, Health and Demographic Surveillance Systems (HDSSs), and Demographic and Health Surveys (DHS) programs to understand population-level health determinants [3,4]. Conducting any survey in those countries requires extensive use of paper-pen and manual processes to manage the data [5]. Paper data collection processes are labor intensive, time-consuming, susceptible to errors, incur high printing and running costs, and are cumbersome and uncomfortable for field data collection [6,7]. The need to support the paper process and the recent advanced popularity of mobile devices fortified the development and use of electronic data collection methods in community health and clinical research works. Electronics devices such as personal digital assistants [8-10], mobile phones [11,12], and tablet computers are noticeably tested for their potential role in replacing the standard paper-based tools [7]. The anticipation is that electronic data capturing tools may overcome substantial limitations of the paper-based system through saving time, improving the data quality, and minimizing overall cost [6,7,9,13,14]. Due the fast evolution of information technology and the heterogenicity of infrastructures in different countries, evaluations of such systems require periodic and setting context evidence to support the growing claims on their efficiency, effectiveness, and impacts [15,16].

Rationale for the Study

The recent emerging research outputs are expediting the use of electronic data collection methods in lower- and middle-income countries. However, most of the available evidence cannot surpass the common critics on the quality of mobile health (mHealth) evidence. The critic shares 2 major points: First, there is little to no field-based study for quantifying the interaction of data quality and data capture technologies. Second, rigorous scientific research designs such as randomized controlled trials are few in number and type [17-20]. The majority of the research papers are work experience reports and simple descriptive one-arm studies [21-23]. Moreover, the comparative studies are from research conducted at a different time [7,9,11] or surveys that are not conducted on field and rather conducted in the hospital and clinic settings [10,12,24]. The presence of an interviewer adds an additional breadth to the multifaceted interaction among the survey questionnaire, the respondent, and the survey delivery approach, which can result in different approach effects. Therefore, use of smart devices such as a tablet computer to support interviewer-based survey questionnaires would be explored separately. There are experiences of using mHealth for data collection in Ethiopia [7,25,26]. According to the authors' literature review and expertise, there is a lack of comparative trial evaluation studies on the effect of data capture tools on the quality of the data in interviewer-administered surveys conducted in demographic and surveillance sites.

Therefore, in this study, we evaluated mobile electronic data capture (EDC) tools and compared them with the traditional standard paper-based system on the quality of recorded responses from rural community household respondents.

Objectives

The purpose of this study was to compare data quality parameters in the data collected using mobile electronic and standard paper-based data capture tools from Dabat Health and demographic surveillance sites in northwest Ethiopia.

The study's primary research question was:

- Is the error rate of data collected by an EDC tool less compared with paper and pen recorded data?

The secondary questions were:

- Is there a learning effect in the use of EDC through the course of the data collection period?
- Is an electronic mobile-based data collection tool usable to the data collectors in Dabat health and demographic surveillance sites?

Study Context

Organizational Setting

This evaluation research was implemented at Dabat HDSS, also called the Dabat Research Center, in northwest Ethiopia. The surveillance site is a member of the INDEPTH global network of HDSSs, which has 42 health research centers as members and 47 HDSS field sites in 18 low- and middle-income countries in Africa, Asia, and Oceania [27]. The surveillance center established in 1996 aims to generate community-based representative evidence through a continuous longitudinal data collection process. The surveillance is run by the College of Medicine and Health Sciences, which is one of the

colleges/faculties of the University of Gondar in Ethiopia. Dabat district, the surveillance site, is one of the 21 districts in North Gondar Administrative Zone of Amhara Region in Ethiopia. According to the INDEPTH Network, Dabat HDSS is working with 69,468 participants [28].

System Details and System in Use

A 30-page paper questionnaire, in 8 subthematic sections, was developed by a team of researchers at the Institute of Public Health at the University of Gondar, Ethiopia. The comprehensive survey was called "integrated survey on maternal, childhood, nutrition, and disability in Dabat Health and Demographic Surveillance System (HDSS)." The themes addressed pregnant women, young females, children aged less than 5 years, and people with disabilities. The number of questions in each subthematic questionnaire varies from 15 to 63, and each interviewer can use at least one of them or a combination of subthematic questionnaires based on the respondent's category of cases. Most of the overall items (88.25%) were closed-ended questions.

Open Data Kit Questionnaire Development

Open Data Kit (ODK) was chosen to develop the EDC system using an open-source app. ODK has an open suite of tools used to design the form, collect, and aggregate the data. Similar experience was observed in a couple of recent studies [7,29]. The exact copy of the paper questionnaire was converted to its electronic replica form using Excel and XLSForm. After passing the technical validity, the form was uploaded to the server. Each data collector can get a new blank form for the first time with an authorized log-in to the server. For multiple-choice questions, single question per screen is displayed on the tablet screen, whereas for questions in a tabular form, many questions per screen are shown on the screen. From overall questionnaire items, only the respondents' household identification (ID), personal ID, and data collectors' and their supervisor's name were required. The questionnaire is designed to filter numbers, assist date function, and check for skip patterns to prevent significant typing and data format errors. No other error controlling functionalities were incorporated because of the 2 data collectors per interview design nature of this study. The data entry fields were restricted to option buttons, check boxes, or empty fields (Figure 1).

After completing each interview, the data collector can use the "Edit Saved Form function for any valid correction and can use the "Send Finalized Form function to send the form to the server instantaneously. The third generation mobile internet network was used to connect each tablet computer to the server. In the case of limited network connectivity, submissions of the saved data were transferred when the data collectors were in a good network coverage area.

For this survey, we used the *TechnoPhantem7* tablet computer. The device is locally manufactured in Ethiopia with an inbuilt Ethiopian national working language called Amharic. A fully charged tablet has a battery life of approximately 48 hours. We have restricted the functionalities of many unnecessary apps in the devices to save the battery lifetime. To test the natural course of electricity infrastructure in the areas, we have not used extra battery or power banks as a reserve power source to charge the tablets.

Figure 1. Screenshot examples of the type of questions; multiple choices (A), number (B), single select (C), and date (D) presented in the electronic data capture tool used for the survey in Dabat Demographic and Health Surveillance site in June 2016, northwest Ethiopia. Translation of the Screen: (A) 217: Which of the signs of pregnancy complications or danger sign of pregnancy have you encountered during postnatal period? bleeding, fever, vaginal gush of fluid, incontinency, and other (describe); (B) 204: Birth weight for the last child (if it was measured); (C) 213: Where did you give birth for your last child? home, hospital, health center, health post, and other (please mention it here); and (D) 218: When was the last child delivered?

Methods

Study Design

We conducted a field-based, prospective, randomized, controlled, crossover trial to investigate the error rate as an indicator of data quality. Moreover, usability evaluation using System Usability Scale (SUS) and semistructured questionnaire (Multimedia Appendix 1) interviews were conducted to observe the user impression in use of EDC.

The evaluation study started after the approval by the Ethical Committee of the University of Oldenburg (vote-number 148/2016, chair Professor Griesinger) and the ethical vote of the University of Gondar, Research and Community Service Gondar, Ethiopia (vote-number O/V/P/RCS/05/501/2015, signed by TT Adefris). The data are kept confidentially and the electronic copy of the data is stored in the Department of Medical Informatics of the University of Oldenburg

Study Participants

Oral consent was obtained from the participants and data collectors after the objective of the research was explained briefly. Data collectors, who were permanently employed to Dabat DHS site and contractually employed for this integrated survey, were our research participants. The data collectors' main task was to travel house to house and conduct a face-to-face interview with the persons living in selected Kebeles (smallest administrative unit) of Dabat district.

From 13 towns chosen for the overall survey, 6 towns were selected for our specific research based on accessibility of internet coverage and electric power supply in the town or nearby towns. On the basis of this, 12 data collectors, who were assigned to work in 6 towns, were involved. On each survey day, the interviewer was randomly selected to have either a tablet computer or a paper questionnaire based on computer-generated orders of tool users in the study period.

Study Flow

The evaluation trial started after the technical team at the University of Gondar, Department of Health Informatics in Ethiopia managed the form creation, ODK Collect app installation on the tablet, and server configuration. Following this, we trained data collectors for 2 days on the basic tablet computer use, EDC functionalities, and questionnaire content. After incorporating comments from the participants on the functionality of the system, we performed pretests of the system for 2 days in towns outside from our trial area. The pretest phase showed some critical deficits of the system, and we have discovered unseen system errors, such as unnecessarily required items that significantly delay the data collection time. They were removed, and many decimal point items were corrected. A mini user manual using the local language was prepared and given to the interviewer along with working procedures.

The 2-interviewer pair, one with tablet computer and the other with paper questionnaire, went together to each household and sat next to each other to conduct the interviews (Multimedia Appendix 2). In a given interview, one of the data collectors leads and asks the questions, whereas both fill in the data independently and concurrently. For each interview day, the schedule of field data collectors were generated according to a computer-generated random order. The order defined the date and the data collection tool—EDC or paper and pen data capture (PPDC)—for each data collector in a team. The interviewers did not know the generated order in advance. Upon the supervisor notification, the paper and tablet devices were exchanged between the interviewers.

To evaluate the electronic tool, we used the first month of the comprehensive survey meant to address a larger population for approximately 4 months. Our actual data collection period was from May 10, 2016, to June 3, 2016. At the end of the field-based data collection period, the usability questionnaire was given to each data collector. Moreover, their satisfaction and opinion during the system use were asked by using a semistructured interview questionnaire (Multimedia Appendix 3).

Outcome Measures or Evaluation Criteria

We chose the overall data quality errors (missing and inaccuracy) as indicators of data quality. It indicates the existing, missing, or inaccuracy errors from the data collected using either paper or electronic tools. The following are the evaluation criteria:

- Overall errors: This indicates the existence of one and more error types from the total records in a questionnaire (how many of the records have one or more items with errors from the total records). For the pen and paper questionnaire, we use the original responses recorded by the data collectors before correction is given from the supervisor or data passes to cleaning and data entry phases.
- Missing: It represents missing answers or questions with no answers in questionnaire completed by the interviewer and submitted to the supervisor or to the server.
- Inaccuracy: It represents any problematic items or incompatible values in the data. It includes decimal point errors, invalid date, or text-unreadable values.
- Quantitative SUS evaluation: Response from the 10-item questionnaire based on a 5-point Likert scale was summarized. The system was considered usable if the overall SUS score was >67 (ranging from 0-100) [30].
- Qualitative isometric: The analysis is based on the International Organization for Standardization (ISO) 9241-110 dialogue principle. The 7 dialogues were based on the description in the studies [31,32]. Transcribed interview response was related to the 6 dialogue principles and summarized with a "+" and "−" notation if the interview response fits positively or negatively, respectively.

Methods for Data Acquisition and Measurement

The data collectors with paper questionnaires hand over completed questionnaires for their respective supervisors daily. To identify the potential errors in a given paper-based questionnaire, the supervisors used error extraction sheets and recorded all identified errors before giving correction, comments, and suggestions back to the data collectors. For the EDC system, the completed questionnaires were directly submitted to the server. The administrator has privileges to

access and monitors the process daily and checks submissions from all interviewers. Technical inquiries from the data collectors or supervisors are replayed immediately by standby technical members.

Electronically submitted data were downloaded on a weekly basis in Excel format from the server and sent to an actual research site where the counter paper-based questionnaire remained. Furthermore, 2 technical assistants compared the data errors recorded on the extraction sheets and electronically submitted data. Each item in the questionnaire was double checked and the errors were noted accordingly.

Anonymized data from the study site were brought to a department of the Medical Informatics at the Carl von Ossietzky University of Oldenburg. MySQL was used for the cleaning and preparation of data for analysis by 2 researchers.

To evaluate the software usability, we used the Amharic language translated SUS questionnaire [30]. The 10-item questionnaire is based on a 5-point Likert scale and produces a maximum score of 100 on the users' impression on the EDC tool. We chose SUS because of its shortness and the understandable phrasing of questions even for noninformation technology skilled persons such as our data collectors. The results of our SUS questionnaire showed a Cronbach alpha of .841 for reliability.

We also used semistructured questionnaires to perform face-to-face interviews with data collectors who were involved in this trial. Tape-recorded and noted data were then transcribed for further analysis based on the dialogue principles defined in the ISO 9241-110 [31].

Methods for Data Analysis

Univariate analyses, to quantify errors in data capture tools, was performed by SPSS (IBM SPSS Statistics for Windows, Version 24.0) and R statistical software (R version 3.4.3, The R Foundation for Statistical Computing Platform) was used to perform ordinal logistics mixed regression model to see the effect of some variables on the quality of the data. As the dataset contains 2 values for each observation, a number of errors for the tablet questionnaire and a number of errors for the paper version, there are a lot of ties or dependencies between the observations. To reduce these ties, we constructed a new variable as the difference between the 2 numbers of errors. However, this variable contains many zeros. We chose to resolve this problem by categorizing the variable into the following 3 categories: fewer errors in the paper version than in the tablet version, zero difference, and more errors in the paper version than on the tablet. We assumed a dependence between the error rates and the overall length of the questionnaire as well as differences in performance between groups of the 6 interviewers. Consequently, we constructed a mixed ordinal regression model to explain the category of error differences along the length of the questionnaire. The model contains random effects for the individual performance of groups of interviewers. At the same time, we estimated models that specialized on the influence of the different types of questions and their quantities. For every model, we used a regression equation with the following structure:

$$Error\_diff\_cat = b\_1\ covariate + a\_G1\ covariate + \ldots + a\_G6 + e \quad [1]$$

"Covariate" stands for the different measures of the length of the questionnaire and a_G1 to a_G6 denote the random effects for each group. For the estimation, we used a cumulative logistic mixed regression model. Estimation was performed within the software R using the add-on package "ordinal." We tried to include additional covariates in a forward selection process though they were eliminated on the basis of the Akaike Information Criterion. For each regression coefficient, we can construct a test for statistical significance of the influence of the corresponding covariate. We may test the hypotheses H0: $|\beta|=0$ versus H1: $|\beta|\neq 0$ by constructing a confidence interval for the level $1-\alpha$ and check whether it includes the value from H0. Hence, we can conclude whether we find a significant increase in the error rate in either of the 2 methods.

For qualitative analysis, 2 computer scientist experts and 1 public health expert with a professional informatics background independently mapped the qualitative responses to 6 of the 7 dialogue principles, and then later approved the category together. When a difference arose, a physician and medical informatics expert was consulted for an agreement.

The dialogue principles of ISO 9241-110, namely, suitability for the task, suitability of learning, controllability, self-descriptiveness, conformity with user expectation (CUE), and error tolerance (ET) is used to map the transcribed interview result. To include many of the respondent views, we have expanded the dialogue principles, mainly suitability for the tasks, focused from the software and hardware context to include social and environmental contexts.

Results

Data Collectors' Sociodemographics

A total of 1246 respondents' data were recorded through face-to-face interviews using conventional PPDC tools, whereas 1251 respondents' records were observed in the electronic database. During data cleaning, we found that 5 of the records submitted through EDC were found empty. We removed the counter of 5 records from EDC because their records could not be depicted in the paper record files using the same IDs.

A total of 12 interviewers, 4 male and 8 female, aged 21 to 32 years, participated in the study. Among these, 4 were nurses, 3 were information technologists, 1 was a person with a background in social science, and the rest of the data collectors had no specific professional background. Regarding their educational status, the majority (8/12) had completed high school and vocational school certificate or diploma. Their experience as a field worker ranged from 2 months to 7 years with a mean of 2.6 years. Moreover, 4 of the interviewers had experience of using smartphones as a personal phone. None of the interviewers had previous experience in using a smartphone or tablet computers as a data collection tool.

Unexpected Events During the Study

Although we started our survey with 6 pairs of data collector groups, unfortunately, one of the team terminated the study after

2 weeks of the survey period. The interruption occurred because of personal conflicts unrelated to the survey in the community. The conflict caused the termination of data collection in that area as the situation could not be resolved during the survey period. Therefore, we completed the survey with the remaining 5 groups.

Study Findings and Outcome Data

Errors per Questionnaire

From the complete 1246 submitted questionnaires in both PPDC and EDC tools, 522 (41.9%) of the PPDC questionnaires had 1 or more errors. From this, the majority with 175 (33.5%) of the questionnaires that had only 1 error followed by 112 (21.4%) questionnaires with 2 errors and 55 (10.5%) with 3 errors.

At EDC, 385 (30.9%) of the questionnaires had one or more errors. Thus, the majority of 241 (62.5%) had only 1 error followed by 76 (19.7%) questionnaires with 2 errors and 41 (10%) with 3 errors (Figure 2 and Multimedia Appendix 1).

Error Rate

The overall error rate, computed from the total error count over the total number of asked items, was 1.14%, from which 0.73% were missing and 0.4% of the rates were inaccuracy errors. The PPDC error rate was 1.67% (missing 0.92% and inaccuracy 0.75%), whereas the EDC error rate was 0.6%, of which 0.54% were missing and 0.064% were inaccuracy errors (Table 1).

Error Rate Over Time/Learning Effect

Though Figure 3 shows no smooth pattern of mean error rate increasing or decreasing over the study time, there is a visible difference observed between the 2 tools regarding the error rate. EDC has a constant error rate ranging below 1%, whereas the overall and PPDC error rates swing between 1% and 1.5%, respectively. There were random peaks of the error rates at different points of time in the study period; the overall error rate does not show a constant trend of decrease or increase over time (Figure 3). The statistical trend analysis also showed that trends are not detected at the $P<.005$ level.

Figure 2. Frequency of error comparison among the electronic data capture (EDC) tools using tablet computer and paper and pen data capture (PPDC) tools during a survey in the demographic survey site in 2016, Dabat, northwest Ethiopia.

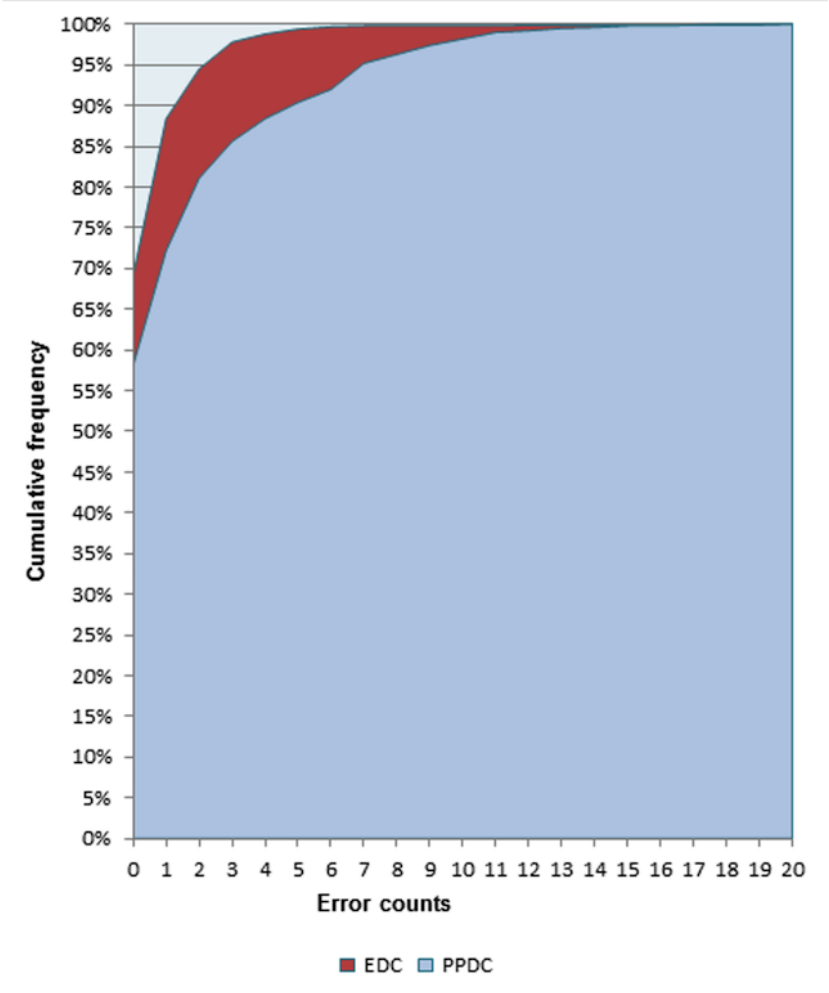


Table 1. Error rate by types of errors and the tools used during a survey in a demographic survey site in 2016, in Dabat, northwest Ethiopia.

Error type	Error type		
	Overall asked items (N=221,106), n (%)	Paper-asked items (N=110,691), n (%)	Tablet-asked items (N=110,415[a]), n (%)
Missing	1620 (0.73)	1020 (0.92)	600 (0.54)
Inaccuracy	901 (0.40)	830 (0.75)	71 (0.10)
All errors	2521 (1.14)	1850 (1.67)	671 (0.60)

[a]The item difference results from the extra asked items in the paper questionnaire because of asking items that have to be skipped accordingly. This applies to other tables, too.

Figure 3. Mean values of the overall error rates trend of electronics data capture tool (EDC) using tablet and paper and pen data capture (PPDC) tools used during the survey in the demographic survey site in 2016, Dabat, northwest Ethiopia.

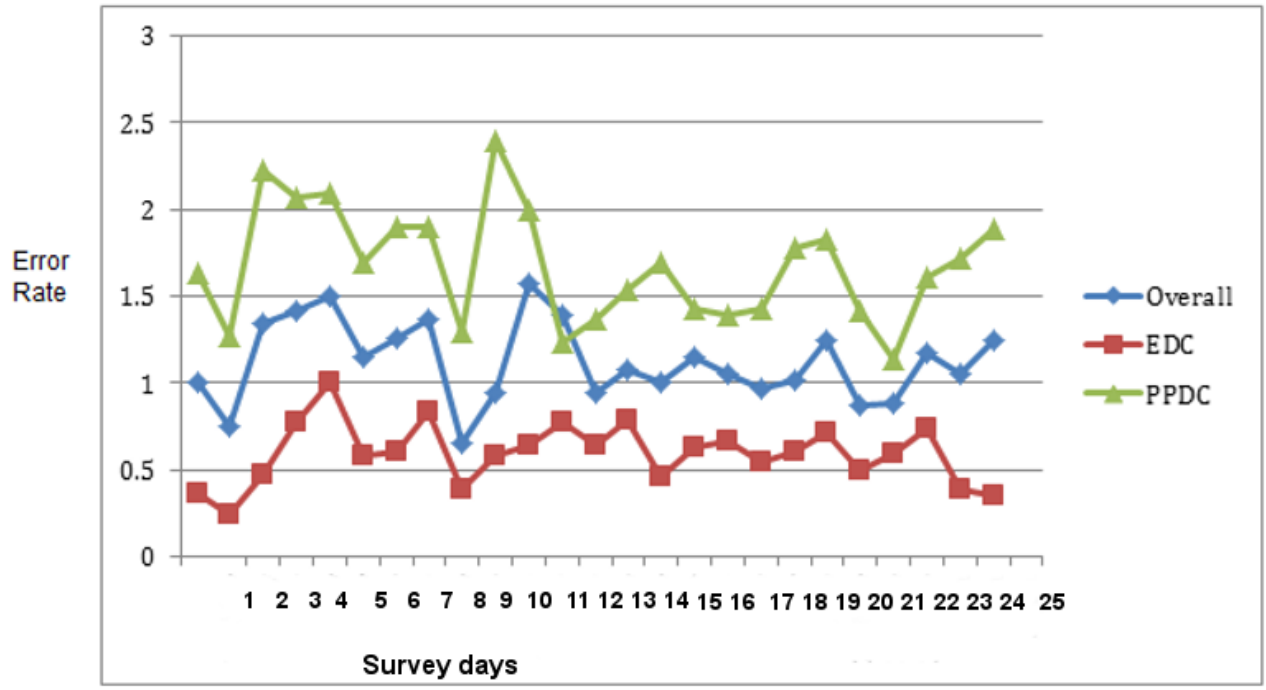


Table 2. Error by item type and the tools used during a survey in a demographic survey site in 2016, in Dabat, northwest Ethiopia.

Item types	Paper		Electronics	
	Total asked items (N=110,691), n	Number of errors (n=1850), n (%)	Total asked items (N=110,415), n	Number of errors (n=671), n (%)
Single select	46,660	466 (0.99)	46,427	145 (0.31)
Single select tabular	19,062	84 (0.44)	19,062	42 (0.22)
Multiple selects	24,303	877 (3.60)	24,302	226 (0.92)
Numbers and dates	18,022	343 (1.90)	17,980	202 (1.12)
Free text	2644	80 (3.02)	2644	56 (2.11)

Frequency of Errors per Item Type

Nearly half (47.4%) of the errors in PPDC were found in questions which had multiple option answers followed by questions with single choice answers (25.2%). At EDC, two-thirds of the errors were shared by multiple options, number and date type question with rates of 33.6% and 30.1%, respectively.

In the paper questionnaire, questions that had multiple answer options had a relatively higher rate of errors (3.65%), followed by free text and the number and date answers with 3% and 1.9%, respectively. In EDC, questions with free-text options had a relatively higher error rate with 2.1%, followed by numbers and date options answers (Table 2).

Regression Results

In each estimated model, we found that the chance for a higher number of errors in the paper version (in comparison with the tablet version) increases with the length of the overall questionnaire. For each additional question in the interview, the chances of more errors by using the paper tool were multiplied by 1.015. We have obtained similar findings for the different subsets of the questionnaire, where we found an increase in the chances of additional errors in the paper version in each type of the questions. The increase is minimal for single-item questions. The increase is strongest in the time and date questions, with an odds ratio of 1.084 per additional question. None of the 95% CIs include an estimate of 0 and an odds ratio of 1, respectively (Table 3). Hence, we can see that each variable measuring a number of items has a significant effect on the comparative error rate.

System Usability Scale Test

The usability of the system among the data collectors was measured based on SUS. The global scores can range between 0 and 100, with 100 reflecting the highest usability. The analysis yielded individual SUS scores between 67.5 and 100, and a total average score of 85.6. The graph for each SUS question depicted that majority of the data collectors' responses categorized from "agree" to "strongly agree" for positively articulated questions. Similarly, "strongly disagree" to "disagree" for the negatively articulated questions (Figures 4 and 5).

Linked Qualitative Interview Themes With the International Organization for Standardization 9241—110 Dialogue Principles

As described in the Methods section, the qualitative interview responses of the system users' perception were transcribed, thematized, and linked to one of the 6 ISO 9241-110 dialogue principles (suitability for the task, suitability of learning, controllability, self-descriptiveness, conformity with user expectation, and error tolerance). Users reported their perception by comparing the advantage of the EDC tool with respect to the disadvantage of paper-based questionnaire and vice versa.

According to the users' perception, EDC tools increased the efficiency of the data collection process through faster data collection time, the possibility of instant data submission to the research center server, and real-time work progress. Users also perceived that EDC reduces potential data quality errors with its automatic data validity checks and skip errors rules. Furthermore, the EDC device was portable, lighter, and less bulky than paper questionnaires; users claim that exposure to EDC technology improved the skills of data collectors for information technology use in the survey. On the other hand, users also perceived the following about paper questionnaire: (1) they are prone to readability errors during correction/editing of respondent answers, (2) they can become out of stock while in the remote survey area, (3) paper is heavy and easily wasted during improper handling or storage, and (4) they can incur higher duplication and operational costs for larger surveys. The perceived benefits of EDC combined with limitations of paper questionnaire was interpreted as—EDC had good suitability of task, ET, controllability, and suitability of learning for the above listed tasks.

Users also perceived that EDC possesses a high risk of losing data because of intentional and unintentional technical mistakes from the user or device failures. The fact that EDC needed sustainable electricity and internet infrastructures for charging the tablet computers and for facilitating instant data transfers was also considered as a treat to work in underserved deep rural towns. The data collectors raised a concern that the deep rural towns usually have neither of them. However, users' familiarity with the paper questionnaire, easiness of paper to correct questionnaire related mistakes, lesser risk of losing data for technical reasons, and less dependency on electricity or internet infrastructures for daily operations were referred to as the advantages of the PPDC tool. Thus, the disadvantages of EDC coupled with perceived benefits of PPDC characterize EDC to be less conformable with user expectation, controllability, ET, and suitability for the above tasks.

The views of the interviewee were also further mapped based on aspects explored during the interview, which include, efficiency/speed, data recording and entry processes, data management/operation, logistics and operation, concern, learning, infrastructure, reported and solved challenges, satisfaction, and confidence and future improvement (Table 4).

Table 3. Mixed model effect, ordinal logistic regression for the data collected by electronic data capture (EDC) and paper and pen data capture (PPDC) tools used during a survey in the demographic survey site in 2016, in Dabat, northwest Ethiopia.

Covariate	Regression coefficient b_1	Exponential(b_1) multiplicative change in chances	SD (95% CI)
Total number of items	0.015	1.015	0.0035 (0.008-0.021)
Number of single items	0.029	1.030	0.0074 (0.014-0.044)
Number of multiple items	0.042	1.043	0.0094 (0.024-0.060)
Number of single table items	0.023	1.023	0.011 (0.001-0.045)
Number of time and date items	0.081	1.084	0.017 (0.048-0.114)

Figure 4. Usability response for negatively articulated questions in the System Usability Scale (SUS) used during a survey in the demographic survey site in 2016, Dabat, northwest Ethiopia.

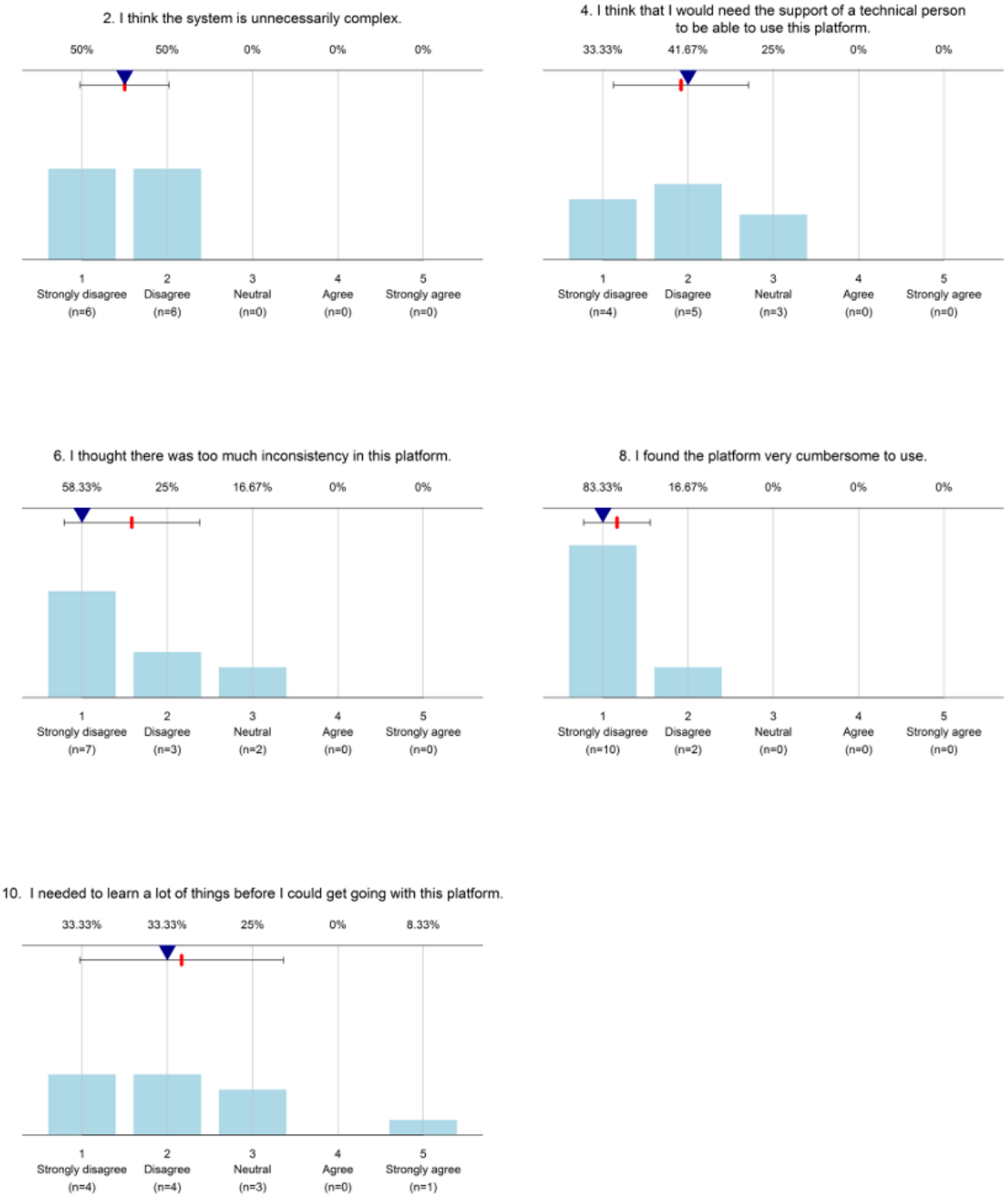


Figure 5. Usability questionnaire response for positively articulated questions in the System Usability Scale (SUS) used during a survey in a demographic survey site in 2016, Dabat, northwest Ethiopia.

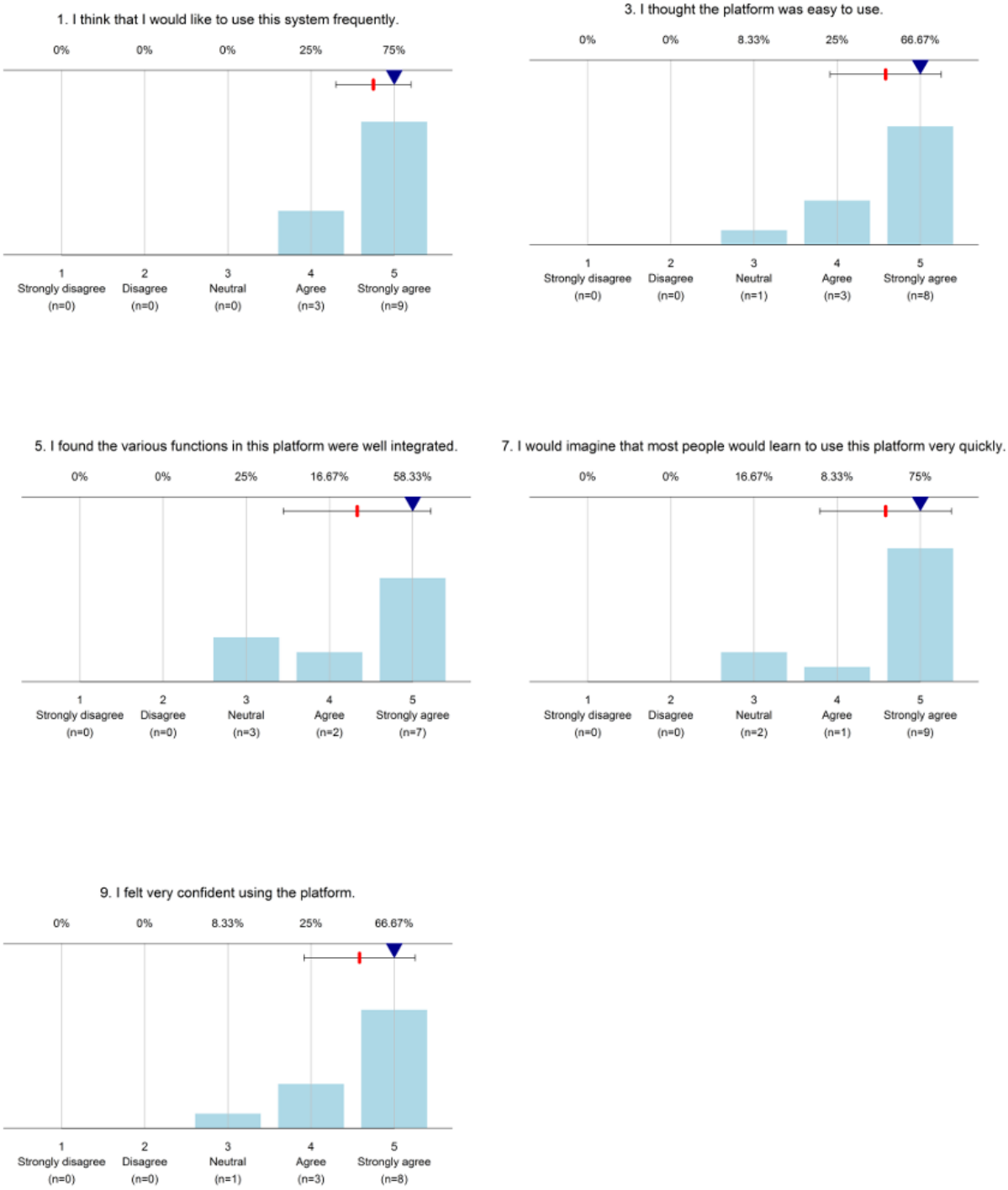


Table 4. Data recorders' perceptions and its isometric dialogue mapping for data capture tools used during a demographic survey site in 2016, in Dabat, northwest Ethiopia.

Aspect explored and data collectors' or users' perceptions	EDC[a] dialogue mapping (+/−)[b]
Efficiency or speed	
Paper questionnaire takes less time than the electronic questionnaire, more data can be collected in less time, as it is more familiar and easier to write on.	ST[c]−
If the device is free of internal software or hardware errors, working with tablet computer is relatively faster than paper, and this can save overall working time.	ST+
Data recording and entry processes	
In the paper questionnaire, deletion or editing respondent response while editing written mistakes or incorrect response can cause readability errors. In tablet computers, we can easily delete or edit responses without causing readability errors.	ET[d]+ and ST+
With the paper questionnaire, we may forget the skip question pattern and ask the follow-up questions, but in the case of using tablet device, there is an automatic skip pattern function to lead the questioning order	C[e] or SDS[f]+
In EDC, the data can reach the concerned bodies faster by jumping the double data entry clerk and other data management processes compared with the paper-based system.	ST+
Data collected with EDC has fewer errors than paper questionnaires as the electronic form has an error controlling validity mechanism.	ET+
Mistakes in questionnaire items were possible to erase or correct in EDC form, whereas in paper form, it is possible to edit or erase the items manually.	C− or ST+/−
A simple technical error in the table computer can affect the whole work. For example; in the case of unknown or unintentional error in the device, there is a risk of losing the whole data, and reinterviewing the respondent is a must.	C− and ET−
Writing on tablet virtual keyboard, moving the cursor, and editing text fields were a challenge and took a longer time. On paper questionnaire, the interviewer can easily write or circle the chosen item and pass to the next question.	C− and ST−
ODK[g] collect app menus, such as get a blank form, fill the blank form, edit saved form, and send finalized form, that were easily understandable.	ST+ and SL+
Data management or operation	
Well-documented paper questionnaires or files can stay longer and are not prone to file corruption or deletion errors as in the case of electronic data collection system.	ST− and ET
In the paper form, if mistakes were found a while after completing the interview, it can be corrected before submitting the questionnaire to the supervisor; however, in EDC, we do not have a chance to correct later as the data are sent instantly to the server while we are in the study field.	ST+
There is a high possibility of data loss in paper questionnaires because sheets of paper suffer from wear and tear during transit and storage.	ST+
Actual work progress of the data collectors was better monitored with daily and instant submission of the completed questionnaire through EDC.	ST+
Logistics and operation	
The paper questionnaire runs out while we are in the field and sometimes it takes longer to reach, which is not the case in EDC.	ST+
The tablet computer is small and easy to handle, whereas the paper questionnaire pack was heavy and bulky. Carrying paper packs house to house for longer distance was difficult and had physical stress.	ST+
Security concern: Data collectors feel insecure to work in rural areas alone with this expensive device. Furthermore, the device or the data recorders might be forcefully stolen or may be attacked in case of theft attempt.	ST−
Learning opportunity: Data collector's skill of tablet computer utilization was increasing day to day as familiarity with the system was growing.	SL[h]+
Infrastructure	
The battery capacity lasts a maximum of a one and a half days; during the study period, overnight charging of the tablet computer in or around the research town was possible.	ST+
Data collectors appreciated the role of paper-based system when their work was interrupted because their tablet battery was low or switched off or when the mobile data connection was interrupted.	ET−
Internet collection was adequate to send the data once a day during the study period.	ST+

Aspect explored and data collectors' or users' perceptions	EDC[a] dialogue mapping (+/−)[b]
Sometimes in deep rural areas, electricity may not be available to charge the battery or adequate enough for the internet network coverage to be available.	ST−
Reported and solved challenges	
Downloading a new form at the beginning of the study or during form updating was challenging.	SDS and C−
Deleting partially filled data was challenging.	C −
Some questions in EDC form were refusing to accept the decimal number during data collection in the first 2 days.	CUE[i]−
Unintentional and accidental touching of some settings in the table computer changed the screen resolution settings, the date and time, and interrupted communication with the server.	C−
Mobile internet data balance was exhausted while the data collectors were in the remote area where charging the balance was not possible.	ST−
Satisfaction and confidence: The data collectors have better satisfaction with EDC and have confidence to work using this device in future surveys.	ST+
Future improvement	
The questionnaire details such as point, number, and legal value, should be well refined in EDC before implementation.	CUE−
As frequent phone consultation for technical help may not be possible in case of network problems, strong training is needed to troubleshoot simple errors by EDC users without technical team consultation.	CUE−UE

[a]EDC: electronic data capture.

[b]+ corresponds to more and − corresponds to less.

[c]ST: suitability of task.

[d]ET: error tolerance.

[e]C: controllability.

[f]SDS: self-descriptiveness.

[g]ODK: open data kit.

[h]SL: suitability for learning.

[i]CUE: conformity with user expectation.

Unexpected Observations

Most of the household face-to-face interviews went smoothly except for a few incidences such as loss of internet connectivity, unexpected electric power interruption in the area, and software bugs. Recording geographic coordinates were part of the EDC questionnaire and was a required input. Detecting the global positioning system (GPS) signals was sometimes a challenge for our 3 survey sites. This required option had created a substantial delay in the data collection process for 3 groups, as further questions could not be asked without detecting the GPS signals. As a quick solution, we removed the required option from the app to resume data collection processes.

Discussion

Answer to the Study Questions

This study aimed to evaluate and compare the magnitude of data quality errors in the data collected with paper and pen as well as EDC technologies. In addition, the usability of the tablet computer to capture survey respondent data was analyzed from the perspective of data collectors. The answer for these study questions are arranged in magnitude of errors, learning effect, quantitative SUS evaluation, and isometric qualitative evaluation for system usability.

Magnitude of Data Quality Errors

In DHS, data collection systems are expected to be sustainable, easy to use, provide timely data, be reliable, and maintain data quality. Our evaluation study demonstrated that a smart device data collection system using the ODK platform outperformed pen-and-paper systems across each data quality parameter for DHS in Ethiopia. Data collected using tablet computers were more likely to have fewer errors compared with the conventional paper questionnaire. Nearly half of the PPDC and one-third of EDC questionnaires exhibited one or more errors in the respondent answers.

The data collection mode features can affect response quality. Face-to-face or personal-visit surveys have a high degree of interaction with the respondent irrespective of the technology use. Interviewers can also support in clarification, probing, and encourage the respondents to provide complete and accurate responses [33]. The conceptual reasons why errors might be less likely to occur when using technology, particularly when human-technology interaction is inevitable, need further explanation. Data quality in general possesses completeness, plausibility, and validity dimensions. Addition of technological support for face-to-face interview mode targeted on the improvement of the completeness and plausibility data quality dimensions. EDC timely avoids completeness and plausibility errors through automated routing in complex questionnaires

and detecting errors using range and consistency checks and warning interviewers of invalid entries. Though improving completeness and plausibility contribute to validity dimension, further research is needed to assess the validity dimension. In addition to improving the design of the instrument with principles of human-computer interaction and user-centered design, appropriate training on error recovery procedures to avoid potentially deleterious consequences (eg, accidental deletion of files) is important [33,34].

It is a very challenging task to find a theoretical base for the superiority of technological support. We can think of the Charles Friedman's fundamental theorem for biomedical informatics. It states that "a person working in partnership with an information resource is 'better' than that same person unassisted." The literal translation to us is "a field-based interviewer who use electronics data collection system for the survey will make fewer data quality errors than the interview administrator without the electronic system." Though full consensus has not been reached on interpreting the theorem, Friedman's theorem is the most arguable applied theorem in the informatics community [35].

Computing error rate, which considers the total number of questions, is important in case of bulky questionnaires such as ours. For example, the error rates for PPDC and EDC tools were 1.67% and 0.6%, respectively. This shows that most of the submitted questionnaires in both tools had very few error counts compared with the total number of asked items in a questionnaire. The majority, that is, 85% of EDC and 55% of PPDC had less than 3 errors from 38 questions in the average questionnaire.

The low rate might be from 2 possible speculated effects; the first is "being observed or supervised effect on the data collectors and the second might be "inter-tool/interaction among the data collectors. The data collectors feel that their performance can be watched indirectly and making mistakes can have an unintended consequence in their job. Extra care might have been taken by the data collectors to avoid potential mistakes while working. The intertool interaction between the paired data collectors might affect the performance of one of the data collector over the other pair. Though we have strictly instructed the interviewers to avoid any side communication during the interview, they might communicate indirectly in the sense of helping each other during technical challenges. This unwanted communication might neutralize the data quality of impact of a tool over the other. Further analysis of paired datasets in the same team supported this claim, where 3 of the 6 paired groups have similar magnitudes of errors between the paired individuals. The effects were unavoidable because of our study design nature where the paired data collectors must conduct the interview.

Missing and inaccuracy errors were analyzed separately, and the percentages indicate both types of errors were more prevalent in the paper-based system. Missing errors are more prevalent than inaccuracy errors in both tools. However, inaccuracy errors in electronic-based data are quite low, 10.5% (71/671), compared with a paper questionnaire, 81.37% (830/1020). This might be because electronic form had partial programmed checking mechanisms such as skip pattern. On the other hand, from the total errors count (671) observed in electronic data, 600 (89.5%) errors were because of missing items. This accounts for 37% of the total errors in both tools. Considering previous experience and the nature of electronic device for minimal tolerance of missing values, the magnitude seems quite high. This is expected because of the study design nature. In a research design such as ours where 2 data collectors sit next to each other, they start and finish the study at approximately the same time, implementing all possible validation rules and complicate the study design. For example, in case the person with the tablet is leading the interview and the person with the paper follows or vice versa, one of them does not have to wait for the other for any mismatch reasons. Implementing strict validation rules will create interdependency or confusion among them. This has been observed in our pretest phase, and hence, we have deliberately omitted some of the validation rules from the EDC.

The possession of more errors in paper compared with EDC were also supported by our qualitative interview data. The majority of the data collectors perceived EDC had minimized the potential errors through inbuilt error controlling validity tool and systematic routing with skip logic.

Learning Effect

A learning curve usually describes technological progress (measured generally in terms of decreasing error rate for EDC) as a function of accumulating experience with EDC over time. It assumes that a technology's performance improves with experience when the technology accumulates. In this study, a single data collector had a cumulative 2 weeks exposure in the 4-week study period. In this short time exposure, performing early assessment is open to criticism as the users are not fully proficient. Late assessment runs the risk that the data collectors will draw their own conclusions about the pros and cons of the new technology. In our case, both the statistical tests and graphical illustrations did not depict any meaningful time pattern on the error rate through the study period. This might be because the interviewers were exposed equally to each tool at random order to avoid any unnecessary learning effect on the quality of the data, which can be sourced from uneven prolonged exposure of the user to one of the tools.

System Usability Scale and Isometric Usability

SUS scores and impression of the data collectors from the interview showed high usability and satisfaction in using EDC for data collection with a usability score of 85.4. Data collectors stated that the system is easy to carry, they were able to directly submit the data to the server, and let them be familiar with the information technology; this encourages the implementation of such systems in resource-limiting settings.

Analysis of the qualitative interview based on the isometric dialog principles for usability identified perceived strengths and areas of improvements to implement the electronic data collection system. For example, the system had better ET because of its error controlling mechanism and less ET for its high risk to lose the data.

The qualitative responses were triangulated with isometric dialogue principles. The conceptual focus of these dialogues is

basically about the hardware and software capability system for the given task. However, we tried to force some of the dialogue principles to accommodate environmental and social contexts that are of particular importance to the research sites settings. For example, responses such as, "we had a constant supply of electricity and internet," "I work relatively faster when I use a tablet computer," and "When I use a tablet, I will not worry about finishing the questionnaire" were categorized under suitability for task dialogue principle. The responses are neither software interface issues nor hardware issues, rather, they are social, economic, and infrastructural setting context to facilitate the data collectors' tasks. This may imply that there could be room to extend the dialogue principles on the importance of local context issues where the system evaluation takes place. Though mapping the qualitative interview cannot explicitly show all the dialogue principles into the level of equal share, relevant information can be drawn even with its limitations. For the future, a qualitative research design where the interview questions intended to include all the 6 dialogue principles can yield a better understanding of the system usability.

Tablet computers are valuable items in resource-limited settings; thus, the device can be stolen or data collectors may be targeted when conducting household surveys in insecure areas. The potential incidence of theft and subsequent physical attacks or fear of losing the device were frequently mentioned concerns by the data collectors. Similar concerns were raised by other researchers in South Africa; such types of insecure feelings may create unintended resistance from the field interviewer. If the system planned for scaled-up implementation in the routine work, addressing their concern is an important step for successful integration of the system in the area.

Strengths and Weaknesses of the Study

This study adopted an existing open source platform to develop the electronic questionnaire and data submission and storage configuration. This showed we could easily adopt the available resources and save the human and material resources to develop the new system from scratch.

The research design in this study is a crossover, randomized trial, which helps to eliminate unnecessary learning trends and shows the real effect of the devices accordingly.

The other possible strength of this study is the fact that the survey showed the usability of the system using the quantitative (SUS) and the qualitative interview supported by isometric dialogues principles.

Pairing data collectors for a single interview may affect the interdata collectors' performance or tools may affect each other, hence lowering the magnitude of actual errors. Due to the inability to implement all the functions of the electronic error controlling mechanism, such as required items for all, value range determination in EDC might have contributed to some preventable errors that existed in EDC.

Though interviewer effects appear not to affect most survey items, literatures showed that interviewer characteristics effects such as race, ethnicity, and gender can nonetheless occur in all interviewer-administered survey modes and can affect survey findings. If these effects are replicated because of use of small number of interviews, it might result in a bias. For our study, data collectors were randomly exposed to either an electronic device or a paper questionnaire. Though the existence of bias is unavoidable, we believe the effect is similar in both tools [36].

Results in Relation to Other Studies

The overall error rate in this study, 1.67% for PPDC tool and 0.6% for EDC is comparable with a study in Kenya that had 1% and 0.1% missing rate in paper and electronics, respectively [11]. In addition, a study with a similar design in India revealed 2% PPDC and 1.99% EDC error rates [29]. However, this magnitude is lower compared with other similar studies (7% and 1%, with no omission in EDC) [9], 3.6% PPDC and 5.2% in EDC [37]. The difference might be an item selection mechanism: the above studies calculated their error rate from selected items in a given questionnaire, whereas this study considered all items in the questionnaire, which is highly likely to inflate nonerror denominators for such types of typical huge number of asked items (over 220,000). Furthermore, random assignation of data captures tools for data collectors, whereas the studies compare data collected from 2 different points in time by various data collectors and nonrandomized methods. Nonrandomized studies might possess inflated error influenced by nonhomogeneous factors distributions among data collectors and other characteristics of the study.

Many EDC studies that use the maximum potential of error control mechanisms such as required items for all, logical skip, or logical input values in EDC, reported 0 to less than 1% of missing responses [14,38]. In our study, because of the design nature mentioned earlier, we could not fully implement all the error-controlling mechanisms in EDC. Under possible circumstances, implementing the maximum potential of the error preventing mechanism in the software is mandatory.

Our study and other related findings revealed there is a statistically significant difference in the chance of having more errors in paper-based records than in electronic data records [7,9,11]. However, an equivalent research report from rural India claims that there was no statistical difference in error magnitude between the 2 tools, where EDC is as accurate as PPDC [29]. In our study, the chance of having more errors increased as the number of questions asked increased. Our study was unique in depicting this finding, because of the nature of the survey. In this survey, all the household respondents were not necessarily given the same type of questionnaires and number of questions. The number of questions and the type of the questionnaire depends on the type of public health cases that existed in the household members interviewed by the data collectors. The implication might be the length of the questionnaire, and this might be an important factor to be considered in the decision of choosing the type of data capture tools during design. Bulky questionnaires might be more suitable for electronic records than a paper-based system.

The SUS score value 85.4 and the qualitative interview isometric dialogue mapping shows the system users, in our case data collectors, use the system comfortably. These findings were shared by other EDC users' impressions in related researchers [11,26,29,37,38]. At ask-technology fit study claims that users

understand their system as tools helping or hindering them from performing their tasks. Users react positively to system features that appreciate their mission-related demands [39].

Meaning and Generalizability of the Study

Our study findings are in line with similar studies that used the ODK app platform to develop the electronic form and the data management settings and implemented in rural and urban community household surveys, though there is a methodological difference. The quantitative and qualitative analysis depicted that EDC is a usable and preferable data capture tool for the field-based survey.

We believe the study might be generalized to rural community setting research sites, with limited internet and electricity access.

Unanswered and New Questions

Working in a pair might put an "intersupervisor effect on the data collector's performance. In a further study, we can assign a data collector independently to work with EDC and avoid the intersupervisor effect. Moreover, measuring the learning effect of the technology after longer exposure with the system and using appropriate statistical techniques might give a different result.

Moreover, research considering organizational readiness evaluation, working culture alteration, and its outcome measurement in using this system in resource-limited settings are tasks on the table. Further research is also necessary to validate qualitative-based isometric usability questions in different linguistic and communication culture settings.

Conclusions

The objectivist approaches of this study conclude that data collected by using an electronic-based data collection system had a significantly better quality compared with a paper-pen data collection system explained by fewer errors. Implementation of electronic data collection tools such as the ones tested in this comparative trial were found to be usable by data collectors in the rural resource-limited settings. Implementation of a full error-controlling function exhaustively, setting a standby technical and monitoring team, and assuring security concerns on the device will contribute to better implementation of the electronic data collection system in the resource limited-settings. Stakeholders of the health information system particularly in a demographic and surveillance site can adapt and use the existing open source mobile device platforms in their routine data collection and management practices.

Acknowledgments

The authors would like to thank the University of Gondar for funding the research expenses. The authors would also like to thank Deutscher Akademischer Austauschdienst for partially funding the research. They would also like to acknowledge Temesgen Azemeraw (Data manager in Dabat Health and Demographic Surveillance System) for his technical contribution in open data kit form development. The authors also thank Brendan Whelan (Dublin) and Stefanie Gacek as well as Lily and Udo Lauter for copyediting the manuscript.

Authors' Contributions

AAZ conceptualized and designed this evaluation study. He participated in the development of the electronic data capture (EDC) tool and performed data collection. He analyzed the data and prepared the first draft of the manuscript. AGW participated in the development of the general proposal, supervised the data collection, and revised the draft manuscript. AD participated in the data collection, the analysis, and revised the draft manuscript. FOS participated in the data analysis and revised the draft manuscript. ML and MW participated in the data analysis and categorization of the interviews and revised the draft of the manuscript. BT initiated and conceptualized the overall project, the design and the development of the EDC tool, participated in the data collection, and analyzed and revised the draft. He led the overall project. RR supervised the study design, data collection and analysis, manuscript writing, and revised the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary table for the frequency of error in electronics data capture and the paper data capture tools.
[PDF File (Adobe PDF File), 22 KB - mhealth_v7i2e10995_app1.pdf]

Multimedia Appendix 2
Sample picture of data collection setting where the 2 data collectors with tablet computer and paper questionnaire conduct the survey simultaneously.
[PNG File , 272 KB - mhealth_v7i2e10995_app2.png]

Multimedia Appendix 3
A semistructured questionnaire used to interview the system users.

[PDF File (Adobe PDF File), 39 KB - mhealth_v7i2e10995_app3.pdf]

References

1. Setel PW, Macfarlane SB, Szreter S, Mikkelsen L, Jha P, Stout S, Monitoring of Vital Events. A scandal of invisibility: making everyone count by counting everyone. Lancet 2007 Nov 03;370(9598):1569-1577. [doi: 10.1016/S0140-6736(07)61307-5] [Medline: 17992727]
2. Phillips DE, AbouZahr C, Lopez AD, Mikkelsen L, de Savigny D, Lozano R, et al. Are well functioning civil registration and vital statistics systems associated with better health outcomes? Lancet 2015 Oct 03;386(10001):1386-1394. [doi: 10.1016/S0140-6736(15)60172-6] [Medline: 25971222]
3. Vital Events Registration Agency. Vital Events Registration Agency. 2014. CRVS Indicative Investment Plan for Ethiopia URL: https://www.getinthepicture.org/sites/default/files/resources/eth_crvs_investment_plan_2014.pdf [WebCite Cache ID 73DCv5dBz]
4. Ethiopia Federal Ministry Of Health. Health Finance and Governance Project. 2014. Ethiopia's Fifth National Health Accounts 2010/2011 URL: https://www.hfgproject.org/wp-content/uploads/2014/04/Ethiopia-Main-NHA-Report.pdf [accessed 2018-10-16] [WebCite Cache ID 73DCVBr3T]
5. Central Statistical Agency [Ethiopia] and ICF International. The DHS Program. 2012. Ethiopia Demographic and Health Survey 2011 URL: https://dhsprogram.com/pubs/pdf/fr255/fr255.pdf [accessed 2018-10-02] [WebCite Cache ID 72rdJXw7J]
6. Okwaraji YB, Cousens S, Berhane Y, Mulholland K, Edmond K. Effect of geographical access to health facilities on child mortality in rural Ethiopia: a community based cross sectional study. PLoS One 2012;7(3):e33564 [FREE Full text] [doi: 10.1371/journal.pone.0033564] [Medline: 22428070]
7. King JD, Buolamwini J, Cromwell EA, Panfel A, Teferi T, Zerihun M, et al. A novel electronic data collection system for large-scale surveys of neglected tropical diseases. PLoS One 2013;8(9):e74570 [FREE Full text] [doi: 10.1371/journal.pone.0074570] [Medline: 24066147]
8. Fletcher LA, Erickson DJ, Toomey TL, Wagenaar AC. Handheld computers. A feasible alternative to paper forms for field data collection. Eval Rev 2003 Apr;27(2):165-178. [doi: 10.1177/0193841X02250527] [Medline: 12703341]
9. Thriemer K, Ley B, Ame SM, Puri MK, Hashim R, Chang NY, et al. Replacing paper data collection forms with electronic data entry in the field: findings from a study of community-acquired bloodstream infections in Pemba, Zanzibar. BMC Res Notes 2012;5:113 [FREE Full text] [doi: 10.1186/1756-0500-5-113] [Medline: 22353420]
10. Koop A, Mösges R. The use of handheld computers in clinical trials. Control Clin Trials 2002 Oct;23(5):469-480. [Medline: 12392861]
11. Njuguna HN, Caselton DL, Arunga GO, Emukule GO, Kinyanjui DK, Kalani RM, et al. A comparison of smartphones to paper-based questionnaires for routine influenza sentinel surveillance, Kenya, 2011-2012. BMC Med Inform Decis Mak 2014;14:107 [FREE Full text] [doi: 10.1186/s12911-014-0107-5] [Medline: 25539745]
12. Ganesan M, Suma P, Mary VP, Janakiraman N, Jhunjhunwala A, Waidyanatha N. The use of mobile phone as a tool for capturing patient data in southern rural Tamil Nadu, India. JHIDC 2011;5(2) [FREE Full text]
13. Ali M, Deen J, Khatib A, Enwere G, von Seidlein L, Reyburn R, et al. Paperless registration during survey enumerations and large oral cholera mass vaccination in Zanzibar, the United Republic of Tanzania. Bull World Health Organ 2010 Jul 01;88(7):556-559. [doi: 10.2471/blt.09.070334]
14. Zhang S, Wu Q, van Velthoven MH, Chen L, Car J, Rudan I, et al. Smartphone versus pen-and-paper data collection of infant feeding practices in rural China. J Med Internet Res 2012;14(5):e119 [FREE Full text] [doi: 10.2196/jmir.2183] [Medline: 22989894]
15. Luna D, Almerares A, Mayan JC, de Quirós FG, Otero C. Health informatics in developing countries: going beyond pilot practices to sustainable implementations: a review of the current challenges. Healthc Inform Res 2014 Jan;20(1):3-10 [FREE Full text] [doi: 10.4258/hir.2014.20.1.3] [Medline: 24627813]
16. Nykänen P, Brender J, Talmon J, de Keizer N, Rigby M, Beuscart-Zephir M, et al. Guideline for good evaluation practice in health informatics (GEP-HI). Int J Med Inform 2011 Dec;80(12):815-827. [doi: 10.1016/j.ijmedinf.2011.08.004] [Medline: 21920809]
17. Labrique AB, Vasudevan L, Kochi E, Fabricant R, Mehl G. mHealth innovations as health system strengthening tools: 12 common applications and a visual framework. Glob Health Sci Pract 2013 Aug;1(2):160-171 [FREE Full text] [doi: 10.9745/GHSP-D-13-00031] [Medline: 25276529]
18. Aranda-Jan CB, Mohutsiwa-Dibe N, Loukanova S. Systematic review on what works, what does not work and why of implementation of mobile health (mHealth) projects in Africa. BMC Public Health 2014;14:188 [FREE Full text] [doi: 10.1186/1471-2458-14-188] [Medline: 24555733]
19. Braun R, Catalani C, Wimbush J, Israelski D. Community health workers and mobile technology: a systematic review of the literature. PLoS One 2013;8(6):e65772 [FREE Full text] [doi: 10.1371/journal.pone.0065772] [Medline: 23776544]
20. Tomlinson M, Rotheram-Borus MJ, Swartz L, Tsai AC. Scaling up mHealth: where is the evidence? PLoS Med 2013;10(2):e1001382 [FREE Full text] [doi: 10.1371/journal.pmed.1001382] [Medline: 23424286]

21. Tomlinson M, Solomon W, Singh Y, Doherty T, Chopra M, Ijumba P, et al. The use of mobile phones as a data collection tool: a report from a household survey in South Africa. BMC Med Inform Decis Mak 2009;9:51 [FREE Full text] [doi: 10.1186/1472-6947-9-51] [Medline: 20030813]
22. Performance Monitoring and Accountability 2020. 2014. Detailed Indicator Report: Ethiopia 2014 URL: https://www.pma2020.org/sites/default/files/DIR-PMA2014-ET-V4-2015-04-14.pdf [accessed 2018-10-16] [WebCite Cache ID 73D0JYdn7]
23. Rajput ZA, Mbugua S, Amadi D, Chepngeno V, Saleem JJ, Anokwa Y, et al. Evaluation of an Android-based mHealth system for population surveillance in developing countries. J Am Med Inform Assoc 2012;19(4):655-659 [FREE Full text] [doi: 10.1136/amiajnl-2011-000476] [Medline: 22366295]
24. Diero L, Rotich JK, Bii J, Mamlin BW, Einterz RM, Kalamai IZ, et al. A computer-based medical record system and personal digital assistants to assess and follow patients with respiratory tract infections visiting a rural Kenyan health centre. BMC Med Inform Decis Mak 2006 Apr 10;6:21 [FREE Full text] [doi: 10.1186/1472-6947-6-21] [Medline: 16606466]
25. Little A, Medhanyie A, Yebyo H, Spigt M, Dinant GJ, Blanco R. Meeting community health worker needs for maternal health care service delivery using appropriate mobile technologies in Ethiopia. PLoS One 2013;8(10):e77563 [FREE Full text] [doi: 10.1371/journal.pone.0077563] [Medline: 24204872]
26. Medhanyie AA, Moser A, Spigt M, Yebyo H, Little A, Dinant G, et al. Mobile health data collection at primary health care in Ethiopia: a feasible challenge. J Clin Epidemiol 2015 Jan;68(1):80-86. [doi: 10.1016/j.jclinepi.2014.09.006] [Medline: 25441699]
27. INDEPTH Network. URL: http://www.indepth-network.org/about-us [accessed 2018-10-16] [WebCite Cache ID 73D8pXIgn]
28. INDEPTH Network. Dabat HDSS - University of Gondar URL: http://www.indepth-network.org/member-centres/dabat-hdss [accessed 2018-10-16] [WebCite Cache ID 73D87waXy]
29. Giduthuri JG, Maire N, Joseph S, Kudale A, Schaetti C, Sundaram N, et al. Developing and validating a tablet version of an illness explanatory model interview for a public health survey in Pune, India. PLoS One 2014;9(9):e107374 [FREE Full text] [doi: 10.1371/journal.pone.0107374] [Medline: 25233212]
30. Brooke J. SUS-A quick and dirty usability scale. In: Jordan PW, Thomas B, McClelland IL, Weerdmeester B, editors. Usability evaluation in industry. London, Great Britain: Taylor & Francis; 1996.
31. Kai-Christoph H, Brigitte V, Hans-Bernd B. Questionnaire based usability evaluation of hospital information systems. EJISE 2004;7(1):21-30 [FREE Full text]
32. Safdari R, Dargahi H, Shahmoradi L, Farzaneh NA. Comparing four softwares based on ISO 9241 part 10. J Med Syst 2012 Oct;36(5):2787-2793 [FREE Full text] [doi: 10.1007/s10916-011-9755-5] [Medline: 21755414]
33. Couper MP, Hansen SE, Sadosky SA. Evaluating Interviewer Use of CAPI Technology. In: Lyberg L, Biemer P, Collins M, de Leeuw E, Dippo C, Schwarz N, et al, editors. Survey Measurement and Process Quality. New Jersey: John Wiley & Sons, Inc; 2012:267-285.
34. Robling MR, Ingledew DK, Greene G, Sayers A, Shaw C, Sander L, et al. Applying an extended theoretical framework for data collection mode to health services research. BMC Health Serv Res 2010 Jun 24;10:180 [FREE Full text] [doi: 10.1186/1472-6963-10-180] [Medline: 20576131]
35. Friedman CP. A "fundamental theorem" of biomedical informatics. J Am Med Inform Assoc 2009;16(2):169-170 [FREE Full text] [doi: 10.1197/jamia.M3092] [Medline: 19074294]
36. Davis RE, Couper MP, Janz NK, Caldwell CH, Resnicow K. Interviewer effects in public health surveys. Health Educ Res 2010 Feb;25(1):14-26 [FREE Full text] [doi: 10.1093/her/cyp046] [Medline: 19762354]
37. Walther B, Hossin S, Townend J, Abernethy N, Parker D, Jeffries D. Comparison of electronic data capture (EDC) with the standard data capture method for clinical trial data. PLoS One 2011;6(9):e25348 [FREE Full text] [doi: 10.1371/journal.pone.0025348] [Medline: 21966505]
38. Fritz F, Balhorn S, Riek M, Breil B, Dugas M. Qualitative and quantitative evaluation of EHR-integrated mobile patient questionnaires regarding usability and cost-efficiency. Int J Med Inform 2012 May;81(5):303-313. [doi: 10.1016/j.ijmedinf.2011.12.008] [Medline: 22236957]
39. Goodhue DL. Understanding user evaluations of information systems. Manage Sci 1995 Dec;41(12):1827-1844. [doi: 10.1287/mnsc.41.12.1827]

Abbreviations

C: controllability
CUE: conformity with user expectation
DHS: Demographic and Health Surveys
EDC: electronic data capture
ET: error tolerance
GPS: global positioning system
HDSS: Health and Demographic Surveillance System
ID: identification

ISO: The International Organization for Standardization
mHealth: mobile health
ODK: open data kit
PPDC: paper and pen data capture
SDS: self-descriptiveness
SL: suitability for learning
ST: suitability of task
SUS: System Usability Scale

Edited by G Eysenbach; submitted 08.05.18; peer-reviewed by C Basch, W Kernan; comments to author 03.09.18; revised version received 25.10.18; accepted 26.10.18; published 11.02.19.

Please cite as:
Zeleke AA, Worku AG, Demissie A, Otto-Sobotka F, Wilken M, Lipprandt M, Tilahun B, Röhrig R
Evaluation of Electronic and Paper-Pen Data Capturing Tools for Data Quality in a Public Health Survey in a Health and Demographic Surveillance Site, Ethiopia: Randomized Controlled Crossover Health Care Information Technology Evaluation
JMIR Mhealth Uhealth 2019;7(2):e10995
URL: http://mhealth.jmir.org/2019/2/e10995/
doi:10.2196/10995
PMID:30741642

Original Paper

Using Twitter to Detect Psychological Characteristics of Self-Identified Persons With Autism Spectrum Disorder: A Feasibility Study

Yulin Hswen[1,2*], MPH; Anuraag Gopaluni[3*], BA; John S Brownstein[2,4,5], PhD; Jared B Hawkins[2,4], MMSc, PhD

[1]Department of Social and Behavioral Sciences, Harvard TH Chan School of Public Health, Boston, MA, United States

[2]Computational Health Informatics Program, Boston Children's Hospital, Boston, MA, United States

[3]Department of Mathematics and Statistics, Boston University, Boston, MA, United States

[4]Department of Pediatrics, Harvard Medical School, Boston, MA, United States

[5]Department of Biomedical Informatics, Harvard Medical School, Boston, MA, United States

[*]these authors contributed equally

Corresponding Author:
Yulin Hswen, MPH
Department of Social and Behavioral Sciences
Harvard TH Chan School of Public Health
677 Huntington Ave
Boston, MA,
United States
Phone: 1 617 775 1889
Fax: 1 617 730 0267
Email: yuh958@mail.harvard.edu

Abstract

Background: More than 3.5 million Americans live with autism spectrum disorder (ASD). Major challenges persist in diagnosing ASD as no medical test exists to diagnose this disorder. Digital phenotyping holds promise to guide in the clinical diagnoses and screening of ASD.

Objective: This study aims to explore the feasibility of using the Web-based social media platform Twitter to detect psychological and behavioral characteristics of self-identified persons with ASD.

Methods: Data from Twitter were retrieved from 152 self-identified users with ASD and 182 randomly selected control users from March 22, 2012 to July 20, 2017. We conducted a between-group comparative textual analysis of tweets about repetitive and obsessive-compulsive behavioral characteristics typically associated with ASD. In addition, common emotional characteristics of persons with ASD, such as fear, paranoia, and anxiety, were examined between groups through textual analysis. Furthermore, we compared the timing of tweets between users with ASD and control users to identify patterns in communication.

Results: Users with ASD posted a significantly higher frequency of tweets related to the specific repetitive behavior of counting compared with control users (P<.001). The textual analysis of obsessive-compulsive behavioral characteristics, such as fixate, excessive, and concern, were significantly higher among users with ASD compared with the control group (P<.001). In addition, emotional terms related to fear, paranoia, and anxiety were tweeted at a significantly higher rate among users with ASD compared with control users (P<.001). Users with ASD posted a smaller proportion of tweets during time intervals of 00:00-05:59 (P<.001), 06:00-11:59 (P<.001), and 18:00-23.59 (P<.001), as well as a greater proportion of tweets from 12:00 to 17:59 (P<.001) compared with control users.

Conclusions: Social media may be a valuable resource for observing unique psychological characteristics of self-identified persons with ASD. Collecting and analyzing data from these digital platforms may afford opportunities to identify the characteristics of ASD and assist in the diagnosis or verification of ASD. This study highlights the feasibility of leveraging digital data for gaining new insights into various health conditions.

(JMIR Mhealth Uhealth 2019;7(2):e12264) doi:10.2196/12264

KEYWORDS

autism; digital data; emotion; mobile phone; obsessive-compulsive disorder; social media; textual analysis; tweets; Twitter; infodemiology

Introduction

Autism spectrum disorder (ASD) is a group of developmental disorders that affect >3.5 million Americans [1]. The increasing prevalence of ASD requires greater investigation and methods to diagnose this disorder [2-5]. Diagnosing ASD is difficult, as there is no medical test to diagnose it [6]. Inadequate screening practices [7] and a general lack of awareness of symptoms can lead to delays or misdiagnoses [8]. Symptoms of ASD can sometimes remain unnoticed in childhood as social and language impairments can be subtle and difficult to detect [9].

The Diagnostic and Statistical Manual of Mental Disorders (DSM-5), a widely used classification system, describes ASD as a disorder characterized by marked impairments in social interaction and communication accompanied by a pattern of restricted, repetitive behaviors and activities [10]. However, many of these behaviors may be mistaken for other disorders. For instance, anxiety symptoms may be present in response to social impairments and, as a result, may be considered social anxiety without recognizing the underlying behaviors or the presence of ASD [9]. Reportedly, anxiety is a source and consequence of ASD behaviors and may be a signal to aid in diagnoses [9,11]. In addition, research shows that persons with ASD exhibit higher states of fear and phobia compared with matched patients without ASD [12]. The presence and identification of these symptoms may be markers for persons with ASD coping with judgment faced by social situations [9,13].

Furthermore, common repetitive behaviors observed among persons with ASD may be confused with similarities of having obsessive-compulsive disorder (OCD) [7]. Time-consuming obsessions and compulsions frequently characterize OCD-related behaviors; however, patients with ASD can display repetitive behavior, such as repetitive checking, washing, and counting, which may appear to be compulsive obsessions [14]. Research has found that patients with ASD have significantly fewer compulsions than persons with OCD [15]. Other studies have shown that behaviors like counting, checking, cleaning, and ordering were also displayed less frequently in patients with ASD compared with OCD individuals [16].

Appropriate and timely diagnosis of ASD is necessary to provide proper treatment and symptom management. Numerous diagnostic tools exist today, yet many of these tools face limitations and serve slightly different purposes [17,18]. For example, the Autism Diagnostic Observation Schedule-Generic (ADOS-G) is a widely used diagnostic tool that allows for the inclusion of direct observational data that, for example, the Autism Diagnostic Interview-Revised (ADI-R) lacks [19,20]. However, the ADOS-G often tends to result in the overdiagnosis of ASD [17]. Observational methods and parental interviews have been deemed as helpful to further identify behaviors of ASD, yet these symptoms may often go unnoticed [21,22]. Although multidisciplinary approaches have been considered most effective at producing an accurate diagnosis, these require a significant amount of time, resources, and information, which is not always available [21,23].

Moreover, the use of multidisciplinary approaches extends to different domains, as well as novel sources of information [24]. Recently, exploring digital data has emerged as a promising method for examining human behavior and identifying symptoms of various mental and behavioral conditions [25-27]. Dubbed as the “digital phenotype,” the collection of data from social media, Web-based searches, and use of smartphones and other digital devices can serve as a source of observational information that can yield insights into the lived experiences and everyday realities of persons living with various health conditions [27]. These digital information sources can be further harnessed to potentially detect symptomatic signals to aid in identifying and diagnosing ASD.

Twitter has emerged as a potentially valuable platform for examining a wide variety of behavioral and mental health conditions [25-26,28]. Exploring textual differences in Web-based communication on Twitter is a commonly used method of analyzing affective states and behavioral patterns [29,30]. Previous studies have documented that Web-based data may capture real-time projections of emotional states and health behaviors [26,31]. For example, drawing from analyses of conversations on Twitter, it may be possible to identify the onset of depressive symptoms and detect mood and affective states [32,33]. Given this promising evidence base, there may be opportunities to leverage data from Twitter for identifying the characteristics of ASD.

To date, few studies have used social media for studying ASD, though it may be possible to leverage these popular platforms and novel sources of Web-based data to contribute to a more in-depth understanding of these disorders. A prior study extracted data from Twitter to examine conversations about ASD-related information, thereby demonstrating that this is a topic of frequent discussion on this platform [28]. It may be especially important to consider the use of social media for studying ASD given that a recent report highlighted that roughly 80% of adults living with ASD use popular social media platforms [34]. In this study, our overarching objective is to expand on this prior work and further our understanding about whether publicly available social media data captured from Twitter can yield insights into the presence of digital diagnostic signals for ASD [27,35,36]. In addition, we aim to determine the feasibility of informing a digital phenotype for ASD using social media.

Twitter is a social networking service through which users can post tweets that can contain up to 140 characters (increased to 280 characters in November 2017). Twitter is a prominent social medium, as evidenced by its >336 million monthly active users [37]. Twitter is a source of observational data that can yield rich data, including real-time thoughts, feelings, and attitudes, which traditional surveys are typically unable to capture. Therefore,

this study aims to compare the textual patterns of communication of Twitter users who self-identify as having ASD to a general population of control users to identify textual signals related to fear, anxiety, and paranoia, which have often been recognized as a response to symptoms of ASD among individuals living with the disorder [12]. Investigating discussion surrounding these emotions for Twitter users with ASD can potentially enable the detection of Web-based emotional signals that can provide more information to better inform diagnoses of ASD. In addition, this study aims to explore common OCD behaviors using the Yale-Brown Obsessive-Compulsive Scale (Y-BOCS) to understand whether these symptoms are present among Twitter users with ASD and whether the presence or absence of digital obsessive-compulsive symptoms could assist in reducing misdiagnoses of ASD. Finally, we investigated the timing of social media communications to explore differences in patterns of social communication between Twitter users with ASD and control users. Information collected from Twitter as a novel digital data source could potentially guide and assist in the clinical diagnoses and treatment of ASD.

Methods

Data Collection

In this study, we retrieved data from the Twitter application programming interface and collected meta-data such as Twitter users' total number of tweets, number of followers, number of friends, and favorites. We collected data from Twitter between March 22, 2012 and July 20, 2017. The trajectories of ASD symptoms are not stable and can fluctuate over the course of a few months or even years, especially during developmental periods. A 5-year period was selected to ensure that symptoms would be detected even with fluctuations to capture these detectable differences between the 2 groups. This study was considered exempt from ethical review because only publicly available Web-based data collected from the Twitter platform were analyzed. We acknowledge the ethics of drawing in data from users who self-identify as ASD. To ensure the privacy of data from this group, data collected were aggregated and analyzed at the population level and were anonymous in this format. Furthermore, this study only used data from users who consented on Twitter to disclose their data publicly (ie, no privacy settings were selected by users) and are completely public.

Twitter Users

We identified the cohort of Twitter users with ASD by self-identification and explicit Twitter discussion of having ASD. If users described or labeled themselves as autistic or having ASD, then they were determined to be Twitter users with ASD. We searched for terms relating to ASD in users' tweets, and users were considered Twitter users with ASD if they tweeted or wrote a caption about their condition or diagnosis of ASD. We used a previously validated approach to recruit participants with behavioral disorders by Hswen et al [25]. We retrieved tweets that contained one or more keywords associated with ASD; these keywords were searched for in the content of users' tweets and information in their Twitter profile. The keyword list is as follows: autism, autistic, Spectrum Disorders, #ASD, #ActuallyAutistic, #AutismSpeaks, #AutismSpeaks10, #AutismParent, #Autchat, #AutismAwareness, #LIUB, aspergers, and aspies. To validate if users did have ASD, 2 independent researchers coded users if they had ASD or not and an agreement from 2 researchers was necessary for a user to be labeled as a person with ASD. If a disagreement occurred, a third researcher was used to tie break the labeling of the user. To the best of our knowledge, this is the first study to identify ASD users through Twitter, and it is an exploratory study for the identification and labeling of ASD users by machine learning classification and manual curation. A random sample was selected that did not have a self-identified user caption or tweet mentioning ASD and served as the control group. Originally, 220 users with ASD and 220 control users were selected. After filtering out bots and spam accounts and inaccessible accounts (ie, accounts that were private, deleted, deactivated, or banned), we enrolled 152 users with ASD and 182 control users for analysis.

Textual Analysis of Tweets Related to Fear, Anxiety, and Paranoia

We examined the tweets related to the emotions of fear, anxiety, and paranoia through textual analysis. These emotions were selected on the basis of previous research studies that have identified these emotions as key differentiating emotions of persons with ASD compared with matched populations without an ASD diagnosis [12]. We aimed to collect all discussion about these 3 emotional states. The root keyword of the emotion was expanded to all tenses and pluralities to capture the breath of tweets related to that emotion (Table 1). Manual selection of these terms was used to ensure that users were explicitly referring to the emotion in their conversations, a procedure we have applied in prior studies [25]. In addition, the terms "scare" and "scared" were used in the category of fear as a qualitative exploration of discussions around fear to express fearful states. Tweets from users with ASD and control users were queried for all tweets that contained keywords for each category. The number of tweets that contained each keyword was summed to create the total number of tweets for each emotion category. A tweet that contained instances of multiple words, such as "anxious" or "anxiety," was counted as 1 tweet. Overlapping keyword categories, such as a tweet that contained "anxious" and "fear," were counted as 1 tweet per category.

Tweets Containing Obsessive-Compulsive Disorder–Related Keywords

To examine the presence of OCD-related tweets, we compiled a list of key terms from symptom categories within the Y-BOCS. First developed by Goodman et al, this widely used checklist comprehensively covers significant OCD-related behaviors [38]. From the Y-BOCS, 13 categories were selected on the basis of previous research that identified these specific behaviors to be confused with symptoms of OCD and can be mislabeled as an OCD diagnosis [14-16]. These OCD symptom categories included the following: obsess; fixate; repeat; routine; freak; clean; check; count; hoard; wash; worry; excess; and concern. Stem words were expanded with the same methodology as applied for emotion keywords to conduct textual analysis (Table 2). For each category, the number of tweets containing these

terms was collected and summed to generate the total number of tweets per OCD category.

Table 1. All terms related to fear, anxiety, and paranoia.

Root keyword	All forms of keyword
fear	fear, feared, fearful, fearing, fears, scare, and scared
anxiety	anxieties, anxiety, anxious, and anxiousness
paranoia	paranoia, paranoiac, paranoiacs, and paranoid

Table 2. All terms related to obsessive-compulsive disorder–related keywords.

Root keyword	All forms of keyword
obsess	obsessed, obsessing, obsession, obsessions, obsessive, obsessively
fixate	fixated, fixation, fixations
repeat	repeat, repeated, repeatedly, repeating, repeats, repetition, repetitions
routine	routinely, routines
freak	freaked, freaking, freakout, freakouts
clean	cleaned, cleaning
check	checking, checked, checks, recheck, rechecks, rechecked, rechecking
count	counted, counting, counts, recount, recounted, recounting, recounts
hoard	hoarded, hoarder, hoarders, hoards
wash	washed, washing, washes
worry	worried, worrying
excess	excessive, excessively, excessiveness
concern	concerned, concerning, concern

Mean of Proportions for Each Category

As a method to standardize the variation of the frequency of tweets for each user, a "mean of proportions" was calculated for each category. This measure considers the mean number of tweets by category and the proportion of users that tweeted the keywords in each category for emotions and OCD terms. For each user, the total number of tweets for each category was divided by the users' total number of tweets during the full study period between March 22, 2012 and July 20, 2017 (a total of 1946 days). This measure aimed to control the volume of tweets by users to ensure that the mean number of tweets by category was not skewed by individual users' frequency of tweets. The mean of each proportion for users with ASD and the control user group was calculated by dividing the proportion of tweets by category over the number of users within each group. Of note, we used these means of proportions values solely for comparison between the ASD user group and the control user group; we do not know or study the clinical significance of the values.

Timing of Tweets

To examine differences in social communication patterns between users with ASD and control users, we analyzed the timing of tweets between the 2 groups to identify whether tweets occurred at different times of the day between the 2 groups. Meta-data were collected on the local time of users' tweets using the universal time code (UTC) offset data. Twitter users can choose to select whether to include their local time zone in their account settings, and tweets with UTC offset data were included in this analysis. These tweets were further classified into the following time intervals: 00:00-05:59; 06:00-11:59; 12:00-17:59; and 18:00-23:59.

Statistical Analyses

We conducted 2-tailed Welch's *t* tests to compare the "mean of proportions" for each category of emotion and OCD keywords between users with ASD and control users. A chi-square test was used to compare the proportion of users in each category of the timing of tweet data between the 2 groups. For all tests, a $P<.05$ was considered statistically significant. All analyses were performed with the Python Programming Language Version 3.6 (Python Software Foundation).

Results

Characteristics of Twitter Users

In total, 4,093,559 tweets were collected between March 22, 2012 and July 20, 2017. Twitter users with ASD (n=152) tweeted a total of 1,700,841 tweets, and control users (n=182) tweeted a total of 2,392,718 tweets. We observed no statistically significant differences between the group of users with ASD and the group of control users in the mean number of tweets per user or for meta-data, including the number of friends, followers, or favorites (Table 3).

Mean Proportion of Tweets for Fear, Anxiety, and Paranoia

Users with ASD posted a greater number of tweets compared with control users for all 3 emotions—fear (3.11e-03 vs 1.98e-03; t_{225}=4.410, P<.001), anxiety (1.91e-03 vs 6.49e-04; t_{172}=3.529, P=.001), and paranoia (1.90e-03 vs 8.12e-05; t_{211}=3.021, P=.003; Table 4; Figure 1). Aggregating all 3 emotions together, Twitter users with ASD posted a higher number of tweets compared with control users (5.15e-03 vs 2.68e-03; t_{190}=4.981, P<.001; Table 4; Figure 1).

Mean Proportion of Tweets for Obsessive-Compulsive Disorder–Related Terms

Users with ASD posted a higher number of tweets compared with control users for 4 OCD-related keyword categories—fixate (3.64e-05 vs 1.23e-0.5; t_{230}=2.356, P=.02), count (1.43e-03 vs 9.69e-04; t_{219}=3.457, P<.001), excessive (1.15e-04 vs 4.82e-05; t_{201}=2.807, P=.005), and concern (9.26e-04 vs 4.55e-04; t_{219}=3.959, P<.001; Table 5; Figure 2). No statistically significant differences between users with ASD and control users were observed for the remaining 8 OCD-related categories (Table 5; Figure 2). In addition, upon aggregating all OCD-related terms, no statistically significant difference between groups was observed (1.20e-02 vs 1.44e-02; t_{193}=1.292, P=.20; Table 5; Figure 2).

Timing of Tweets

The number of Twitter users who had UTC offset data necessary to retrieve the timing of tweets was 69.2% (231/334). We observed no statistically significant difference in the availability of time zone information between users with ASD (111/152, 73.0%) compared with control users (120/182, 65.9%; χ^2_1=2.0, P=.16). A lower proportion of tweets from users with ASD compared with control users was observed during the time intervals of 00:00-05:59 (222,810/1,700,841, 13.10% vs 330,915/2,392,718, 13.83%; P<.001), 06:00-11:59 (294,245/1,700,841, 17.30% vs 468,972/2,392,718, 19.60%; P<.001), and 18:00-23:59 (612,302/1,700,841, 36.00% vs 890,091/2,392,718, 37.20%; P<.001; Table 6). A higher proportion of tweets from users with ASD compared with control users was observed during the time interval 12:00-17:59 (571,483/1,700,841, 33.60% vs 703,459/2,392,718, 29.40%; P<.001; Table 6).

Table 3. The characteristics of Twitter users with autism spectrum disorder (ASD) and control users.

Twitter user characteristics	Twitter users with ASD (n=152), mean (SD)	Control twitter users (n=182), mean (SD)	t (df)[a]	P value
Mean overall tweets per user	11,189 (23,019)	13,146 (20,159)	–0.818 (303)	.41
Mean retweets per user	2237 (4886)	4942 (9680)	–3.299 (277)	.001
Mean original tweets per user	8952 (20,663)	8204 (12,369)	0.391 (237)	.70
Number of friends	1460 (5422)	1193 (5559)	0.442 (324)	.66
Number of followers	1778 (5753)	1891 (7752)	0.153 (327)	.88
Number of favorites	15,556 (31,966)	12,515 (19,978)	1.018 (243)	.31

[a]2-tailed Welch's t tests were used to compare the "mean of proportions" for each category.

Table 4. Tweets containing emotion-related keywords among Twitter users with autism spectrum disorder (ASD) and control users.

Emotion terms category	Twitter users with ASD (n=152), mean (SD)	Control twitter users (n=182), mean (SD)	t (df)[a]	P value
fear	3.105e-03 (2.822e-03)	1.975e-03 (1.553e-03)	4.410 (225)	<.001
paranoid	1.902e-04 (4.057e-04)	8.120e-05 (1.993e-04)	3.021 (211)	.003
anxious	1.914e-03 (4.275e-03)	6.486e-04 (1.230e-03)	3.529 (172)	.001
Tweets with any of the 3 emotional categories' terms	5.154e-03 (5.760e-03)	2.679e-03 (2.282e-03)	4.981 (190)	<.001

[a]2-tailed Welch's t tests were used to compare the "mean of proportions" for each category.

Figure 1. Differences in emotions between autism spectrum disorder (ASD) and control users.

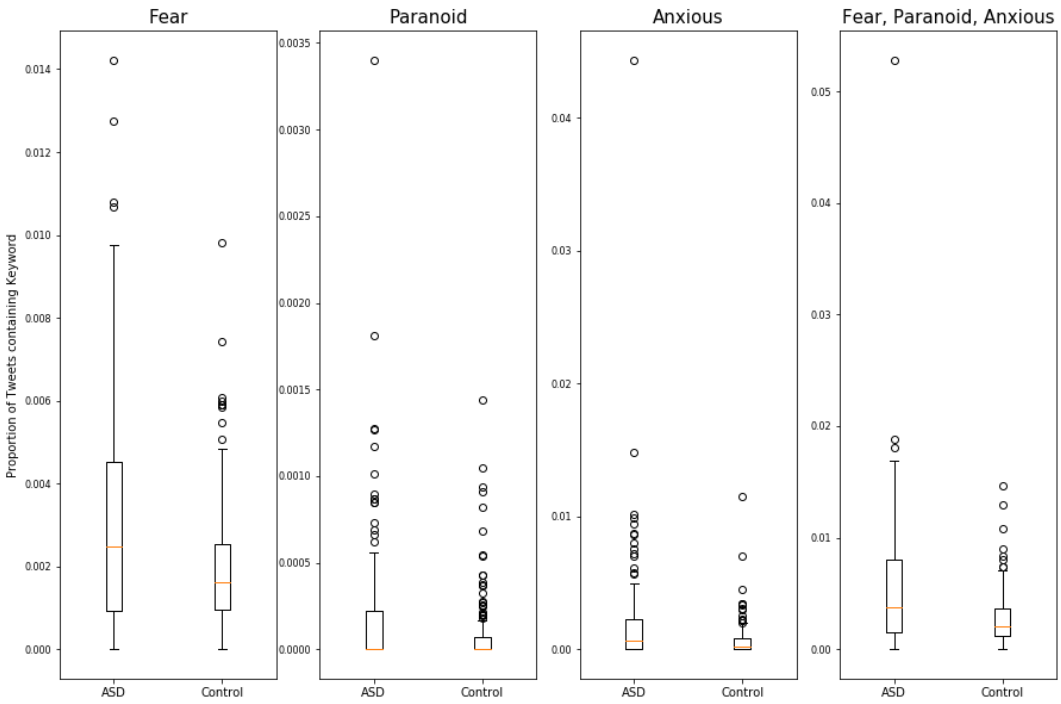


Table 5. Tweets containing obsessive-compulsive disorder (OCD)–related keywords among Twitter users with autism spectrum disorder (ASD) and control users.

OCD keyword category	Twitter users with ASD (n=152), mean (SD)	Control twitter users (n=182), mean (SD)	*t* (*df*)[a]	*P* value
obsess	5.826e-04 (9.807e-04)	7.086e-04 (9.377e-04)	–1.193 (316)	.23
fixate	3.640e-05 (1.120e-04)	1.226e-05 (6.404e-05)	2.356 (230)	.02
repeat	8.711e-04 (1.381e-03)	7.239e-04 (1.775e-03)	0.852 (330)	.40
routine	1.027e-03 (7.171e-03)	2.866e-04 (1.114e-03)	1.260 (157)	.21
freak	7.584e-04 (1.066e-03)	8.827e-04 (1.342e-03)	–0.943 (331)	.35
clean	1.068e-03 (1.534e-03)	9.481e-04 (1.026e-03)	0.821 (255)	.41
check	1.061e-02 (4.840e-02)	6.944e-03 (1.982e-02)	0.874 (193)	.38
count	1.427e-03 (1.471e-03)	9.693e-04 (7.744e-04)	3.457 (219)	.001
hoard	5.608e-05 (1.876e-04)	4.938e-05 (2.616e-04)	0.272 (325)	.79
wash	5.766e-04 (1.176e-03)	5.986e-04 (8.416e-04)	–0.193 (267)	.85
worry	2.071e-03 (2.167e-03)	1.878e-03 (1.454e-03)	0.934 (256)	.35
excessive	1.152e-04 (2.728e-04)	4.816e-05 (1.215e-04)	2.807 (201)	.005
concern	9.260e-04 (1.322e-03)	4.548e-04 (6.964e-04)	3.959 (219)	<.001
Tweets with any of the 13 categories' terms	1.997e-02 (4.940e-02)	1.445e-02 (2.023e-02)	1.292 (193)	.20

[a]2-tailed Welch's *t* tests were used to compare the "mean of proportions" for each category.

Figure 2. Differences in obsessive-compulsive disorder-related discussion between autism spectrum disorder (ASD) and control users.

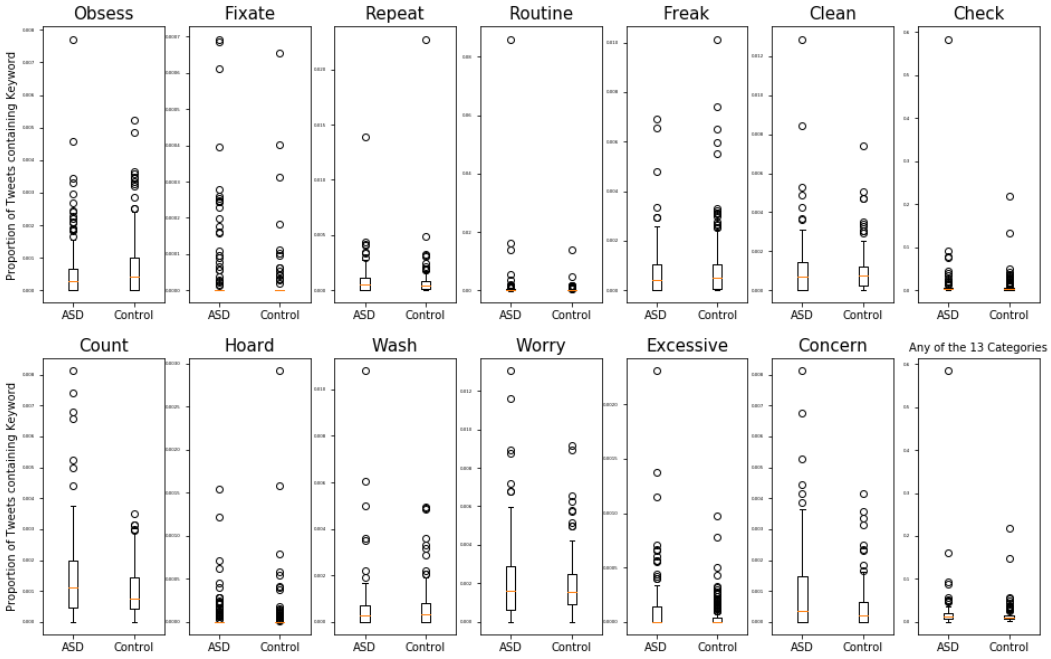


Table 6. Timing of users' tweets between Twitter users with autism spectrum disorder (ASD) and control users.

Time interval	Proportion of tweets (%)		χ^2 (*df*)	*P* value
	Among twitter users with ASD	Among control twitter users		
00:00-05:59	13.1	13.8	351.9 (*1*)	<.001
06:00-11:59	17.3	19.6	2952.9 (*1*)	<.001
12:00-17:59	33.6	29.4	6713.7 (*1*)	<.001
18:00-23:59	36.0	37.2	474.3 (*1*)	<.001

Discussion

Principal Findings

The diagnosis of ASD is complex. Often a multidisciplinary approach is required to produce the most accurate diagnosis. This study investigated the feasibility of using Twitter as a novel digital platform for generating additional insights to assist in the diagnosis of ASD. The findings revealed that through the textual analysis of tweets, discussions about commonly heightened emotional states associated with ASD are expressed among Twitter users who self-identify as having ASD. Specifically, Twitter users with ASD tweeted more frequently about fear, anxiety, and paranoia compared with a randomly generated control group of Twitter users. In addition, these results suggest that Web-based emotional patterns expressed among Twitter users with ASD may parallel offline emotional symptoms known to be associated with the condition and, therefore, may be used as a signal for detecting or confirming a diagnosis of ASD.

As misdiagnoses of ASD for OCD are common because of many overlapping and similar symptoms [39,40], in this study, we also explored whether there may be differences in the expression of OCD-related behaviors among Twitter users with ASD compared with control users. Of 13 selected OCD-related categories, users with ASD discussed only 4 categories—fixate, count, excessive, and concern—significantly more than the control user group. Furthermore, when aggregating all 13 categories, we observed no statistically significant difference. These findings indicate that Twitter users with ASD may not display a wide number of symptoms of OCD and may highlight the most salient symptoms that could be explored further as a method to distinguish symptoms of ASD from individuals with OCD.

Although research on the presence of some of these behaviors may steer clinicians toward a diagnosis of OCD, misattribution to a behavior or misdiagnosis can lead to the inappropriate treatment for ASD [7,41]. Therefore, it is important to consider the spectrum of symptoms of OCD and determine whether Web-based discussions of these behaviors are consistent with offline representations of OCD. However, this highlights an

important area for future investigation because offline observational information of this variety of symptoms can be difficult to obtain. Ideally, there may be opportunities to leverage Web-based social media data that allow us to capture minute-by-minute thoughts and behaviors to provide clinicians a more objective illustration of patients' behaviors toward better understanding the daily occurrences of these symptoms and inform diagnosis and treatment recommendations.

Finally, we sought to determine differences in patterns of the timing of Twitter communication between users with ASD and control users. As ASD is characterized as a social disorder [41], there could be differences in patterns of the Web-based communication timing among individuals with this condition compared with other mental disorders. We found some differences in the proportion of tweets between groups at each time-point throughout the day; however, these differences were very small and are likely not indicative of any diagnostically relevant differences between groups. This is consistent with previous studies that have found no differences in the timing of tweets between a general population of Twitters users and users with mental disorders such as schizophrenia [25]. Additional research is needed to explore whether potential differences in the timing of social media use and posts between users with ASD and control users could inform the development of targeted clinical and public health interventions to reach this patient group online.

Limitations

Although Web-based social media offers an opportunity to observe a wide range of thoughts and behaviors, our data have some limitations. First, Twitter users who self-identify as having ASD may not be representative of the general population of individuals with ASD. The Twitter population, on average, is composed of a larger proportion of younger adults aged 18-29 years with higher levels of education compared with those who do not use Twitter [42]. Therefore, it is more likely that users with ASD in our study are younger adults and with higher education levels than persons with ASD who do not use social media. Second, identification of users with ASD was through self-identification on Twitter. Many individuals with ASD who are on Twitter may not openly self-identify as having ASD, thereby making the sample included in this study potentially more comfortable discussing their disorder over the Web with others. This might mean that our study population is different compared with Twitter users with ASD who choose not to self-identify as having ASD online. Although no formal clinical diagnosis was confirmed for Twitter users included in the analyses conducted in this study, the nature of the stigmatizing diagnosis of ASD makes it highly unlikely that individuals would be dishonest about publicly self-identifying as having ASD on Twitter. There is a possibility, however, that there are some who might have self-diagnosed themselves as having ASD without a clinical consultation, and we, therefore, cannot verify the clinical diagnosis of ASD in our study population. Of note, this was an exploratory study that used textual analyses to explore whether common emotions and OCD-related behaviors could be detected among Twitter users with ASD and whether these discussions may differ compared with control users. As a result, we can only report on publicly available discussions captured from Twitter and, therefore, we are unable to confirm whether these are also occurring offline. In addition, we did not identify a cohort of OCD users and did not conduct a comparison between ASD and OCD users. Therefore, we are not able to confirm the differences in Web-based patterns between the 2 groups and further examination is necessary. Furthermore, we note that we developed an adaptive list for the emotion categories related to autism but did not expand this list for the OCD-related categories. We specifically used the terminology used in the Y-BOCS to mimic the survey and the terms used to diagnose in the clinical setting. However, we acknowledge that these OCD terms could have been adapted; therefore, we potentially missed OCD-related behaviors in our analysis of tweets. Finally, we acknowledge that small yet detectable differences were yielded from our analyses. Owing to the size and breadth of our dataset, although small, these differences can detect important differences in the characteristics between ASD users and the general population and are worth exploring further as to what is a meaningful change to influence a diagnosis.

The strength of this study is that this is the first known study to leverage Web-based social media data from Twitter to compare emotional and OCD-related discussions between users with ASD and a randomly select group of control users. A comprehensive picture of a digital footprint of emotions and behaviors can be gleaned from Twitter through the examination of Web-based social media data. This study highlights the feasibility of conducting this type of analysis through these emerging digital streams to further contribute to our understanding of a "digital phenotype" for ASD. This study yields additional insights that contribute to a growing body of evidence depicting digital phenotyping and the potential to draw from Web-based data sources to detect signals and symptoms for a range of disorders such as ASD [27,35,36]. The data captured in this way were in a naturalistic and unsolicited format, thereby avoiding the concerns of social desirability bias or recall bias that are present with traditional survey data collection methods. Therefore, the methods applied in this study highlight the potential of collecting novel data streams that could be used to supplement the existing data collection methods. Furthermore, it is important to note that we are dealing with individuals from vulnerable population groups, and they are contributing their conversations and thoughts for observational research studies such as ours; however, all of these publicly available data are always aggregated, and we purposefully did not identify the individuals.

Conclusions

Digital signals from Web-based social media may offer an additional resource to capture information not always available for direct observation. This Web-based data source is not meant to replace traditional forms of diagnosis but to supplement these existing approaches by helping to identify symptoms or confirm diagnoses of ASD. Multidisciplinary approaches are often required to gain a full breadth of information and appear to be the gold standard of diagnostic methods. The accuracy of diagnoses is necessary to ensure that persons with ASD receive proper treatment and care in a timely manner so that they can live a fulfilling and healthy life [21,43-45]. Findings from this

study can contribute to the development of multidisciplinary clinical assessments that draw from novel digital sources, as well as traditional surveys and clinical interviews, and takes steps toward illustrating the potential feasibility of digital detection methods for ASD.

Acknowledgments

This study was funded by the Robert Wood Johnson Foundation Grant 73495 (to YH and JBH), National Institutes of Health/National Human Genome Research Institute Grant 5U54HG007963-04 (to JSB and JBH), the Weatherhead Center for International Affairs (to YH).

Conflicts of Interest

None declared.

References

1. Buescher AVS, Cidav Z, Knapp M, Mandell DS. Costs of autism spectrum disorders in the United Kingdom and the United States. JAMA Pediatr 2014 Aug;168(8):721-728. [doi: 10.1001/jamapediatrics.2014.210] [Medline: 24911948]
2. Autism and Developmental Disabilities Monitoring Network Surveillance Year 2008 Principal Investigators, Centers for Disease Control and Prevention. Prevalence of autism spectrum disorders–Autism and Developmental Disabilities Monitoring Network, 14 sites, United States, 2008. MMWR Surveill Summ 2012 Mar 30;61(3):1-19 [FREE Full text] [Medline: 22456193]
3. Baird G, Simonoff E, Pickles A, Chandler S, Loucas T, Meldrum D, et al. Prevalence of disorders of the autism spectrum in a population cohort of children in South Thames: the Special Needs and Autism Project (SNAP). Lancet 2006 Jul 15;368(9531):210-215. [doi: 10.1016/S0140-6736(06)69041-7] [Medline: 16844490]
4. Baron-Cohen S, Scott FJ, Allison C, Williams J, Bolton P, Matthews FE, et al. Prevalence of autism-spectrum conditions: UK school-based population study. Br J Psychiatry 2009 Jun;194(6):500-509 [FREE Full text] [doi: 10.1192/bjp.bp.108.059345] [Medline: 19478287]
5. Brugha TS, McManus S, Bankart J, Scott F, Purdon S, Smith J, et al. Epidemiology of autism spectrum disorders in adults in the community in England. Arch Gen Psychiatry 2011 May;68(5):459-465. [doi: 10.1001/archgenpsychiatry.2011.38] [Medline: 21536975]
6. Centers for Disease Control and Prevention. 2012. Screening and diagnosis for healthcare providers URL: https://www.cdc.gov/ncbddd/autism/hcp-screening.html [WebCite Cache ID 72Evv3U4G]
7. Mandell DS, Ittenbach RF, Levy SE, Pinto-Martin JA. Disparities in diagnoses received prior to a diagnosis of autism spectrum disorder. J Autism Dev Disord 2007 Oct;37(9):1795-1802 [FREE Full text] [doi: 10.1007/s10803-006-0314-8] [Medline: 17160456]
8. Volkmar FR, Lord C, Bailey A, Schultz RT, Klin A. Autism and pervasive developmental disorders. J Child Psychol Psychiatry 2004 Jan;45(1):135-170. [Medline: 14959806]
9. Cath DC, Ran N, Smit JH, van Balkom AJLM, Comijs HC. Symptom overlap between autism spectrum disorder, generalized social anxiety disorder and obsessive-compulsive disorder in adults: a preliminary case-controlled study. Psychopathology 2008;41(2):101-110. [doi: 10.1159/000111555] [Medline: 18033980]
10. Diagnostic and statistical manual of mental disorders. 5th edition. Arlington, VA: American Psychiatric Association; 2013.
11. Bryson SE, Rogers SJ, Fombonne E. Autism spectrum disorders: early detection, intervention, education, and psychopharmacological management. Can J Psychiatry 2003 Sep;48(8):506-516. [doi: 10.1177/070674370304800802] [Medline: 14574826]
12. Evans DW, Canavera K, Kleinpeter FL, Maccubbin E, Taga K. The fears, phobias and anxieties of children with autism spectrum disorders and Down syndrome: comparisons with developmentally and chronologically age matched children. Child Psychiatry Hum Dev 2005;36(1):3-26. [doi: 10.1007/s10578-004-3619-x] [Medline: 16049642]
13. Grant BF, Hasin DS, Blanco C, Stinson FS, Chou SP, Goldstein RB, et al. The epidemiology of social anxiety disorder in the United States: results from the National Epidemiologic Survey on Alcohol and Related Conditions. J Clin Psychiatry 2005 Nov;66(11):1351-1361. [Medline: 16420070]
14. Rumsey JM, Rapoport JL, Sceery WR. Autistic children as adults: psychiatric, social, and behavioral outcomes. J Am Acad Child Psychiatry 1985 Jul;24(4):465-473. [Medline: 4019976]
15. Zandt F, Prior M, Kyrios M. Repetitive behaviour in children with high functioning autism and obsessive compulsive disorder. J Autism Dev Disord 2007 Feb;37(2):251-259. [doi: 10.1007/s10803-006-0158-2] [Medline: 16865546]
16. Colasanto M. A Comparative Study of Repetitive Behaviours in Pediatric Obsessive-compulsive Disorder and Autism Spectrum Disorder. University of Toronto 2015 Nov [FREE Full text]
17. Bishop DVM, Norbury CF. Exploring the borderlands of autistic disorder and specific language impairment: a study using standardised diagnostic instruments. J Child Psychol Psychiatry 2002 Oct;43(7):917-929. [Medline: 12405479]

18. Risi S, Lord C, Gotham K, Corsello C, Chrysler C, Szatmari P, et al. Combining information from multiple sources in the diagnosis of autism spectrum disorders. J Am Acad Child Adolesc Psychiatry 2006 Sep;45(9):1094-1103. [doi: 10.1097/01.chi.0000227880.42780.0e] [Medline: 16926617]
19. Lord C, Risi S, Lambrecht L, Cook EH, Leventhal BL, DiLavore PC, et al. The autism diagnostic observation schedule-generic: a standard measure of social and communication deficits associated with the spectrum of autism. J Autism Dev Disord 2000 Jun;30(3):205-223. [Medline: 11055457]
20. Lord C, Rutter M, Le Couteur A. Autism Diagnostic Interview-Revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. J Autism Dev Disord 1994 Oct;24(5):659-685. [Medline: 7814313]
21. Filipek PA, Accardo PJ, Ashwal S, Baranek GT, Cook EH, Dawson G, et al. Practice parameter: screening and diagnosis of autism: report of the Quality Standards Subcommittee of the American Academy of Neurology and the Child Neurology Society. Neurology 2000 Aug 22;55(4):468-479. [Medline: 10953176]
22. Ozonoff S, Heung K, Byrd R, Hansen R, Hertz-Picciotto I. The onset of autism: patterns of symptom emergence in the first years of life. Autism Res 2008 Dec;1(6):320-328 [FREE Full text] [doi: 10.1002/aur.53] [Medline: 19360687]
23. Westman AG, Miniscalco C, Gillberg C. Autism in preschoolers: does individual clinician's first visit diagnosis agree with final comprehensive diagnosis? Scientific World Journal 2013;2013:716267 [FREE Full text] [doi: 10.1155/2013/716267] [Medline: 24082856]
24. Grzadzinski R, Huerta M, Lord C. DSM-5 and autism spectrum disorders (ASDs): an opportunity for identifying ASD subtypes. Mol Autism 2013 May 15;4(1):12 [FREE Full text] [doi: 10.1186/2040-2392-4-12] [Medline: 23675638]
25. Hswen Y, Naslund JA, Brownstein JS, Hawkins JB. Online Communication about Depression and Anxiety among Twitter Users with Schizophrenia: Preliminary Findings to Inform a Digital Phenotype Using Social Media. Psychiatr Q 2018 Dec;89(3):569-580. [doi: 10.1007/s11126-017-9559-y] [Medline: 29327218]
26. Hswen Y, Naslund JA, Chandrashekar P, Siegel R, Brownstein JS, Hawkins JB. Exploring online communication about cigarette smoking among Twitter users who self-identify as having schizophrenia. Psychiatry Res 2017 Dec;257:479-484 [FREE Full text] [doi: 10.1016/j.psychres.2017.08.002] [Medline: 28841509]
27. Jain SH, Powers BW, Hawkins JB, Brownstein JS. The digital phenotype. Nat Biotechnol 2015 May;33(5):462-463. [doi: 10.1038/nbt.3223] [Medline: 25965751]
28. Beykikhoshk A, Arandjelović O, Phung D, Venkatesh S, Caelli T. Using Twitter to learn about the autism community. Soc Netw Anal Min 2015 Jun 2;5(1). [doi: 10.1007/s13278-015-0261-5]
29. Chen X, Sykora M, Jackson T, Elayan S. What about Mood Swings? Identifying Depression on Twitter with Temporal Measures of Emotions. 2018 Presented at: Proceedings of The Web Conference 2018; April 2018; Lyon, France p. 23-27. [doi: 10.1145/3184558.3191624]
30. Coppersmith G, Dredze M, Harman C. Quantifying mental health signals in twitter. : Association for Computational Linguistics; 2014 Presented at: Proceedings of the Workshop on Computational Linguistics and Clinical Psychology: From Linguistic Signal to Clinical Reality; June 2014; Baltimore, Maryland p. 51-60 URL: http://aclweb.org/anthology/W14-3207 [doi: 10.3115/v1/W14-3207]
31. McIver DJ, Hawkins JB, Chunara R, Chatterjee AK, Bhandari A, Fitzgerald TP, et al. Characterizing Sleep Issues Using Twitter. J Med Internet Res 2015 Jun 08;17(6):e140 [FREE Full text] [doi: 10.2196/jmir.4476] [Medline: 26054530]
32. De Choudhury M, Gamon M, Counts S. Happy, nervous or surprised? Classification of human affective states in social media. 2012 Presented at: Proceedings of the Sixth International AAAI Conference on Weblogs and Social Media; June 2012; Dublin, Ireland p. 4-7.
33. De Choudhury M, Gamon M, Counts S, Horvitz E. Predicting depression via social media. 2013 Presented at: Proceedings of the Seventh International AAAI Conference on Weblogs and Social Media; 2013; Cambridge, Massachusetts p. 8-11.
34. Mazurek MO. Social media use among adults with autism spectrum disorders. Computers in Human Behavior 2013 Jul;29(4):1709-1714. [doi: 10.1016/j.chb.2013.02.004]
35. Insel TR. Digital Phenotyping: Technology for a New Science of Behavior. JAMA 2017 Oct 03;318(13):1215-1216. [doi: 10.1001/jama.2017.11295] [Medline: 28973224]
36. Onnela J, Rauch SL. Harnessing Smartphone-Based Digital Phenotyping to Enhance Behavioral and Mental Health. Neuropsychopharmacology 2016 Jun;41(7):1691-1696. [doi: 10.1038/npp.2016.7] [Medline: 26818126]
37. Statista. Number of monthly active Twitter users worldwide from 1st quarter 2010 to 3rd quarter 2018 (in millions) URL: https://www.statista.com/statistics/282087/number-of-monthly-active-twitter-users/ [accessed 2018-09-09] [WebCite Cache ID 72IBsyybY]
38. Goodman WK, Price LH, Rasmussen SA, Mazure C, Fleischmann RL, Hill CL, et al. The Yale-Brown Obsessive Compulsive Scale. I. Development, use, and reliability. Arch Gen Psychiatry 1989 Nov;46(11):1006-1011. [Medline: 2684084]
39. McDougle CJ, Kresch LE, Goodman WK, Naylor ST, Volkmar FR, Cohen DJ, et al. A case-controlled study of repetitive thoughts and behavior in adults with autistic disorder and obsessive-compulsive disorder. Am J Psychiatry 1995 May;152(5):772-777. [doi: 10.1176/ajp.152.5.772] [Medline: 7726318]
40. Russell AJ, Mataix-Cols D, Anson M, Murphy DGM. Obsessions and compulsions in Asperger syndrome and high-functioning autism. Br J Psychiatry 2005 Jun;186:525-528. [doi: 10.1192/bjp.186.6.525] [Medline: 15928364]

41. American Psychiatric Association. In: Diagnostic and statistical manual of mental disorders, Fifth Edition. Arlington, VA: American Psychiatric Publishing; 2013.
42. Pew Research Center. 2018 Feb 5. Social Media Fact Sheet URL: http://www.pewinternet.org/fact-sheet/social-media/ [accessed 2018-09-10] [WebCite Cache ID 72Jv3sDCB]
43. Dawson G, Rogers S, Munson J, Smith M, Winter J, Greenson J, et al. Randomized, controlled trial of an intervention for toddlers with autism: the Early Start Denver Model. Pediatrics 2010 Jan;125(1):e17-e23 [FREE Full text] [doi: 10.1542/peds.2009-0958] [Medline: 19948568]
44. Kasari C, Gulsrud AC, Wong C, Kwon S, Locke J. Randomized controlled caregiver mediated joint engagement intervention for toddlers with autism. J Autism Dev Disord 2010 Sep;40(9):1045-1056 [FREE Full text] [doi: 10.1007/s10803-010-0955-5] [Medline: 20145986]
45. Rogers SJ, Vismara L, Wagner AL, McCormick C, Young G, Ozonoff S. Autism treatment in the first year of life: a pilot study of infant start, a parent-implemented intervention for symptomatic infants. J Autism Dev Disord 2014 Dec;44(12):2981-2995 [FREE Full text] [doi: 10.1007/s10803-014-2202-y] [Medline: 25212413]

Abbreviations

ASD: autism spectrum disorder
OCD: obsessive-compulsive disorder
UTC: universal time code
Y-BOCS: Yale-Brown Obsessive-Compulsive Scale

Edited by C Dias; submitted 24.09.18; peer-reviewed by X Chen, S Henderson, L Geneviève, D Mciver; comments to author 02.11.18; revised version received 16.11.18; accepted 18.11.18; published 12.02.19.

Please cite as:
Hswen Y, Gopaluni A, Brownstein JS, Hawkins JB
Using Twitter to Detect Psychological Characteristics of Self-Identified Persons With Autism Spectrum Disorder: A Feasibility Study
JMIR Mhealth Uhealth 2019;7(2):e12264
URL: http://mhealth.jmir.org/2019/2/e12264/
doi:10.2196/12264
PMID:30747718

Original Paper

Change in Patient Comfort Using Mobile Phones Following the Use of an App to Monitor Tuberculosis Treatment Adherence: Longitudinal Study

Diana Do[1], MPH; Richard S Garfein[2], MPH, PhD; Jazmine Cuevas-Mota[2], MPH; Kelly Collins[2], MPH, PhD; Lin Liu[2], PhD

[1]San Diego State University, San Diego, CA, United States

[2]University of California San Diego, School of Medicine, La Jolla, CA, United States

Corresponding Author:
Richard S Garfein, MPH, PhD
University of California San Diego
School of Medicine
9500 Gilman Drive, MC-0507
La Jolla, CA, 92093-0507
United States
Phone: 1 1 858 822 3018
Fax: 1 858 534 7566
Email: rgarfein@ucsd.edu

Abstract

Background: As mHealth apps proliferate, it is necessary for patients to feel capable and comfortable using devices that run them. However, limited research is available on changes in comfort level before and after the use of an mHealth app.

Objective: The objective of this study was to determine whether patients with tuberculosis who used an mHealth app called Video Directly Observed Therapy (VDOT) to monitor their antituberculosis treatment became more comfortable using mobile phones after the intervention and to identify factors associated with change in comfort.

Methods: We analyzed data from a longitudinal study assessing the feasibility and acceptability of the VDOT app among patients receiving antituberculosis treatment from public health departments in San Diego, San Francisco, and New York City. Comfort levels on six domains of mobile phone use (making phone calls, taking pictures, recording videos, text messaging, internet and email use on the phone) were measured on a 10-point scale (1=very uncomfortable; 10=very comfortable) at the start and end of treatment using VDOT via telephone interviews. The main outcomes were change in comfort level on each domain (recoded as binary measures) and an overall change score (sum of individual measures). Linear and logistic regression analyses were performed to assess whether sociodemographics, risk factors, and VDOT perceptions were associated with change of comfort measures.

Results: Among 120 participants with complete data, mean age was 39.8 years (SD 14.8, range 18-87 years), 46.7% (56/120) were female, and 76.7% (92/120) were foreign born. The combined comfort level at baseline was high overall (mean 48.8, SD 14.2, interquartile range 43.0-60.0) and the mean comfort score increased by 1.92 points at follow-up (P=.07). Statistically significant increases in comfort on individual domains included taking pictures (P=.02) and recording videos (P=.002). Females were more likely to have increased comfort in using the internet on the phone compared to males (odds ratio [OR] 3.03, 95% CI 1.08-8.52, P=.04). Participants who worked less hours per week were more likely to have increased comfort recording videos although this did not meet statistical significance (OR 1.03, 95% CI 1.00-1.05, P=.06).

Conclusions: Findings suggest that, despite a high level of comfort using mobile phones at baseline, experience using the VDOT app was associated with increased comfort using mobile phone features. Additional research involving participants with lower baseline mobile phone experience is needed. An implication of these findings is that as patients begin to use mHealth apps for one health condition, they could acquire skills and confidence to more quickly adapt to using mHealth apps for other conditions.

(JMIR Mhealth Uhealth 2019;7(2):e11638) doi:10.2196/11638

KEYWORDS

mHealth; medication adherence monitoring; mobile phone; video technology; tuberculosis

Introduction

Mobile phone apps play an increasingly important role in health care as evidence supporting these interventions accumulates [1]. These mHealth (mobile health) interventions have been applied to monitor postoperative care [2,3], diabetes self-management [4,5], smoking cessation [6,7], care and prevention of HIV and other sexually transmitted diseases [8], and more. As these interventions begin to spread throughout health care, there has been a growing interest in the response to these apps, including topics such as acceptability and feasibility [2], satisfaction by patients and providers [5], and efficiency [4]. However, there is limited research measuring users' comfort in operating mobile phone functions and whether comfort increases as a result of engagement with these apps.

The World Health Organization and Centers of Disease Control recommend the use of directly observed therapy (DOT) for monitoring treatment of both tuberculosis and sometimes latent tuberculosis infection. DOT is a case management system developed for monitoring antituberculosis treatment [9]. Patients with tuberculosis have to travel to a health care facility or a trained DOT worker travels to the patient's home to observe them taking their medication. DOT improves adherence because it requires close supervision, but it has some barriers as well [10]. One example of a barrier to DOT is that it requires both a patient's and DOT worker's presence; both schedules may interfere with time and place of when the medication is taken. These and other barriers prompted the development of an mHealth intervention called Video Directly Observed Therapy (VDOT). VDOT was developed as an alternative to DOT to help reduce the barriers of DOT and increase adherence [11]. The VDOT process allows patients to record themselves taking the medication and send the date- and time-stamped video to a DOT worker who watches the video and documents each dose taken. These recorded videos allow patients the flexibility of choosing the best time and place to take their medication. Patients receive daily medication reminders by short message service (SMS) text message, and may use the phone to text, email, or call their provider throughout treatment.

A study was conducted from August 2013 through July 2014 to assess the feasibility and acceptability of VDOT for monitoring tuberculosis treatment adherence in three US health departments (San Diego, San Francisco, and New York); patient comfort using mobile phones was also measured during baseline and follow-up interviews [11]. In this paper, we assess the change in comfort using mobile phone features and apps and look at factors associated with change. Understanding patient's comfort and use of mHealth apps will assist in the design and development of other health-based mobile phone apps.

Methods

Participant Recruitment

This analysis used data from a three-city study of VDOT among patients with active tuberculosis. Patients 18 years or older undergoing treatment for active tuberculosis in San Diego, San Francisco, and New York City Health Departments were invited to participate in the study by health department staff. Research staff conducted phone interviews before patients started to use the VDOT app (baseline) and after their treatment or study participation ended (follow-up). The interviews included questions on sociodemographics, risk behaviors, VDOT perceptions, and comfort using mobile phone apps and phone features. Participants were provided phones with the VDOT app installed and were told that they could use the phones for anything related to their tuberculosis treatment: making calls, text messaging, emailing, taking pictures, and accessing the internet. The study was approved by the University of California San Diego Human Research Protection Program. All participants provided written informed consent.

Outcomes

The main outcome for this analysis was change in comfort using mobile phone features between the baseline and follow-up interviews. Comfort with phone features was stratified into six categories: making phone calls, taking pictures, recording videos, text messaging, internet browsing, and using email. Participants were asked, "On a scale of 1 to 10, how comfortable are you using a cell phone to (category)? 1=very uncomfortable and 10=very comfortable." This scale was developed by the authors and was first used in our study; therefore, we assessed internal reliability using robust alpha coefficients [12]. The estimated alpha was .93 (95% CI .87-.96), which showed excellent reliability. These questions were asked during both the baseline and follow-up interviews. Change in comfort was calculated by subtracting the baseline scores from the follow-up scores. "Not applicable" and "refuse to answer" were potential response options and were treated as missing for this analysis. Overall change in comfort was calculated by summing the six individual technology comfort scores (range 1-10) to produce a cumulative score for baseline and follow-up (range 6-60).

Data Analysis

The overall change in technology comfort and the change in six separate technology comfort variables were summarized by descriptive statistics (mean, median, quartiles, standard deviation) and assessed by Wilcoxon signed rank test. Univariate tests (Spearman correlation coefficient, Wilcoxon rank sum test, and Kruskal-Wallis test) and simple linear regression analysis were performed to assess the bivariate association between overall comfort change score and each potential covariate (sociodemographics, site, risk behaviors, VDOT perceptions, and baseline comfort score). For individual item comfort change measures, we analyzed them as binary outcomes (improved vs not improved) since normal assumption of residuals for a linear regression model was not satisfied; a

change score greater than zero was defined as an improvement. Univariate tests (Wilcoxon rank sum test and Fisher exact test) and simple logistic regression analysis were performed to assess bivariate associations between each individual item change and baseline covariates. Results from the univariate test and regression analysis were consistent; therefore, estimated coefficients, odds ratios (ORs), and *P* values from regression analyses were reported. Variables that were significant with $P<.10$ in the simple linear/logistic regression were considered for inclusion in the multivariable regression analysis. Backward elimination was used to remove nonsignificant variables from the model and the variable with the largest *P* value was removed from the model at each step, until only variables with $P<.10$ remained in the final model. A normal probability plot was used to assess the normality assumption of linear model residuals and influential observations were assessed by residual and Cook's distance. Sensitivity analyses were performed by excluding the influential observations from the analysis and comparing it to the original model. Since the results remained robust, we report regression analysis results from the original model. A $P<.05$ was interpreted as statistically significant. All data analyses were performed using R (version 3.3.0) and SAS 14 software (Cary, NC, USA).

Results

Participant Characteristics

Overall, 151 participants were enrolled in the parent study, of which 126 completed the follow-up interview. Since participants could refuse to answer any question in the interviews, this analysis includes only those who provided responses to at least three comfort questions at both visits, leaving 120 participants for analysis. The mean age was 39.8 years (SD 14.8, range 18-87), 46.7% (56/120) were female, 47.5% (57/120) were Asian, 26.7% (32/120) were Hispanic, and 12.5% (15/120) were African American. Overall, 76.7% (92/120) were foreign born, 55.4% (62/112) reported an annual income of more than US $10,000/year, 79.0% (94/119) had health insurance, and 47.0% (55/117) had education beyond high school. Behavior characteristics showed that 27.7% (33/119) consumed alcohol at least once a month and 40.0% (48/120) had ever smoked cigarettes. Before starting VDOT, 72.5% (87/120) of participants owned a mobile phone and 80.8% (97/120) of participants preferred communicating with their health care provider through text, phone call, or email (text message: 11.7%, 14/120; phone call: 63.3%, 76/120; 14/120; email: 5.8%, 7/120). The mean duration of daily VDOT use was 5.46 months (SD 3.25, interquartile range [IQR] 3.09-7.47).

Overall, mean baseline comfort scores for the six individual domains of mobile phone use were high (range 7.43-9.02; Table 1). The average comfort scores were lowest for email use on phone and highest for making phone calls at both baseline and follow-up. The average change in comfort score was highest for recording videos (mean 0.79, SD 3.04) and lowest for making phone calls (mean 0.11, SD 2.03) and email uses (mean 0.11, SD 2.72); however, the changes were only statistically significant for taking pictures ($P=.02$) and recording video ($P=.002$). Although not statistically significant, an increase was also observed for the combined overall comfort score ($P=.07$).

Table 1. Self-reported comfort score using mobile phone features before and after using an mHealth app to monitor antituberculosis treatment adherence among patients in San Diego, San Francisco, and New York City (N=120).

Comfort variables	Baseline	Follow-up	Change	Improvement (change score>0)[a], n (%)	*P* value[b]
Making phone calls				25 (23.6)	.30
Median (IQR[c])	10 (9, 10)	10 (9, 10)	0 (0, 0)		
Mean (SD)	9.02 (1.95)	9.34 (1.52)	0.11 (2.03)		
Taking pictures				39 (32.8)	.02
Median (IQR)	10 (7, 10)	10 (8, 10)	0 (0, 1.5)		
Mean (SD)	8.21 (2.65)	8.76 (2.29)	0.55 (3.24)		
Recording videos				39 (34.2)	.002
Median (IQR)	9 (6, 10)	10 (8, 10)	0 (0, 1)		
Mean (SD)	7.69 (3.08)	8.48 (2.38)	0.79 (3.04)		
Text messaging				30 (25.9)	.26
Median (IQR)	10 (8, 10)	10 (8, 10)	0 (0, 1)		
Mean (SD)	8.23 (2.87)	8.52 (2.55)	0.27 (2.55)		
Internet use				28 (26.2)	.58
Median (IQR)	10 (7, 10)	10 (7, 10)	0 (0, 1)		
Mean (SD)	7.90 (3.15)	7.95 (3.07)	0.19 (2.64)		
Email use				24 (22.9)	.87
Median (IQR)	9 (5.5, 10)	9 (6, 10)	0 (–1, 0)		
Mean (SD)	7.43 (3.28)	7.47 (3.13)	0.11 (2.72)		
Overall comfort				N/A	.07
Median (IQR)	54 (43, 60)	56 (50, 60)	0 (–2, 6)		
Mean (SD)	48.8 (14.2)	51.3 (12.2)	1.92 (11.1)		

[a]Improvement indicates the number and percentage of participants whose change scores were greater than zero for each individual comfort variable.

[b]*P* values are from Wilcoxon signed rank tests.

[c]IQR: interquartile range.

Combined Overall Comfort

From simple linear regression analysis (Table 2), we found that older age was associated with an increase in overall comfort score change (*P*<.001). Those without a mobile phone at baseline (*P*=.01) and those who had lower comfort at baseline (*P*<.001) were more likely to have increased comfort at follow-up. Using multivariable linear regression, we found only the baseline overall score stayed significant with a higher overall baseline score (ie, more comfortable at baseline) associated with a smaller increase at follow-up (beta=–0.45, *P*<.001).

Comfort in Six Individual Comfort Domains

To identify factors independently associated with change in comfort level for each of the six individual phone functions, we performed simple (Table 2) and multivariable (Table 3) logistic regression analyses. Variables significantly associated with comfort changes on individual items in the simple logistic regression analysis included age, country of birth, education, hours worked per week, previous mobile phone ownership, ever recording VDOT videos while away from home, and baseline comfort score (all *P*'s<.05).

Variables significantly associated with comfort change for each individual item in simple logistic regression were assessed in multivariable logistic regression models. Baseline comfort levels were the strongest predictors in the individual item change measures (all *P'* s<.001). After controlling for baseline scores, only female participants (OR 3.03, 95% CI 1.08-8.52, *P*=.04) were more likely to have increased comfort in internet use. Participants with fewer hours worked per week (OR 1.03, 95% CI 1.00-1.05, *P*=.06) were also more likely to have increased comfort in recording video although this did not meet statistical significance.

Table 2. Simple linear/logistic regression analysis[a] of change in comfort using mobile phone functions among Video Directly Observed Therapy (VDOT) study participants (N=120).

Variables	n (%)[b]	Overall change, beta	Individual comfort measure, odds ratio					
			Making phone calls	Taking pictures	Recording videos	Text messaging	Internet use	Email use
Sociodemographics								
Age (years), mean (SD)	39.8 (14.8)	0.24[c]	1.02	1.03[d]	1.04[c]	1.04[c]	1.05[c]	1.04[c]
Hours/week worked at current job, mean (SD)[e]	25.7 (20.4)	0.05	1.04[c]	1.02[f]	1.03	1.01	1.01	1.02
Country of birth								
Other (ref[g])	75 (62.5)							
United States	28 (23.3)	–2.94	0.39	0.34[d]	0.52	0.43	0.45	1.53
Mexico	17 (14.2)	–2.62	1.28	0.65	1.21	1.56	0.86	0.93
Site								
San Diego (ref)	43 (35.8)							
San Francisco	44 (36.7)	0.54	1.00	1.24	1.73	0.50	0.72	1.03
New York	33 (27.5)	–1.70	0.63	1.15	1.88	1.31	0.66	1.55
Race/ethnicity								
Asian (ref)	57 (47.5)							
Hispanic	32 (26.7)	–0.89	1.00	1.20	2.02	1.18	1.35	1.85
Other	31 (25.8)	–1.03	0.71	1.26	1.44	1.32	0.90	1.03
Female sex (ref: male)	56 (46.7)	0.46	0.62	0.95	1.28	0.84	2.15[f]	1.25
>High school education (ref: <high school)	55 (47.0)	–1.53	0.83	0.52	0.42[d]	0.57	0.56	0.39[f]
Have health insurance (ref: no)	94 (79.0)	2.82	0.66	1.25	1.07	1.15	2.68	1.45
Household income (US$) >$10,000/year (ref: <$10,000/year)	62 (55.4)	–2.84	0.45[f]	0.50[f]	0.62	0.91	0.71	0.54
Behavioral characteristics								
Alcohol consumed ≥once/month (ref: <once/month)	33 (27.7)	0.97	0.52	1.02	1.45	0.43	0.91	0.62
Ever smoked cigarettes (ref: never smoked)	48 (40.0)	–2.92	0.46	0.57	0.99	0.63	0.95	1.04
VDOT characteristics								
Days of practice needed to learn VDOT process, mean (SD)	1.53 (1.51)	–0.76	1.27	1.01	1.02	1.14	0.99	0.88
Miles to tuberculosis clinic from home, mean (SD)	13.0 (10.1)	–0.12	0.98	0.99	1.00	1.00	1.01	0.97
Owned a mobile phone (ref: no)	87 (72.5)	–6.22[c]	0.37[f]	0.20[c]	0.33[d]	0.23[c]	0.64	0.49
Used VDOT phone while away from home-ever (ref: never)	84 (71.2)	–3.43	0.82	0.36[d]	0.76	0.67	0.30[d]	0.54
Preferred communicating with physician in person (ref: text/call/email)	23 (19.2)	1.39	2.01	1.78	1.12	0.81	0.57	0.81
Baseline comfort score[h]		–0.45[c]	0.06[c]	0.43[c]	0.57[c]	0.61[c]	0.75[c]	0.74[c]

[a]Overall change score was analyzed as a continuous measure and the beta coefficient from simple linear regression was reported for overall change score; individual item change scores were analyzed as binary measures since the normal assumption of residuals in linear regression was not satisfied

when analyzed as continuous measures, odds ratios from simple logistic regression were reported for individual item change scores.

[b]Values represent n (%) unless otherwise indicated.

[c]$P<.01$.

[d]$P<.05$.

[e]Beta and odds ratio are interpreted for per 1 hour less.

[f]$P<.10$.

[g]Ref: Reference.

[h]Mean (SD) for baseline comfort scores were presented in Table 1.

Table 3. Multivariable regression model[a] for associations with change in comfort using mobile phone functions among VDOT participants.

Comfort variable	Variables, odds ratio (95% CI)		
	Baseline comfort score	Hours/week worked at current job (per 1 hour less)	Sex (female vs male)
Making phone calls	0.06 (0.02 to 0.18)[b]	N/A[c]	N/A
Taking pictures	0.43 (0.31 to 0.59)[b]	N/A	N/A
Recording videos	0.58 (0.47 to 0.72)[b]	1.03 (1.00 to 1.05)[d]	N/A
Text messaging	0.61 (0.50 to 0.74)[b]	N/A	N/A
Internet use	0.73 (0.62 to 0.85)[b]	N/A	3.03 (1.08 to 8.52)[e]
Email use	0.74 (0.64 to 0.85)[b]	N/A	N/A
Overall score	–0.45 (–0.56 to –0.33)[f]	N/A	N/A

[a]Overall change score was analyzed as a continuous measure and the beta coefficient from multivariable linear regression was reported; individual item change scores were analyzed as binary measures and the odds ratios from multivariable logistic regression were reported.

[b]$P<.01$.

[c]N/A: not applicable.

[d]$P<.10$.

[e]$P<.05$.

[f]Beta coefficient (95% CI) from linear regression analysis.

Discussion

The main function of the VDOT app was to record and send videos of each medication ingestion event, and we found a significantly positive change in comfort with using a mobile phone to take videos and photos. We also observed a nonstatistically significant increase in comfort using all other mobile phone functions following the use of the VDOT app for monitoring tuberculosis treatment adherence. Except for baseline comfort, sex, and hours worked, we did not find baseline sociodemographics to affect comfort change. This indicates that the use of VDOT has the potential to increase patient skills with various mobile phone functions across a range of patient characteristics, which should increase their willingness and ability to utilize future mHealth apps.

Other mHealth studies considered the usability and receptiveness of app-based tools using different forms of Likert scales [2,7] or qualitative measures from patient responses [5,6] to find important associations related to app quality and user opinions. These studies assessed a variety of health interventions and most had positive results when investigating topics such as usability [3] and acceptability [2,7]. Unlike previous studies, this study quantified comfort with using mobile phone functions and examined correlates of those changes. These measures of change in comfort help to elucidate how patients perceive the use of an mHealth app and how its use can impact comfort using mobile phones. Finding that comfort increased after using an mHealth app suggests that introducing patients to one mHealth app can increase the likelihood that they will be willing and able to engage with future mHealth interventions. The factors that contribute to comfort using mobile phone features have important relevance to whether patients have a good experience using the different tools and benefit from the intended purpose of the app.

Some study limitations must be considered. All participants in this study were undergoing treatment for tuberculosis; thus, this sample is not representative of all patient populations. Tuberculosis in the United States most commonly occurs among immigrants from high tuberculosis-burden countries; persons who are homeless, inject illegal drugs, or are HIV infected; and people who live or work in hospitals, homeless shelters, correctional facilities, nursing homes, and residential homes for those with HIV [13]. The study took place in three major metropolitan cities, so the results might not reflect those of patients in rural areas. Also, the modest sample size might have limited our ability to detect statistically significant associations. Since most participants owned mobile phones and had high comfort levels at baseline, they had little room to increase comfort, which also limited our ability to detect changes. Future

studies involving patients with less mobile phone experience at baseline should be considered. Lastly, the data were all self-reported, which may have introduced information bias.

This study has several strengths. The longitudinal study design supports a causal relationship between the observed change in comfort and use of the VDOT app. Baseline and follow-up questionnaires encompassed a wide range of variables including sociodemographic, behavioral, and treatment perceptions, which allowed us to assess multiple potential correlates of change in comfort. The study assessed comfort using six different functions of a mobile phone rather than mobile phone use in general, which allowed for a more comprehensive investigation of the effect of using an mHealth app on change in comfort using mobile phones. Lastly, this is the first study to measure change in comfort using an mHealth app, which helps to address gaps in existing literature.

This study shows that after using a mobile phone for one mHealth intervention, patients became more comfortable using the device, which suggests that they will more quickly embrace other mHealth interventions (eg, smoking cessation or diabetes management). With expanded use of mHealth apps like VDOT, it is important for patients to be comfortable with using mobile phone functions. Measuring the comfort level that patients have while using mHealth apps can help to determine if these apps are improving patient skills using technology. This will help prepare patients to adopt the use of other forms of mHealth, as well as accessing health information through their mobile phone.

Acknowledgments

The authors would like to thank the study participants and tuberculosis control program staff who contributed their time and information for this study. Funding for the parent study was provided by a gift from the Verizon Foundation to the University of California San Diego. RSG and LL are also funded by the National Institutes of Health (U01-AI116392).

Authors' Contributions

DD contributed to analysis of data, interpretation of results, and drafted the manuscript. RSG conceptualized the research questions, advised on study implementation, and contributed to manuscript preparation and final review. JCM assisted in data collection, data quality assurance, and manuscript preparation. KC contributed to the questionnaire design, data collection, and topic conceptualization. LL led the study implementation, designed data analysis plan, supervised data analysis, interpretation of results, manuscript writing, and final review. All authors critically reviewed and approved the manuscript.

Conflicts of Interest

RSG and KC are cofounders of SureAdhere Mobile Technology, Inc—a VDOT service provider. No funding, software, or other resources were provided by SureAdhere for the study. To mitigate potential conflicts of interest, interpretation and reporting of the study findings were approved by coauthors who are unaffiliated with SureAdhere. The terms of this arrangement have been reviewed and approved by the University of California, San Diego, in accordance with its conflict of interest policies.

References

1. Whittaker R. Issues in mHealth: findings from key informant interviews. J Med Internet Res 2012 Oct;14(5):e129 [FREE Full text] [doi: 10.2196/jmir.1989] [Medline: 23032424]
2. Semple JL, Sharpe S, Murnaghan ML, Theodoropoulos J, Metcalfe KA. Using a mobile app for monitoring post-operative quality of recovery of patients at home: a feasibility study. JMIR Mhealth Uhealth 2015 Feb;3(1):e18 [FREE Full text] [doi: 10.2196/mhealth.3929] [Medline: 25679749]
3. Scott AR, Alore EA, Naik AD, Berger DH, Suliburk JW. Mixed-methods analysis of factors impacting use of a postoperative mHealth app. JMIR Mhealth Uhealth 2017 Feb 08;5(2):e11. [doi: 10.2196/mhealth.6728] [Medline: 28179215]
4. Wu Y, Yao X, Vespasiani G, Nicolucci A, Dong Y, Kwong J, et al. Mobile app-based interventions to support diabetes self-management: a systematic review of randomized controlled trials to identify functions associated with glycemic efficacy. JMIR Mhealth Uhealth 2017 Mar 14;5(3):e35 [FREE Full text] [doi: 10.2196/mhealth.6522] [Medline: 28292740]
5. Holtz BE, Murray KM, Hershey DD, Dunneback JK, Cotten SR, Holmstrom AJ, et al. Developing a patient-centered mHealth app: a tool for adolescents with type 1 diabetes and their parents. JMIR Mhealth Uhealth 2017 Apr 19;5(4):e53 [FREE Full text] [doi: 10.2196/mhealth.6654] [Medline: 28428167]
6. 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]
7. Finkelstein J, Cha EM. Using a mobile app to promote smoking cessation in hospitalized patients. JMIR Mhealth Uhealth 2016 May 06;4(2):e59 [FREE Full text] [doi: 10.2196/mhealth.5149] [Medline: 27154792]
8. Muessig KE, Pike EC, Legrand S, Hightow-Weidman LB. Mobile phone applications for the care and prevention of HIV and other sexually transmitted diseases: a review. J Med Internet Res 2013 Jan 04;15(1):e1 [FREE Full text] [doi: 10.2196/jmir.2301] [Medline: 23291245]

9. Centers for Disease Control and Prevention, National Center for HIV/AIDS, Viral Hepatitis, STD, and TB Prevention Division of Tuberculosis Elimination. Core Curriculum on Tuberculosis: What the Clinicians Should Know. 2015. URL: https://www.cdc.gov/tb/education/corecurr/pdf/corecurr_all.pdf [accessed 2018-11-27] [WebCite Cache ID 74FCo3leT]
10. Hopewell PC, Pai M, Maher D, Uplekar M, Raviglione MC. International standards for tuberculosis care. Lancet Infect Dis 2006 Nov;6(11):710-725. [doi: 10.1016/S1473-3099(06)70628-4] [Medline: 17067920]
11. Garfein RS, Collins K, Muñoz F, Moser K, Cerecer-Callu P, Raab F, et al. Feasibility of tuberculosis treatment monitoring by video directly observed therapy: a binational pilot study. Int J Tuberc Lung Dis 2015 Sep;19(9):1057-1064 [FREE Full text] [doi: 10.5588/ijtld.14.0923] [Medline: 26260824]
12. Zhang Z, Yuan K. Robust coefficients alpha and omega and confidence intervals with outlying observations and missing data: methods and software. Educ Psychol Meas 2016 Jun;76(3):387-411 [FREE Full text] [doi: 10.1177/0013164415594658] [Medline: 29795870]
13. Centers for Disease Control and Prevention. TB Risk Factors. 2016 Mar 18. URL: https://www.cdc.gov/tb/topic/basics/risk.htm [accessed 2018-06-19] [WebCite Cache ID 70HbIUfFj]

Abbreviations

DOT: directly observed therapy
OR: odds ratio
VDOT: Video Directly Observed Therapy

Edited by G Eysenbach; submitted 28.07.18; peer-reviewed by S Li, K Murray, P Di Carlo, V Dhikav; comments to author 30.08.18; revised version received 17.10.18; accepted 23.10.18; published 01.02.19.

Please cite as:
Do D, Garfein RS, Cuevas-Mota J, Collins K, Liu L
Change in Patient Comfort Using Mobile Phones Following the Use of an App to Monitor Tuberculosis Treatment Adherence: Longitudinal Study
JMIR Mhealth Uhealth 2019;7(2):e11638
URL: https://mhealth.jmir.org/2019/2/e11638/
doi:10.2196/11638
PMID:30707103

Original Paper

The MomConnect Nurses and Midwives Support Platform (NurseConnect): A Qualitative Process Evaluation

Alex Emilio Fischer[1], BSc; Jane Sebidi[2], Dip Nurs, BCur, PHC service management; Peter Barron[2,3], BCom, FFCH (SA); Samanta Tresha Lalla-Edward[1], PhD

[1]Wits Reproductive Health and HIV Institute, University of Witwatersrand, Johannesburg, South Africa

[2]HIV/AIDS, TB and MCWH, National Department of Health, Government of South Africa, Pretoria, South Africa

[3]School of Public Health, University of Witwatersrand, Johannesburg, South Africa

Corresponding Author:
Alex Emilio Fischer, BSc
Wits Reproductive Health and HIV Institute
University of Witwatersrand
Hillbrow Health Precinct
22 Esselen Street, HIllbrow
Johannesburg, 2001
South Africa
Phone: 27 073 776 2705
Email: afischer@wrhi.ac.za

Abstract

Background: Over the past decade, mobile health has steadily increased in low-income and middle-income countries. However, few platforms have been able to sustainably scale up like the MomConnect program in South Africa. NurseConnect was created as a capacity building component of MomConnect, aimed at supporting nurses and midwives in maternal and child health. The National Department of Health has committed to expanding NurseConnect to all nurses across the country, and an evaluation of the current user experience was conducted to inform a successful scale up.

Objective: This study aims to evaluate the perception and use of NurseConnect by nurses and midwives to produce feedback that can be used to optimize the user experience as the platform continues to scale up.

Methods: We conducted focus group discussions and in-depth interviews with 110 nurses and midwives from 18 randomly selected health care facilities across South Africa. Questions focused on mobile phone use, access to medical information and their experience with NurseConnect registration, as well as the content and different platforms.

Results: All participants had mobile phones and communication through calls and messaging was the main use in both personal and work settings. Of 110 participants, 108 (98.2%) had data-enabled phones, and the internet, Google, and apps (South African National Department of Health Guidelines, iTriage, Drugs.com) were commonly used, especially to find information in the work setting. Of 110 participants, 62 (56.4%) were registered NurseConnect users and liked the message content, especially listeriosis and motivational messages, which created behavioral change in some instances. The mobisite and helpdesk, however, were underutilized because of a lack of information surrounding these platforms. Some participants did not trust medical information from websites and had more confidence in apps, while others associated a “helpdesk” with a call-in service, not a messaging one. Many of the unregistered participants had not heard of NurseConnect, and some cited data and time constraints as barriers to both registration and uptake.

Conclusions: Mobile and smartphone penetration was very high, and participants often used their phone to find medical information. The NurseConnect messages were well-liked by all registered participants; however, the mobisite and helpdesk were underutilized owing to a lack of information and training around these platforms. Enhanced marketing and training initiatives that optimize existing social networks, as well as the provision of data and Wi-Fi, should be explored to ensure that registration improves, and that users are active across all platforms.

(JMIR Mhealth Uhealth 2019;7(2):e11644) doi:10.2196/11644

KEYWORDS

evaluation; mHealth; mobile phone; MomConnect; NurseConnect; South Africa

Introduction

As a middle-income country, South Africa has committed to achieving the maternal and child health targets set out in both the Millennium Development Goals and Sustainable Development Goals (SDGs) [1,2]. Over the past decade, great strides have been made, which include the introduction of MomConnect by the South African National Department of Health (NDOH) in 2014 [3,4]. MomConnect is a mobile health (mHealth) solution created to improve and promote maternal health services in South Africa by providing pregnant mothers with free stage-based messaging and a helpdesk, while also creating a universal pregnancy registry [5,6].

Strong government leadership and high mobile penetration [7] have led to over 2 million registered MomConnect users; this success prompted the development of NurseConnect, a complementary platform for health care workers in the maternal and child health space [6]. Like MomConnect, NurseConnect provides its users with short message service (SMS) text messages and access to a helpdesk. In addition, users have access to a mobile website (mobisite) with comprehensive information on themes introduced by the SMS text messages [5,8]. Since its launch in 2016 [9], NurseConnect has >23,000 registered users (Personal communication Jane Sebidi; based on extract from District Health Information System, July 5, 2018), leading the NDOH to consider expanding the platform to engage nurses from all departments, not just maternal and child health [5].

Globally, the World Health Organization (WHO) has identified the health workforce as one of its 6 building blocks for health system strengthening [10], while South Africa's National Development Plan 2030 highlighted primary health care, with a focus on training and mentoring nurses to build capacity and increase job satisfaction [11]. In 2015, the NDOH incorporated the need for health system strengthening into the South African mHealth Strategy 2015-2019 and recommended that mHealth should be utilized to strengthen health human resources [12], which is what NurseConnect has been developed to do. Many studies have shown the positive effects of mHealth interventions on health care workers in low- and middle-income countries [13-15]; however, many of these platforms were pilot projects that failed to scale up successfully [16,17]. With the NDOH committed to expanding NurseConnect to all nurses across the country [5], and the initial high enrollment numbers [18], NurseConnect has the potential to be successfully scaled up, but the current platform and user experience must be evaluated to identify possible challenges going forward [13,18]. Therefore, a mixed-method process evaluation was undertaken. In this paper, we present the qualitative aspects only aimed at evaluating the perceptions and use of NurseConnect by health care workers to provide feedback to optimize the user experience to inform the expansion of NurseConnect.

Methods

The NurseConnect evaluation in this paper specifically refers to the collection and evaluation of qualitative data from focus group discussion (FGDs) and in-depth interviews (IDIs) with nurses and midwives from selected facilities across South Africa. Of note, this paper does not evaluate any quantitative data pertaining to the outcomes or effectiveness of the NurseConnect platform, as these finding will be the focus of a separate publication.

NurseConnect Content

The NurseConnect platform was based on the Integrated Behavioral Model and Adult Learning Theory of change, where evidence indicates that engagement is the key to absorbing information from Web-based learning situations. Health care worker training and mentoring specialists compiled and designed the content with input from local doctors and nurses in the maternal health field. SMS text messages were presented in concise, simple language 2-3 times a week and often contained links to expanded papers on the mobisite. The mobisite could also be reached directly from the internet, and the helpdesk could be activated by responding to any of the SMS text messages. More recently, where nurses have smartphones, SMS text messages have been replaced by WhatsApp messages.

The content was divided into 2 main categories—informational and motivational. The informational content aimed to improve users' knowledge of maternal and child health, while the motivational content aimed to inspire users to make small actionable changes to increase productivity and happiness in their work. Figure 1 displays example messages, as well as a screenshot of the NurseConnect landing page.

Setting

To minimize bias, we randomly chose 18 facilities to equally represent a national population by ensuring that all provinces, types of facilities (ie, hospitals, clinics, and community health centers [CHCs]), and regions (ie, urban, periurban, and rural) were included. Figure 2 details the names and locations of these 18 facilities.

After provincial and district approval were obtained, facility visits were scheduled. Each facility was asked to provide a group of 6-8 registered staff to participate in an FGD or IDI, as well as a private room for the discussions to take place. Each site visit lasted from half a day to 3 days. Each FGD and IDI took between 15 minutes and 1 hour, depending on the operational demands of the facility and the active participation of the staff.

Figure 1. NurseConnect sample messages and mobisite.

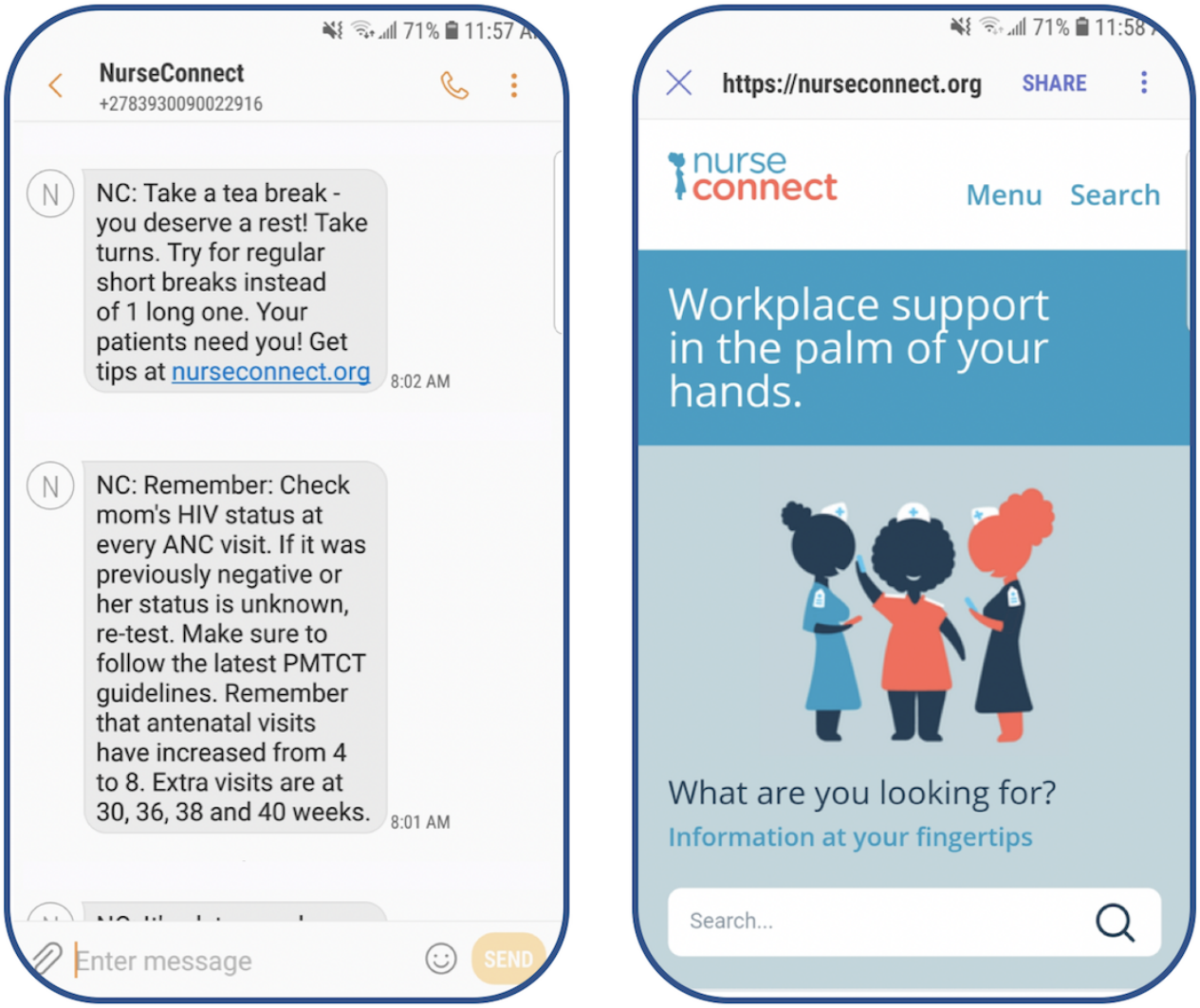
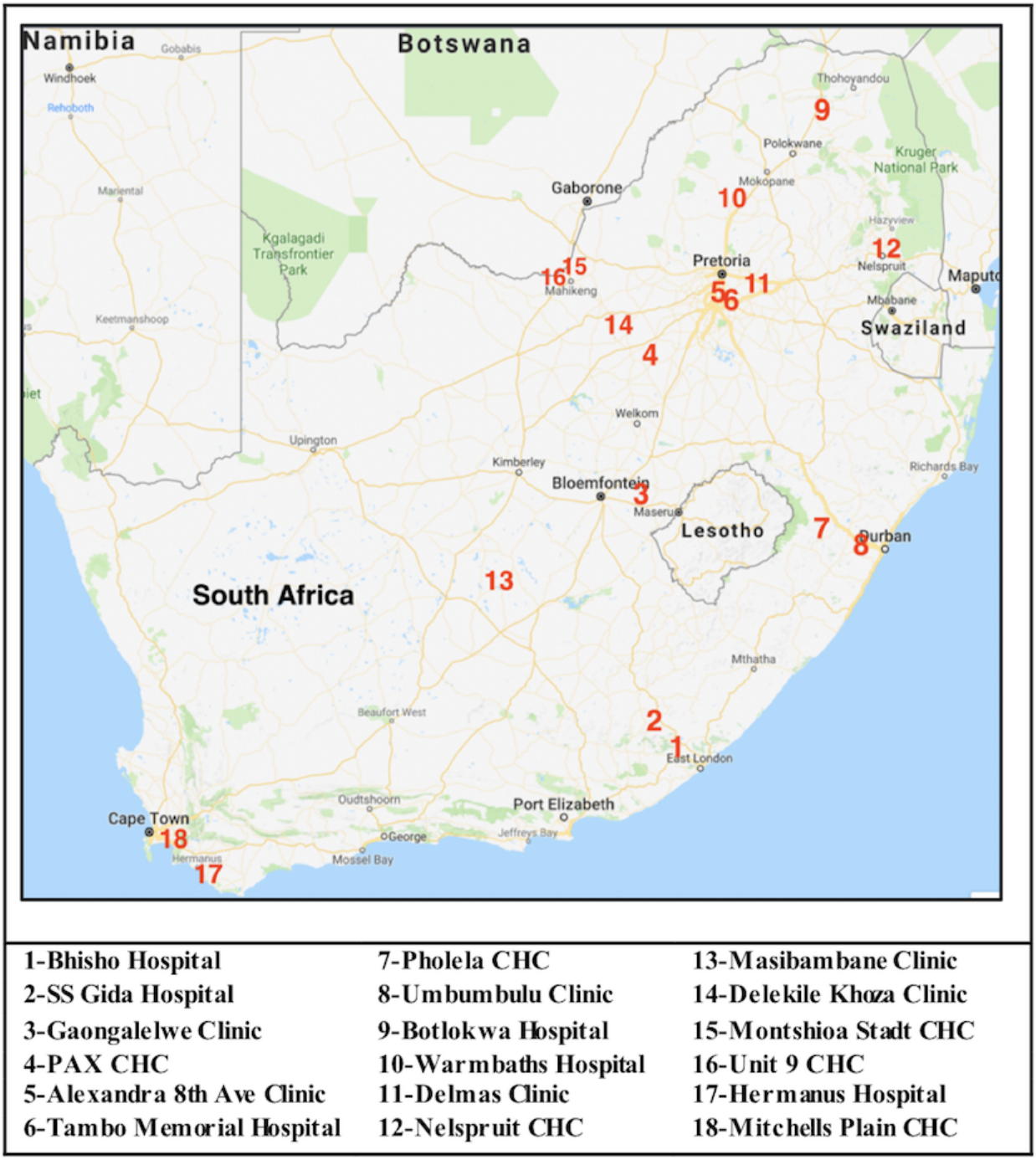


Figure 2. Facility locations. Ave: avenue; CHC: community health center.



Data Collection

All 18 approached facilities consented to participate in the research, and the site visits were conducted between December 12, 2017 and April 10, 2018. Convenience sampling provided a total of 110 nurses and midwives, who participated in the focus groups and interviews.

Using a pretested and piloted guide, 2 experienced moderators facilitated the FGDs and IDIs, with the principal investigator being present for 15 of the 18 site visits. The decision to conduct an FGD or IDI depended on the number of available NurseConnect registered participants present at the facility during the data collection visit. The FGDs and IDIs were audiorecorded so that the conversations could later be transcribed, and the facilitator also documented notes after discussions. Although an option of answering in vernacular was presented, participants elected to speak English.

Data Analysis

A total of 18 FGDs and 9 IDIs were conducted during the site visits. All these recordings were transcribed by the principal investigator into Microsoft Word documents. A selection of transcripts was randomly selected and independently verified by another team member to ensure accuracy. A code list agreed upon by the evaluation team was constructed to define relevant codes and emergent themes. The transcripts were then uploaded

to MAXQDA V 1.2 (Verbi Software) and investigated with the code list. All transcripts were coded by the principal investigator, while a selection of transcripts was independently coded by another study team member to ensure consistency. Upon completion, each code was concentrated into data reduction tables, then further refined into summary tables for reporting. Data were summarized into the themes of mobile use, registration, platforms, and user experience and content.

Ethical Consideration and Approval

Ethics approval for this evaluation was obtained from the University of the Witwatersrand Human Research Ethics Committee (M106976) on October 21, 2016. Participation in the data collection was voluntary. Consent forms were signed for both participation and voice recording.

Results

Mobile Phone Use

All 110 participants stated that they had personal mobile phones; 4 also had access to a mobile phone provided by work. Two participants had basic phones, some had feature phones (access to data and some apps like the internet and WhatsApp), and the majority used smartphones.

The most common personal use of phones was for communication, with the majority of participants using WhatsApp (often in group chats), SMS text messages, and phone calls to stay in touch with their friends and family, as well as their work and their children's schools. In addition, Facebook was a popular form of communication and networking, as participants could interact with nursing communities on a national or global scale such as the International Midwives Page. Furthermore, participants used the internet and Google, as well as a variety of apps, which included the camera, banking, GPS or navigation, voice or video recordings, educational apps, gaming, and YouTube (Figure 3).

Some participants stated that high data costs prevent them from using their phones for work, but most participants were active phone users and took advantage of this resource to help them with their health care in a variety of ways (Figure 3). Communication was the main use, with many participants calling or messaging their colleagues when they needed assistance. Most interactions on WhatsApp were made in groups though, with many participants belonging to a number of health-related groups; these groups consisted of colleagues, as well as a variety of other health care professionals, and these different clusters allowed participants to share media, ask questions, and support each other.

> *So, we use WhatsApp. We just shoot the picture, neh, of the X-ray, then we just send it to that doctor, maybe the specialist, and then that specialist will respond.* [Nurse, rural hospital]

In addition, participants frequently used their phones to look up clinical information about symptoms, diagnoses, and medications, stating that they were generally able to find the information that they required. The main source of information was the internet through Google; however, some participants were weary of the information found over the Web and preferred to use trusted medical apps, like the NDOH guidelines, iTriage, Medscape, Drugs.com, and Drugs Dictionary.

> *Websites, you know, the internet is full of…it's not authentic, because they are not journals per se. So, you need things that are authentic because we are dealing with people, you know what I mean, not robots here. We don't do trial and error here. We need to be 100% correct.* [Nurse, urban clinic]

NurseConnect

Registration

Of 110 participants, 62 (56.4%) were registered to NurseConnect, and many of them had been registered for over a year (Table 1). The absence of more registered participants was mainly attributed to a lack of awareness of the NurseConnect platform, which included a number of health care workers who thought that NurseConnect and MomConnect were the same product.

> *With me, this is the first time. I even though you wanted to write about MomConnect, when the sisters said it was NurseConnect, so this was the first time.* [Nurse, rural CHC]

Figure 3. Common mobile phone uses. SMS: short message service.

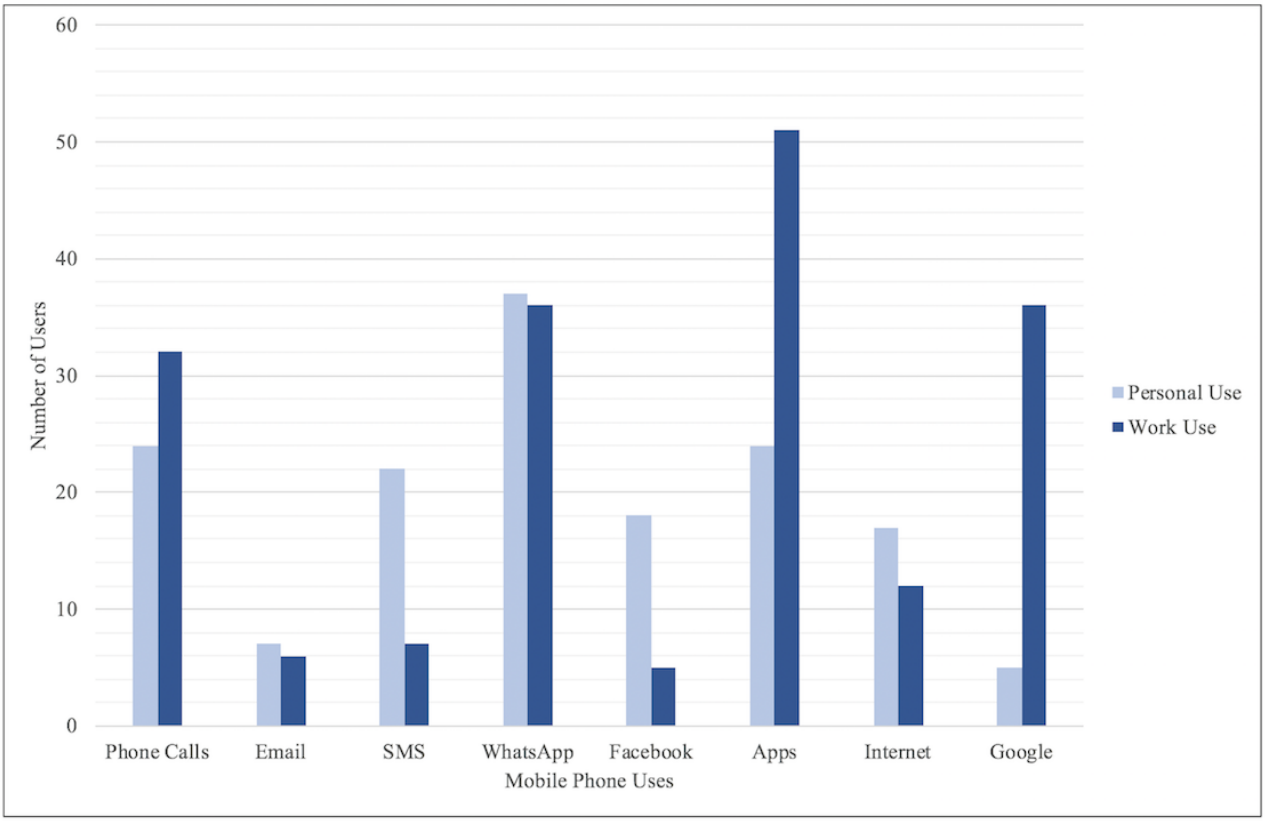


Table 1. Facilities sampled and the number of participants per facility.

Facility	Province	Total participants	NurseConnect registered participants, n (%)
Bhisho Hospital	Eastern Cape	2	2 (100.0)
SS Gida Hospital	Eastern Cape	15	6 (40.0)
Gaongalelwe Clinic	Free State	8	6 (75.0)
PAX Community Health Center	Free State	1	1 (100.0)
Alexandra 8th Avenue Clinic	Gauteng	1	1 (100.0)
Tambo Memorial Hospital	Gauteng	7	0 (0.0)
Pholela Community Health Center	KwaZulu-Natal	7	5 (71.4)
Umbumbulu Clinic	KwaZulu-Natal	13	12 (92.3)
Botlokwa Hospital	Limpopo	6	6 (100.0)
Warmbaths Hospital	Limpopo	7	7 (100.0)
Delmas Clinic	Mpumalanga	3	3 (100.0)
Nelspruit Community Health Center	Mpumalanga	2	2 (100.0)
Masibambane Clinic	Northern Cape	1	1 (100.0)
Delekile Khoza Clinic	North West	7	2 (28.5)
Montshioa Stadt Community Health Center	North West	6	0 (0.0)
Unit 9 Community Health Center	North West	8	2 (25.0)
Hermanus Hospital	Western Cape	13	3 (23.0)
Mitchells Plain Community Health Center	Western Cape	3	3 (100.0)
Total	Not applicable	110	62 (56.4)

The NurseConnect registration process was described as quick and easy by the majority of participants, with only a few of them stating that they had complications because of phone issues or difficulties with the clinic code. When asked about the initial

sign-up process, there seemed to be 2 main ways that participants became aware of NurseConnect and were signed up. The first was during visits from an external representative familiar with NurseConnect (ie, District Clinical Specialist Team Midwife, NDOH representative, Ideal Clinic Representative) who shared the registration process with staff during their visits. The second method of NurseConnect registration was when facilities had an internal NurseConnect representative that went to a NurseConnect training workshop or in-service, then brought the information back to connect their fellow staff members.

In addition, unregistered participants were probed about why they had never signed up, and most of them stated that they had never heard of NurseConnect before, or thought that it was referring to MomConnect. For those familiar with the platform, some of them were not working the day that their facility was registered, while others cited data and time constraints as the main barriers. Some participants were not allowed to have phones at work, while others mentioned message fatigue, lack of personal motivation, or being too junior as reasons not to register.

> *Like for instance, we are not being given, offered data. Like when you read message, it consumes data, and we are getting this every day. I would appeal to the Department to give us the data to those that are NurseConnected.* [Nurse, rural clinic]

Platforms and User Experience

All registered participants stated that they received the NurseConnect messaging, and all but one stated that they received as SMS text message. Participants liked receiving the SMS text messages because they were free and easy to access, even for people who were not very technology-savvy or did not have adequate data. Some participants saved the SMS text messages so that they could refer to them at a later date, while others forwarded the SMS text messages to friends and colleagues who were not registered.

> *I will say SMS, you just read it, you don't need data, you don't need anything. Your phone needs to be just charged, that's all.* [Nurse, urban hospital]

> *I sent it to my sister, I think it is the one she is talking about. The one with after work, the legs, my sister is a nurse as well, and working in maternity. So, I forwarded it to her.* [Nurse, periurban hospital]

Only 1 participant was receiving messages through WhatsApp, and none of the other participants knew that this was an option. When asked about the possibility of accessing NurseConnect through WhatsApp, participants were evenly divided, with some against the idea, while others were open to it. Participants who were open to using WhatsApp for NurseConnect stated that they liked the user-friendly interface and suggested that WhatsApp groups could be leveraged to create discussions and ask questions about topics, instead of the one-way flow of information from SMS text messages.

> *So, if we can then get connected on it, and then maybe open up a platform like a WhatsApp where people can then also interact about those messages and then take discussions further.* [Nurse, urban clinic]

Participants who did not want to use WhatsApp for NurseConnect stated that data costs were the main barrier, while others viewed WhatsApp as a personal platform and did not want to use it for work.

> *My WhatsApp is more like my personal, personal people only, like close family and friends. I prefer SMS.* [Nurse, urban hospital]

Of 62 registered NurseConnect users, only 2 stated that they had logged on to the NurseConnect mobisite. One participant stated that she enjoyed the mobisite and searched it for information when they were bored. Another participant stated that she followed the blue link from the SMS text message to get more information; however, she was not familiar with the term "mobisite" and had never explored the site. Four other participants stated that they had clicked on the NurseConnect links in the SMS text messages but were also not aware that the linked paper was on a NurseConnect mobisite. Some other participants stated that even if they were familiar with the site, data costs would still prevent them from using it.

> *We didn't hear about the mobisite. That means at the time when she went for the in-service, I don't think that the mobisite was mentioned.* [Nurse, rural clinic]

> *That's the big thing. That's the issue why we don't go to those "dot com." The data.* [Nurse, rural clinic]

None of the registered NurseConnect users who participated had ever used the helpdesk, and only a few of them knew that it existed. One person mentioned that she had received an SMS text message explaining that they could respond but had never tried. Another person had heard of the helpdesk but was under the impression that it would be a call center and did not know that it was active or that it could be accessed via SMS text messages.

> *We don't know whether to respond back, so we are not responding back. That is why I am saying it is one-sided. It is giving you the information, you don't have any chance to, to share your problem.* [Nurse, rural hospital]

The main barrier to the helpdesk and mobisite was that participants did not know they existed; however, once informed, participants liked the idea and said that they would like to try them. Another barrier to use that many participants mentioned was the data costs of accessing a mobisite or sending an SMS text message.

> *The problem is like sister said, we don't have data. That's the problem, it is why we don't respond to the messages. Because of money.* [Nurse, rural hospital]

> *I think that if we had free Wi-Fi here in the hospital, like if you want to access any information online you don't have to use your data because it's work related.* [Nurse, urban hospital]

Content

Participants generally liked the language, structure, and frequency of the NurseConnect messages, stating that they were

easy to understand and relevant to their daily work. In addition, they preferred English messages, as the medical content is easier to read and understand in English than traditional languages.

> *And the language is simple and straight forward. There is no bombastic words like when we were at schools. So, the messages are straight up.* [Nurse, rural clinic]

> *It is easy for me. Because it is English. And all of us speak English and it is, it's the terms that we are using in our profession.* [Nurse, rural hospital]

> *If it was like 20 messages a week, then I would get a bit upset because I would think it's too much. And it would just be taking up all the space on my phone, and I don't think I would read all of it then. Where if it was less messages every day, or every second day or third day, I would actually take the time and read it, and then maybe even forward it to someone.* [Nurse, urban hospital]

Messages about the listeriosis outbreak and motivation or stress relief were favored by many, and some participants stated that these messages created behavioral changes. The listeriosis information prompted one facility to conduct a full in-service based on the topic, while another facility improved their temperature-taking practices after receiving the messages. Some participants stated that the motivational messages led them to eat and rest properly during long shifts and before nightshifts, while others greatly appreciated the encouraging messages.

> *I can give an example, the other day, and I was so happy. I was, the whole week I was teaching patients on HIV, and when the message comes and say your patients are happy about the information that you give them, I was like wow, thank you, talking to me. And I was happy!* [Nurse, periurban hospital]

Some participants suggested that the topics be expanded to include pediatrics, chronic diseases, geriatrics, and mental health; however, they generally found the messages very informative and helpful. Furthermore, participants suggested that the government could share updated policies, guidelines, and nursing standards through the NurseConnect platform.

> *What I noticed is that I think you mostly on babies and pregnant women, but there's nothing about paediatrics, chronic diseases, of geriatrics.* [Nurse, rural clinic]

> *There is no need for us to be going on apps and downloading things, let the government give us what they want us to have. The latest information, you know.* [Nurse, urban hospital]

Discussion

Principal Findings

This is the first report describing and evaluating the rollout, uptake, and utilization of NurseConnect in South African public health facilities and the findings show that there were no marked variations between nurses from different provinces, facility types, or regions. High smartphone saturation, coupled with the nurses' use of WhatsApp, the internet, and various medical apps, reinforced the need to provide quality mHealth solutions that are tailored to the local context. The smartphone infiltration and extensive use of WhatsApp for personal and medical communication suggested that WhatsApp should be encouraged as a NurseConnect channel. To distinguish this platform from SMS text messaging, media messages (voice notes, pictures, and video clips) could be offered as a way to enhance the user experience.

Only 56.4% (62/110) of participants were registered on NurseConnect. The nurses who were not registered exposed gaps that should be addressed to successfully scale up the platform toward 2020 [5,9]. An extensive mixed media marketing campaign should be used to create awareness for the platform, while existing social networks on Facebook and WhatsApp should also be leveraged to spread awareness [5]. The need for marketing initiatives and platform training was reinforced by the lack of mobisite and helpdesk use. Many participants stated that they were not aware of these platforms before the site visits were conducted. Participants thought these platforms would be useful; however, the term "mobisite" was unfamiliar with participants, and "helpdesk" was associated with call-in services, not a 2-way messaging service. Some participants did not trust information found over the Web and preferred using reliable apps, so rebranding the mobisite as a mHealth app, could increase user activity.

> *I think for an example, in Facebook, there are so many groups including those who are nurses, midwives, primary health care. So, maybe one could paste on Facebook to say there is NurseConnect, this is what they say guys, sign in and let's get, let's make it viral.* [Nurse, urban CHC]

Similar to the internal NurseConnect representatives, who registered their colleagues, high-level users can be identified on these networks and trained as ambassadors or mentors to create demand for NurseConnect, facilitate discussions, and assist with registrations [11]; this bottom-up approach should foster peer-to-peer interactions, while also providing users with a contact to facilitate awareness and training for new platforms [16].

Aside from the lack of awareness of the NurseConnect platform, the biggest barrier to NurseConnect registration, uptake, and utilization remains the cost of data. Data cost was mentioned in almost every part of the discussions, as South Africa has the highest data charges out of the 6 largest Telcom markets in Africa. Data charges in South Africa have been the cause of much debate and protests by the majority of citizens not just these health care workers, making this a very relevant practical (possibly longer term) challenge, which needs to be addressed [19]. As the NDOH prepares to scale up the provision of facility mobile phones, subsidized data or airtime or free Wi-Fi in health care facilities will need to be considered.

The expansion of NurseConnect to all nurses also provides an opportunity to develop and tailor new message streams beyond maternal and child health. Participants suggested incorporating information about noncommunicable diseases and mental health,

which are in-line with the WHO country strategy for South Africa [20].

Limitations

This study has some limitations. While we endeavored to collect information from as many NurseConnect registered users as possible, we ended up with a smaller number than anticipated owing to the health care worker confusion around NurseConnect and MomConnect. Although the identification of the confusion was a valuable finding in itself, it did limit the amount of information we could collect about the NurseConnect implementation. Second, we did not collect sociodemographic information, which may have assisted in contextualizing some of the results or opinions presented. Third, although we collected information from different types of health care facilities across the country, the results presented are not conclusively representative of the entire NurseConnect registered health care population and may lack some differing opinions and experiences of users from facilities that did not participate in the evaluation. Finally, this was a qualitative evaluation that independently investigated health care workers' perceptions and uses of NurseConnect, without examining the efficacy and outcomes of the platform.

Conclusions

After discussing NurseConnect with nurses from all 9 provinces of South Africa, it is evident that the platform is not just useful but well-liked by nurses and midwives who use it. Participants found the messages informative, and their suggestions on expanding the content showed that they have been engaging with the platform and felt that it would continue to help them learn and grow. While this positive feedback reinforced the fact that this platform works, this evaluation has also identified gaps and areas of improvement that can be turned into strengths as the NDOH continues to rollout NurseConnect to all nurses across the country by 2020.

Across many facilities, it was discovered that the NurseConnect brand needs strengthening. The lack of awareness surrounding the WhatsApp messages, mobisite, and helpdesk presents an excellent opportunity for growth, as many participants liked these ideas but were previously unaware that platforms other than SMS text messaging existed. The other main barriers to these platforms were time constraints and data costs. The promotion of a zero-rated app and messaging, as well as free Wi-Fi for facilities, would help overcome the data barrier. By combating these barriers and leveraging existing mobile social networks to increase registration and awareness, NurseConnect should continue to develop as the platform becomes available to all nurses across the country.

Acknowledgments

The authors thank all nurses and midwives from each of the involved health care facilities who provided their time to participate in the evaluation. We also acknowledge the managers and administrators from the district and provincial Departments of Health who facilitated the regulatory processes and scheduling of the site visits. In addition, we thank Siphamandla Gumede and Sithembiso Ndlovu from Wits Reproductive Health and HIV Institute for their assistance with FGD facilitation, as well as Lauren Kotze and Perez Mnkile from Praekelt.

NurseConnect is a program of the NDOH. NurseConnect has also been funded by the Discovery Foundation, Johnson and Johnson, and the Elma Foundation.

Authors' Contributions

STLE designed the study and was involved in the data cleaning and analysis. AEF collected data, was involved in the data cleaning and analysis, and wrote the initial draft of the manuscript. All authors critically reviewed and approved of the final draft.

Conflicts of Interest

PB is a technical assistant in the NDOH and JS also works for NDOH. All other authors have no conflicts of interest to declare.

References

1. United Nations. Transforming our world: The 2030 agenda for sustainable development. 2015 Oct 27 Presented at: United Nations Headquarters; 2015/10/27; New York URL: https://sustainabledevelopment.un.org/content/documents/21252030%20Agenda%20for%20Sustainable%20Development%20web.pdf
2. National Department of Health, South Africa. Strategic Plan 2015-2020. Pretoria: National Department of Health, South Africa; 2015.
3. Chola L, Pillay Y, Barron P, Tugendhaft A, Kerber K, Hofman K. Cost and impact of scaling up interventions to save lives of mothers and children: taking South Africa closer to MDGs 4 and 5. Glob Health Action 2015;8:27265 [FREE Full text] [doi: 10.3402/gha.v8.27265] [Medline: 25906769]
4. Mehl G, Vasudevan L, Gonsalves L, Berg M, Seimon T, Temmerman M. Harnessing mHealth in low-resource settings to overcome health system constraintsachieve universal access to healthcare. In: Marsch L, Lord S, Dallery J, editors. Behavioural Healthcare and Technology: Using Science-Based Innovations to Transform Practice. Oxford, UK: Oxford University Press; Dec 08, 2014:239-259.

5. Praekelt. MomConnect for Nurses and Midwives Programme Status Report (Quarter 2) 1 April 2017- 30 June. Johannesburg: Praekelt; 2017.
6. Seebregts C, Dane P, Parsons AN, Fogwill T, Rogers D, Bekker M, et al. Designing for scale: optimising the health information system architecture for mobile maternal health messaging in South Africa (MomConnect). BMJ Glob Health 2018 Apr;3(Suppl 2):e000563 [FREE Full text] [doi: 10.1136/bmjgh-2017-000563] [Medline: 29713506]
7. GSMA. The Mobile Economy Sub-Saharan Africa. London: GMSA; 2017.
8. National Department of Health, South Africa. Johannesburg: Praekelt Foundation; 2016. NurseConnect Fact Sheet URL: https://www.knowledgehub.org.za/system/files/NurseConnect%20Fact%20Sheet.pdf?file=1&type=node&id=34663&force=0 [accessed 2018-03-20]
9. National Department of Health, South Africa. 2018. NurseConnect URL: http://www.health.gov.za/index.php/mom-connect?id=369 [accessed 2018-07-20] [WebCite Cache ID 71307rVAE]
10. World Health Organization. Monitoring the Building Blocks of Health Systems: A Handbook of Indicators and Their Measurement Strategies. Geneva: WHO Press; 2010.
11. National Planning Commission, South Africa. National Development Plan 2030 Our Future-make it work. Pretoria: Sherino Printers; 2012:329-359.
12. National National Department of Health, South Africa. mHealth Strategy 2015-2019. Pretoria: National Department of Health; 2015.
13. Agarwal S, LeFevre AE, Lee J, L'Engle K, Mehl G, Sinha C, WHO mHealth Technical Evidence Review Group. Guidelines for reporting of health interventions using mobile phones: mobile health (mHealth) evidence reporting and assessment (mERA) checklist. BMJ 2016 Mar 17;352:i1174. [doi: 10.1136/bmj.i1174] [Medline: 26988021]
14. Anstey Watkins JOT, Goudge J, Gómez-Olivé FX, Griffiths F. Mobile phone use among patients and health workers to enhance primary healthcare: A qualitative study in rural South Africa. Soc Sci Med 2018 Dec;198:139-147 [FREE Full text] [doi: 10.1016/j.socscimed.2018.01.011] [Medline: 29335160]
15. Free C, Phillips G, Watson L, Galli L, Felix L, Edwards P, et al. The effectiveness of mobile-health technologies to improve health care service delivery processes: a systematic review and meta-analysis. PLoS Med 2013;10(1):e1001363 [FREE Full text] [doi: 10.1371/journal.pmed.1001363] [Medline: 23458994]
16. Agarwal S, Perry HB, Long L, Labrique AB. Evidence on feasibility and effective use of mHealth strategies by frontline health workers in developing countries: systematic review. Trop Med Int Health 2015 Aug;20(8):1003-1014 [FREE Full text] [doi: 10.1111/tmi.12525] [Medline: 25881735]
17. Hall CS, Fottrell E, Wilkinson S, Byass P. Assessing the impact of mHealth interventions in low- and middle-income countries – what has been shown to work? Global Health Action 2014 Oct 27;7. [doi: 10.3402/gha.v7.25606]
18. National Department of Health, South Africa. MomConnect Nurses and Midwives Performance Support Platform; Country Programme: South Africa; Monitoring, Evaluation and Reporting Plan. Pretoria: National Department of Health; 2015.
19. researchICTafrica.net. Cape Town: Research ICT Africa; 2017. State of prepaid market in South Africa: Submission to parliament of South Africa on the Cost to Communicate in South Africa URL: https://www.ellipsis.co.za/wp-content/uploads/2016/09/cost_to_communicate_research_ict_africa.pdf [accessed 2018-12-07] [WebCite Cache ID 74UoUUa5i]
20. World Health Organization. WHO Country Cooperation Strategy 2016-2020 South Africa. Congo: WHO Regional Office for Africa; 2016.

Abbreviations

CHC: Community Health Center
FGD: focus group discussion
IDI: in-depth interviews
NDOH: National Department of Health
SMS: short message service
WHO: World Health Organization

Edited by G Eysenbach; submitted 20.07.18; peer-reviewed by H Chen, P Zhang; comments to author 08.10.18; revised version received 26.10.18; accepted 26.10.18; published 13.02.19.

Please cite as:
Fischer AE, Sebidi J, Barron P, Lalla-Edward ST
The MomConnect Nurses and Midwives Support Platform (NurseConnect): A Qualitative Process Evaluation
JMIR Mhealth Uhealth 2019;7(2):e11644
URL: http://mhealth.jmir.org/2019/2/e11644/
doi:10.2196/11644
PMID:30758298

Original Paper

Women Using Mobile Phones for Health Communication Are More Likely to Use Prenatal and Postnatal Services in Bangladesh: Cross-Sectional Study

Shangfeng Tang[1], PhD; Bishwajit Ghose[1,2], PhD, MPH; Md Rakibul Hoque[3], PhD; Gang Hao[4], PhD; Sanni Yaya[2], PhD

[1]School of Medicine and Health Management, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China

[2]School of International Development and Global Studies, University of Ottawa, Ottawa, ON, Canada

[3]Department of Management Information Systems, Faculty of Business Studies, University of Dhaka, Dhaka, Bangladesh

[4]Zunyi Medical and Pharmaceutical College, Guizhou, China

Corresponding Author:
Bishwajit Ghose, PhD, MPH
School of International Development and Global Studies, University of Ottawa
120 University Private
Ottawa, ON, K1N5J1
Canada
Phone: 1 6132624856
Email: brammaputram@gmail.com

Abstract

Background: The public health system in Bangladesh has been struggling to provide coverage and utilization of basic maternal health care services in pursuit of achieving maternal and child mortality-related goals. Interestingly, the rapid popularization of mobile technology in the country is transforming the landscape of health care access and delivery. However, little is known regarding the use of mobile phones from the perspective of maternal health care service utilization.

Objective: In this study, we aimed to investigate the prevalence and sociodemographic pattern of mobile phone use for health services among women and relationship between the use of mobile phone use and the uptake of essential maternal health services (MHSs).

Methods: Cross-sectional data from the Bangladesh Demographic and Health Survey on 4494 mothers aged between 15 and 39 years were used in the analysis. Using mobile phones to get health services or advice was hypothesized to have a positive association with the uptake of basic MHSs (antenatal care, ANC, facility delivery services, postnatal care) and postnatal care for the newborn. Data were analyzed using bivariate and multivariable techniques.

Results: More than a quarter (1276/4494, 28.4%; 95% CI 26.8-30.3) of the women aged 15-39 years reported using mobile phones to get health services with significant sociodemographic variations in the use of mobile phones. Analysis of the specific purposes revealed that, in most cases, mobile phones were used to contact service providers and consult with the same about what to do, whereas a smaller proportion reported using mobile phone for the purposes of arranging money and transportation. Multivariable analysis showed that compared with respondents who reported not using mobile phones for health care services, those who used them had higher odds of making 3+ ANC visits and delivering at a health facility. The odds were slightly higher for rural residents than for those in the urban areas.

Conclusions: The findings of this study conclude that women who use mobile phones are more likely to use ANC and professional delivery services than those who do not. More in-depth studies are necessary to understand the mechanism through which mobile phone-based services enhance the uptake of maternal health care.

(JMIR Mhealth Uhealth 2019;7(2):e10645) doi:10.2196/10645

KEYWORDS

antenatal care; facility delivery services; postnatal care; mHealth; Bangladesh

Introduction

The demography of Bangladesh is usually characterized by total fertility rate and maternal and child mortality [1,2]. Like all other south Asian nations, widespread poverty, poor health care infrastructure, inequality in access to care, and low health literacy constitute some of the major causes of underutilization of essential maternal health care services (eg, antenatal care, ANC, and professional services for childbirth) that translate to higher burden of maternal and child mortality, which is recognized as a serious public health problem in the country [1,3,4]. Over the past few decades, the Ministry of Health and Family Welfare (MoHFW) has developed and implemented strategies for tackling high maternal and child mortality rates through programmatic intervention within national (Population and Health Policy), international (Millennium Development Goal, 5) [5], and joint (Maternal and Neonatal Health Initiatives in Bangladesh) [6] policy frameworks. Millennium Development Goals progress reports reveal significant gains in terms of decline in both maternal and child mortality rates with an estimated reduction of 57% in child mortality and 66% in maternal mortality [1]. The statistics on essential maternal health service uptake (Millennium Development Goal 5) on the other hand appear to be less encouraging because the figures remain well behind the internationally agreed targets: increase in facility delivery between 2001 (9%) and 2010 (23%) and a 25% increase in ANC uptake between 1996-1997 (29%) and 2010 (54%) [7]. Despite the growing public and private sector initiatives to promote the utilization of maternal health services (MHSs), progress is slackened by persistent issues, for example, the lack of human resources in health care, high out-of-pocket expenditure, and inequitable access to quality services, especially among the marginalized and rural population [1].

Amid this multitude of challenges, the hope is that the wide availability and progressive expansion of mobile and internet technology could open windows for more expansive and equitable care delivery to communities otherwise deprived from accessing basic health care. It is no longer a new concept that telemedicine or the application of information and communication technology in the public health sector can substantially improve the quality of care and efficiency in administrative and managerial tasks in a cost-effective manner [8-10]. In developed countries, telemedicine has so far earned massive popularity as a convenient, cost-effective, and time-saving platform for almost all aspects of preventive, curative, and rehabilitative care including obstetrics and gynecology, psychiatry, and self-management of chronic noncommunicable diseases. In contrast, mobile health (mHealth) technology is still in its infancy in Bangladesh and other south Asian nations but has been gaining growing attention among health and policy experts. Evidence suggests that mHealth is highly effective in reducing financial and transportation barriers and facilitates basic MHSs as well as emergency care during emergency obstetric referrals in low-income settings [11,12]. Therefore, mHealth is increasingly seen as a key strategy to promoting maternal and child health in the developing countries [13-16].

So far, countries in south Asia, especially India and Bangladesh, have embarked on various electronic health (eHealth) initiatives (eg, electronic health records at the managerial level) and provision mHealth services as a means to ameliorate primary health care services at the population level [17,18]. Although the first eHealth initiative was taken by MoHFW in 1998, since then, the adoption of this technology has been proliferating mainly in the private sector where the common services include teleconsultation, prescription, and referral [19]. One of the pioneering private eHealth institution in Bangladesh called Telemedicine Reference Center Limited was established in 1999 with the goal of using mobile phones for health care delivery [20]. Services including computerization of health facilities with internet servers and mHealth service for communicating with health care providers were introduced in 2008 [17]. Mobile phone-based initiative for maternal and child care is more recent and was launched by Mobile Alliance for Maternal Action (locally known as Aponjon) in collaboration with MoHFW [21]. Service includes biweekly short message service (SMS) text messages or voice messages from the beginning of conception through the first birthday of the child on vital health information targeted at expectant and new mothers and their relatives [21]. The project became functional in 2012 with the vision of addressing the dire situation of maternal and child health services in the country.

Despite the emerging interests in this field, there is a noticeable absence of large-scale research studies that are necessary to generate the evidence base for informing investors in preventing the initiation of major pilot projects. In a previous study, we have shown the regional disparities among urban women in using mobile phone services for delivery services [1]. In this study, we aimed to investigate what percentage of women are using mobile phones for health services and whether seeking care through mobile phones has any relationship with their utilization of essential maternal and child health care services namely antenatal, facility delivery, and postnatal care for the mother and the newborn.

Methods

Survey and Data Source

We obtained the data for this study from the seventh round of Bangladesh Demographic and Health Survey (BDHS) 2014. This survey was implemented through a collaborative effort of the National Institute of Population Research and Training, Inner City Fund International, United States, and Mitra & Associates under the financial auspices of the United States Agency for International Development, Bangladesh. The main objectives of the surveys are to provide quality and nationally representative data on crucial health indicators needed for the monitoring and evaluation of national health programs and thereby assisting in policy making, designing public health programs in the country. BDHS adopted a 2-stage stratified technique for selecting sample households. In the first stage, 600 enumeration areas (EAs; primary sampling units with an average of about 120 households) were selected with probability proportional to the EA size with 207 EAs in urban areas and 393 in rural areas. In the second stage, a systematic sample of

30 households on average was selected per EA. This selection was expected to result in completed interviews with about 18,000 individual women, of whom 17,863 were finally interviewed yielding a response rate of 98%. For this study, we included only those participants who reported having a completed pregnancy in the years preceding the survey (Figure 1). Participants gave informed consent before taking part in the survey. All Demographic and Health Surveys are approved by Inner City Fund international and an Institutional Review Board in the host country to make sure that the protocols are in compliance with the US Department of Health and Human Services regulations for the protection of human subjects. A more detailed version of the methodology is published elsewhere [1,3].

Variables

The outcomes variables included 3 maternal variables: ANC, place of delivery, postnatal care, and 1 newborn care variable (postnatal care for the baby). These variables are described in Table 1 along with the list of independent variables. The timing of first ANC is an important aspect of MHSs, but it was not available for the 2014 survey.

Data Analysis

Data analyses were carried out using IBM SPSS Statistics version 24. Before analysis, the dataset was cleaned, tested for multicollinearity, and checked for cases that fulfilled the inclusion criteria, experience of at least 1 childbirth in the preceding 5 years, providing information on mobile phone usage. Owing to the clustered nature of the survey, we used a complex sample analysis technique that takes into account the sampling weight, strata, and clusters. Participants' demographic characteristics were presented using frequencies and percentages. Pearson chi-square tests were performed to examine the bivariate tests of association between the dependent and explanatory variables. Variables that showed association at $P<.1$ were retained for multivariable analysis. Binary logistic regression analyses were used to calculate the odds ratios of using the 3 types of maternal and postnatal care for newborns. Both crude and adjusted odds ratios were calculated. At this stage, P values (two-tailed) were considered statistically significant only when below .05.

Figure 1. Flowchart of the sampling procedure.

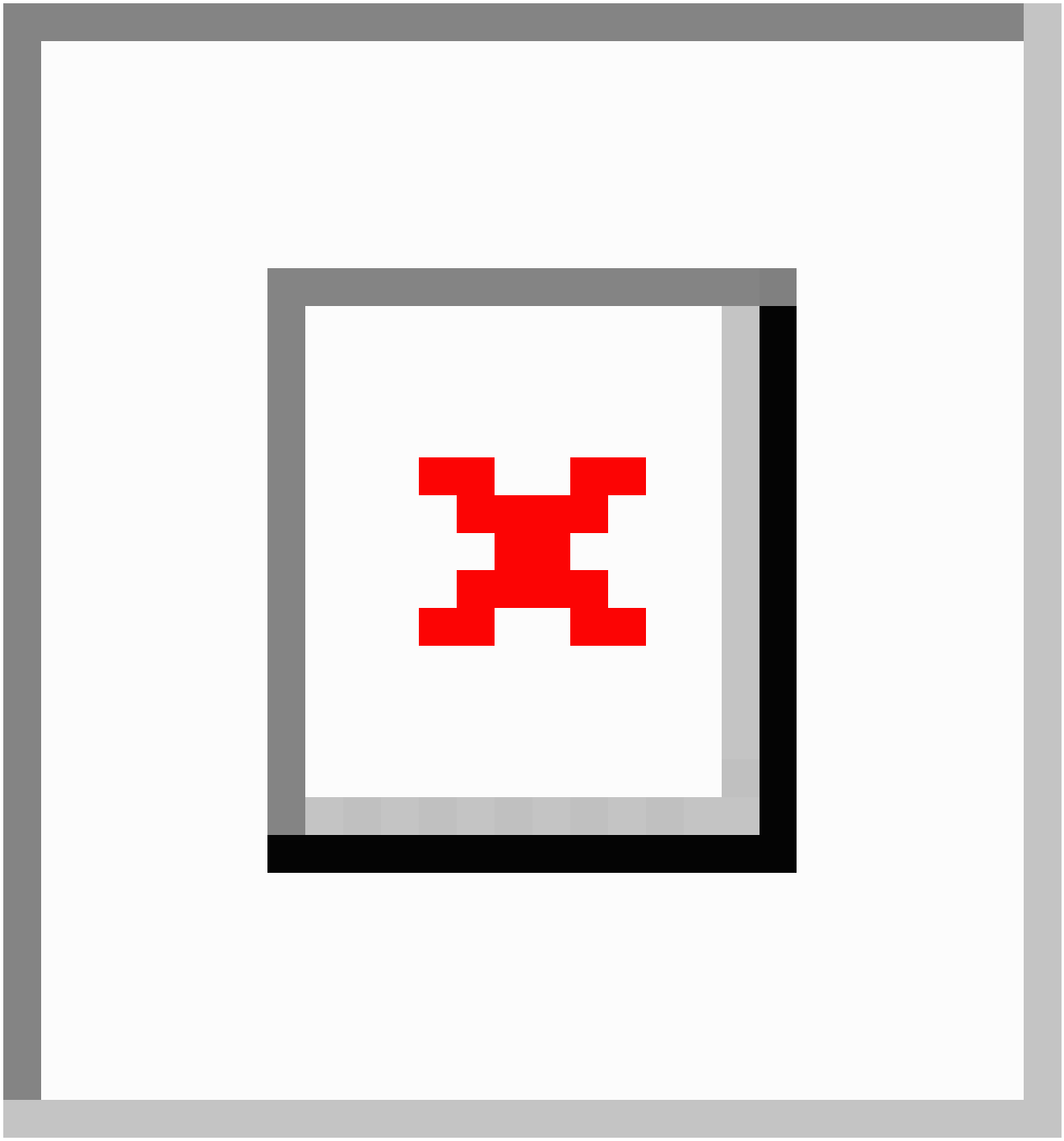


Table 1. Description of dependent and independent variables.

Variables	Description
Dependent variables	
Antenatal care visits	Number of antenatal care visits during the last pregnancy. It was categorized as adequate (4 and 4+) and inadequate (0-3)
Place of delivery	Place where respondents delivered the most. It was categorized as Home (home of the respondent or relatives) and Heath facility (eg, hospitals, clinics, maternal health centers)
Postnatal checkup for mother	Mother's health checkup by a professional after delivery (Yes or No)
Postnatal checkup for baby	Baby's health checkup by a professional after delivery (Yes or No)
Independent variables	
Age	Age of the respondent. The survey year was categorized into 5-year groups: 15-19, 20-24, 25-29, 30-34, and 35-39.
Division	Seven administrative regions: Barisal, Chittagong, Dhaka, Khulna, Rajshahi, Rangpur, and Sylhet
Residency	Type of place of current residence: Urban or Rural
Education	Educational attainment based on the number of years of formal schooling: Nil, Primary, Secondary, or Higher
Religion	Islam, Hinduism, or others
Wealth index	Calculated based on household possession of durable goods (eg, TV[a], radio, bicycle) and housing quality (eg, type of floor, wall, and roof). Each item is assigned a factor score generated through principal component analysis, which is then summed and standardized for the households. These standardized scores place the households in a continuous scale based on relative wealth scores. The scores were thus obtained from a continuous scale and subsequently categorized into quintiles to rank the household as Poorest, Poorer, Middle, Richer, or Richest.
Employed	Employment status of the respondent: Yes or No
Reads newspaper	Frequency of reading newspaper: weekly or less than weekly: yes, and do not read at all: No
Uses TV or radio	Frequency of using TV or radio: weekly or less than weekly (Yes) and do not use at all (No)

[a]TV: television.

Results

The analysis included 4494 women aged 15 to 39 years. The basic sociodemographic profile of the participants is presented in Table 2. The prevalence of attending 4+ ANC visits was 32.0% (1440/4494; 95% CI 29.5-33.2), facility delivery 40.1% (1801/4494; 95% CI 38.3-42.1), and postnatal checkup for the mother 65.7% (2953/4494; 95% CI 64.0-67.7) and for the baby 64.9% (2917/4494; 95% CI 64.0-67.6; not shown in the table). More than a quarter (1276/4494, 28.4%; 95% CI 26.8-30.3) of the women reported using mobile phones to get health services. Those who reported ever using a mobile phone to get health services had significantly higher (P=.01) percentages of making at least 4 ANC visits or higher (2543/4494, 56.6%; 95% CI 53.0-60.1) and delivering at a health facility (2435/4494, 54.2%; 95% CI 51.8-58.5). Percentages were also higher for postnatal checkup among mothers and babies but were not statistically significant.

Respondents were further asked to describe the specific purposes for using a mobile phone. As shown in Figure 2, in most cases, they used mobile phones to contact service providers and consult with the same about what steps to take, whereas a smaller proportion reported on the use of mobile phone for the purposes of arranging money and transportation. The percentage of use varied slightly between urban and rural residents; rural residents made phone calls more frequently than their urban counterparts.

Results of multivariate analysis on the association between the use of mobile phone and uptake of essential MHSs and postnatal care for newborn are presented in Table 3. Results indicate that compared with women who reported not using mobile phones for health care services, those who used had higher odds (adjusted) of achieving an adequate number (4 and 4+) of ANC visits (adjusted odds ratio, AOR 1.612, 95% CI 1.309-1.985) in both urban and rural areas. The odds were slightly higher for rural residents (AOR 1.700, 95% CI 1.317-2.194) than those in the urban areas (AOR 1.454, 95% CI 1.034-2.045). The same was true for delivering at the health facility. Users of mobile phone for health services in rural areas and urban areas had 1.7 times and 1.5 times higher odds, respectively, of delivering at a health facility.

Table 2. Sociodemographic profile of the participants (N=4494). Bangladesh Demographic and Health Survey 2014. All the percentages are weighted.

Variables	Participants (N=4494), n (%)	Ever used a mobile phone to get health services or advice				*P* value
		Yes		No		
		%	95% CI	%	95% CI	
Number of antenatal care visits						.001
0-3	3054 (68.00)	43.40	39.90-47.00	73.60	71.30-75.70	
≥4	1440 (32.00)	56.60	53.0-60.10	26.40	24.30-28.70	
Place of delivery						.001
Home	2693 (59.90)	45.80	42.50-48.20	61.40	58.10-64.70	
Heath facility	1801 (40.10)	54.20	51.80-58.50	38.60	35.30-41.90	
Postnatal checkup for mother						.13
No	1541 (34.30)	35.70	32.60-39.10	33.50	31.20-35.90	
Yes	2953 (65.70)	64.30	60.90-67.40	66.50	64.10-68.80	
Postnatal checkup for baby						.14
No	1577 (35.10)	36.50	33.30-39.70	33.30	31.10-35.50	
Yes	2917 (64.90)	63.50	60.30-66.70	66.70	64.50-68.90	
Age in years						.07
15-19	1531 (34.10)	21.00	18.50-23.70	21.0	19.10-23.00	
20-24	1147 (25.50)	33.20	30.10-36.60	33.70	31.60-35.90	
25-29	608 (13.50)	26.90	24.00-29.90	25.20	23.20-27.30	
30-34	939 (20.90)	14.50	11.50-18.20	13.40	12.10-15.00	
35-39	269 (60.00)	4.40	3.30-5.80	6.60	5.50-8.00	
Division						<.001
Barisal	532 (11.80)	6.30	4.90-8.00	5.60	4.70-6.60	
Chittagong	862 (19.20)	26.30	23.20-29.70	20.10	17.90-22.40	
Dhaka	795 (17.70)	32.50	28.60-36.70	36.40	33.30-39.70	
Khulna	531 (11.80)	8.30	7.00-9.90	7.90	6.90-9.00	
Rajshahi	546 (12.10)	9.70	8.10-11.50	10.20	8.90-11.60	
Rangpur	550 (12.20)	10.10	8.20-12.30	9.60	8.30-11.10	
Sylhet	678 (15.10)	6.70	5.10-8.80	10.30	8.50-12.30	
Residency						<.001
Urban	1451 (32.30)	31.20	27.90-34.60	24.10	22.00-26.40	
Rural	3043 (67.70)	68.80	65.40-72.10	75.90	73.60-78.00	
Education						<.001
Nil	607 (13.50)	6.50	5.00-8.40	17.20	15.40-19.20	
Primary	1235 (27.50)	200	17.30-22.90	31.10	28.90-33.50	
Secondary	2130 (47.40)	56.60	53.10-60.10	44.10	41.60-46.80	
Higher	522 (11.60)	16.90	14.50-19.60	7.50	6.50-8.70	
Religion						<.001
Islam	4134 (92.00)	91.40	89.20-93.20	91.80	90.10-93.30	
Hinduism or others	360 (8.00)	8.60	6.80-10.80	8.20	6.70-9.90	
Wealth index						<.001
Poorest	940 (20.90)	13.70	11.10-16.90	24.80	22.50-27.30	
Poorer	855 (19.00)	15.00	12.80-17.50	20.50	18.80-22.40	

Variables	Participants (N=4494), n (%)	Ever used a mobile phone to get health services or advice				*P* value
		Yes		No		
		%	95% CI	%	95% CI	
Middle	860 (19.10)	19.40	16.90-22.10	18.90	16.80-21.30	
Richer	946 (21.10)	22.40	19.50-25.70	19.90	18.10-21.90	
Richest	893 (19.90)	29.50	26.50-32.70	15.80	14.20-17.60	
Employed						.01
No	3511 (78.10)	79.80	76.60-82.70	74.90	72.70-77.10	
Yes	983 (21.90)	20.20	17.30-23.40	25.10	22.90-27.30	
Reads newspapers						<.001
No	3776 (84.00)	77.00	74.10-79.70	88.50	86.80-90.00	
Yes	718 (16.00)	23.00	20.30-25.90	11.50	10.00-13.20	
User of television or radio						<.001
No	1796 (40.00)	31.00	27.70-34.50	43.80	41.20-46.30	
Yes	2698 (60.00)	69.00	65.50-72.30	56.20	53.70-58.80	

Figure 2. Self-reported purposes (%) for using mobile phone among participants. BDHS 2014.

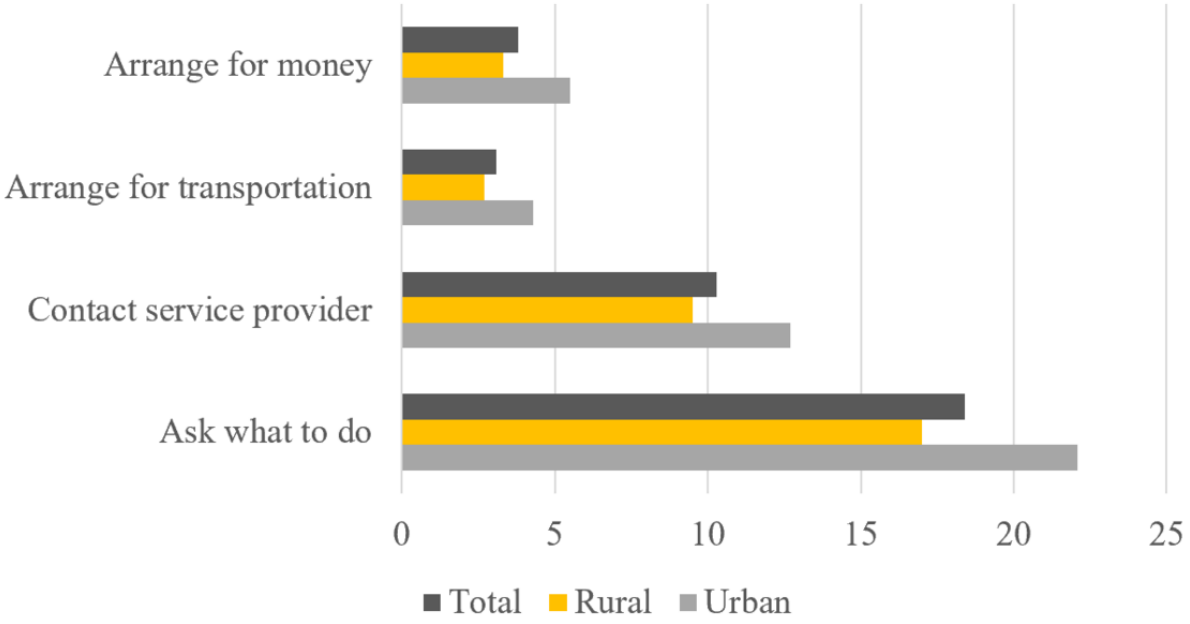


Table 3. Odds ratios of attending at least 4 antenatal care visits and facility delivery among those who reported using mobile phone for health care services.

Sample group	3+ antenatal care visits		Delivery at the health facility		Postnatal care (mother)		Postnatal care (newborn)	
	COR[a] (95% CI)	AOR[b] (95% CI)	COR (95% CI)	AOR (95% CI)	COR (95% CI)	AOR (95% CI)	COR (95% CI)	AOR (95% CI)
Overall	2.032 (1.771-2.567)	1.612 (1.309-1.985)	2.132 (1.771-2.567)	1.612 (1.309-1.985)	0.905 (0.757-1.081)	1.144 (1.001-1.308)	1.134 (0.992-1.295)	0.868 (0.728-1.036)
Urban	1.862 (1.387-2.502)	1.454 (1.034-2.045)	1.862 (1.387-2.502)	1.454 (1.034-2.045)	0822 (0.613-1.104)	0.931 (0.690-1.255)	0.904 (0.671-1.217	1.002 (0.737-1.360)
Rural	2.021 (1.686-2.746)	1.700 (1.317-2.194)	2.152 (1.686-2.746)	1.680 (1.210-2.094)	0.974 (0.781-1.214)	1.062 (0.842-1.339)	0.898 (0.723-1.116)	1.006 (0.837-1.208)

[a]COR: Unadjusted odds ratios.

[b]AOR: Adjusted for Age, Division, Residency, Education, Religion, Wealth index, Employment, Newspaper, TV or radio use.

Discussion

Principal Findings

In this study, we aimed to measure the prevalence of mobile phone use for health care services among adult women in Bangladesh and the association between its use and uptake of essential maternal health care services and postnatal care for newborns. Our findings indicate that little higher than a quarter of the respondents reported using mobile phones to get health services. Important sociodemographic variations were observed in the use of mobile phones. The percentage was noticeably higher among those aged between 20 and 29 years, those who were from Dhaka and Chittagong division, those who lived in rural areas, and those who had primary-secondary level education. The reason behind this could be that women of this age bracket are more likely to use mobile and internet technologies, be concerned about reproductive health, and take proactive measures to ensure better pregnancy outcomes. Presumably, the higher percentage in the rural areas can be indicative of lower concentration of health facilities or lower provider to patient ratio and consequently higher dependency on distant consultations over mobile phones. Even those with no educational attainment reported using mobile phones for seeking health care advices. Another interesting finding was that the rate of using mobile phones increased linearly with the wealth quintile of the households with the rate being lowest in poorest and highest among the wealthiest households, indicating the mediating role of financial well-being in the association. Mobile phone use may also play a strong, enabling role among women lacking access to urban facilities. As the findings further suggested, the odds of having 3+ ANC visits and health facility delivery in relation to mobile phone use were higher among rural women than urban women. Women who were employed had lower rates of mobile phone use for health care seeking purposes, perhaps because they are more empowered or likely to make direct contacts with health centers. Users of electronic and print media (TV, radio, and newspaper) were also more likely to use mobile phones for seeking health care. This relationship is an intuitive one as print and electronic media have become important tools for health communication especially in the areas of maternal and child health.

Bivariate analysis revealed that those who reported ever using mobile phones to get health services had significantly higher rates of attending the minimum recommended number of ANC visits, as well as delivering at a health facility, but not for postnatal checkup among mothers and babies. These findings were reaffirmed by multivariable analysis adjusted for potential confounding variables. The odds of having 3+ ANC visits and delivering at a health facility were significantly higher among the users of mobile phones for health care purposes. These associations were further investigated by stratifying for urban and rural areas. It appeared that the odds of availing ANC and facility delivery services were higher among rural residents than urban residents.

Past Studies and Future Research Directions

Mobile-based health service delivery strategies have attracted considerable attention as a means of supporting maternal, neonatal, and child health in developing countries [15,22,23]. There is a growing body of literature documenting the effectiveness of mHealth in improving the utilization of MHSs. However, there remains a dearth of observational and cohort studies on the effectiveness of mHealth in promoting MHSs in Bangladesh. Similar to the findings of this study, in a previous research, we have shown the disparities in the use of mobile phones for seeking childbirth services in urban Bangladesh and found that women using mobile phones for health care were more likely to deliver at a health facility [1]. A recent trial study in Ethiopia reported a positive contribution of SMS text message-based mobile phone intervention in most of the selected maternal and child health service indicators, such as improvement in the percentage of the recommended number of ANC visits, percentage of deliveries attended by health workers, and facilitating the work processes of the health workers in rural areas [24]. Another trial study made similar conclusions based on the evidence on the effectiveness of voice messages for the early initiation of MHSs [25]. However, to date, there is no evidence on the effectiveness of mHealth in promoting all the components of maternal and neonatal care. Future studies on mHealth within the context of maternal health care in Bangladesh should focus on a broader category of services and the impact of SMS text messaging services in promoting maternal health literacy and utilization of essential MHSs.

General Discussion

Despite increasing local and international commitments to reducing maternal and child mortality rates, each year, a remarkably high percentage of women are dying owing to pregnancy and childbirth complications, a large proportion of which could have been avoided through providing essential MHSs [26-28]. In low-income countries such as Bangladesh where health care systems are fraught with infrastructure and human resource crises, the application of mHealth technologies remains a huge untapped resource. Health care experts in Bangladesh are beginning to realize the potential of health technologies and are developing policies leveraging the widespread penetration of telecommunication market to address key public health issues especially in the areas of family planning, sexual, and reproductive health care. However, shifting from traditional mode of care delivery and management to technology-driven environments will require an investment in training telemedicine facilitators and a supply of skilled professionals [29-31]. One review article mentioned that limited internet bandwidth, high cost of infrastructure, and software development are some of the main barriers to the adoption of telemedicine in Bangladesh [17]. For capacity building in this sector, university and community clinic-based telemedicine training programs for providers and community dwellers can prove highly beneficial. Finally, reaping the benefits of these programs to promote health care especially among the marginalized population will require addressing the digital divide by making sure that people have physical access and the necessary skills to properly utilize the technologies [32,33].

Strengths and Limitations

As far as we are concerned, this has been the very first study to report the sociodemographic pattern of mobile phone use for health services among women and its relationship with the uptake of essential MHSs in any developing country. The survey was representative of the country because the sample population was selected from all the districts in the country. Given the low uptake of maternal health care services and expansion of the telecommunications sector in Bangladesh, studies of this kind are crucial to designing mHealth interventions for maternal and reproductive health.

Among the limitations was the cross-sectional and secondary nature of the data. There was no detailed information regarding the subjective report on the benefits of using mobile phone for service uptake. It is also possible that women using mobile phones were more aware of reproductive health and were socioeconomically more empowered, which are strong determinants of using health care services. There is also no indication regarding the quality of the services and whether the proportion of women who use mobile phones for health care purposes experience better reproductive outcomes. Variables were self-reported and therefore remain subject to reporting or recall error. Because the information on outcome and explanatory variables were collected at the same time, our analysis cannot suggest any causal relationship between the use of mobile phones and uptake of maternal health care services. Despite the limitations, this study offers several important insights for maternal health care programs in Bangladesh and calls for further research to investigate the quality of pregnancy outcomes among mobile phone users.

Conclusion

Based on the analysis of BDHS (2014), this study found that, currently, little higher than a quarter of women aged 15-39 years are using mobile phones to get health services. However, there are important sociodemographic patters in the use of mobile phones. The findings conclude that women who are using mobile phones are more likely to use antenatal and professional delivery services than those who do not. Currently, there is not enough evidence to confirm any strong connection between mobile phone use and uptake of maternal health care services. More studies are necessary to replicate these findings. Future studies should focus on measuring the potential of mHealth technologies in meeting national maternal health care objectives and health priorities of the population.

Acknowledgments

This study was supported by the Science and Technology Department of Guizhou Province “The service package research and survey for elderly in pension institutions of Zunyi City” and “The comparative research of public health performance among different health systems in rural area,” Grant no BS2016001.

Conflicts of Interest

None declared.

References

1. Bishwajit G, Hoque M, Yaya S. Disparities in the use of mobile phone for seeking childbirth services among women in the urban areas: Bangladesh Urban Health Survey. BMC Med Inform Decis Mak 2017 Dec 29;17(1):182 [FREE Full text] [doi: 10.1186/s12911-017-0578-2] [Medline: 29284477]
2. Bishwajit G. Nutrition transition in South Asia: the emergence of non-communicable chronic diseases. F1000Res 2015;4:8 [FREE Full text] [doi: 10.12688/f1000research.5732.2] [Medline: 26834976]
3. Ghose B, Feng D, Tang S, Yaya S, He Z, Udenigwe O, et al. Women's decision-making autonomy and utilisation of maternal healthcare services: results from the Bangladesh Demographic and Health Survey. BMJ Open 2017 Sep 07;7(9):e017142 [FREE Full text] [doi: 10.1136/bmjopen-2017-017142] [Medline: 28882921]

4. Bishwajit G, Tang S, Yaya S, Ide S, Fu H, Wang M, et al. Factors associated with male involvement in reproductive care in Bangladesh. BMC Public Health 2017 Dec 03;17(1):3 [FREE Full text] [doi: 10.1186/s12889-016-3915-y] [Medline: 28049458]
5. Sack DA. Achieving the millennium development goals in Bangladesh. J Health Popul Nutr 2008 Sep;26(3):251-252 [FREE Full text] [Medline: 18831221]
6. Haider MR, Rahman MM, Moinuddin M, Rahman AE, Ahmed S, Khan MM. Impact of maternal and neonatal health initiatives on inequity in maternal health care utilization in Bangladesh. PLoS One 2017 Jul;12(7):e0181408 [FREE Full text] [doi: 10.1371/journal.pone.0181408] [Medline: 28742825]
7. Chowdhury S, Banu LA, Chowdhury TA, Rubayet S, Khatoon S. Achieving Millennium Development Goals 4 and 5 in Bangladesh. BJOG 2011 Sep;118 Suppl 2:36-46 [FREE Full text] [doi: 10.1111/j.1471-0528.2011.03111.x] [Medline: 21951501]
8. Park Y, Han D. Current Status of Electronic Medical Record Systems in Hospitals and Clinics in Korea. Healthc Inform Res 2017 Jul;23(3):189-198 [FREE Full text] [doi: 10.4258/hir.2017.23.3.189] [Medline: 28875054]
9. White LA, Krousel-Wood MA, Mather F. Technology meets healthcare: distance learning and telehealth. Ochsner J 2001 Jan;3(1):22-29 [FREE Full text] [Medline: 21765713]
10. Park Y, Lee J. Factors affecting electronic medical record system adoption in small korean hospitals. Healthc Inform Res 2014 Jul;20(3):183-190 [FREE Full text] [doi: 10.4258/hir.2014.20.3.183] [Medline: 25152831]
11. Destigter K. mHealth and developing countries: a successful obstetric care model in Uganda. Biomed Instrum Technol 2012;Suppl:41-44. [doi: 10.2345/0899-8205-46.s2.41] [Medline: 23039775]
12. Colaci D, Chaudhri S, Vasan A. mHealth Interventions in Low-Income Countries to Address Maternal Health: A Systematic Review. Ann Glob Health 2016 Sep;82(5):922-935 [FREE Full text] [doi: 10.1016/j.aogh.2016.09.001] [Medline: 28283147]
13. Fiordelli M, Diviani N, Schulz PJ. Mapping mHealth research: a decade of evolution. J Med Internet Res 2013;15(5):e95 [FREE Full text] [doi: 10.2196/jmir.2430] [Medline: 23697600]
14. Chen H, Chai Y, Dong L, Niu W, Zhang P. Effectiveness and Appropriateness of mHealth Interventions for Maternal and Child Health: Systematic Review. JMIR Mhealth Uhealth 2018 Jan 09;6(1):e7 [FREE Full text] [doi: 10.2196/mhealth.8998] [Medline: 29317380]
15. Nurmatov U, Lee S, Nwaru B, Mukherjee M, Grant L, Pagliari C. The effectiveness of mHealth interventions for maternal, newborn and child health in low- and middle-income countries: Protocol for a systematic review and meta-analysis. J Glob Health 2014 Jun;4(1):010407 [FREE Full text] [doi: 10.7189/jogh.04.010407] [Medline: 24976966]
16. Feroz A, Perveen S, Aftab W. Role of mHealth applications for improving antenatal and postnatal care in low and middle income countries: a systematic review. BMC Health Serv Res 2017 Nov 07;17(1):704 [FREE Full text] [doi: 10.1186/s12913-017-2664-7] [Medline: 29115992]
17. Islam SMS, Tabassum R. Implementation of information and communication technologies for health in Bangladesh. Bull World Health Organ 2015 Nov 01;93(11):806-809 [FREE Full text] [doi: 10.2471/BLT.15.153684] [Medline: 26549909]
18. Powell AC, Ludhar JK, Ostrovsky Y. Electronic health record use in an affluent region in India: Findings from a survey of Chandigarh hospitals. Int J Med Inform 2017 Dec;103:78-82. [doi: 10.1016/j.ijmedinf.2017.04.011] [Medline: 28551005]
19. Ahmed T, Lucas H, Khan AS, Islam R, Bhuiya A, Iqbal M. eHealth and mHealth initiatives in Bangladesh: a scoping study. BMC Health Serv Res 2014;14:260 [FREE Full text] [doi: 10.1186/1472-6963-14-260] [Medline: 24934164]
20. Changemakers. 2011 Mar 11. Healthcare on demand URL: https://www.changemakers.com/innovations4health/nominations/telemedicine-reference-centre-ltd-healthcare-demand [accessed 2018-11-28] [WebCite Cache ID 74HbLy1sq]
21. HFG. 2015. Mobile Money for Health Case Study: Aponjon (MAMA Bangladesh) Internet URL: https://www.hfgproject.org/mobile-money-for-health-case-study-aponjon-mama-bangladesh/ [accessed 2018-11-28] [WebCite Cache ID 74HbKWQgO]
22. Tamrat T, Kachnowski S. Special delivery: an analysis of mHealth in maternal and newborn health programs and their outcomes around the world. Matern Child Health J 2012 Jul;16(5):1092-1101. [doi: 10.1007/s10995-011-0836-3] [Medline: 21688111]
23. Amoakoh-Coleman M, Borgstein A, Sondaal S, Grobbee D, Miltenburg A, Verwijs M, et al. Effectiveness of mHealth Interventions Targeting Health Care Workers to Improve Pregnancy Outcomes in Low- and Middle-Income Countries: A Systematic Review. J Med Internet Res 2016 Dec 19;18(8):e226 [FREE Full text] [doi: 10.2196/jmir.5533] [Medline: 27543152]
24. Atnafu A, Otto K, Herbst C. The role of mHealth intervention on maternal and child health service delivery: findings from a randomized controlled field trial in rural Ethiopia. Mhealth Internet 2017 Sep:14.
25. Gelano TF, Assefa N, Bacha YD, Mahamed AA, Roba KT, Hambisa MT. Effect of Mobile-health on maternal health care service utilization in Eastern Ethiopia: study protocol for a randomized controlled trial. Trials 2018 Feb 12;19(1):102 [FREE Full text] [doi: 10.1186/s13063-018-2446-5] [Medline: 29433537]
26. Koblinsky M, Anwar I, Mridha MK, Chowdhury ME, Botlero R. Reducing maternal mortality and improving maternal health: Bangladesh and MDG 5. J Health Popul Nutr 2008 Sep;26(3):280-294 [FREE Full text] [Medline: 18831225]
27. Zureick-Brown S, Newby H, Chou D, Mizoguchi N, Say L, Suzuki E, et al. Understanding global trends in maternal mortality. Int Perspect Sex Reprod Health 2013 Mar;39(1):32-41 [FREE Full text] [doi: 10.1363/3903213] [Medline: 23584466]

28. Sajedinejad S, Majdzadeh R, Vedadhir A, Tabatabaei M, Mohammad K. Maternal mortality: a cross-sectional study in global health. Global Health 2015 Feb 12;11:4 [FREE Full text] [doi: 10.1186/s12992-015-0087-y] [Medline: 25889910]
29. Medicine (, Grossmann C, Powers B, McGinnis J. Technical Issues for the Digital Health Infrastructure Internet. In: Digital Infrastructure for the Learning Health System: The Foundation for Continuous Improvement in Health and Health Care: Workshop Series Summary. Washington, DC, USA: National Academies Press (US); 2011.
30. Lennon M, Bouamrane M, Devlin A, O'Connor S, O'Donnell C, Chetty U, et al. Readiness for Delivering Digital Health at Scale: Lessons From a Longitudinal Qualitative Evaluation of a National Digital Health Innovation Program in the United Kingdom. J Med Internet Res 2017 Dec 16;19(2):e42 [FREE Full text] [doi: 10.2196/jmir.6900] [Medline: 28209558]
31. Kostkova P. Grand challenges in digital health. Front Public Health 2015;3:134 [FREE Full text] [doi: 10.3389/fpubh.2015.00134] [Medline: 26000272]
32. van Deursen AJAM, van Dijk JAGM. Toward a Multifaceted Model of Internet Access for Understanding Digital Divides: An Empirical Investigation. The Information Society 2015 Sep 11;31(5):379-391. [doi: 10.1080/01972243.2015.1069770]
33. van Deursen AJAM, van Dijk JAGM, ten Klooster PM. Increasing inequalities in what we do online: A longitudinal cross sectional analysis of Internet activities among the Dutch population (2010 to 2013) over gender, age, education, and income. Telematics and Informatics 2015 May;32(2):259-272. [doi: 10.1016/j.tele.2014.09.003]

Abbreviations

AOR: adjusted odds ratio
ANC: antenatal care
BDHS: Bangladesh Demographic and Health Survey
EA: enumeration area
eHealth: electronic health
MDG 5: Millennium Development Goal 5
mHealth: mobile health
MHS: maternal health service
MoHFW: Ministry of Health and Family Welfare
SMS: short message service

Edited by G Eysenbach; submitted 29.03.18; peer-reviewed by M Fiordelli, E Neter, H Lucas; comments to author 05.08.18; revised version received 17.09.18; accepted 22.10.18; published 28.02.19.

Please cite as:
Tang S, Ghose B, Hoque MR, Hao G, Yaya S
Women Using Mobile Phones for Health Communication Are More Likely to Use Prenatal and Postnatal Services in Bangladesh: Cross-Sectional Study
JMIR Mhealth Uhealth 2019;7(2):e10645
URL: http://mhealth.jmir.org/2019/2/e10645/
doi:10.2196/10645
PMID:30816850

Original Paper

Research-Tested Mobile Apps for Breast Cancer Care: Systematic Review

Chiara Jongerius[1], MSc; Selena Russo[2], PhD; Ketti Mazzocco[3], PhD; Gabriella Pravettoni[3], PhD

[1]Department of Medical Psychology, Amsterdam UMC, University of Amsterdam, Amsterdam, Netherlands

[2]Kids Cancer Centre, Sydney Children's Hospital, University of New South Wales, Sydney, Australia

[3]Department of Oncology and Hemato-oncology, IEO, European Institute of Oncology IRCCS, Milan, Italy

Corresponding Author:
Chiara Jongerius, MSc
Department of Medical Psychology
Amsterdam UMC
University of Amsterdam
Meibergdreef 9
Amsterdam, 1100 DD
Netherlands
Phone: 31 205667787
Email: c.jongerius@amc.uva.nl

Abstract

Background: The use of mobile health (mHealth) apps in clinical settings is increasing widely. mHealth has been used to promote prevention, improve early detection, manage care, and support survivors and chronic patients. However, data on the efficacy and utility of mHealth apps are limited.

Objective: The main objective of this review was to provide an overview of the available research-tested interventions using mHealth apps and their impact on breast cancer care.

Methods: A systematic search of Medline, PsycINFO, Embase, and Scopus was performed to identify relevant studies. From the selected studies, the following information was extracted: authors, publication date, study objectives, study population, study design, interventions' features, outcome measures, and results.

Results: We identified 29 empirical studies that described a health care intervention using an mHealth app in breast cancer care. Of these, 7 studies were about the use of an mHealth application in an intervention for breast cancer prevention and early detection, 12 targeted care management, and 10 focused on breast cancer survivors.

Conclusions: Our results indicate consistent and promising findings of interventions using mHealth apps that target care management in breast cancer. Among the categories of mHealth apps focusing on survivorship, mHealth-based interventions showed a positive effect by promoting weight loss, improving the quality of life, and decreasing stress. There is conflicting and less conclusive data on the effect of mHealth apps on psychological dimensions. We advocate further investigation to confirm and strengthen these findings. No consistent evidence for the impact of interventions using mHealth apps in breast cancer prevention and early detection was identified due to the limited number of studies identified by our search. Future research should continue to explore the impact of mHealth apps on breast cancer care to build on these initial recommendations.

(JMIR Mhealth Uhealth 2019;7(2):e10930) doi:10.2196/10930

KEYWORDS

breast cancer care; breast cancer management; breast cancer prevention; breast cancer survivorship; mobile applications; mHealth applications

Introduction

Mobile health (mHealth) is a method to deliver health care or related services through portable devices [1] and is broadly accessible and often freely delivered in the app stores of app providers. The online app market is open to developers and allows them to sell or provide their apps free of charge. The current estimates suggest that there are more than 40,000 mHealth apps [2]. According to a recent report by Grand View Research, Inc., the global mHealth market is expected to reach

US $111.8 billion by 2025, and there is a growing need to reduce long waiting periods to access health care services, which is the primary driver responsible for the adoption of mHealth [3].

mHealth apps are extremely relevant for both industrialized nations and developing countries, as they provide extended access to health care and health-pertinent information in a cost-effective way [4]. Scholars and health care professionals have shown interest in this new technology, and the use of mHealth in clinical settings is increasing widely [5]. Recent reviews showed that in cancer care, mHealth apps have been employed to promote prevention, improve early detection, manage cancer care, and support cancer survivors [5,6]. Furthermore, research-tested apps offer the unique possibility of providing accessible information and education at minimal costs throughout the cancer care continuum [5].

Despite their impressive and promising potential, the utility and effectiveness of this e-health technology remain unclear. Scholars have reviewed the use of mHealth apps in several fields [4,7-10]. Mobasheri et al [4] reviewed 158 mHealth apps for breast cancer and found that there is a lack of evidence on their utility, effectiveness, and safety. Most mHealth apps lack expert involvement and do not adhere to relevant medical evidence or reflect patients' needs [11]. In addition, given the large number of available mHealth apps, it seems unrealistic to test each app thoroughly and scientifically before its release [4,5]. A recently developed mHealth app had negative effects on patients, augmenting their anxiety after breast cancer surgery [12]. The mHealth app aimed at providing extensive information to patients; however, the control patients, who did not use the application, had lower anxiety levels than the test patients, which correlated to higher quality of life in the former group [12].

The possibilities to enhance positive health outcomes and promote patients' feelings of control over their health through the use of mHealth technology should be fully explored and tailored to patients' needs [13-15]. Scholars and stakeholders advocate for more medical professional involvement, inclusion of patients' preferences, and specific regulations [4].

The need for integration between research-tested and privately developed mHealth apps has been widely stressed, as such integration plays a crucial role in achieving the necessary fundamentals of effective, evidence-based care [10,16,17]. Research should provide a scientific basis to build elements that can be effective for patient care. It is especially important to know which mHealth app has been scientifically tested and provides an overview of the interventions using mHealth apps. Our focus is on mHealth apps in breast cancer care. The present review aimed to provide an overview of the available evidence on research-tested interventions using mHealth apps in breast cancer care. Our findings aim to provide valuable information to health care professionals, mHealth apps developers, breast cancer patients, and other stakeholders on the characteristics of existing research-tested apps used in breast cancer care, including their advantages and pitfalls, in order to correctly, effectively, and safely support breast cancer patients.

Methods

Search Strategy and Selection of Articles

An extended bibliographic search was conducted in the Medline, PsycINFO, Embase, and Scopus databases. The following search string was used: ((phone OR mobile OR smartphone*) AND (app* OR application*)) AND ((breast* OR mammary) AND (cancer* OR neoplasm* OR carcinom* OR tumor*)).

The search was limited to studies published from January 2008 (when the first mobile phone app was created) to September 2018. The review was registered in PROSPERO with the registration number CRD42017056239. All detected articles were screened according to the following inclusion criteria: original studies, English language, studies that include the use of an mHealth app, studies that include patients' use of an mHealth app, and studies that concern the prevention/detection/care management/survivorship of breast cancer. We excluded studies that used websites, text messaging, emails, or other technological interventions that did not include mobile apps and studies that used mobile apps without applying them to breast cancer patients (eg, health care professionals that use mHealth apps).

Data Collection and Extraction Process

A data-extraction form was developed on basis of the Centre for Reviews and Dissemination templates [18]. Two reviewers independently extracted the data from the included studies by using the extraction form. Disagreements in data extraction were resolved through discussions between the authors until an agreement was reached. Relevant articles were then selected by cross-examining and reviewing the articles. Data collected included information on authors, publication date, intervention's features, study design, sample size, outcome measures, and results. The quality of the quantitative studies was evaluated independently by two researchers using the Effective Public Health Practice Project Quality Assessment Tool for Quantitative Studies [19]. This tool provides a standardized means to assess the quality of a quantitative study, leading to an overall methodological rating of strong, moderate, or weak in eight sections: selection bias, study design, confounders, blinding, data-collection methods, withdrawals and dropouts, intervention integrity, and analysis. The quality of the qualitative studies was evaluated using the Joanna Briggs Institute Critical Appraisal Checklist for Qualitative Research [20]. This instrument is underpinned by a multi-dimensional concept of quality in research, and the 10 items assess quality according to several domains including quality of reporting, methodological rigor, and conceptual depth and bread. Discordances in quality rating were resolved through discussion between the researchers.

Results

Study Selection

The results of the systematic search are summarized in Figure 1 in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses [21]. We identified a total of 256 articles (100 from Medline, 9 from PsycINFO, 230

from Embase, and 129 from Scopus), from which we excluded 212 duplicates. Abstracts and titles were screened to identify articles discussing an intervention using an mHealth app for the care of patients with breast cancer, and 69 such articles were found. Further screening on the basis of the entire text according to our inclusion and exclusion criteria led to a final selection of 29 articles.

Study Characteristics

The 29 studies identified were published between 2011 and 2018. Majority of the studies (n=12) focused on apps for breast cancer care management, 10 focused on survivorship, and 7 focused on prevention and early detection.

Ten of the 29 studies were randomized controlled trials, 8 were prospective cohort studies, 7 were cross-sectional studies, and 4 used qualitative analysis. Using the Quality Assessment Tool for Quantitative Studies [19], 10 studies were rated as strong; 9, as moderate; and 6, as weak. The 4 qualitative studies assessed using the Critical Appraisal Checklist for Qualitative Research [20] revealed a methodological quality that allowed their inclusion in the review. Two of the 29 studies focused on the use of an mHealth app in breast cancer prevention, 5 targeted early detection, 12 were on care management, and 10 focused on survivors of breast cancer. The sample included in the studies comprised adult patients in different countries with different ethnic backgrounds: 14 studies in North America, 8 in Europe, and 6 in Asia. Table 1 provides a summary of the features of the included studies.

Figure 1. Flow diagram of identification, screening, eligibility, and inclusion of studies.

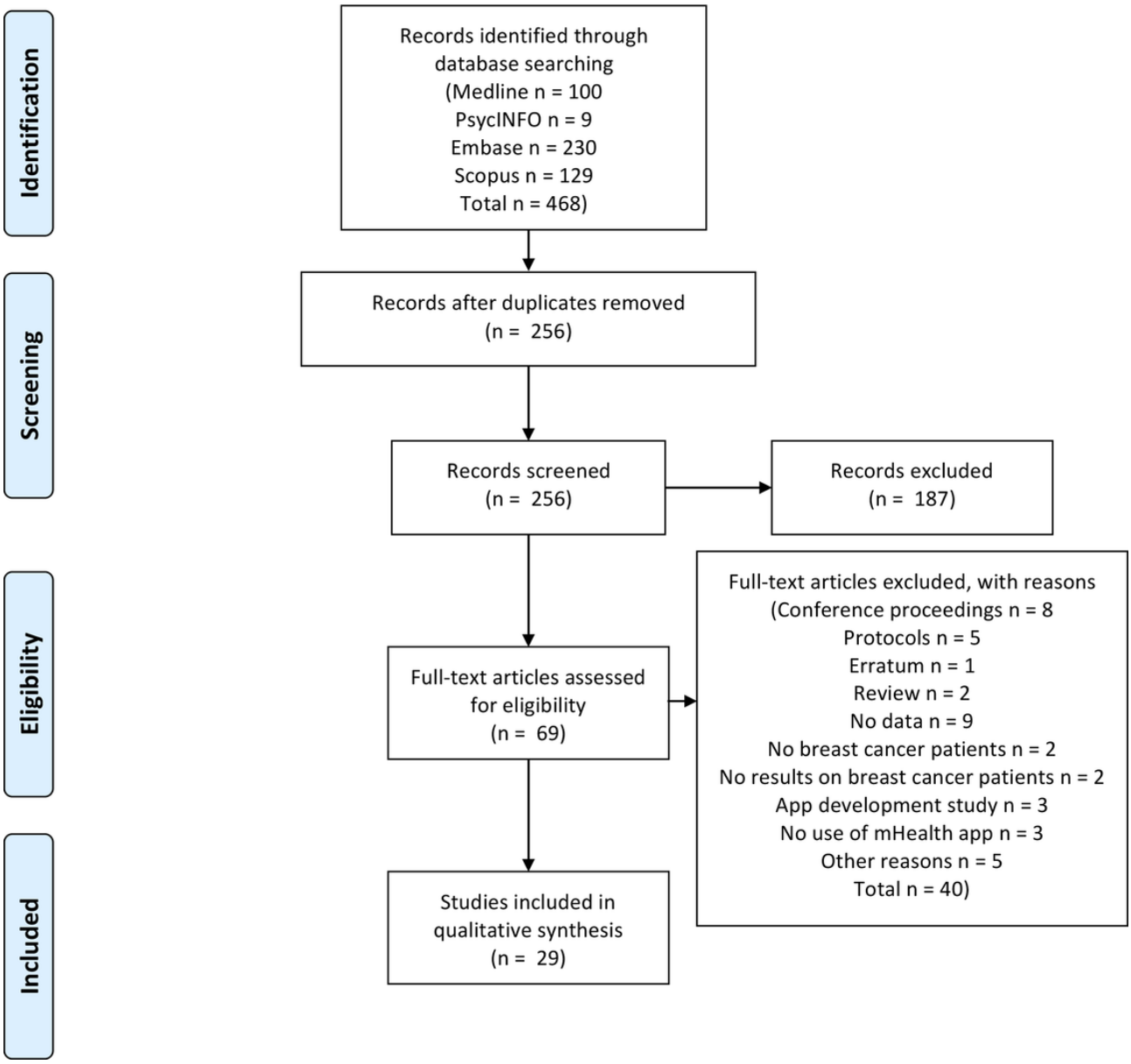


Table 1. Description of the characteristics of the studies presented in the review[a].

Characteristics and studies	Intervention target	Sample size	Duration & follow-up	Intervention components	Outcome measures	Study quality[b]
Prevention						
Alanzi et al [23] (2018)	To create awareness about breast cancer	Intervention group: N=96; Control group: N=95	4 weeks; Measurements at baseline and at 4-week follow-up	*SnapChat* app: breast cancer awareness information on mobiles, covering knowledge about breast cancer, its symptoms, diagnosis process, and available treatments. New information provided three times each week	18-item questionnaire on breast cancer awareness	Moderate
Hartman et al [22] (2016)	Weight loss	Intervention group: N=36; Usual care group: N=18	6 months	Combined technology-based self-monitoring tools with individualized phone calls: Electronic calorie-counting tool (*MyFitnessPal*); 12 phone calls (30 minutes each) over 6 months; Accelerometer-based activity meter that provides real-time feedback on the number of steps taken and minutes of moderate-intensity activity (*Fitbit*)	Weight and accelerometer-measured physical activity	Strong
Early detection						
Eden et al [24] (2015)	To help women in their 40s gain deeper insights into their priorities for screening and prepare them to discuss mammography screening with their health care providers	N=75	Before and after use of the app (same day)	The decision aid (*Mammopad*) included educational modules on breast cancer, mammography, risk assessment, and priority setting about screening	Decisional conflict; Decision self-efficacy scale	Moderate
Heo et al [25] (2013)	To encourage breast self-examination	N=45	Before and after use of the app	A mobile phone app developed with functions including a breast self-examination date alarm, a reminder to encourage mother and daughter to practice breast self-examination together, record keeping, and educational content with video clips	Survey: increased breast self-examination	Weak
Keohane et al [26] (2017)	To improve risk perception	Intervention group: N=42; Usual care group: N=42	Measurements before the counseling session (T1), after the counseling session (T2), and 6 weeks later (T3)	A mobile phone app displaying data on risk of developing breast cancer as well as risk of carrying the *BRCA* gene	IBIS[c] Breast Cancer Risk Evaluation; Perception of risk: "Patient Survey on Risk Perception of Breast Cancer and Health Literacy"	Strong
Lee et al [27] (2017)	To increase knowledge and awareness and promote mammogram screening	Intervention group N=60 Usual care group N=60	1 week; Measurements at baseline, 1 week, and 6 months	*mMammogram*: Each day participants received 8-21 messages covering various topical areas, including breast cancer, screening guidelines, and types of screening; breast cancer risk factors; individual, structural, and cultural barriers to screening; communication strategies; follow-up for test results; and information on local clinics. Messages followed a trajectory from basic knowledge building to specific strategies aimed to enhance motivation for and access to mammography	Knowledge, attitudes, and beliefs about breast cancer screening, readiness for mammography, and mammogram receipt; Feasibility and acceptability of the mMammogram intervention	Strong
Care management						

Characteristics and studies	Intervention target	Sample size	Duration & follow-up	Intervention components	Outcome measures	Study quality[b]
Lee et al [28] (2018)	To increase knowledge and awareness and promote mammogram screening	N=14	1 week	*mMammogram*: Each day participants received 8-21 messages covering various topical areas including breast cancer, screening guidelines, and types of screening; breast cancer risk factors; individual, structural, and cultural barriers to screening; communication strategies; follow-up for test results; and information on local clinics. Messages followed a trajectory from basic knowledge building to specific strategies aimed to enhance motivation for and access to mammography	Thematic analysis was used to analyze data from focus groups	Included
Egbring et al [29] (2016)	Collection of patient-reported daily functional activity	Control group: N=41; Unsupervised intervention group: N=45; Supervised intervention group: N=41	3 visits	Mobile and Web app to record daily functional activity and adverse events: Patients could report daily functional activity or symptoms with indication of severity	Functional activity; Eastern Cooperative Oncology Group scoring; Common Terminology Criteria for Adverse Events	Strong
Foley et al [12] (2016)	To decrease anxiety levels of patients undergoing surgery through educational materials	Intervention group: N=13; Control group: N=26	Measurement one day before, one day after, and 7 days after surgery	An iPad app containing tailored information on surgery pertaining to individual patients	Mini-Mental Adjustment to Cancer questionnaire; Hospital Anxiety and Depression Scale; Information Satisfaction Questionnaire	Moderate
Harder et al [39] (2017)	To optimize self-management of arm and shoulder exercises for upper-limb dysfunction after breast cancer treatment	N=3	8 weeks; Measurement at the end of the study	*bWell* is an evidence-based program providing progressive exercises for passive and active mobilization, stretching, and strengthening. App features include tailored information, video demonstrations of the exercises, push notifications, and tracking and progress features	Questionnaire capturing users' feedback and evaluating content, functionality (including ratings), and explored areas of improvement	Weak
Hwang [30] (2016)	To communicate and share images of the wound postoperatively	Intervention group: N=35; Control group: N=37	1, 3, 7, and 14 days after surgery	A virtual care platform that consists of a mobile phone app and secure password-protected online account (*Medeo* app). Patients can take photos of their wounds postoperatively and send them to the surgeon using the mobile phone app. The surgeon then responded to each patient message within 24 hours.	Less readmission to the hospital; Use of mobile phone app for question; Improved perceived care	Moderate
Kim et al [31] (2016)	To collect and track daily mental health indicators for depression	N=78	48 weeks	A mobile mental health tracker (part of *Pit-a-Pat* app) that uses three daily mental health ratings (sleep satisfaction, mood, and anxiety) as indicators for depression	Patient Health Questionnaire-9; Adherence level	Moderate

Characteristics and studies	Intervention target	Sample size	Duration & follow-up	Intervention components	Outcome measures	Study quality[b]
Klasnja et al [34] (2011)	Manage care-related information	N=9	4 weeks; Measurement at baseline, 2 weeks, and 4 weeks	A web component (*HealthWeaver* website) to manage personal and health issues. It provides a calendar to manage health-related events; functionality for organizing notes, lists, bookmarks, and care-related files; a tracking system for symptoms, pain, and well-being; and logs for medications, supplements, and the care of post-surgery wounds. A mobile component (*HealthWeaver Mobile*) through which patients can create, edit, and view the information stored in the website; share them with family members and health care professionals; create photo, audio, and text notes and link those notes to related appointments; and synchronize the app calendar to the phone's calendar app	Qualitative thematic analysis of interviews (at 2 weeks and 4 weeks) + demographics and experience with technology questionnaires (at baseline)	Included
Min et al [32] (2014)	Sleep disturbance-related data collection from breast cancer patients receiving chemotherapy	N=30	90 days	App (*Pit-a-Pat*) developed to self-report three health experiences that may be caused by breast cancer diagnosis and treatments: sleep-disturbance symptoms related to mild depression, acute symptoms related to cytotoxic chemotherapeutic agents, and medication diary for antihormonal treatment	Compliance to use of mobile app	Moderate
Rosen et al [37] (2018)	To decrease physical and psychological distress and improve the quality of life	Intervention group: N=57; Control group: N=55	8 weeks; Measurement at baseline, during intervention (5 - weeks), after intervention (9 - weeks), and follow - up (12 - weeks)	Mindfulness meditation training was delivered through a commercially available mindfulness app (*Headspace*) that uses audio and animated video	The Brief Pain Inventory - Short Form-32; The Brief Health Literacy Screening Tool; The eHealth Literacy Scale; Functional Assessment of Cancer Therapy—Breast version; Mindful Attention Awareness Scale; App utilization	Strong
Wallwiener et al [38] (2017)	Tablet-based measurement app for EORTC QLQ-C30[d]	N=106	Not available	e-Patient-Reported Outcome versions of the EORTC QLQ-C30[d] questionnaires	Electronic and paper-based versions of the health-related quality of life EORTC QLQ-C30[d] questionnaire	Strong
Weaver et al [33] (2014)	Real-time symptom monitoring of patients receiving oral chemotherapy	N=26	8 times for 3 weeks	Patients completed a symptom, temperature, and dose diary twice a day using a mobile phone app. This information was encrypted and automatically transmitted in real time to a secure server, with moderate levels of toxicity automatically prompting self-care symptom management messages on the screen of the patient's mobile phone or in severe cases, a call from a specialist nurse to advice on care according to an agreed protocol	Medication dose and monitoring side effects	Weak

Characteristics and studies	Intervention target	Sample size	Duration & follow-up	Intervention components	Outcome measures	Study quality[b]
Zhu et al [35] (2018)	To address women's self-efficacy, social support, symptom distress, quality of life, anxiety, and depression	N=13	12 weeks; Measurements at the end of intervention	Breast Cancer e-Support with four components: a Learning forum, a Discussion forum, an Ask-the-Expert forum, and a Personal Stories forum.	Inductive content analysis of the interviews	Included
Zhu et al [36] (2018)	To address women's self-efficacy, social support, symptom distress, quality of life, anxiety, and depression	Intervention group: N=55; Control group: N=49	12 weeks; Measurements at baseline, after 3 months, and after 6 months	Breast Cancer e-Support with four components: a Learning forum, a Discussion forum, an Ask-the-Expert forum, and a Personal Stories forum.	Self-efficacy: Stanford Inventory of Cancer Patient Adjustment; Social support: Multidimensional Scale of Perceived Social Support; Symptom distress: MD Anderson Symptom Inventory; Quality of life: Functional Assessment of Cancer Treatment-Breast; Anxiety and depression: Hospital Anxiety and Depression Scale	Strong
Survivorship						
Ainsworth et al [44] (2018)	To measure time use to further research testing on how to optimize physical activity promotion	N=40	5 days	Time-use measurement app (*Life in a Day*)	Questionnaire assessing functionality and satisfaction with the time use app and Frequency of Forgetting scale	Moderate
Buscemi et al [48] (2018)	To increase health-related quality of life in Hispanic breast cancer survivors	N=25	4 weeks	Intervention delivered through *MyGuide* app, focused on enhancement of psychosocial adaptation after breast cancer; cancer knowledge; stress awareness and management; social support; and communication with friends, family, and oncology providers. Two to three 15-minute telecoaching calls to facilitate adherence to the *MyGuide* app	Functional Assessment of Cancer Therapy–General Seven Engagement; Acceptability survey; Knowledge about Breast Cancer questionnaire	Moderate
Fazzino et al [40] (2015)	Weight loss	N=186	6 months	Three-phase intervention: a 6-month weight-loss phase (0-6 months) where all participants received weekly group phone sessions, a 12-month weight-loss maintenance phase (6-18 months) in which participants were randomized to continue group phone sessions or a newsletter comparison condition, and a 6-month no-contact follow-up phase (18-24 months) to evaluate the sustained effects	Qualitative thematic analysis	Included

Characteristics and studies	Intervention target	Sample size	Duration & follow-up	Intervention components	Outcome measures	Study quality[b]
Lengacher et al [47] (2018)	To improve psychological and physical symptoms of depression, anxiety, stress, fear of recurrence, cognitive functioning, sleep, fatigue, pain, and quality of life	N=15	6 weeks; Measurements at baseline and at 6 - week follow - up	The mMBSR(BC)[e] providing a 2 - hour session intervention weekly for 6 weeks via iPad. Weekly modules on formal meditative techniques (sitting meditation, walking meditation, body scan, and gentle Hatha yoga) via video files and lectures in informal meditative techniques (integrating mindfulness into daily life activities) via audio files	Fatigue Symptom Inventory; Brief Pain Inventory; Pittsburgh Sleep Quality Index; Center for Epidemiological Studies Depression Scale; State - Trait Anxiety Inventory; Perceived Stress Scale; Concerns about Recurrence Scale; Everyday Cognition; Five Facet Mindfulness Questionnaire; Medical Outcomes Studies Short - Form; Acceptability and usability ratings	Weak
Lozano-Lozano et al [45] (2018)	To assess, monitor, and facilitate adherence to healthy lifestyles	N=20	8 days	The BENECA[f] mHealth app records diet and physical exercise and provides daily notification about energy balance and physical activity and dietary recommendations	Physical activity measured with tri-axial accelerometers; Dietary habits: dietary records and 24-hour dietary recalls	Weak
McCarroll et al [41] (2015)	Lifestyle program on nutrition quality, physical activity, and eating self-efficacy	N=50	4 weeks	A “beta” health care provider version of the app *LoseIt!* to deliver a comprehensive lifestyle program with emphasis on nutrition quality, physical activity, and eating self-efficacy by a multidisciplinary team that provided real-time feedback notifications	Body mass index; Waist circumference; Anthropometrics; FACT-G[f]; Macronutrient consumption; Physical activity patterns; Weight Efficacy Lifestyle questionnaire	Moderate
Pope et al [46] (2018)	To improve physiological, psychosocial, and quality of life outcomes	N=10	10 weeks; Measurement at baseline and after intervention	A combined mobile phone app (*MapMyFitness*) and *Facebook*-delivered health behavior change intervention	Physical Activity Readiness Questionnaire; Acceptability survey; Physical activity levels/energy expenditure via Actigraph GT3X[h] + accelerometer; Weight and body fat percentage; Cardiovascular fitness (YMCA 3-min Step Test); Patient Reported Outcome Measurement Information System; Self-efficacy for exercise scale; Adaptation of the Patient-Centered Assessment and Counseling for Exercise questionnaire; 5-item physical activity enjoyment measure; 14-question measure on physical activity barriers; Outcome expectancy	Weak
Uhm et al [42] (2016)	To improve physical function and quality of life	Intervention group: N=179; Control group: N=177	12 weeks	Mobile phone exercise app (Smart After Care) and an *InBodyBand* pedometer	International physical activity questionnaire short form; EORTC-QLQ-C30[d]; Quality of Life Questionnaire - Breast Cancer Module 23	Strong

Characteristics and studies	Intervention target	Sample size	Duration & follow-up	Intervention components	Outcome measures	Study quality[b]
Valle et al [43] (2017)	Weight gain prevention intervention in African-American breast cancer survivors	Intervention group: N=11; Intervention + activity monitoring group: N=13; Control group: N=11	Baseline, 3 months, and 6 months	Both intervention groups received a face-to-face individual session; a Bluetooth and Wifi-enabled wireless scale with access to a companion mobile app and website with graphs of weight trends; 24 weekly email-delivered behavioral lessons; and 24 weekly emails with tailored feedback on self-weighing and weight data	Weight change and daily self-weighing perception	Strong
Visser et al [49] (2018)	To provide professional and peer support to breast cancer survivors	Intervention group: N=59; Control group: N=50	3 months Measurements at baseline (1 week prior to the visit) (T0), 1 week (T1), 3 months (T2), and 6 months after the visit (T3)	*My-GMC* intervention: a face-to-face group medical consultation combined with an online app, providing three online support group sessions and additional information	Symptom Checklist-90; Dutch Empowerment Questionnaire for breast cancer patients; Cancer Worry Scale; EORTC-QLQ-C30[d]; EORTC-BR23[g]; Medication Adherence Report Scale; Self-reported usage statistics; QUOTE[h] questionnaire; Intervention-specific questions	Strong

[a]A detailed table showing features of the extracted studies is available in Multimedia Appendix 1.

[b]The quality of the study was rated using the Effective Public Health Practice Project Quality Assessment Tool for Quantitative Studies [19] and the Joanna Briggs Institute Critical Appraisal Checklist for Qualitative Research [20].

[c]IBIS: International Breast Cancer Intervention Study.

[d]EORTC-QLQ-C30: European Organization for Research and Treatment Cancer Quality of Life Questionnaire - Core 30.

[e]mMBSR(BC): mobile Mindfulness-Based Stress Reduction for Breast Cancer.

[f]FACT-G: Functional Assessment of Cancer Therapy - General.

[g]EORTC-BR23: European Organization for Research and Treatment Cancer Quality Of Life Questionnaire - Breast Cancer Module.

[h]QUOTE: Quality of Care Through the Patient's Eyes

mHealth in Breast Cancer Prevention and Early Detection

We identified 7 papers in the area of breast cancer prevention and early detection [22-28]. Hartman et al [22] designed a randomized study to test a weight-loss intervention to reduce the risk of breast cancer. The intervention combined technology-based self-monitoring apps with individualized phone calls. In particular, the intervention entailed an electronic calorie-counting tool, an accelerometer-based activity meter that provides real-time feedback of the number of steps taken and minutes of moderate-intensity activity, and 12 phone calls (30 minutes each) over the 6-month intervention period. After 6 months, the 36 participants in the experimental group had lost significantly more weight and a greater percentage of starting weight than the control group.

SnapChat, a social networking mobile app used to deliver educational material on breast cancer through videos, texts, and pictures, increased breast cancer awareness among female students in the Dammam region of Saudi Arabia [23]. Eden et al [24] studied decisional conflict and self-efficacy in 75 women before and after the use of a mammography screening decision aid in the form of an mHealth app and showed that the use of the mHealth app decreased the decisional conflict and increased the decisional self-efficacy. Heo et al [25] developed an mHealth app to encourage breast self-examination. The mHealth app included a breast self-examination reminder, a record-keeping function, and educational features. After using the app, the number of participants who practiced breast self-examination increased from 28 to 32. However, participants below the age of 30 years performed a significantly higher number of breast self-examinations and more appropriate self-examinations than participants aged above 30 years.

An mHealth app developed to convey risk information to patients attending a high-risk breast cancer clinic resulted in increased accuracy in risk perception compared to standard risk counselling [26]. Lee et al [27] tested the efficacy and feasibility of *mMammogram*, an intervention entailing tailored multimedia messages to increase knowledge and awareness and promote mammogram screening among Korean-American immigrant women. The intervention group showed greater knowledge of breast cancer and screening guidelines, readiness for mammography use, and greater satisfaction with the intervention. At the 6-month follow-up, 75% of women who received the *mMammogram* intervention completed the mammograms compared to 30% of the women in the control group. Focus groups with women in the intervention group supported the fact that *mMammogram* enhanced the understanding of breast cancer and screening through *mMammogram* and that health navigators promoted free or

low-cost mammograms and scheduling of mammogram appointments [28].

mHealth and Care Management

A total of 12 papers investigating the use of mHealth apps in breast cancer care management emerged from our review and focused on care management and monitoring of treatment side effects [29-33]; personal data management [34]; psychological aspects, social support, and health-related quality of life [35-38]; physical activity [39]; and educational aspects [12]. The mHealth apps that included the monitoring and care management of treatment side effects covered wound monitoring [30], sleep quality collection [32], daily functional activity stabilization [29], mental health tracker [31], and chemotherapy dose adaptation [33]. These apps were developed for and focused on reliable collection and communication of relevant breast cancer data from patients and oncologists.

In the Hwang's study [30], a postoperative wound e-monitoring app significantly decreased the number of readmissions to the hospital, unscheduled visits to the emergency department, or walk-in clinic visits. The mHealth app intervention is a virtual care platform that consists of a mobile phone app and secure password-protected online account that allowed patients to take a photo of their wounds postoperatively, attach the photos to electronic messages, and send them to the surgeon using the mobile phone app. The surgeon then responded to each patient's message within 24 hours. A vast majority (95%) of patients felt that electronic wound monitoring improved their care and would recommend such technologies to a friend or colleague. Min et al [32] analyzed the feasibility of and compliance with an mHealth app for collection of sleep disturbance-related data from breast cancer patients receiving chemotherapy. Participants self-recorded three health experiences—sleep-disturbance symptoms related to mild depression, acute symptoms related to cytotoxic chemotherapeutic agents, and medication diary for antihormonal treatment—that may be caused by breast cancer diagnosis and treatments. The overall compliance during the 90-day longitudinal collection of daily self-reporting of sleep-disturbance data was approximately 45%, confirming the feasibility of the intervention; women without any form of employment exhibited a higher compliance rate. No other association between patient characteristics and compliance was found. An intervention using an mHealth app on patient-reported daily functional activity in both supervised and unsupervised settings showed that the use of the mHealth app in collaboration with physicians was associated with stabilized daily functional activity and fewer and more precise entries than the unsupervised use of the mHealth app. The mHealth and Web app allows patients to record daily functional activity or symptoms with an indication of severity [29]. Kim et al [31] evaluated the validity and screening performance of an mHealth app for depression, which collected self-reported mental health ratings from patients with breast cancer. The app tracked three mental health indicators for depression (sleep satisfaction, mood, and anxiety) daily, and the results strongly supported the potential of a mobile mental health tracker as a tool for screening depression in patients with breast cancer. An mHealth app developed to collect and monitor real-time symptoms in patients receiving oral chemotherapy showed that both patients and oncologists felt reassured by the use of the mHealth app [33]. In the intervention, patients were required to complete a symptom, temperature, and dose diary twice a day using an mHealth app. This information was encrypted and automatically transmitted in real time to a secure server, with moderate levels of toxicity automatically prompting self-care symptom-management messages on the screen of the patient's mHealth phone or in severe cases, a call from a specialist nurse to advice about care according to an agreed protocol [33].

Klasnja et al [34] tested the feasibility of an mHealth app for data management. The mHealth app included a web component and a mobile component. Using the mHealth app, patients could create, edit, and view the full range of information stored in the website and share them with family members and health care professionals. In addition, patients could create photo, audio, and text notes to quickly capture information and immediately link those notes to related appointments. The mobile app synchronized the app calendar to the phone's native calendar app, allowing patients see their care events alongside their other commitments. Their study highlighted an empowering effect of mHealth analysis.

Harden et al [39] developed *bWell*, a mobile phone application intervention to optimize self-management of arm and shoulder exercises for upper-limb dysfunction after breast cancer treatment. User testing demonstrated ease of use along with clear, engaging, and motivating content.

Zhu et al explored the effectiveness of *Breast Cancer e-Support*, an app-based program, to address women's self-efficacy, social support, symptom distress, quality of life, anxiety, and depression [35,36]. The app-based intervention was found to be more effective than usual care in promoting women's self-efficacy and quality of life and decreasing symptom burden during chemotherapy. The mHealth app did not show any advantage over usual care in terms of social support, symptom severity, anxiety, and depression. In addition, the beneficial effects were not sustained at the 6-month follow-up [36]. Analyzing qualitative interviews, Zhu et al [35] highlighted that participants perceived the program to be helpful in enhancing knowledge, improving confidence level, and promoting emotional well-being. The most prominent benefit was access to tailored advices from experts, whereas the barriers to use were the physical or psychological health status and app instability. Participants suggested implementing message reminders and a search engine to add more interesting and practical knowledge and updating the information more often.

The use of a mindfulness meditation training delivered through a commercially available mindfulness app called *Headspace* that uses audio and animated video improved quality of life and decreased physical and psychological distress [37]. Wallweiner et al [38] analyzed the reliability of a tablet-based measurement app for the European Organization for Research and Treatment Cancer Quality of Life Questionnaire - C30, a measure of health-related quality of life in patients with metastatic breast cancer. The electronic version of this measure was reliable for patients with both adjuvant and metastatic breast cancer, showing a high correlation in almost all questions of the questionnaire.

Foley et al [12] educated patients about breast cancer surgery by using an mHealth app that provided information on the surgery procedure tailored to individual patients. They found that at 7 days after surgery, participants who were more educated about their breast cancer surgery procedure experienced more anxiety than the control participants.

mHealth in Survivorship

In this review, we identified 10 studies focusing on the use of mobile phone apps among breast cancer survivors. All 10 studies aimed at assessing lifestyle changes, in particular, improving physical activity and reinforcing weight loss [40-46]; reducing stress [47]; improving health-related quality of life [46,48]; and providing information, support, and attention for psychosocial issues [49]. Of the 10 studies, 4 found significant effects of the mHealth apps on weight loss [41-43,46]. McCarroll et al [41] assessed a 1-month lifestyle intervention using a Web and mHealth weight-loss app called *LoseIt!* This intervention placed emphasis on the nutritional quality and physical activity and showed significant reductions in body weight and improvements in eating self-efficacy.

Uhm et al [42] tested and compared a 12-week home-based program of aerobic and resistance exercises using a mobile phone exercise app and a pedometer with a conventional exercise program using a brochure. Physical function, physical activity, and quality of life significantly improved regardless of the intervention method, and changes were not significantly different between the two interventions. The feasibility and preliminary efficacy of two self-regulation interventions using the mHealth app and website as well as tailored feedback emails to prevent weight gain among African-American breast cancer survivors were evaluated [43]. The interventions focused on daily self-weighing and used objective monitoring and tailored feedback about weight and physical activity. Both the interventions were successful in controlling weight, and both intervention groups indicated highly positive perceptions about the intervention; in addition, 100% of participants would recommend the program to other breast cancer survivors. Preliminary findings of an intervention combining the use of a mobile phone app and Facebook-delivered health education material showed an increase in physical activity and decrease in weight and body fat percentage [46]. Fazzino et al [40] performed a thematic qualitative analysis and identified themes related to successful weight loss in breast cancer survivors undertaking a group phone-based weight-loss intervention. Examples of themes that emerged are the importance of being part of a group and internal motivation. *BENECA mHealth*, an application for remotely assessing and monitoring energy balance in breast cancer survivors, showed positive agreement with daily, 24-hour dietary recalls and accelerometer data [45]. Ainsworth et al [44] assessed the acceptability of a 5-day trial of a mobile phone app for measuring time use in order to inform physical activity measurement and promotion interventions. Majority of the participants agreed that learning to use the app was easy, and most preferred to use the app over the paper-and-pencil diary method of record. The 6-week mindfulness-based stress reduction for the breast cancer program delivered via a tablet proved to significantly improve the psychological and physical symptoms of depression, state anxiety, stress, fear of recurrence, sleep quality, fatigue, and quality of life [47]. Buscemi et al [48] pilot-tested a mobile phone-based intervention to boost health-related quality of life in Hispanic breast cancer survivors in the United States. Participants' knowledge on breast cancer and health-related quality of life increased over the course of the 4-week intervention. A mobile phone app combined with a social media-delivered health behavioral change intervention promoted the improvement of several physiological, psychosocial, and quality of life outcomes over the course of 10 weeks [46]. For example, it improved physical activity and the ability to engage in social roles [46]. The blended care intervention *My-GMC*, which combines a face-to-face group medical consultation with an online app, positively influenced provision of information, support, and attention for psychosocial themes. However, no significant effects on psychological distress and empowerment were observed as compared to usual care [49].

Discussion

Considering the increasing number of mHealth apps available to patients and their increasing use in breast cancer care, it is important to understand their effects. Therefore, we conducted a systematic review on studies that scientifically tested interventions using an mHealth app for breast cancer care. We identified a total of 29 studies, which encompassed important phases in breast cancer care and addressed prevention and survivorship.

Overall, the results of the studies on the mHealth apps for breast cancer care were promising. Majority of the identified studies in patients' care management showed a positive impact of the use of mHealth apps. Many hopeful opportunities offered by the mHealth apps rely in the amelioration of the communication process between patient and doctors, favoring an effective exchange of information [33,34]. Positive effects of mHealth-based interventions on health-related quality of life [36,38] and stress reduction [37] have been reported among patients receiving treatment. However, the role of mHealth-based interventions in the psychological impact of treatment is unclear. Only two studies addressed this topic [12,36] but showed no advantage of social support, anxiety, and depression [36] and highlighted an alarming detrimental effect of providing educational material about surgery treatments on patients' anxiety levels [12].

The number, quality, and findings of studies identified by our search highlight the fact that mHealth technology may play a relevant role in the care of breast cancer survivors. Evidence points to a positive effect of mHealth-based interventions on promoting weight loss [41-43,46], stemming stress, and sustaining the quality of life [44,46,48,49]. However, no convincing data are available on the benefit of mHealth for enduring adverse psychological sequel [46,49]. The restricted number of studies and methodological differences partially account for the heterogeneity of results.

mHealth technology may result in interesting opportunities to improve the lifestyle of individuals at risk of breast cancer and support the spread of knowledge and awareness of breast cancer. However, there is no consistent evidence of the impact of

interventions using mHealth apps in breast cancer prevention and early detection due to the limited number of studies identified by our search. Despite the high quality of the studies, only one study considered an mHealth-based intervention in breast cancer prevention [22], and therefore, no strong conclusion can be drawn.

Several aspects have to be considered, as they may limit the generalizability of our results. The heterogeneity in the focus of the included studies may affect the strength of our results and conclusions. Regarding the design and quality of the studies included in our study, a considerable number of studies were feasibility or pilot studies [25,32,33,41,43], which restricted the strength of their results. The data indicate that more evidence is needed to clarify the benefit of mHealth technology for breast cancer care. Examples of the need for more evidence are provided in the studies of Foley et al [12] and Heo et al [25], who found that the effects of their mHealth apps were contrary to what they intuitively expected based on the literature on the effect of education in managing treatment side effects and facilitating adaptation in women with breast cancer [50,51]. Foley et al [12] hypothesized that more information regarding breast cancer would decrease anxiety levels before and after breast cancer surgery. However, their preliminary results showed that women who used the mHealth app experienced higher anxiety levels at 7 days after the surgery than women in the control group. Similarly, Heo et al [25] hypothesized that the use of the mHealth app would augment women's intentions to regularly perform breast self-examination. This appeared to be true only for women younger than 30 years of age. In contrast, in women older than 30 years, the intention to perform a breast self-examination decreased. A possible explanation of these results may lay in the study design, the sample size, the method used to deliver the program, or the outcomes' measures.

Despite the aforementioned limitations, the present study shows promising results for the inclusion of mHealth apps in breast cancer care, but also calls for caution when implementing interventions using mHealth technology. The effects of using mHealth apps in the field of breast cancer are only recently explored and may be unpredictable. Few evidence-based interventions are described in the literature, and therefore, there is a need for good-quality clinical studies to guide future implementations.

Our considerations build upon previous evidence highlighting the need for strict regulations in this field and a solid integration between privately tested and research-tested mHealth apps [4,6]. Breast cancer patients should, at all times, be safely assisted with regard to effective management of their health. Therefore, apps need to be extensively research tested before making them available to the public. Scientifically sound data are needed to draw strong conclusions on the utility, effectiveness, and safety of mHealth apps in breast cancer care.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Description of the characteristics of the studies presented in the review.

[PDF File (Adobe PDF File), 42KB - mhealth_v7i2e10930_app1.pdf]

References

1. Whittaker R. Issues in mHealth: findings from key informant interviews. J Med Internet Res 2012;14(5):e129 [FREE Full text] [doi: 10.2196/jmir.1989] [Medline: 23032424]
2. Aitken M, Gauntlett C. Patient apps for improved healthcare: from novelty to mainstream. Parsippany (NJ): IMS Institute for Healthcare Informatics; 2013.
3. Research and Markets. 2017. In mHealth App Market By Type And Segment Forecasts 2014 - 2025 URL: https://www.researchandmarkets.com/reports/4396364/mhealth-app-market-by-type-fitness-lifestyle [accessed 2018-12-05] [WebCite Cache ID 74RZARbua]
4. Mobasheri MH, Johnston M, King D, Leff D, Thiruchelvam P, Darzi A. Smartphone breast applications - what's the evidence? Breast 2014 Oct;23(5):683-689. [doi: 10.1016/j.breast.2014.07.006] [Medline: 25153432]
5. Davis S, Oakley-Girvan I. mHealth Education Applications Along the Cancer Continuum. J Cancer Educ 2015 Jun;30(2):388-394. [doi: 10.1007/s13187-014-0761-4] [Medline: 25482319]
6. Bender JL, Yue RYK, To MJ, Deacken L, Jadad AR. A lot of action, but not in the right direction: systematic review and content analysis of smartphone applications for the prevention, detection, and management of cancer. J Med Internet Res 2013;15(12):e287 [FREE Full text] [doi: 10.2196/jmir.2661] [Medline: 24366061]
7. Arnhold M, Quade M, Kirch W. Mobile applications for diabetics: a systematic review and expert-based usability evaluation considering the special requirements of diabetes patients age 50 years or older. J Med Internet Res 2014;16(4):e104 [FREE Full text] [doi: 10.2196/jmir.2968] [Medline: 24718852]
8. 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]

9. Hamine S, Gerth-Guyette E, Faulx D, Green BB, Ginsburg AS. Impact of mHealth chronic disease management on treatment adherence and patient outcomes: a systematic review. J Med Internet Res 2015;17(2):e52 [FREE Full text] [doi: 10.2196/jmir.3951] [Medline: 25803266]
10. Quanbeck A, Gustafson DH, Marsch LA, Chih M, Kornfield R, McTavish F, et al. Implementing a Mobile Health System to Integrate the Treatment of Addiction Into Primary Care: A Hybrid Implementation-Effectiveness Study. J Med Internet Res 2018 Jan 30;20(1):e37 [FREE Full text] [doi: 10.2196/jmir.8928] [Medline: 29382624]
11. Subhi Y, Bube SH, Rolskov BS, Skou TAS, Konge L. Expert Involvement and Adherence to Medical Evidence in Medical Mobile Phone Apps: A Systematic Review. JMIR Mhealth Uhealth 2015 Jul 27;3(3):e79 [FREE Full text] [doi: 10.2196/mhealth.4169] [Medline: 26215371]
12. Foley N, O'Connell E, Lehane E, Livingstone V, Maher B, Kaimkhani S, et al. PATI: Patient accessed tailored information: A pilot study to evaluate the effect on preoperative breast cancer patients of information delivered via a mobile application. Breast 2016 Dec;30:54-58. [doi: 10.1016/j.breast.2016.08.012] [Medline: 27611236]
13. Nasi G, Cucciniello M, Guerrazzi C. The performance of mHealth in cancer supportive care: a research agenda. J Med Internet Res 2015 Feb 13;17(1):e9 [FREE Full text] [doi: 10.2196/jmir.3764] [Medline: 25720295]
14. McCorkle R, Ercolano E, Lazenby M, Schulman-Green D, Schilling LS, Lorig K, et al. Self-management: Enabling and empowering patients living with cancer as a chronic illness. CA Cancer J Clin 2011;61(1):50-62 [FREE Full text] [doi: 10.3322/caac.20093] [Medline: 21205833]
15. Galligioni E, Piras EM, Galvagni M, Eccher C, Caramatti S, Zanolli D, et al. Integrating mHealth in Oncology: Experience in the Province of Trento. J Med Internet Res 2015;17(5):e114 [FREE Full text] [doi: 10.2196/jmir.3743] [Medline: 25972226]
16. Fiordelli M, Diviani N, Schulz PJ. Mapping mHealth research: a decade of evolution. J Med Internet Res 2013;15(5):e95 [FREE Full text] [doi: 10.2196/jmir.2430] [Medline: 23697600]
17. Wicks P, Chiauzzi E. 'Trust but verify' - five approaches to ensure safe medical apps. BMC Med 2015;13:205 [FREE Full text] [doi: 10.1186/s12916-015-0451-z] [Medline: 26404791]
18. CRD's Guidance for Undertaking Reviews in Health Care. Layerthorpe: University of York, Centre for Reviews & Dissemination; 2009.
19. Armijo-Olivo S, Stiles C, Hagen N, Biondo P, Cummings G. Assessment of study quality for systematic reviews: a comparison of the Cochrane Collaboration Risk of Bias Tool and the Effective Public Health Practice Project Quality Assessment Tool: methodological research. J Eval Clin Pract 2012 Feb;18(1):12-18. [doi: 10.1111/j.1365-2753.2010.01516.x] [Medline: 20698919]
20. Joanna Briggs Institute Reviewers' Manual edition 2014. Adelaide: The Joanna Briggs Institute; 2014.
21. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Int J Surg 2010;8(5):336-341 [FREE Full text] [doi: 10.1016/j.ijsu.2010.02.007] [Medline: 20171303]
22. Hartman SJ, Nelson SH, Cadmus-Bertram LA, Patterson RE, Parker BA, Pierce JP. Technology- and Phone-Based Weight Loss Intervention: Pilot RCT in Women at Elevated Breast Cancer Risk. Am J Prev Med 2016 Nov;51(5):714-721. [doi: 10.1016/j.amepre.2016.06.024] [Medline: 27593420]
23. Alanzi TM, Alobrah A, Alhumaidi R, Aloraifi S. Evaluation of the SnapChat mobile social networking application for breast cancer awareness among Saudi students in the Dammam Region of the Kingdom of Saudi Arabia. Breast Cancer (Dove Med Press) 2018;10:113-119 [FREE Full text] [doi: 10.2147/BCTT.S166135] [Medline: 30034251]
24. Eden KB, Scariati P, Klein K, Watson L, Remiker M, Hribar M, et al. Mammography Decision Aid Reduces Decisional Conflict for Women in Their Forties Considering Screening. J Womens Health (Larchmt) 2015 Dec;24(12):1013-1020 [FREE Full text] [doi: 10.1089/jwh.2015.5256] [Medline: 26360918]
25. Heo J, Chun M, Lee KY, Oh Y, Noh OK, Park RW. Effects of a smartphone application on breast self-examination: a feasibility study. Healthc Inform Res 2013 Dec;19(4):250-260 [FREE Full text] [doi: 10.4258/hir.2013.19.4.250] [Medline: 24523989]
26. Keohane D, Lehane E, Rutherford E, Livingstone V, Kelly L, Kaimkhani S, et al. Can an educational application increase risk perception accuracy amongst patients attending a high-risk breast cancer clinic? Breast 2017 Apr;32:192-198. [doi: 10.1016/j.breast.2017.02.009] [Medline: 28237842]
27. Lee H, Ghebre R, Le C, Jang YJ, Sharratt M, Yee D. Mobile Phone Multilevel and Multimedia Messaging Intervention for Breast Cancer Screening: Pilot Randomized Controlled Trial. JMIR Mhealth Uhealth 2017 Nov 07;5(11):e154 [FREE Full text] [doi: 10.2196/mhealth.7091] [Medline: 29113961]
28. Lee H, Lee M, Gao Z, Sadak K. Development and Evaluation of Culturally and Linguistically Tailored Mobile App to Promote Breast Cancer Screening. J Clin Med 2018 Jul 24;7(8) [FREE Full text] [doi: 10.3390/jcm7080181] [Medline: 30042291]
29. Egbring M, Far E, Roos M, Dietrich M, Brauchbar M, Kullak-Ublick GA, et al. A Mobile App to Stabilize Daily Functional Activity of Breast Cancer Patients in Collaboration With the Physician: A Randomized Controlled Clinical Trial. J Med Internet Res 2016 Sep 06;18(9):e238 [FREE Full text] [doi: 10.2196/jmir.6414] [Medline: 27601354]
30. Hwang H. Electronic wound monitoring after ambulatory breast cancer surgery: Improving patient care and satisfaction using a smart phone app. BCMJ 2016;58(8):448-453.

31. Kim J, Lim S, Min YH, Shin Y, Lee B, Sohn G, et al. Depression Screening Using Daily Mental-Health Ratings from a Smartphone Application for Breast Cancer Patients. J Med Internet Res 2016 Aug 04;18(8):e216 [FREE Full text] [doi: 10.2196/jmir.5598] [Medline: 27492880]
32. Min YH, Lee JW, Shin Y, Jo M, Sohn G, Lee J, et al. Daily collection of self-reporting sleep disturbance data via a smartphone app in breast cancer patients receiving chemotherapy: a feasibility study. J Med Internet Res 2014;16(5):e135 [FREE Full text] [doi: 10.2196/jmir.3421] [Medline: 24860070]
33. Weaver A, Love S, Larsen M, Shanyinde M, Waters R, Grainger L, et al. A pilot study: dose adaptation of capecitabine using mobile phone toxicity monitoring - supporting patients in their homes. Support Care Cancer 2014 Oct;22(10):2677-2685. [doi: 10.1007/s00520-014-2224-1] [Medline: 24771299]
34. Klasnja P, Hartzler A, Powell C, Pratt W. Supporting cancer patients' unanchored health information management with mobile technology. AMIA Annu Symp Proc 2011;2011:732-741 [FREE Full text] [Medline: 22195130]
35. Zhu J, Ebert L, Guo D, Yang S, Han Q, Chan SW. Mobile Breast Cancer e-Support Program for Chinese Women With Breast Cancer Undergoing Chemotherapy (Part 1): Qualitative Study of Women's Perceptions. JMIR Mhealth Uhealth 2018 Apr 11;6(4):e85 [FREE Full text] [doi: 10.2196/mhealth.9311] [Medline: 29643056]
36. Zhu J, Ebert L, Liu X, Wei D, Chan SW. Mobile Breast Cancer e-Support Program for Chinese Women With Breast Cancer Undergoing Chemotherapy (Part 2): Multicenter Randomized Controlled Trial. JMIR Mhealth Uhealth 2018 Apr 30;6(4):e104 [FREE Full text] [doi: 10.2196/mhealth.9438] [Medline: 29712622]
37. Rosen K, Paniagua S, Kazanis W, Jones S, Potter J. Quality of life among women diagnosed with breast Cancer: A randomized waitlist controlled trial of commercially available mobile app-delivered mindfulness training. Psychooncology 2018 Aug;27(8):2023-2030. [doi: 10.1002/pon.4764] [Medline: 29766596]
38. Wallwiener M, Matthies L, Simoes E, Keilmann L, Hartkopf AD, Sokolov AN, et al. Reliability of an e-PRO Tool of EORTC QLQ-C30 for Measurement of Health-Related Quality of Life in Patients With Breast Cancer: Prospective Randomized Trial. J Med Internet Res 2017 Sep 14;19(9):e322 [FREE Full text] [doi: 10.2196/jmir.8210] [Medline: 28912116]
39. Harder H, Holroyd P, Burkinshaw L, Watten P, Zammit C, Harris PR, et al. A user-centred approach to developing bWell, a mobile app for arm and shoulder exercises after breast cancer treatment. J Cancer Surviv 2017 Dec;11(6):732-742 [FREE Full text] [doi: 10.1007/s11764-017-0630-3] [Medline: 28741202]
40. Fazzino T, Sporn N, Befort C. A qualitative evaluation of a group phone-based weight loss intervention for rural breast cancer survivors: Themes and mechanisms of success. Support Care Cancer 2016 Dec;24(7):3165-3173 [FREE Full text] [doi: 10.1007/s00520-016-3149-7] [Medline: 26932848]
41. McCarroll M, Armbruster S, Pohle-Krauza R, Lyzen A, Min S, Nash D, et al. Feasibility of a lifestyle intervention for overweight/obese endometrial and breast cancer survivors using an interactive mobile application. Gynecol Oncol 2015 Jun;137(3):508-515. [doi: 10.1016/j.ygyno.2014.12.025] [Medline: 25681782]
42. Uhm KE, Yoo JS, Chung SH, Lee JD, Lee I, Kim JI, et al. Effects of exercise intervention in breast cancer patients: is mobile health (mHealth) with pedometer more effective than conventional program using brochure? Breast Cancer Res Treat 2017 Dec;161(3):443-452. [doi: 10.1007/s10549-016-4065-8] [Medline: 27933450]
43. Valle CG, Deal AM, Tate DF. Preventing weight gain in African American breast cancer survivors using smart scales and activity trackers: a randomized controlled pilot study. J Cancer Surviv 2017 Feb;11(1):133-148. [doi: 10.1007/s11764-016-0571-2] [Medline: 27631874]
44. Ainsworth MC, Pekmezi D, Bowles H, Ehlers D, McAuley E, Courneya KS, et al. Acceptability of a Mobile Phone App for Measuring Time Use in Breast Cancer Survivors (Life in a Day): Mixed-Methods Study. JMIR Cancer 2018 May 14;4(1):e9 [FREE Full text] [doi: 10.2196/cancer.8951] [Medline: 29759953]
45. Lozano-Lozano M, Galiano-Castillo N, Martín-Martín L, Pace-Bedetti N, Fernández-Lao C, Arroyo-Morales M, et al. Monitoring Energy Balance in Breast Cancer Survivors Using a Mobile App: Reliability Study. JMIR Mhealth Uhealth 2018 Mar 27;6(3):e67 [FREE Full text] [doi: 10.2196/mhealth.9669] [Medline: 29588273]
46. Pope Z, Lee J, Zeng N, Lee H, Gao Z. Feasibility of smartphone application and social media intervention on breast cancer survivors' health outcomes. Transl Behav Med 2018 Feb 17. [doi: 10.1093/tbm/iby002] [Medline: 29471477]
47. Lengacher C, Reich R, Ramesar S, Alinat C, Moscoso M, Cousin L, et al. Feasibility of the mobile mindfulness-based stress reduction for breast cancer (mMBSR(BC)) program for symptom improvement among breast cancer survivors. Psychooncology 2018 Feb;27(2):524-531. [doi: 10.1002/pon.4491] [Medline: 28665541]
48. Buscemi J, Buitrago D, Iacobelli F, Penedo F, Maciel C, Guitleman J, et al. Feasibility of a Smartphone-based pilot intervention for Hispanic breast cancer survivors: a brief report. Transl Behav Med 2018 Jul 07. [doi: 10.1093/tbm/iby058] [Medline: 29986120]
49. Visser A, Prins J, Jansen L, Radema S, Schlooz M, van Dalen T, et al. Group medical consultations (GMCs) and tablet-based online support group sessions in the follow-up of breast cancer: A multicenter randomized controlled trial. Breast 2018 Aug;40:181-188. [doi: 10.1016/j.breast.2018.05.012] [Medline: 29906741]
50. Williams SA, Schreier AM. The effect of education in managing side effects in women receiving chemotherapy for treatment of breast cancer. Oncol Nurs Forum 2004 Feb;31(1):E16-E23. [doi: 10.1188/04.ONF.E16-E23] [Medline: 14722602]

51. Attai DJ, Cowher MS, Al-Hamadani M, Schoger JM, Staley AC, Landercasper J. Twitter Social Media is an Effective Tool for Breast Cancer Patient Education and Support: Patient-Reported Outcomes by Survey. J Med Internet Res 2015;17(7):e188 [FREE Full text] [doi: 10.2196/jmir.4721] [Medline: 26228234]

Abbreviations

EORTC-BR23: European Organization for Research and Treatment Cancer Quality Of Life Questionnaire - Breast Cancer Module
EORTC-QLQ-C30: European Organization for Research and Treatment Cancer Quality of Life Questionnaire - C30
FACT-G: Functional Assessment of Cancer Therapy: General
IBIS: International Breast Cancer Intervention Study
mMBSR(BC): mobile Mindfulness-Based Stress Reduction for Breast Cancer
mHealth: mobile health
QUOTE: Quality of Care Through the Patient's Eyes

Edited by G Eysenbach; submitted 31.05.18; peer-reviewed by M Fiordelli, K Crew; comments to author 18.08.18; revised version received 20.10.18; accepted 30.10.18; published 11.02.19.

Please cite as:
Jongerius C, Russo S, Mazzocco K, Pravettoni G
Research-Tested Mobile Apps for Breast Cancer Care: Systematic Review
JMIR Mhealth Uhealth 2019;7(2):e10930
URL: http://mhealth.jmir.org/2019/2/e10930/
doi:10.2196/10930
PMID:30741644

Original Paper

Empowering Young People Living With Juvenile Idiopathic Arthritis to Better Communicate With Families and Care Teams: Content Analysis of Semistructured Interviews

Stuart W Grande[1], MPA, PhD; Meghan R Longacre[1], PhD; Karin Palmblad[2], MD, PhD; Meera V Montan[3], MPH; Rikard P Berquist[3], MSc, PMP; Andreas Hager[3], LLM; Greg Kotzbauer[1], MS

[1]Division of Health Policy and Management, School of Public Health, University of Minnesota, Minneapolis, MN, United States

[2]Karolinska Institutet, Department of Women and Child Health, Karolinska University Hospital, Stockholm, Sweden

[3]Upstream Dream, Bromma, Sweden

Corresponding Author:
Stuart W Grande, MPA, PhD
Division of Health Policy and Management
School of Public Health
University of Minnesota
420 Delaware St SE
Minneapolis, MN, 55455
United States
Phone: 1 612 625 3790
Email: sgrande@umn.edu

Abstract

Background: Young people living with juvenile idiopathic arthritis (JIA) face a number of communication barriers for achieving optimal health as they transition from pediatric care into adult care. Despite growing interest in mobile or wireless technologies to support health (mHealth), it is uncertain how these engagement tools might support young people, their families, and care teams to optimize preference-based treatment strategies.

Objective: This study aims to examine how an mHealth patient support system (mPSS) might foster partnership between young people living with JIA, their families, and care teams.

Methods: Semistructured interviews with young people (5-15 years old), their families, and JIA care teams were conducted using researcher-developed interviews guides. Transcribed data were qualitatively analyzed using conventional content analysis.

Results: We conducted semistructured interviews with 15 young people, their parents, and 4 care team members. Content analysis revealed the potential of an mPSS to support productive dialogue between families and care teams. We identified four main themes: (1) young people with JIA face communication challenges, (2) normalizing illness through shared experience may improve adherence, (3) partnership opens windows into illness experiences, and (4) readiness to engage appears critical for clinic implementation.

Conclusions: A human-centered mPSS design that offers JIA patients the ability to track personally relevant illness concerns and needs can enhance communication, generate consensus-based treatment decisions, and improve efficiency and personalization of care. Technology that supports continuous learning and promotes better understanding of disease management may reduce practice burden while increasing patient engagement and autonomy in fostering lasting treatment decisions and ultimately supporting personalized care and improving outcomes.

(JMIR Mhealth Uhealth 2019;7(2):e10401) doi:10.2196/10401

KEYWORDS

juvenile arthritis; interviews; health communication; patient participation

Introduction

Young people living with juvenile idiopathic arthritis (JIA) face a number of communication barriers for achieving optimal health. JIA is one of the most common acquired chronic diseases during childhood and affects both short-term and long-term disability. JIA may develop at any age during childhood, and girls are more often affected than boys. Nomenclature as well as classification of JIA have been controversial. In an attempt to develop globally accepted terminology and criteria, the International League against Rheumatism (ILAR) introduced the term JIA [1]. Using the ILAR criteria, a Nordic population-based epidemiological study reported a JIA incidence of 15 per 100,000 [2]. JIA is not a single disease but rather a heterogeneous group of diseases, divided into seven different subgroups, within which three major onset subtypes can be identified: (1) systemic JIA, a systemic form with rash, fever, and commonly perimyocarditis, (2) polyarthritis with 5 or more joints involved, and (3) oligoarthritis (or pauciarticular JIA) with 4 or fewer joints involved.

Among young people, JIA can be quite painful and lasting well into adulthood. While many experience resolution of their disease, 50% will live with chronic arthritis into adulthood [3]. Although the disease is self-limiting in some children, it is not possible at the onset of disease to predict which child will recover and which will have a lifelong challenge. Early diagnosis and active therapeutic interventions are essential to minimize residual deformity and disability due to irreversible consequences of the disease such as joint destructions, asymmetric bone growth, and vision impairment. The treatment for JIA is multifaceted and requires a combination of monitoring, physical therapy, joint injections, medications, and sometimes surgery [4]. Although optimal treatment of JIA varies extensively by individual patient, making it highly preference-sensitive, research on preferences among young people living with JIA is minimal. A review of 27 studies found that children with JIA fear being seen as different from their peers and are interested in seeking health information to manage their own illness [3]. In other words, the preferences of young people reported in these studies appear focused on developing autonomy around care management [5]. For many young people who are optimistic about their future, these tools offer novel approaches that guide a process for making informed decisions about their care and social needs.

As widely recommended, the best way to determine the most appropriate treatment options, where uncertainty is high, are through assessing patient values, priorities, and experiences as part of preference elicitation [6-8]. Research suggests that children who are actively engaged in treatment consultations with parents may improve their confidence in managing JIA into adulthood [9,10]. By continuing to support children, youth, and adolescents with information, social support, and active involvement in the management of their illness, their confidence and long-term health may improve [3]. Over the last 20 years, researchers have been working on building tools to help patients with chronic conditions better communicate preferences with clinicians. Patient engagement tools like decision aids have shown that adult patients are more knowledgeable, better informed, understand their values, and may be more engaged in decision making than previously thought [11], but there is little comparable evidence for young people.

Approaches to minimize the effects of JIA on young people vary widely, and their success is often very personalized to the individual. One study evaluating a decision aid for children with JIA found that the tool had high acceptability but lower efficacy, leading authors to call for more innovative approaches to using decision aids and assessing outcomes among children [12]. A review of mobile phone and tablet apps that support personal management of illness in young people found that apps for diabetes, asthma, and chemotherapy recovery show some impact on monitoring and adherence, but the strength of evidence is weak [13]. Others examined an online decision aid for patients with rheumatoid arthritis and found that decision comfort and knowledge improved [14]. In one example, a group in Canada that designed an iPhone app for adolescents with cancer called Pain Squad+, found that applying a user-centered design strategy led to an acceptable and effective tool for the self-management of pain [15]. While evidence points to strengths of online tools for promoting self-management and guiding patient preference, there appears to be less evidence regarding online interventions or apps designed with the intention of promoting partnership, collaboration, or consensus between young people, families, and clinicians/care teams.

The benefits of using the Internet to search for information related to illness and JIA are well known, yet the documented benefits for online peer-support for young people is less robust [16]. Recent evidence of other applications for engaging young people in mobile apps for self-management demonstrate high acceptability and usability [17]. There is also evidence in other fields that peer-support has emerged naturally in online environments, where the benefits appear based on personal preferences for use, ease of access, and interaction with others who face similar challenges [18]. This suggests that online tools or apps should be meaningful to those who use them and will benefit those who take time to engage with them online [13]. The question remains how these newer technologies may best support young people to develop skills in self-management and treatment decision making that will serve them into adulthood.

In addition, developing tools to support the adolescent transition to long-term self-management in JIA are needed [19]. Research has called for shifting habits of clinician-centered problem solving to an expanded understanding of patient experience [20,21], arguably demonstrating a more general demand for co-produced care plans [22]. Conceptually, co-production reorients traditional models of care delivery towards a more patient-centered or person-centered approach. In practice, this has involved clinicians supporting a standardized assessment of patients, applying guidelines to inform care and offering non-narcotic medication management protocols [23-25]. Other patient-centered approaches include pre-visit planning by giving evidence-based information and advocacy tools to guide preference development to enhance decision making at the point of care [26]. When it comes to developing communication skills for young people as they transition from parent/clinician-supported decision-making models to more autonomous decision-making ones, the best approaches are less

understood. While the challenges of communicating with young people with JIA have been well described [27], a major contributing factor to reduced quality of care outcomes for young people with JIA is poor adherence to treatment [28,29]. Adolescence is a critical time for developing long-term healthy habits as adults, and while online and mHealth skill development for self-management is suggested, there is little evidence on how best to build and evaluate these skills [27].

The purpose of this study was to examine how an mHealth Patient Support System (mPSS) could help improve health communication between young people, their parents, and their care team. This study examined how Genia might ready young people with chronic illness for their inevitable transition to adult care. Specifically, we explored how the mPSS influenced the engagement in self-management of JIA and communication between young people, their families, and their clinical care providers.

Methods

Genia is an iOS-based mPSS designed to establish patient, family, and care team partnerships, with an emphasis on placing the young person at the center of the decision-making process (Figure 1). As a platform and mPSS, Genia was designed with principles of user-based design [30], feed-forward systems [31], and concepts of co-production [22]. Core functionality of the mPSS were operationalized within a learning collaborative associated with the Lund Pediatric Cystic Fibrosis clinical microsystem. This meant that clinicians, researchers, family members, and cystic fibrosis patients were involved in iterative cycles of change from inception to completion of the mPSS. Key functionalities developed within the app were inspired and optimized by patients, viewed as experts in their disease and experience of care [32].

Further details of functionality are published elsewhere but summarized briefly here [32]. There is basic functionality for patient users to chat with other patient users in the app. All information entered into the app is private to users and their network. Users can, however, choose to share information with others by sending reports (ie, structured forms with specific questions for pre-visit planning) that also allow users to share self-selected daily observations and/or graphs of symptom/assessment tracking for discussion (ie, a graph of self-reported pain and fatigue in the last month). Reports can be sent through the app to clinical care teams via application program interfaces that connect to clinic IT systems. Users can also choose to send reports to others (ie, family or friends) using a basic share function within an operating system, enabling them to share via email or instant messaging.

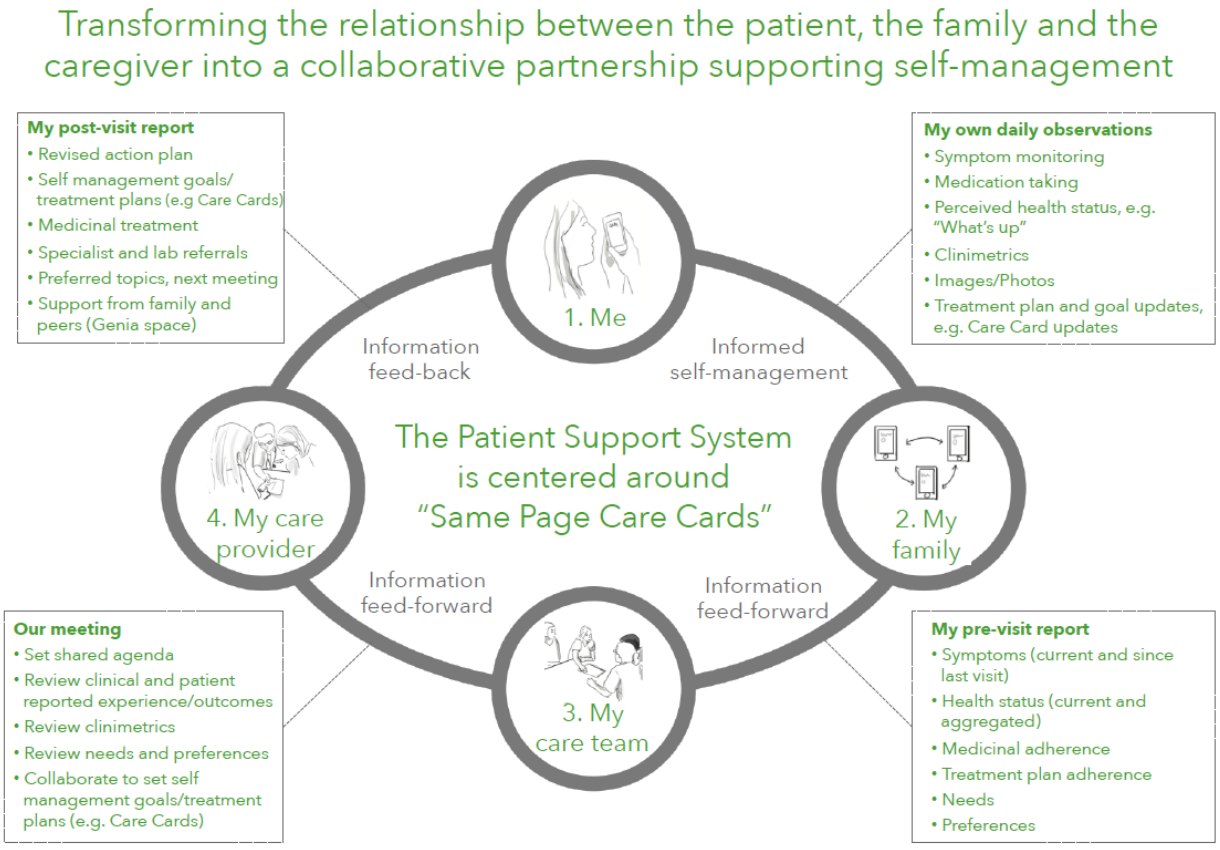
The mPSS also captures the patient experience using a series of daily observation “trackers” that allow the patient to record daily observations and perspectives. For example, one tracker permits patients to identify what they are feeling that day, a feature notable for young people who may not be used to calling attention to or labeling their feelings. Such patterns have been identified as contributing to emotional competence in young people and are supported by parent-child and other socialization models [33]. Another tracker allows a patient to record their level of physical activity, but more unique is a tracker that includes the patient’s social interests and their assessment of life with family and with friends. This distinguishing mPSS feature permits the young person to make note of important preferences, in this case needs, wants, and fears, for subsequent review with friends, family, or care teams. This form of emotional notation and engagement has been linked to childhood development of emotional regulation and positive social engagement [34,35].

Patient-reported daily observations enable young people to document their disease activity and preferences in real-time in between clinical visits. Each of these daily data points is collected into a dashboard that can be shared electronically with a patient-determined list of observers such as family or friends they can invite to their network. They can also include these dashboards in pre-visit reports to clinical care providers. For older children and adolescents, the patient is the locus-of-control, determining who has access, the level of interaction with the app, and the amount of data being shared. For younger pediatric patients, parents can serve as proxies by entering daily observations about their child into the app.

We approached the research aim from the perspective that subjective experiences in the scope of JIA are influenced by place, persons, and institutions. For this study the place was Sweden, the persons were children aged 5-15 years living with JIA and their parents, and the place of care was the largest pediatric rheumatology clinic in Sweden, at Astrid Lindgren Children’s Hospital, within Karolinska University located in Stockholm, Sweden.

Two Swedish research assistants conducted semistructured interviews with Swedish young people, their parents, and JIA care teams about their use of and reactions to the potential benefits of the mPSS. Two researcher-designed interview guides (Multimedia Appendices 1 and 2) were developed (one for families and another for clinical team members) with clinical and project partners to include the following domains: current use of technology, health communication between clinicians and patients, personal health tracking and symptom monitoring, preparation for clinic visits, and current illness management. Data collection and analysis were supported by a multidisciplinary team, which included a medical sociologist, a health care policy expert, and a developmental psychologist. This research team worked closely with Upstream Dream, the mPSS development team in Sweden that created Genia, to gain access to clinical care partners, understand clinic organizational dynamics, and examine product development history. The study was determined to be exempt from ethics review based on Common Rule 2 by the Committee (blinded for review) for the Protection of Human Subjects at Dartmouth during a blinded review. This study was approved by the Stockholm Ethical Committee, Stockholm, Sweden. Written consent was obtained from all patients/families by the clinicians participating in the study.

Figure 1. Genia flow diagram.



Study Participants

Potential study participants were identified by clinical partners who already use the mPSS with some of their patients and families. Clinicians then sent recruitment letters to participants with a Genia brochure inviting them to participate. The JIA clinical care team invited patients to participate in the study. Consent involved reading an informational script including the study purpose, design, and goals. Families who agreed to participate were subsequently interviewed in-person by Swedish researchers at a location and time that were convenient. Interviews were conducted with parents and children together based on preference and ethical considerations. Some of the participants had not used the mPSS prior to being interviewed, while others had already started using the mPSS. All interviews were conducted in Swedish and later translated into English. The clinical care team interviews were led by the US research team and conducted face-to-face in English at the JIA clinics in Stockholm. The clinical care team included a physiotherapist, occupational therapist, and 2 physicians of the JIA clinic. Research team members provided the interview guide to the care team members prior to the interview to give adequate time to prepare responses in English.

Data Analysis

We conducted a conventional content analysis as described by Hsieh and Shannon [36]. Initial codes were developed independently by 2 researchers (ML and SG) using ATLAS.ti version 8.1.2. These first set of codes reflected interactions and observations by the research team from 2016-2017 as recruitment, design, and data collection were occurring. This initial code set was later updated through a second coding process based on the interview transcripts for both young people and families as well as care providers. All members of the research and clinical team who collected data spoke and wrote English fluently. To ensure accuracy of transcripts, 2 Swedish team members reviewed translations and confirmed translation with Swedish-American partners. Secondary codes were substantively related to the interview guide, in part because of how the guide was organized by domains, as well as due to the nature of the conversations that focused on specific JIA-related experiences. Following the second round of coding, the research team discussed emergent codes, which were labeled new language and concepts. Team members (SG and ML) developed memos (ie, a process of writing short descriptive narratives summarizing meaningful aspects of the data) during the coding and discussion process to draw attention to notable relationships in the codes. Team members (SG, ML, and GK) met several times over the course of 2017 to build consensus around grouping codes based on emergent and context-based themes. During the coding and subsequent consensus process, the Genia development team was asked to react to initial codes and offer alternative explanations and suggestions. Data saturation was assessed using an iterative process of constant comparison [37] throughout data collection and analysis phases. Upon a third review of data at 12 participants, the presence of language patterns and repeated concepts alerted team to potential saturation. The team agreed to three additional interviews to provide some methodological assurance that no new patient

experiences were being missed or overlooked [38]. This iterative coding, memo writing, and theme building process mimics a type of data triangulation [39]. Final themes were developed by grouping codes around larger meaning units, whose content sufficiently reflected all data under consideration.

Results

Study Participants

The age, gender, and disease information for the study participants are shown in Table 1.

The young people interviewed for this study ranged in age from 5-15 years old. There were 7 females in the group, and 13 of the 15 had used Genia prior to the interview. The time since diagnosis ranged from 1-14 years. We also interviewed four clinicians who used Genia as part of their routine practice: 2 medical doctors, 1 physiotherapist, and 1 occupational therapist. Patient interviews were conducted with both the young person and at least one parent present. The clinician interviews were conducted individually, except two, which included both the physiotherapist and occupational therapist together.

A content analysis of transcripts generated four themes that characterize attitudes and beliefs as well as fears and expectations on how an mPSS may foster a strategy for improving communication and achieving shared decision making in JIA treatment. Experiences of young people dealing with JIA every day reflected feelings of frustration and confusion, which we termed, “Young people with JIA face communication challenges.” While the data suggest there is a developmental difference between younger and older patients’ feelings about sharing their diagnosis status with friends, there was affirming content that supported a second idea we portrayed as, “Normalizing illness through shared experience may improve adherence.” When asked about the potential of technologies or current care processes, families spoke about the value of routine check-ins between visits and symptom updates. The habit of checking-in and preparing for visits appeared to lead to deeper engagement between young people and their parents. We described this concept as, “Partnership opens a window into illness experiences.” Interview data were further informed by clinician points of view, which included support for more informed patients and families as well as for recognizing the challenge of communicating with young people. More to the point of implementation and use of a novel technology in clinical practice, we observed wide support for an mPSS and called this idea, “Readiness to engage appears critical for clinic implementation.” We present supporting illustrative quotes drawn from transcripts to characterize the categories in the sections that follow.

Young people with Juvenile Idiopathic Arthritis face communication challenges

Young JIA patients, when they experienced pain, described feeling frustrated because people (often classmates) did not believe them, or the classmates believed they were lying to get out of an activity. This appeared to separate the young person from others, both physically and socially.

> *Patient 14 (14 years old): Everyone in my family is aware, so it’s okay. It’s hardest with classmates.*
>
> *Interviewer: What do they say then?*
>
> *Patient 14: One time there was one who said that I always blame everything on my foot.*

We observed a tendency in these interviews for parents to describe interrupting the doctor’s visit to elaborate on details that either the child did not think was important or did not remember. These interruptions were viewed by young people as acceptable and often confirmatory.

> *Patient 7 (15 years old): I usually do it [the clinical visit] but sometimes Mom helps me to make it fair. Sometimes I say it doesn’t hurt at all, though it actually has been hurting also, then Mom jumps in [giggles].*

We noticed that parents also have their own ways of dealing with their children’s pain and social struggles. Some have begun to reach out to other parents using social media platforms like Facebook, while others have talked with other parents in person. This reaching out for help reflects both the complexity and challenges of supporting children with JIA. Parents showed a desire to be helpful by wanting effective tools for their children but often lack the resources. The act of reaching out to other parents through Facebook groups underscored the importance and value of an mPSS designed to support parents to better understand their children’s illness and acquire the skills to better assist their children.

We observed young children talking about the benefits of working with their parents to help them communicate about pain and patterns in their eating and behavior. Some reflected how parents were motivated to help them develop techniques to manage their JIA independently. In one case this approach appeared to work.

> *Patient 10 (15 years old): Ehh...I often do not have so many of my own views. He [the clinician] usually asks me things. Sometimes I have my own questions, for example last year I had trouble with my knees, then my questions could be about that.*
>
> *Mother: It’s mostly about how you’ve had it lately.*
>
> *Interviewer: Do you usually go with him (asking the mother)?*
>
> *Mother: Yes, I usually do. But we try to hand over it more and more to [my son] because he’s getting older. It is important now that he is moving over to adult care.*
>
> *Interviewer: How do you feel about moving over to adult care?*
>
> *Patient 10: Hmm...well it feels good.*

Table 1. Description of participants.

Participant	Age in years[a]	Sex	JIA subtype[b]	Duration of disease[c]
Patient 1	11	Female	Oligo JIA	1 year
Patient 2	10	Male	PsA	4 years
Patient 3	13	Female	ERA	2 years
Patient 4	12	Male	Oligo JIA	3 years
Patient 5	11	Female	Poly JIA	1 year
Patient 6	5	Male	Oligo JIA	2 years
Patient 7	15	Female	PsA	14 years
Patient 8	12	Male	Poly JIA	4 years
Patient 9	15	Male	Undifferentiated	1 year
Patient 10	15	Male	PsA	2 years
Patient 11	11	Male	Oligo JIA	4 years
Patient 12	14	Female	JIA	5 years
Patient 13	13	Female	Oligo JIA	1 year
Patient 14	14	Female	Poly JIA	9 years
Patient 15	14	Male	JAS	4 years

[a]The average age of participants was 12.3 years (SD 2.3).

[b]Subtype for JIA: Oligo JIA=oligoarticular JIA, Poly JIA=polyarticular JIA, PsA=psoriatic JIA, ERA=enthesitis-related arthritis, JAS=juvenile ankylosing spondylitis, and undifferentiated=multisymptom JIA.

[c]The average duration of disease was 3.8 years (SD 3.5).

We further noted that most young people and adolescents in the interviews were reluctant to complain about pain or discomfort. Many felt they were a burden on their parents. Reluctance to talk may also be linked to an inability to communicate feelings effectively or how they experienced pain. As one young person showed, reluctance may be a show of strength. For example, one young person expressed his feelings and attitudes on sharing symptoms, which may have signaled strength or fear. Certainly, the interjection by the parent in the following dialogue called our attention to the challenge of communication for both parents and their children.

> *Patient 9 (15 years old): It feels good.*
>
> *Interviewer: Do you feel that you are telling him [the clinician] exactly how you feel?*
>
> *Patient 9: Yes, I think so.*
>
> *Mother: ...but he always says he's better than he is.*
>
> *Patient 9: No.*
>
> *Mother: Yes.*

Normalizing Illness Through Shared Experience May Improve Treatment Plan Adherence

There were some comments on connecting others with JIA through social media as a means of overcoming feelings of isolation and inability to offer help. The act of reaching out for connection may be a means of feeling normal or at least confirming that one is not alone. As we observed, younger children seemed uninterested in reaching out through social media. Yet, the fact that a 14-year old was attempting to connect virtually with others suggested that as young people age, they become more aware of their friends and social networks as sources of support.

> *Interviewer: Do you know how to get in touch and talk to peers?*
>
> *Patient 12 (14 years old): Yes, those I've met I have on Snapchat and Instagram.*
>
> *Interviewer: Do you have Facebook?*
>
> *Patient 12: No, I don't. But I have their numbers so I can call them.*
>
> *Interviewer: The old way (ha-ha) Mother: ha-ha*
>
> *Mother: (She) is in touch with some from her last rehab trip. Then also from her physical therapy, not just people with rheumatism, even people who need training.*
>
> *Interviewer: How do you think it's [nice] having friends who also have rheumatism?*
>
> *Patient 12: Nice. The other day we sat up and talked and it was very nice.*

We saw that young people seemed reliant on communicating with their parents, particularly when it came to navigating their relationship with their doctor or care provider. We noticed that this type of communication was not only about symptoms but about being empowered to communicate comfortably and build trust. Further, parents reflected on the power of sharing in a way their children felt confident when hearing.

> *Father of 11-year-old patient: Yes, absolutely. Since it is so reoccurring...and to be able to share how (she) feels, and how bad it is. If this could be possible*

through an app it would be great, both for us and for the doctors.

In one example, we saw parents being challenged to think about ways the mPSS may help support behaviors like monitoring pain or symptoms over time. This was seen widely as a way to improve their child's awareness and self-management skills.

Interviewer: Do you have a diary about how you feel?

Mother: No, we do not actually. But I have it very much in my head all the time, but it's the same all the time. But on the other hand, I think the app [Genia] can help because we do not remember the days that actually are good. It's the misery most of the time. I think it would be great for [my daughter] to see her situation over time. So, she can see that she has better days.

The young people interviewed talked about the challenges of describing their pain and frustration with JIA, whether with physicians or friends. When asked about how Genia might contribute to talking about their JIA issues, several young people referred to the value of keeping a daily diary as a way to have better or more productive visits with their doctor.

Mother: He was so terribly afraid that it would not last (reduced pain). But then he started playing sports...living a normal life.

Interviewer: Do you remember how you felt?

Patient 15: It felt very strange at the beginning.

Mother: You came back to life.

Interviewer: Does your conversation help you manage your rheumatism?

Patient 15: Yes, it does.

Interviewer: Is there something good or less good?

Patient 15: I would say everything is good.

Interviewer: Do you keep a diary about how you feel?

Patient 15: No, only when I take syringes, then I fill it in in Genia. I think that kind of documentation is good for remembering when I talk to my doctor.

Overall feedback from young people related to the value of Genia, or potential of Genia's functionality, suggested that diary entries and note-taking improved conversations with care teams, which helped families and young people remember benefits of certain medications and treatment decisions.

Partnership Opens a Window Into Illness Experiences

We heard from young people and their families that using a patient support tool to keep daily observations and reminders for routine updates was very helpful for developing strategies to communicate with providers.

Interviewer: In Genia you can send a report before a health care meeting, do you think that could be something for you?

Patient 12 (14 years old): Yes, for me at least.

Interviewer: In what way?

Patient 12: I have difficulties remembering. I mix things up. I cannot tell when something was, it could have been two weeks ago or two months ago.

Mother: Although we try to be a bit prepared, it is incredibly difficult.

At the same time, clinicians also believed there was real, meaningful value in young people and their parents being better prepared for their visit.

The data is helpful, and the registry really helps me see the patient condition over time – but it only shows data at a point in time (at the time of visits). It doesn't help me understand what is going on in the patient's life between visits. And if they aren't communicative during the visit, or their feelings contradict those of their parents, then it is hard for me to truly understand the patient needs; it is hard for me to understand why the patients' perceptions of their pain and quality of life is different than my perception. [Physician 1]

While having informed young people and families might improve visit efficiency by targeting interests and generating focused questions, additional behavior changes appeared to guide strategies for improved clinical conversations. The clinicians we interviewed argued that the mPSS presented a connection into the lives of patients that was meaningfully different from routine practice: "For the first time, I feel part of her [patient's] team" (Physician 1).

Readiness to Engage Appears Critical for Clinic Implementation

Patient monitoring following a care visit was expressed as a gap in care by clinicians. This has particular significance given that patients contend with a majority of their care challenges outside the clinic. The clinicians we interviewed shared an interest in being able to track or follow patient progress over time, which permits new and advantageous support systems: "With Genia, if we can get patients to send us status reports in between scheduled visits, then we can confirm what is working and try new ideas much more frequently" (Physician 2).

Besides the importance of tracking, clinicians reflected that engaging patients outside the clinic in meaningful and proactive ways was essential for improving communication during the clinic visit. In this way, clinicians recognized that the pre-visit data sharing feature of the mPSS app enabled more targeted and thoughtful interactions: "To improve pre-visit planning, to provide more time for dialogue with the patient during the visit, to help patients better understand what causes flare-ups and what therapy seems to work, and to provide a mechanism to track patients between visits" (Physician 1).

Other reactions highlighted the potential of the mPSS to modify routine practice. This was seen as a major improvement compared to standard care. One clinician pointed to gaps in current communication where patients are unable to speak to their symptoms, pain, or status: "When I asked how are you today, everyone says 'fine' but I don't really have a clear picture. To know if there is something more I can do" (Physician 1).

Another clinician also emphasized the need for more robust efforts to guide patients prior to coming to their clinical visit. When asked if she understood the needs of her patients, one clinician suggested there was a need for improvement: "No, not really. Patients complete a long questionnaire when they get to the office, which captures responses based on industry standard instruments, but they don't tell me [their] goals or key concerns" (Physiotherapist).

We noticed that the mPSS clarified the role of the patient, which fostered guided support and strengthened self-management skills. In fact, we saw that use of the mPSS in practice, with patients over time, appeared to change their behavior. In the case of one young person, a clinician commented that they would normally rely on the parent to share symptoms. After using the mPSS, this parent felt their child was able to do much more with the doctor: "The mother used to bring in a clipboard and do all of the talking, now the patient herself leads the conversation, leaving the mother a more bystander role" (Physician 1).

Discussion

Principal Findings

A co-designed mobile patient support system that meaningfully engaged users beyond the clinic visit expanded opportunities for improving treatment strategies between young people, their families, and care teams. Through the promotion of consensus-building design features, we observed improved communication in young people and their parents about symptom recognition and pain characterization. As an mHealth app, Genia, appeared to hone communication skills that parents and clinicians believe ensure a healthy transition from pediatric to adult care—a space with clear unmet need. The insights for young people, families, and clinicians to enable more substantive, targeted, and authentic communication, where young people learn to reflect on symptoms without being prompted by their parents, was a novel finding. As a mobile phone app, the routine monitoring of symptoms and pain over time normalizes a type of sharing process that appears to ease disclosure and improve efficacy in young people and parents. This was particularly relevant for young people who felt isolated by this illness and struggled to disclose their experiences. While evidence has shown how integration of novel interventions into clinical practice can be difficult, clinicians who were ready and prepared to use a novel approach appeared to mitigate barriers to Genia implementation.

Comparison With Previous Research

With the explosive growth of and interest in mHealth initiatives, there seem to be endless opportunities for patients to engage the health system in new ways. This is one of the first studies to identify an mHealth app designed to modify communication strategies between patients, families, and care teams. Most mHealth interventions reviewed focus on self-management skill development rather than improving or optimizing patient-provider communication strategies [40]. For mHealth technology to advance behavior change strategies in young people living with JIA, more should be done to build skills that enhance consensus-building frames between patients, families, and care teams. Such an approach is consistent with the literature on the benefits of building a shared mind as a treatment strategy [41].

This study is a step towards being able to identify how an mHealth app, co-designed with patients, families, and care teams, helps identify what matters most, as well as ways of translating patient goals into actionable solutions. Focusing on principles of continuous learning and consensus building, the patient support system prioritizes processes of monitoring, reporting, and pre-visit planning to inform clinical communication. Despite the wide growth and interest in these novel technologies for improving access, there is mixed evidence on the long-term impact on patient health outcomes or behaviors [42]. This form of co-development has been well supported by health services research, which has pointed to the benefits of co-designed tools for behavior modification strategies [43-45].

A recent paper detailed the user-centered development and evaluation of an app, JIApp, which applied this same person-centered development framework and strongly supports our findings [17]. A key difference between JIApp and Genia is the co-design strategy focused on improving consensus for optimizing treatment decision making. The creators of JIApp aimed their intervention on self-management and engagement, primarily focused on "recording symptoms and encouraging self-management." While collaboration is a key feature of patient-provider communication and self-management skills are associated with a higher quality of engagement, Genia as an mPSS was designed to modify patient-provider interaction and behaviors that likely influence consensus-building skills between young people, families, and care teams for long-term management of treatment. As mentioned previously, there is opportunity for users to chat with other users directly in the app. However, this functionality was rarely used by the participants in this study and we believe that there is more work to be done to explore how this peer-to-peer connection could be enhanced to support patients. While this is still a burgeoning area of investigation, more research is needed on understanding the co-design and implementation process of mHealth apps. Additional research is also needed to identify and measure ways that mHealth apps like Genia or JIApp modify communication responses for young people in transition from pediatric to adult care services. When compared against apps developed in isolation and often not in partnership with care providers, there is a clear benefit of co-designed apps that foster evidence-based approaches and direct input from end-users like patients and care providers [13]. Apps that support self-management and symptom management have higher potential for success, which supports the design and implementation of Genia as an mPSS [46].

What appears unique about the Genia approach in comparison to other published apps is the intentional strategy used to guide both patient and family engagement. The use of qualitative data to inform evidence on platforms that are most effective for patients has been suggested elsewhere [47]. Much of the current literature points to the strengths of mHealth to enhance self-management [42,48], tracking and monitoring [49], as well as routine behavior modification [50], and medicine adherence [42,51]. Usability testing of an online self-management health

portal with adolescents with JIA and their parents underscored the importance of building these self-management platforms with user feedback, a key feature of any person-centered design [52]. Beyond the JIApp [17], there was no other evidence of mHealth apps developed to improve communication strategies or consensus-based strategies to enhance patient, family, and care team integration, particularly around the concept of shared reasoning [42]. Failures or limitations of other mHealth apps and strategies appear driven by a one-dimensional approach, which prioritizes biomarkers and tracking and fails to incorporate consensus building and communication skills. There is no question that mHealth initiatives are the way of the future; they are cheap, easy to build, and potentially highly accessible to large populations. The benefits of mHealth technology are that they are personal, adaptive, and sustainably designed. Yet, these benefits are hindered by a lack of clear standards of evaluation and measurement [53].

Limitations

Although our study had 15 participants and involved both families and young people, the data analysis included triangulated data sources to help clarify and validate responses. Due to the number of participants and their variation in ages and experiences, findings must be read with caution. Qualitative approaches have limited generalizability outside the scope of participants' lived experiences and other clinical settings. The unique responses of parents, care team physicians, and young people discussed here provide insight into expectations and hopes for mHealth interventions, especially interventions designed with and for end-users.

While Genia was developed in partnership with a Karolinska clinical team along with patients and parents, there were many new processes and functionalities that should be considered "new" as they were not extracted or modified from a pre-existing mobile phone app. While other research on this mPSS had been done within cystic fibrosis, the unique nature of JIA required new design and functionalities determined within a pilot phase, which has been reflected in some of the work presented here. These data and findings provide a rich and context-specific perspective of how an mPSS was adopted, integrated, and utilized by a small sample of patients and clinicians in Sweden. The nature of the research question reviewed here and its supporting methodological approach contributed to valuable context-driven findings. However, these same methods are unable to provide insight into the long-term impact of mHealth apps like Genia on changes in self-management or communication skills for young people transitioning into adult care as well among adults who have experienced transition. Moving forward, designing a study to answer long-term impacts of young people's use of mHealth apps for self-management of JIA will be an important contribution to determine the overall efficacy of an mPSS in improving self-management and associated communication skills.

Conclusions

A technology-enabled mPSS that meaningfully engages care providers to partner with patients, families, and their support networks permits novel care planning through the formation of a consensus-building strategy. We believe that offering patients the opportunity to engage friends, family, and their care team in developing treatment solutions may provide the emotional and clinical support they need to meet their personal health goals and sense of well-being.

Acknowledgments

We would like to thank the generosity of families who agreed to participate in this project and the hard work of Nathalie Ebel and Madeleine Grimhusen for conducting the interviews. We appreciate the scientific input at various phases of the study and editorial comments provided by Sarah Donelson and the program management provided by Raya Koury.

This work was funded as a collaborative research project with Genentech, Inc.

Conflicts of Interest

AH, MM, and RB are employees of Upstream Dream, a private company that developed the Genia app.

Multimedia Appendix 1

Semistructured interview guide.

[PDF File (Adobe PDF File), 35KB - mhealth_v7i2e10401_app1.pdf]

Multimedia Appendix 2

JIA care team interview guide.

[PDF File (Adobe PDF File), 38KB - mhealth_v7i2e10401_app2.pdf]

References

1. Petty RE. Classification of childhood arthritis: a work in progress. Baillieres Clin Rheumatol 1998 May;12(2):181-190. [Medline: 9890092]

2. Berntson L, Andersson Gäre B, Fasth A, Herlin T, Kristinsson J, Lahdenne P, Nordic Study Group. Incidence of juvenile idiopathic arthritis in the Nordic countries. A population based study with special reference to the validity of the ILAR and EULAR criteria. J Rheumatol 2003 Oct;30(10):2275-2282. [Medline: 14528529]
3. Tong A, Jones J, Craig JC, Singh-Grewal D. Children's experiences of living with juvenile idiopathic arthritis: a thematic synthesis of qualitative studies. Arthritis Care Res (Hoboken) 2012 Sep;64(9):1392-1404 [FREE Full text] [doi: 10.1002/acr.21695] [Medline: 22504867]
4. Giancane G, Consolaro A, Lanni S, Davì S, Schiappapietra B, Ravelli A. Juvenile Idiopathic Arthritis: Diagnosis and Treatment. Rheumatol Ther 2016 Dec;3(2):187-207 [FREE Full text] [doi: 10.1007/s40744-016-0040-4] [Medline: 27747582]
5. Shaw KL, Southwood TR, McDonagh JE, British Paediatric Rheumatology Group. User perspectives of transitional care for adolescents with juvenile idiopathic arthritis. Rheumatology (Oxford) 2004 Jun;43(6):770-778. [doi: 10.1093/rheumatology/keh175] [Medline: 15039498]
6. Mulley AG, Trimble C, Elwyn G. Stop the silent misdiagnosis: patients' preferences matter. BMJ 2012 Nov 08;345:e6572. [doi: 10.1136/bmj.e6572] [Medline: 23137819]
7. Politi MC, Street RL. The importance of communication in collaborative decision making: facilitating shared mind and the management of uncertainty. J Eval Clin Pract 2011 Aug;17(4):579-584. [doi: 10.1111/j.1365-2753.2010.01549.x] [Medline: 20831668]
8. Elwyn G, Lloyd A, May C, van der Weijden T, Stiggelbout A, Edwards A, et al. Collaborative deliberation: a model for patient care. Patient Educ Couns 2014 Nov;97(2):158-164. [doi: 10.1016/j.pec.2014.07.027] [Medline: 25175366]
9. Dures E. I90 Well-Being and the Ability to Self-Manage: The Support needs of Patients with Inflammatory Arthritis. Rheumatology 2016;55(Issue suppl_1):i18 [FREE Full text] [doi: 10.1093/rheumatology/kew097.001]
10. Lipstein EA, Muething KA, Dodds CM, Britto MT. "I'm the one taking it": adolescent participation in chronic disease treatment decisions. J Adolesc Health 2013 Aug;53(2):253-259. [doi: 10.1016/j.jadohealth.2013.02.004] [Medline: 23561895]
11. Stacey D, Légaré F, Lewis K, Barry MJ, Bennett CL, Eden KB, et al. Decision aids for people facing health treatment or screening decisions. Cochrane Database Syst Rev 2017 Dec 12;4:CD001431. [doi: 10.1002/14651858.CD001431.pub5] [Medline: 28402085]
12. Brinkman WB, Lipstein EA, Taylor J, Schoettker PJ, Naylor K, Jones K, et al. Design and implementation of a decision aid for juvenile idiopathic arthritis medication choices. Pediatr Rheumatol Online J 2017 Jun 05;15(1):48 [FREE Full text] [doi: 10.1186/s12969-017-0177-x] [Medline: 28583183]
13. Majeed-Ariss R, Baildam E, Campbell M, Chieng A, Fallon D, Hall A, et al. Apps and Adolescents: A Systematic Review of Adolescents' Use of Mobile Phone and Tablet Apps That Support Personal Management of Their Chronic or Long-Term Physical Conditions. J Med Internet Res 2015 Dec 23;17(12):e287 [FREE Full text] [doi: 10.2196/jmir.5043] [Medline: 26701961]
14. Li LC, Adam PM, Backman CL, Lineker S, Jones CA, Lacaille D, et al. Proof-of-concept study of a Web-based methotrexate decision aid for patients with rheumatoid arthritis. Arthritis Care Res (Hoboken) 2014 Oct;66(10):1472-1481 [FREE Full text] [doi: 10.1002/acr.22319] [Medline: 24623617]
15. Jibb L, Cafazzo J, Nathan P, Seto E, Stevens B, Nguyen C, et al. Development of a mHealth Real-Time Pain Self-Management App for Adolescents With Cancer: An Iterative Usability Testing Study [Formula: see text]. J Pediatr Oncol Nurs 2017;34(4):283-294. [doi: 10.1177/1043454217697022] [Medline: 28376666]
16. van Pelt PA, Drossaert CHC, Kruize AA, Huisman J, Dolhain RJEM, Wulffraat NM. Use and perceived relevance of health-related Internet sites and online contact with peers among young people with juvenile idiopathic arthritis. Rheumatology (Oxford) 2015 Oct;54(10):1833-1841. [doi: 10.1093/rheumatology/kev193] [Medline: 25998449]
17. Cai RA, Beste D, Chaplin H, Varakliotis S, Suffield L, Josephs F, et al. Developing and Evaluating JIApp: Acceptability and Usability of a Smartphone App System to Improve Self-Management in Young People With Juvenile Idiopathic Arthritis. JMIR Mhealth Uhealth 2017 Aug 15;5(8):e121 [FREE Full text] [doi: 10.2196/mhealth.7229] [Medline: 28811270]
18. Naslund JA, Grande SW, Aschbrenner KA, Elwyn G. Naturally occurring peer support through social media: the experiences of individuals with severe mental illness using YouTube. PLoS One 2014;9(10):e110171 [FREE Full text] [doi: 10.1371/journal.pone.0110171] [Medline: 25333470]
19. McDonagh JE. Has the gap been bridged yet? Young people in paediatric and adult rheumatology. Rheumatology (Oxford) 2013 Aug;52(8):1349-1351. [doi: 10.1093/rheumatology/kes407] [Medline: 23353646]
20. Dures E, Hewlett S, Ambler N, Jenkins R, Clarke J, Gooberman-Hill R. Rheumatology clinicians' experiences of brief training and implementation of skills to support patient self-management. BMC Musculoskelet Disord 2014 Mar 28;15:108 [FREE Full text] [doi: 10.1186/1471-2474-15-108] [Medline: 24678645]
21. Parker SJ, Chen EK, Pillemer K, Filiberto D, Laureano E, Piper J, et al. Participatory adaptation of an evidence-based, arthritis self-management program: making changes to improve program fit. Fam Community Health 2012;35(3):236-245 [FREE Full text] [doi: 10.1097/FCH.0b013e318250bd5f] [Medline: 22617414]
22. Batalden M, Batalden P, Margolis P, Seid M, Armstrong G, Opipari-Arrigan L, et al. Coproduction of healthcare service. BMJ Qual Saf 2016 Dec;25(7):509-517 [FREE Full text] [doi: 10.1136/bmjqs-2015-004315] [Medline: 26376674]
23. Lee Y. Effect and treatment of chronic pain in inflammatory arthritis. Curr Rheumatol Rep 2013 Jan;15(1):300 [FREE Full text] [doi: 10.1007/s11926-012-0300-4] [Medline: 23292816]

24. Tzelepis F, Sanson-Fisher RW, Zucca AC, Fradgley EA. Measuring the quality of patient-centered care: why patient-reported measures are critical to reliable assessment. Patient Prefer Adherence 2015;9:831-835 [FREE Full text] [doi: 10.2147/PPA.S81975] [Medline: 26150703]
25. DeWitt EM, Kimura Y, Beukelman T, Nigrovic PA, Onel K, Prahalad S, Juvenile Idiopathic Arthritis Disease-specific Research Committee of Childhood Arthritis RheumatologyResearch Alliance. Consensus treatment plans for new-onset systemic juvenile idiopathic arthritis. Arthritis Care Res (Hoboken) 2012 Jul;64(7):1001-1010 [FREE Full text] [doi: 10.1002/acr.21625] [Medline: 22290637]
26. Savarino JR, Kaplan JL, Winter HS, Moran CJ, Israel EJ. Improving Clinical Remission Rates in Pediatric Inflammatory Bowel Disease with Previsit Planning. BMJ Qual Improv Rep 2016;5(1) [FREE Full text] [doi: 10.1136/bmjquality.u211063.w4361] [Medline: 27559471]
27. Swallow V. Special issue on children, young people, and families living with long-term conditions. J Pediatr Nurs 2015;30(1):1-3. [doi: 10.1016/j.pedn.2014.10.015] [Medline: 25448470]
28. Bugni VM, Ozaki LS, Okamoto KYK, Barbosa CMPL, Hilário MOE, Len CA, et al. Factors associated with adherence to treatment in children and adolescents with chronic rheumatic diseases. J Pediatr (Rio J) 2012 Dec 20;88(6):483-488. [doi: 10.2223/JPED.2227]
29. Coda A, Sculley D, Santos D, Girones X, Brosseau L, Smith DR, et al. Harnessing interactive technologies to improve health outcomes in juvenile idiopathic arthritis. Pediatr Rheumatol Online J 2017 May 16;15(1):40 [FREE Full text] [doi: 10.1186/s12969-017-0168-y] [Medline: 28511689]
30. Bate P, Robert G, Maher L. Bringing User Experience to Healthcare Improvement : The Concepts, Methods and Practices of Experience-Based Design. Oxford: Radcliffe Publishing; 2018.
31. Hvitfeldt H, Carli C, Nelson E, Mortenson D, Ruppert B, Lindblad S. Feed forward systems for patient participation and provider support: adoption results from the original US context to Sweden and beyond. Qual Manag Health Care 2009;18(4):247-256. [doi: 10.1097/QMH.0b013e3181bee32e] [Medline: 19851232]
32. Longacre M, Grande SW, Hager A, Montan M, Bergquist R, Martensson M, et al. Clinical Adoption of mHealth Technology to Support Pediatric Cystic Fibrosis Care in Sweden: Qualitative Case Study. JMIR Pediatr Parent 2018 Dec 10;1(2):e11080 [FREE Full text] [doi: 10.2196/11080]
33. Denham S. Dealing with feelings: How children negotiate the worlds of emotions and social relationships. Cognition, Brain, and Behavior 2007 Mar;11(1):1-48 [FREE Full text]
34. Denham S. Social-Emotional Competence as Support for School Readiness: What Is It and How Do We Assess It? Early Education & Development 2006 Jan 03;17(1):57-89. [doi: 10.1207/s15566935eed1701_4]
35. Koivula M, Huttunen K, Mustola M, Lipponen S, Laakso ML. The Emotion Detectives Game: Supporting the Social-emotional Competence of Young Children. In: Ma M, Oikonomou A, editors. Serious Games And Edutainment Applications: Volume II. New York: Springer; Mar 04, 2017:29-53.
36. Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
37. Glaser BG. The Constant Comparative Method of Qualitative Analysis. Social Problems 1965 Apr;12(4):436-445. [doi: 10.2307/798843]
38. Saunders B, Sim J, Kingstone T, Baker S, Waterfield J, Bartlam B, et al. Saturation in qualitative research: exploring its conceptualization and operationalization. Qual Quant 2018;52(4):1893-1907 [FREE Full text] [doi: 10.1007/s11135-017-0574-8] [Medline: 29937585]
39. Patton MQ. Enhancing the quality and credibility of qualitative analysis. Health Serv Res 1999 Dec;34(5 Pt 2):1189-1208 [FREE Full text] [Medline: 10591279]
40. Voruganti T, Grunfeld E, Makuwaza T, Bender JL. Web-Based Tools for Text-Based Patient-Provider Communication in Chronic Conditions: Scoping Review. J Med Internet Res 2017 Dec 27;19(10):e366 [FREE Full text] [doi: 10.2196/jmir.7987] [Medline: 29079552]
41. Epstein RM, Street RL. Shared mind: communication, decision making, and autonomy in serious illness. Ann Fam Med 2011;9(5):454-461 [FREE Full text] [doi: 10.1370/afm.1301] [Medline: 21911765]
42. Hamine S, Gerth-Guyette E, Faulx D, Green BB, Ginsburg AS. Impact of mHealth chronic disease management on treatment adherence and patient outcomes: a systematic review. J Med Internet Res 2015 Feb 24;17(2):e52 [FREE Full text] [doi: 10.2196/jmir.3951] [Medline: 25803266]
43. Montori VM, Breslin M, Maleska M, Weymiller AJ. Creating a conversation: insights from the development of a decision aid. PLoS Med 2007 Aug;4(8):e233 [FREE Full text] [doi: 10.1371/journal.pmed.0040233] [Medline: 17683195]
44. Breslin M, Mullan RJ, Montori VM. The design of a decision aid about diabetes medications for use during the consultation with patients with type 2 diabetes. Patient Educ Couns 2008 Dec;73(3):465-472. [doi: 10.1016/j.pec.2008.07.024] [Medline: 18771876]
45. Durand M, Alam S, Grande SW, Elwyn G. 'Much clearer with pictures': using community-based participatory research to design and test a Picture Option Grid for underserved patients with breast cancer. BMJ Open 2016 Feb 02;6(2):e010008 [FREE Full text] [doi: 10.1136/bmjopen-2015-010008] [Medline: 26839014]

46. Whitehead L, Seaton P. The Effectiveness of Self-Management Mobile Phone and Tablet Apps in Long-term Condition Management: A Systematic Review. J Med Internet Res 2016 May 16;18(5):e97 [FREE Full text] [doi: 10.2196/jmir.4883] [Medline: 27185295]
47. Mechael P, Nemser B, Cosmaciuc R, Cole-Lewis H, Ohemeng-Dapaah S, Dusabe S, et al. Capitalizing on the characteristics of mHealth to evaluate its impact. J Health Commun 2012;17 Suppl 1:62-66. [doi: 10.1080/10810730.2012.679847] [Medline: 22548600]
48. Knight E, Stuckey MI, Petrella RJ. Health promotion through primary care: enhancing self-management with activity prescription and mHealth. Phys Sportsmed 2014 Sep;42(3):90-99. [doi: 10.3810/psm.2014.09.2080] [Medline: 25295771]
49. Lentferink AJ, Oldenhuis HK, de Groot M, Polstra L, Velthuijsen H, van Gemert-Pijnen JEWC. Key Components in eHealth Interventions Combining Self-Tracking and Persuasive eCoaching to Promote a Healthier Lifestyle: A Scoping Review. J Med Internet Res 2017 Dec 01;19(8):e277 [FREE Full text] [doi: 10.2196/jmir.7288] [Medline: 28765103]
50. Zhao J, Freeman B, Li M. Can Mobile Phone Apps Influence People's Health Behavior Change? An Evidence Review. J Med Internet Res 2016 Dec 31;18(11):e287 [FREE Full text] [doi: 10.2196/jmir.5692] [Medline: 27806926]
51. Anglada-Martinez H, Riu-Viladoms G, Martin-Conde M, Rovira-Illamola M, Sotoca-Momblona JM, Codina-Jane C. Does mHealth increase adherence to medication? Results of a systematic review. Int J Clin Pract 2015 Jan;69(1):9-32. [doi: 10.1111/ijcp.12582] [Medline: 25472682]
52. Stinson J, McGrath P, Hodnett E, Feldman B, Duffy C, Huber A, et al. Usability testing of an online self-management program for adolescents with juvenile idiopathic arthritis. J Med Internet Res 2010 Jul 29;12(3):e30 [FREE Full text] [doi: 10.2196/jmir.1349] [Medline: 20675293]
53. Nilsen W, Kumar S, Shar A, Varoquiers C, Wiley T, Riley WT, et al. Advancing the science of mHealth. J Health Commun 2012;17 Suppl 1:5-10. [doi: 10.1080/10810730.2012.677394] [Medline: 22548593]

Abbreviations

ERA: enthesitis-related arthritis
ILAR: International League against Rheumatism
JAS: juvenile ankylosing spondylitis
JIA: juvenile idiopathic arthritis
mHealth: mobile health
mPSS: mHealth Patient Support System
Oligo JIA: oligoarticular JIA
Poly JIA: polyarticular JIA
PsA: psoriatic JIA

Edited by G Eysenbach; submitted 13.03.18; peer-reviewed by R Cai, A Coda, T Jiang; comments to author 18.08.18; revised version received 02.11.18; accepted 10.11.18; published 22.02.19.

Please cite as:
Grande SW, Longacre MR, Palmblad K, Montan MV, Berquist RP, Hager A, Kotzbauer G
Empowering Young People Living With Juvenile Idiopathic Arthritis to Better Communicate With Families and Care Teams: Content Analysis of Semistructured Interviews
JMIR Mhealth Uhealth 2019;7(2):e10401
URL: http://mhealth.jmir.org/2019/2/e10401/
doi:10.2196/10401
PMID:30794202

Original Paper

Use, Perspectives, and Attitudes Regarding Diabetes Management Mobile Apps Among Diabetes Patients and Diabetologists in China: National Web-Based Survey

Yiyu Zhang[1,2,3*], MMed; Xia Li[1,2,3*], MD; Shuoming Luo[1,2,3], MD; Chaoyuan Liu[4], MMed; Yuting Xie[1,2,3], MD; Jia Guo[5], PhD, RN; Fang Liu[1], MNurs, RN; Zhiguang Zhou[1,2,3], MD

[1]Department of Metabolism and Endocrinology, The Second Xiangya Hospital, Central South University, Changsha, China

[2]Key Laboratory of Diabetes Immunology, Ministry of Education, Changsha, China

[3]National Clinical Research Center for Metabolic Diseases, Changsha, China

[4]Department of Oncology, The Second Xiangya Hospital, Central South University, Changsha, China

[5]Xiangya School of Nursing, Central South University, Changsha, China

[*]these authors contributed equally

Corresponding Author:
Zhiguang Zhou, MD
Department of Metabolism and Endocrinology
The Second Xiangya Hospital
Central South University
No 139, Renmin Road
Changsha, 410011
China
Phone: 86 073185292154
Fax: 86 073185367220
Email: zhouzhiguang@csu.edu.cn

Abstract

Background: The diabetes disease burden in China is heavy, and mobile apps have a great potential for diabetes management. However, there is a lack of investigation of diabetes app use among Chinese diabetes patients and diabetologists. The perspectives and attitudes of diabetes patients and diabetologists regarding diabetes apps are also unclear.

Objective: Our objectives were to investigate diabetes patients' and diabetologists' use, attitudes, and perspectives, as well as patients' needs, with respect to diabetes apps to provide information regarding the optimal design of diabetes apps and the best strategies to promote their use.

Methods: Diabetes patients and diabetologists across China were surveyed on the WeChat (Tencent Corp) network using Sojump (Changsha ran Xing InfoTech Ltd) from January 23, 2018, to July 30, 2018. In total, 2 survey links were initially sent to doctors from 46 Latent Autoimmune Diabetes of Adults Study collaborative hospitals in China in 25 major cities and were spread on their WeChat contacts network. We also published the patient survey link on 3 WeChat public accounts and requested diabetes patients to fill out questionnaires. A multivariate regression analysis was used to identify associations of demographic and basic disease information with app usage among adult patients.

Results: Overall, 1276 individuals from 30 provincial regions responded to the patient survey; among them, the overall app awareness rate was 29.94% (382/1276) and usage was 15.44% (197/1276). The usage was higher among patients with type 1 diabetes (T1DM) than among patients with type 2 diabetes (T2DM; 108/473, 22.8% vs 79/733, 10.8%; $P<.001$). The multivariate regression analysis showed that diabetes type, age, education, family income, and location were associated with app use in adult patients ($P<.05$). The need for and selection of diabetes apps differed slightly between patients with T1DM and patients with T2DM. The reasons why patients discontinued the use of an app included limited time (59/197, 29.9%), complicated operations (50/197, 25.4%), ineffectiveness for glycemic control (48/197, 24.4%), and cost (38/197, 19.3%). Of the 608 responders to the diabetologist survey, 40.5% (246/608) recommended diabetes apps to patients and 25.2% (153/608) used diabetes apps to manage patients. The greatest obstacles to the diabetologists' use of apps to manage diabetes patients include limited time (280/608, 46.1%), legal issues (129/608, 21.2%), patients' distrust (108/608, 17.8%), and billing issues (66/608, 10.9%).

Conclusions: The awareness and use of diabetes apps in Chinese people with diabetes and the proportion of diabetologists using diabetes apps to manage patients are low. Designing apps targeting different patient needs and conducting high-quality randomized controlled trials will improve the effectiveness of the apps, provide evidence for patients to choose suitable apps, and be conducive to the promotion of app use.

(JMIR Mhealth Uhealth 2019;7(2):e12658) doi:10.2196/12658

KEYWORDS

diabetes mellitus; mobile applications; surveys and questionnaires

Introduction

Background

The prevalence of diabetes has been increasing worldwide. In 2017, it was estimated that there were 451 million people aged 18 to 99 years with diabetes. Furthermore, 5.0 million deaths were attributable to diabetes [1]. Among the Chinese adult population, the estimated prevalence of diabetes was 11.6%, representing an estimated 113.9 million adults in China with diabetes. However, only 39.7% of those treated had adequate glycemic control [2]. Poor glycemic control can cause various complications [3] and bring heavy economic burden to the world. In 2015, the global cost of diabetes was estimated to be US $1.31 trillion or 1.8% of the global gross domestic product (GDP) [4].

Diabetes self-management education and support (DSMES) is a critical element of care for patients with diabetes [5]. The AADE7 Self-Care Behaviors defined by the American Association of Diabetes Educators is a framework for patient-centered diabetes education and care [6]. The 7 self-care behaviors are eating healthy, being active, monitoring, taking medications, solving problems, healthy coping, and reducing risks, and these are essential for diabetes self-management. Due to an imbalance of medical resources in China [7], patients flock from rural areas to urban areas seeking medical resources. However, doctors in tertiary hospitals are overloaded with work [8,9]. Patients receive only a few minutes for outpatient consultation and receive little self-management knowledge in such a limited time. Furthermore, many outpatients do not have a record of their blood sugar; therefore, doctors cannot give accurate guidance for their treatment. Thus, a different type of health service might be needed to supplement traditional outpatient consultations.

Mobile apps can record, transmit, and receive feedback anytime and anywhere, facilitating remote monitoring and delivery of timely recommendations for health care. Mobile apps could increase the capacity for self-management, help sustain necessary lifestyle modifications, and improve communication between patients, family members, and health care professionals (HCPs). Furthermore, there were 1.32 billion mobile phone users in China in 2016 [10]. With the popularity of smart phones in China, mobile apps have great potential for managing chronic diseases, especially diabetes. A systematic investigation revealed that diabetes apps are the most common disease-specific apps in China's mobile health (mHealth) market [11]. Many studies have demonstrated that diabetes apps can potentially help control glycemia [12-16]. Mobile apps are also recommended by the American Diabetes Association guidelines for DSMES [17]. A meta-analysis of 21 randomized controlled trials (RCTs) showed that diabetes apps were associated with a mean reduction of 0.57% in glycosylated hemoglobin (HbA_{1c}) among patients with type 2 diabetes mellitus (T2DM) and 0.49% among patients with type 1 diabetes mellitus (T1DM). However, the results had significant heterogeneity [18]. Several studies also suggested that diabetes app use can increase blood glucose monitoring frequency [13,19], reduce feelings of loneliness, help patients gain knowledge and skills to manage diabetes [20], improve hypoglycemic fears and behavioral scores [15], and strengthen the perception of self-care by offering better information and health education to patients [16].

Diabetes app usage varies among different countries [21-25], from 3% in Latin countries in 2015 [21] to 19.6% in New Zealand in 2016 [22]. China has the largest absolute disease burden of diabetes in the world and the greatest potential diabetes app market. However, there is a lack of large-scale investigations of the usage of diabetes apps in China. Patients' perspectives and attitudes regarding diabetes apps are also not very clear. In addition, user requirements are very important for app design. A survey by Boyle et al revealed that the most favored feature of patients was a glucose diary, and an insulin calculator was the most desirable function for a future app [22]. A survey by Trawley et al showed carbohydrate counting was the most common purpose among adults with T1DM and glucose monitoring was the most common purpose among adults with T2DM [23]. A recent meta-analysis of diabetes apps revealed that the reduction in HbA_{1c} is explained by the frequency of HCP feedback [18]. However, few surveys have been conducted to investigate the feature of patient-doctor communication. Our previous study found that both patients and diabetologists believed that doctor-patient communication and diabetes diaries were the most important functions of a diabetes app [26]. However, our previous study focused only on patients with T1DM and the samples were relatively small. Diabetologists' use, attitudes, and perspectives concerning diabetes management apps in China are poorly understood.

Objectives

The objective of our study was to investigate the use, perspectives, attitudes, and associated factors of diabetes patients and diabetologists regarding diabetes management apps, as well as patient needs for these apps, to provide information for the design of diabetes apps and how to best promote their use.

Methods

Questionnaire Design

The questionnaire design was described in a previous study [26]. An expert panel consisting of 5 diabetologists (YZ, SL, YX, XL, and ZZ) and a diabetes education nurse (FL) searched apps from the iOS and Android platforms and designed the questionnaires according to the functions of current diabetes apps [25,27-31], diabetes guidelines [32], and the problems they encountered during clinical practice. The questions were presented in a choice format. If responders did not agree with the listed options, they could select the option *others* and write their answers in the remarks column. The diabetologist questions covered their basic information and use, attitudes, and perspectives regarding diabetes management apps. The patient questions covered their use, perspectives, and attitudes regarding and needs for diabetes apps, demographic information, and basic disease information.

To establish the validity of the content, the survey items were rated based on their relevance and clarity using a 4-point ordinal scale from 1 (irrelevant) to 4 (highly relevant) by 15 experts: 12 diabetologists and 3 diabetes education nurses with at least 5 years of experience treating patients with diabetes. The content validity indexes of the patient and diabetologist questionnaires were 0.91 and 0.93, respectively. Before administering the questionnaires, we performed pilot tests with 20 patients with diabetes and 10 diabetologists from Second Xiangya Hospital. As 3 patients were unwilling to disclose their exact income and 1 doctor was reluctant to reveal his age, we revised our questionnaires such that patients were not required to answer the question regarding income and the doctors' ages were grouped into ranges. Cronbach alpha values for the patient and diabetologist questionnaires were .98 and .79, respectively.

Introduction of the WeChat Survey Platform

WeChat provides many services, including messaging, free phone calls, browsing and posting for information sharing on moments, and mobile payments [33]. It is installed in over 90% of mobile phones and is integrated into most people's daily lives [34]. As the most widely and frequently used social communication tool in China, WeChat has a powerful contact network. As of 2016, over 44% of WeChat users had more than 200 contacts on the social networking service. Approximately 90% of WeChat users had more than 50 contacts [35]. This network makes it possible to administer questionnaires via WeChat.

Samples and Survey Methods

The participants were diabetes patients and diabetologists across China. Doctors at other departments who did not treat diabetes patients were excluded from our investigation.

Diabetes patients and diabetologists were surveyed through snowball sampling via the WeChat contacts network and convenience sampling through WeChat public accounts using the Web-based survey tool, Sojump, [36] from January 23, 2018, to July 30, 2018. The patient and diabetologist survey links were initially sent to doctors from 46 latent autoimmune diabetes of adults (LADA) Study China collaborative hospitals in 25 representative major cities in China [37]. We asked these doctors to spread the survey links on their WeChat contacts network. In addition to using snowball sampling through the WeChat contacts network, we published the patient survey link on 3 WeChat public accounts concerning diabetes, which have 50,000 subscribed followers, and asked diabetes patients to complete the questionnaires. The parents of juvenile patients answered the questions for their children. We introduced the background of our survey, and the questionnaires were completed voluntarily without any compensation. A part of the survey results concerning patients with T1DM has been reported in a previous study [26].

Ethical Approval

The study was approved by the ethics committee of the Second Xiangya Hospital, Central South University (ID: 2017-S107).

Statistics

The data were analyzed using SPSS version 23.0 (IBM Corp). Q-Q plots were used to check the normality of all the continuous variables, which are expressed as the means (SDs) or medians (interquartile ranges [IQRs]) where appropriate. Categorical variables are expressed as percentages (numbers, n). Differences among groups were assessed using Chi-square tests. The generalized logistic model was used to obtain odds ratios (OR) and their 95% CIs in a simultaneous manner. First, we performed a univariable analysis to obtain unadjusted ORs of potential correlates of app use with demographic factors and disease characteristics in adult patients (aged ≥18 years). We then entered all the significant factors in the multivariate analysis to obtain the multivariable adjusted ORs. Questionnaires with missing values were excluded from the multivariate analysis. Statistical significance was indicated with $P<.05$.

Results

Patient Survey

Sample Characteristics

A total of 1276 patients with diabetes (414 from North China and 862 from South China) distributed among 30 provinces in China (Figure 1) responded to the patient survey. The responder characteristics are shown in Table 1. Of the responders, 50.31% (642/1276) were male, with a mean age of 41.3 years (SD 18.5). The mean disease duration was 6.8 years (SD 6.9). Overall, 37.07% (473/1276) were patients with T1DM and, of these, 178 were juveniles; 57.45% (733/1276) were patients with T2DM; 2.12% (27/1276) of the patients had gestational diabetes; and 3.45% (43/1276) did not know their diabetes type.

Figure 1. Distribution of the diabetic patient sample in China by province.

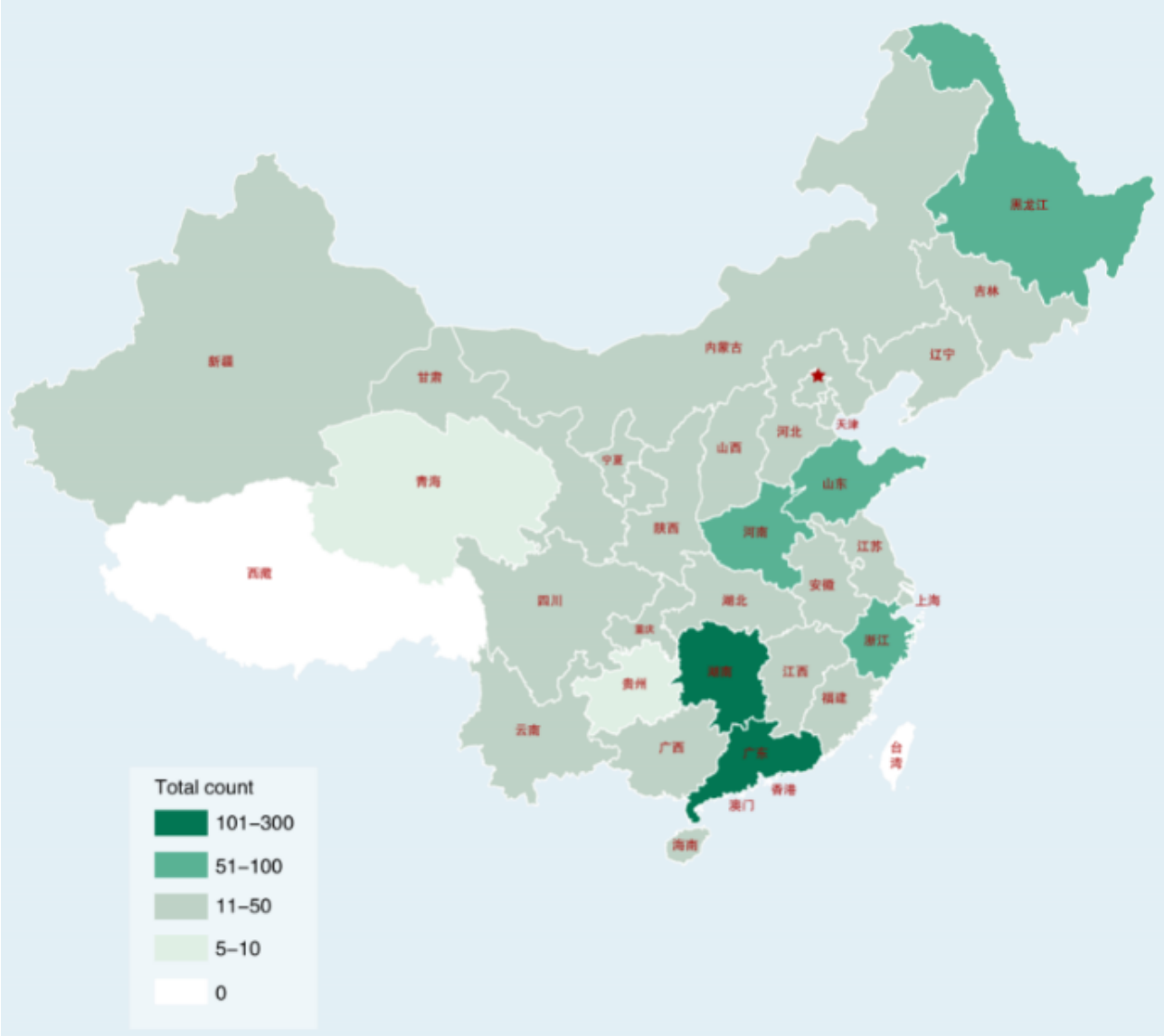


Table 1. Characteristics of patients with diabetes.

Characteristics	T1DM[a] (n=473)		T2DM[b] (n=733)	Total[c] (N=1276)
	Adults (n=295)	Juveniles (n=178)		
Gender, n (%)				
Male	115 (39.0)	81 (45.5)	426 (58.1)	642 (50.31)
Age (years), mean (SD)	33.5 (11.9)	10.3 (4.2)	52.2 (12.0)	41.3 (18.5)
Disease duration (years), mean (SD)	7.9 (8.0)	3.0 (3.2)	7.7(6.8)	6.8(6.9)
Residence, n (%)				
Urban	206 (69.8)	110 (61.8)	572 (78.0)	933 (73.12)
Rural	89 (30.2)	68 (38.2)	161 (22.0)	343 (26.88)
Education, n (%)				
Junior middle school or below	55 (18.6)	—[d]	193 (26.3)	422 (33.07)
High school	93 (31.5)	—	244 (33.3)	373 (29.23)
University or above	147 (49.8)	—	296 (40.4)	481 (37.70)
Treatment, n (%)				
Oral medicine	14 (4.7)	1 (0.6)	379 (51.7)	416 (32.60)
Insulin injection	196 (66.4)	128 (71.9)	281 (38.3)	636 (49.84)
Insulin pump	83 (28.1)	49 (27.5)	9 (1.2)	144 (11.29)
Untreated	2 (0.7)	0	64 (8.7)	80 (6.27)
Occupation, n (%)				
Student	44 (14.9)	—	3 (0.4)	206 (16.14)
Institutional staff	45 (15.3)	—	115 (15.7)	171 (13.40)
Employee of state-owned enterprise	19 (6.4)	—	83 (11.3)	109 (8.54)
Employee of foreign or private company	43 (14.6)	—	69 (9.4)	121 (9.48)
Private enterprise owner or self-employed	26 (8.8)	—	77 (10.5)	113 (8.86)
Retired	19 (6.4)	—	206 (28.1)	231 (18.10)
Farmer	19 (6.4)	—	75 (10.2)	103 (8.07)
Unemployed	47 (15.9)	—	70 (9.5)	131 (10.27)
Others	33 (11.2)	—	35 (4.8)	9 (7.05)

[a]T1DM: type 1 diabetes mellitus.

[b]T2DM: type 2 diabetes mellitus.

[c]Total including patients with T1DM, patients with T2DM, 27 patients with gestational diabetes, and 43 patients with an unknown type of diabetes.

[d]Indicates that there is no value.

Diabetes App Use and Associated Factors Among Diabetes Patients

The overall diabetes app awareness rate was 29.94% (382/1276), the usage rate was 15.44% (197/1276), and 43.7% (86/197) of the patients who use the apps used them every day. The app usage of patients with T1DM was higher than that of patients with T2DM (108/473, 22.8% vs 79/733, 10.8%; *P*<.001). The utilization rate of adult patients with T1DM was higher than that of juvenile patients with T1DM (80/295, 27.1% vs 28/178, 15.7%; *P*=.004). A comparison of patients located in the 10 provinces of top GDP per capita [38] with the other 20 provinces showed that the former had higher diabetes app usage than the latter (79/401, 19.7% vs 118/875, 13.5%; *P*=.004).

The app usage of juvenile patients with T1DM treated with an insulin pump was higher than that of those treated with a subcutaneous injection (13/49, 26.5% vs 15/128, 11.7%; *P*=.009). The comparison of the 3 groups of juvenile patients with different parental educational levels (junior middle school and below, senior high school, and university and above) revealed that a higher education level of parents was associated with a higher app usage (4/64, 6.3% vs 6/41, 14.6% vs 17/65, 26.2%; *P*=.008). App usage in children was higher than that in adolescents, but the difference was not statistically significant (21/106, 19.8% vs 7/72, 9.7%; *P*=.07).

A univariate regression analysis showed that app use by adult patients was significantly associated with age, annual family income, occupation, education, locus, type of diabetes, and

treatment (P<.05). Gender, disease duration, and rural or urban residence had no correlation with app use. Variables with statistical significance were included in the multivariate regression analysis. Diabetes type, age, education, annual family income, and location remained statistically significant (P<.05). Compared with low-income patients, patients with high income had higher app usage (OR=1.73, 95% CI 1.07-2.81; P=.03). The use of apps decreased with patient age. Notably, the higher the education level of the patients, the higher was the usage of apps (see Table 2).

Patients' Need for and Selection of Diabetes Apps

Patients with both T1DM and T2DM believed that the most important functions of a diabetes app were diabetes diaries (blood sugar, diet, exercise, and medication records) and doctor-patient communication (Table 3). Patients with T2DM had a greater demand for doctor-patient communication (299/733, 40.8% vs 151/473, 31.9%; P=.002), and patients with T1DM had a greater demand for an insulin dose calculator (51/473, 10.8% vs 11/733, 1.5%; P<.001). Almost all patients believed that the listed functions were important or very important (Figure 2).

Table 2. Factors associated with app use by multivariate logistic regression analysis (N=1008).

Characteristics	App usage rate, n (%)	Adjusted odds ratio (95% CI)	*P* value
Age (years)			
18-39[a]	97 (27.2)	—[b]	—
40-59	47 (10.7)	0.42 (0.27-0.65)	<.001
≥60	19 (9.0)	0.40 (0.22-0.72)	.002
Annual family income[c]			
¥<50,000[a]	34 (10.9)	—	—
¥50,000-100,000	50 (16.1)	1.38 (0.84-2.28)	.20
¥>100,000	79 (20.5)	1.73 (1.07-2.81)	.03
Diabetes type			
T2DM[a,d]	78 (11.2)	—	—
T1DM[e]	76 (29.7)	2.0 (1.30-3.05)	.001
Gestational diabetes	3 (11.1)	0.47 (0.13-1.67)	.25
Unknown type	6 (16.2)	0.15 (0.45-2.93)	.77
Location			
The other 20 provinces[a]	96 (14.1)	—	—
10 provinces of the top GDP[f] per capita	67 (20.5)	1.5 (1.04-2.17)	.03
Education			
Junior middle school or below[a]	18 (7.6)	—	—
High school	49 (15.0)	1.85 (1.02-3.37)	.04
University or above	96 (21.7)	2.40 (1.34-4.25)	.003

[a]Reference group.

[b]Not applicable.

[c]90 samples with missing data on family income were excluded from the logistic regression analysis.

[d]T2DM: type 2 diabetes mellitus.

[e]T1DM: type 1 diabetes mellitus.

[f]GDP: gross domestic product.

Table 3. App functions considered to be most important by patients with both T1DM and T2DM.

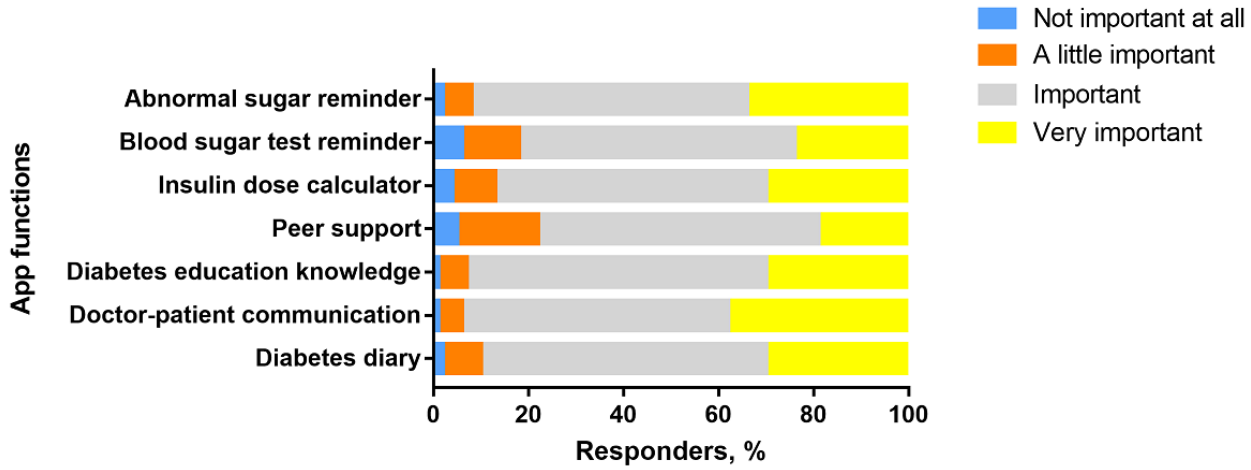
Features	T1DM[a] (N=473), n (%)	T2DM[b] (N=733), n (%)	*P* value [c]
Diabetes diaries	109 (23.0)	192 (26.2)	.22
Doctor-patient communication	151 (31.9)	299 (40.8)	.002
Diabetes education knowledge	54 (11.2)	86 (11.7)	.87
Peer support	13 (2.7)	10 (1.4)	.09
Insulin dose calculator	51 (10.8)	11 (1.5)	<.001
Abnormal blood sugar reminder	60 (12.7)	85 (11.6)	.57
Blood sugar test reminder	8 (1.7)	15 (2.0)	.66
Others	27 (5.7)	35 (4.8)	.47

[a]T1DM: type 1 diabetes mellitus.

[b]T2DM: type 2 diabetes mellitus.

[c]A Chi-square test was used to calculate *P* values.

Figure 2. Importance of different app functions reported by patients with diabetes (N=1276).



Differences in app selection were found between patients with T1DM and patients with T2DM (see Table 4). The most popular app among patients with T1DM was Diabetes Circle (30/108, 27.8%), which is targeted for patients with T1DM, followed by Diabetes Nurse (28/108, 25.9%) to which blood sugar tested by a Dnurse glucometer (Sinocare Inc) can be directly transmitted. The most popular app among patients with T2DM was Diabetes Nurse (34/79, 43.0%).

Only 19.3% (38/197) of the apps were recommended by HCPs. Most patients selected diabetes apps as recommended by other patients (55/197, 27.9%) or selected randomly because they did not know which one was best (54/197, 27.4%). The rest were recommended by the media (30/197, 15.2%) or other channels (20/197, 10.2%).

Table 4. Differences in the selection of diabetes apps between patients with type 1 diabetes mellitus (T1DM) and patients with type 2 diabetes mellitus (T2DM).

App name	T1DM (N=108), n (%)	T2DM (N=79), n (%)	*P* value[a]
Welltang (Shanghai Geping Information Technology Co, Ltd)	24 (22.2)	5 (6.3)	.003
Diabetes Circle (Aibaowei Biotechnology Co, Ltd)	30 (27.8)	3 (3.8)	<.001
Control Diabetes (Fuzhou Kangwei Network Technology Co, Ltd)	1 (0.9)	5 (6.3)	.08
Diabetes Doctor (Shanghai Huima Medical Technology Co, Ltd)	6 (5.6)	14 (17.7)	.008
Diabetes Nurse (Beijing Dnurse Technology Co,Ltd)	28 (25.9)	34 (43.0)	.01
Others	19 (17.6)	18 (22.8)	.38

[a]A Chi-square test was used to calculate the *P* values.

Patients' Perspectives on Diabetes Apps

Among the patients who had used diabetes apps, the reasons for discontinuation of use included limited time (59/197, 29.9%), complicated operations (50/197, 25.4%), ineffectiveness for glycemic control (48/197, 24.4%), cost (38/197, 19.3%), and others (48/197, 24.4%). Most patients thought that the diabetes app was very helpful (89/197, 45.2%) or helpful (67/197, 34.0%) to them. Overall, 14.2% (28/197) of the patients thought it was a little helpful. Only 6.6% (13/197) of them thought it was unhelpful.

Of the patients, 58.4% (115/197) indicated that their apps had a function for consulting HCPs and, of these patients, 50.4% (58/115) had used the function to consult an HCP. The proportion of patients with T2DM who consulted HCPs was higher than that of patients with T1DM, but this difference was not statistically significant (26/46, 57% vs 28/61, 46%; *P*=.28). A total of 67% (39/58) of their consultations were free. Most patients thought that these consultations were helpful (33/58, 57%) or very helpful (4/58, 7%). The reasons cited by the patients as affecting the effectiveness of the consultation included short consultation time (25/58, 43%), delayed response (15/58, 26%), unqualified HCPs (10/58, 17%), and others (8/58, 14%).

Patients' Attitudes Toward Diabetes Apps

Only 34.87% (445/1276) of patients believed that consulting HCPs via apps should cost money. However, 59.72% (762/1276) of these indicated that they would continue to use this function if the consultation effect was good. Of these, 8.30% (106/1276) said they would certainly continue to use and 31.97% (408/1276) said they would not continue to use the app if they had to pay. Almost all patients said they were in need (889/1276, 69.67%) or in great need (335/1276, 26.25%) of a good app to help with their glycemic control. Only 4.08% (52/1276) of them said they did not need a diabetes app.

Diabetologist Survey

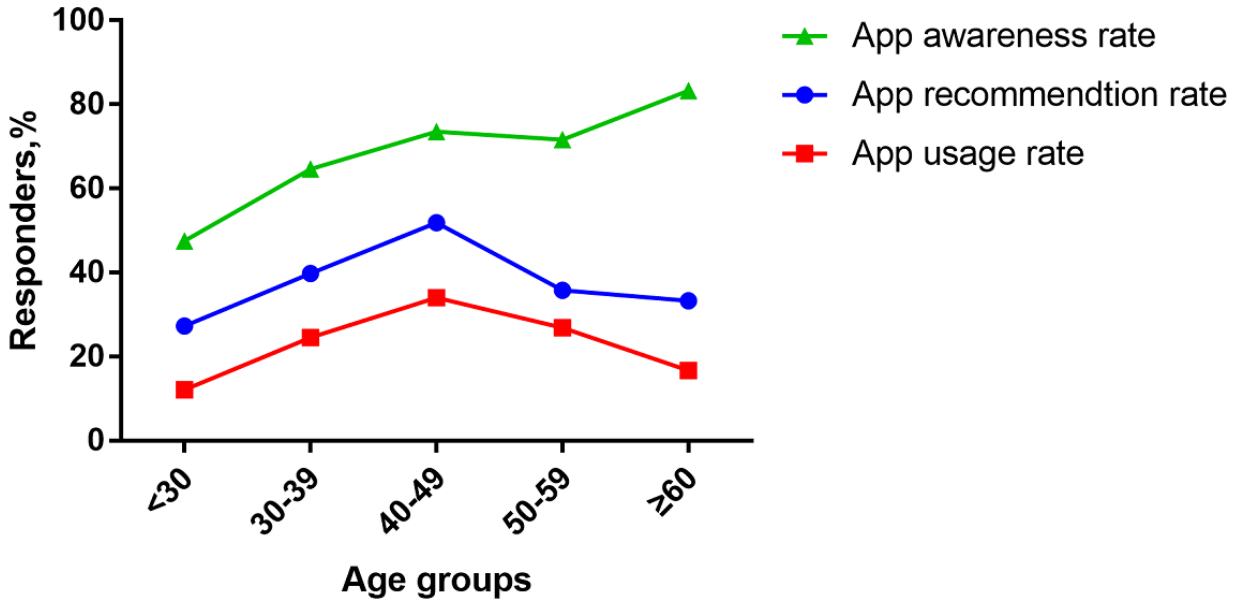
Diabetologists' Recommendation and Use of Diabetes Management Apps and Associated Factors

In total, 608 diabetologists (223 from North China and 385 from South China) from 21 provinces in China responded to the diabetologist survey. Table 5 shows the characteristics of the surveyed diabetologists.

Table 5. Characteristics of the surveyed diabetologists (N=608).

Characteristics	Statistics
Gender, n (%)	
Male	197 (32.4)
Female	411 (67.6)
Age (years), n (%)	
≤30	99 (16.3)
30-39	274 (45.1)
40-49	162 (26.6)
50-59	67 (11.0)
≥60	6 (1.0)
Title, n (%)	
Resident	141 (23.2)
Attending specialist	239 (39.3)
Associate chief doctor	125 (20.6)
Chief doctor	103 (16.9)
Hospital level, n (%)	
Tertiary hospital	419 (68.9)
Secondary hospital or lower	189 (31.0)

Figure 3. App awareness rate, recommendation rate, and usage among different age groups of diabetologists (N=608).



Of the diabetologists, 43.8% (266/608) had downloaded a diabetes management app, 40.5% (246/608) of them had recommended diabetes apps to their patients, and 25.2% (153/608) used diabetes apps to manage their patients. The app awareness rate was lowest in the doctor group younger than 30 years (47/99, 47%) and it gradually increased with age (P<.001). App recommendation rate and usage increased gradually with the age of doctors. The highest recommendation rate and usage were found in the 40 to 49 age group (84/162, 51.9% and 55/162, 34.0%, respectively), and then decreased again with age (P=.002 and P=.003, respectively; Figure 3). The app recommendation rate and usage among doctors in tertiary hospitals were higher than those in secondary or lower hospitals (190/419, 45.3% vs 56/189, 29.6%; P<.001 and 119/419, 28.4% vs 34/189, 18%; P<.001).

The app most recommended by diabetologists was Diabetes Doctor (56/246, 22.8%), which enables doctors to follow their patients, followed by Welltang (45/246, 18.3%), which is the only app tested via an RCT in China. The most important factors that influenced diabetologists' recommendations of apps to their patients included not knowing of a suitable app (296/608, 48.7%), not knowing of the existence of diabetes apps (212/608, 34.9%), no time to recommend (182/608, 29.9%), no evidence demonstrating their effectiveness (90/608, 14.8%), no effect on blood sugar (47/608, 7.7%), and others (49/608, 8.1%).

The greatest obstacles to diabetologists' use of apps to manage patients with diabetes include limited time (280/608, 46.1%), legal issues (129/608, 21.2%), patients' distrust (108/608, 17.8%), uncertainty on how to bill patients (66/608, 10.9%), and others (25/608, 4.1%). The proportion of diabetologists in tertiary hospitals who thought the largest obstacle was limited time was higher than that in secondary or lower hospitals (212/419, 50.6% vs 68/189, 36.0%; P=.001). The proportion of doctors who believed that the greatest obstacle was patient distrust was higher in doctors from secondary or lower hospitals and junior doctors than in doctors from tertiary hospitals and among senior doctors (47/189, 24.9% vs 61/419, 14.6%; P=.002; and 38/141, 27.0% vs 70/467, 15.0%; P=.001, respectively).

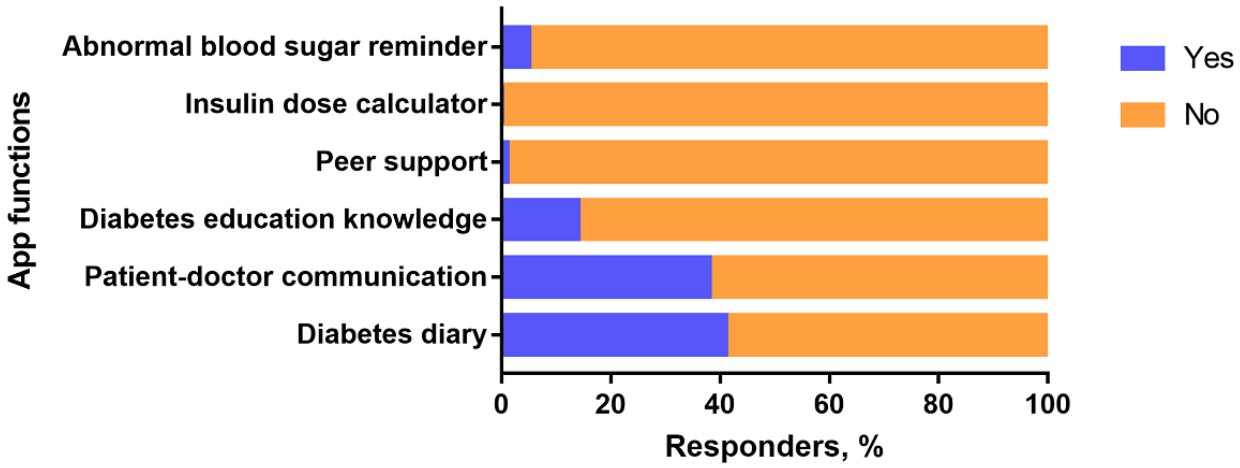
In all, 94% (141/150) of the diabetologists use apps to manage patients free of charge. Most of them managed less than 50 patients (125/150, 83.3%). Most diabetologists who had managed patients with an app thought that the app had some effect (83/150, 55.3%) or minimal effect (58/150, 38.7%) on blood sugar, whereas 2.0% (3/150) thought it had no effect and 4% (6/150) thought it was very effective. Most diabetologists did not know whether it was legal to use apps to manage patients (311/608, 51.2%), and 33.6% (204/608) and 15.3% (93/608) thought that using an app for this purpose was legal and illegal, respectively.

Diabetologists' Perspectives of Diabetes Apps

The diabetologists believed that the most important functions of a diabetes app were diabetes diaries (247/608, 40.6%) and doctor-patient communication (233/608, 38.4%; Figure 4). Of the diabetologists, 71.5% (435/608) believed that patients with T1DM and patients with T2DM needed different apps.

Diabetologists believed that the reasons for the poor effect of a diabetes app on blood sugar included the following: patients could not adhere to the use of apps (431/608, 70.9%); HCPs did not participate in apps or gave too little guidance for patients (383/608, 63%), diabetes education knowledge on an app was not systematic (280/608, 46.1%), apps lacked comprehensive functions (225/608, 37.0%), and others (16/608, 2.6%).

Figure 4. Diabetes app functions considered to be most important by diabetologists (N=608).



Diabetologists' Attitudes Toward Diabetes Apps

Most diabetologists said they might (380/608, 62.5%) or would certainly (219/608, 36%) use a diabetes app to manage patients in the future. Only 1.5% (9/608) indicated that they would not manage diabetes patients using an app. Most said that they might (349/608, 57.4%) or would definitely (253/608, 41.6%) recommend diabetes apps to their patients, whereas only 1.0% (6/608) said that they would not recommend one. Most diabetologists believed that diabetes apps showed good (325/608, 53.5%) or very good (127/608, 20.9%) potential.

Discussion

Principal Findings

Diabetes App Use among Diabetes Patients and Its Associated Factors

App usage was 22.8% among patients with T1DM and 10.8% among patients with T2DM, which was comparable with surveys conducted in New Zealand [22] and Australia [23], and higher than the 7% found in a Scottish survey in 2016 [25] and the 3% found among Latinos in 2015 [21]. The app usage of patients with T2DM in China was lower than that of patients with T1DM and was associated with age, education, family income, and location. The possible reasons are that younger patients and well-educated patients are more likely to acquire and accept new technology, and those with higher household income are more likely to focus on their glycemic control and actively seek new ways to control blood sugar. China's economic development is unbalanced, and medical resources are unevenly distributed and relatively concentrated in economically developed areas [7]. There are obvious regional differences in glycemic control in China [39]. The regional difference in app usage might be related to these differences in glycemic control. Furthermore, the app usage among patients with T2DM was lower than that among patients with T1DM, which is consistent with former studies [22,23]. This finding may be because the blood sugar of patients with T1DM is more difficult to control and their need for an app is greater.

We found that the app usage in children was higher than that in adolescents, possibly because children's blood sugar is always managed by their parents, as their parents use diabetes apps to help with glycemic control. However, adolescents gradually withdraw support from their parents and take over the management tasks.

Suggestions for Promoting Diabetes App Use

App use among patients with diabetes in China is low, largely because of the low awareness of diabetes apps. Only 29.94% of the patients knew that diabetes apps existed, but only half of these patients who knew of diabetes apps would use one. Additionally, many doctors, particularly younger doctors, had no information on diabetes apps. Specifically, younger doctors had less awareness of diabetes apps than did senior doctors, which might be related to younger doctors having less awareness of the progress made in diabetes treatment. Therefore, public awareness of diabetes apps must be increased.

Our study found that only a small number of patient apps were recommended by doctors. Most apps were recommended by patients or were casually chosen. In total, 40.5% of the diabetologists in China recommended diabetes apps to their patients, which was lower than the 60.1% found in the New Zealand survey [22] and the 62% found in the US survey [40]. The most important factor that influenced diabetologists' recommendation of apps to patients was that they had no idea of a suitable one among the numerous apps. Although there are thousands of diabetes apps, only a small number of them were tested for efficacy [31]. The quality of the studies was not high, and the effects of the apps on blood sugar were inconsistent [41]. Thus, it is difficult for HCPs to recommend a suitable app to their patients. Therefore, it is very important to carry out high-quality RCTs to test app efficacy [42].

Barriers to Doctor-Patient Communication and Suggestions for Improvement

Both diabetologists and patients believed that doctor-patient communication and diabetes diaries were the most important functions of a diabetes app, which was consistent with our previous report [26]. The app most recommended by

diabetologists was one that doctors could use to follow-up with their patients. However, only 25.2% of the diabetologists managed diabetes patients with apps.

The main reasons given by diabetologists as affecting patient management with apps were limited time, issue of legality, and patients' distrust. More doctors in tertiary hospitals than in primary hospitals thought that the largest obstacle to using an app to manage patients with diabetes was limited time, and more younger doctors and doctors from primary hospitals believed that patients' distrust was the largest obstacle. Therefore, improving the specialty of young doctors and doctors from primary hospitals can effectively improve patients' trust in these doctors, reduce the burden on senior doctors from tertiary hospitals, and effectively promote using diabetes apps to manage patients. Fortunately, the Chinese government is actively promoting standardized training for residents and a hierarchical medical system, which will effectively reduce the burden on doctors from tertiary hospitals, improve the specialties of young doctors and doctors from primary hospitals, and enhance patients' trust in them. Owing to a lack of face-to-face physical examinations and complete medical histories, most doctors did not know whether using an app to manage patients was legal. China is vigorously promoting internet health care and improving relevant legislations [43]. Therefore, it may become possible for doctors to manage their patients with an app.

The approach used to bill patients was also a reason given by diabetologists that affected patient management. At present, most diabetologists use apps to manage patients free of charge, which affects the HCP's enthusiasm. Medical insurance should be included, and an effective billing system should be established.

Suggestions for the Design of Diabetes Apps

There were a few differences in the needs of an app between patients with T1DM and patients with T2DM. More patients with T1DM believed that the insulin dose calculator was the most important function of diabetes apps. Patients with T1DM rely on insulin therapy, and insulin dosage must be adjusted according to diet and exercise. Thus, these patients have a greater need for this function. There were also differences in the choice of apps between the 2 groups. The most common choice of patients with T1DM was an app targeted for patients with T1DM (Diabetes Circle). The most common choice of patients with T2DM was Diabetes Nurse, to which blood sugar tested by a Dnurse glucometer can be directly transmitted. Additionally, most doctors believed that patients with T1DM and patients with T2DM need different apps. Therefore, apps should be designed according to different types of diabetes patients' demands.

The data entry burden is the major reason why patients cannot persist in using an app [44]. Data transferred directly from a glucometer to an app will reduce the data entry burden, which is why most patients chose the Diabetes Nurse app. Diabetologists believed that the main reason why an app was ineffective was that patients could not persist in using it. The main reasons why patients did not want to continue to use an app were lack of time and complicated operations. Therefore, app design should enable blood sugar data from glucometers to be automatically transmitted to the app, which will greatly increase patient compliance.

Of the diabetologists, 46.1% believed that a lack of systematic and standardized diabetes education knowledge was a reason for the poor efficacy of apps. Most apps do not have educational information cited from accredited sources [45]. Previous studies indicated that a mobile app was preferable to receiving DSMES in a hospital [26,46]. Therefore, diabetes education knowledge compiled by a multidisciplinary team will effectively improve patients' self-management ability and improve the effectiveness of an app.

Prospect of Diabetes Apps

Almost all patients indicated that they needed an effective app to manage blood sugar. Almost all diabetologists said they would or could use diabetes apps to manage patients and recommend diabetes apps to their patients. China has the largest number of patients with diabetes in the world [2], and our results suggest that diabetes apps have a very good future in China.

Comparison With Prior Work

To our knowledge, no large-scale investigations have been conducted on diabetes app use among people with diabetes in China. An Australian survey suggested that disease duration was related to the use of app in diabetes patients [23,24]. Our study did not find such an association, probably because of different samples. The Australian survey only investigated patients with a disease duration of more than 1 year, and our study included patients of all disease durations. Furthermore, the disease durations were self-reported and might be inaccurate. Our study found that diabetes app users tended to be younger, have higher incomes, and be more educated, which was consistent with a health app survey in the United States [44].

Boyle et al investigated both diabetes patients and HCPs for their use and beliefs about diabetes apps. They also found that diabetes diaries were most useful for diabetes patients [22]. Our strength was that we recruited a large sample throughout China, and we also investigated the patient-doctor communication feature from the perspectives of both patients and doctors, which is considered to be the most important function by both patients and diabetologists [26].

Few studies have surveyed diabetes app use from the point of view of HCPs. Karduck and Chapman-Novakofski investigated clinicians' perspectives on mHealth apps and found that most clinicians (62%) recommended mobile phone apps to their patients to track their diet and physical activity [40]. Our study found a 40.5% recommendation rate. However, most of their samples were nutritionists and diabetes educators. We recruited only diabetologists because diabetes patients in China are mostly treated by these doctors.

In addition to investigating patients' app use, we also investigated diabetologists' use, perspectives, and attitudes regarding diabetes apps, as well as patients' needs, perspectives, and attitudes regarding diabetes apps. Therefore, we can provide more effective information for app design and promotion of app usage and better understand the diabetes app market in China.

Limitations

A strength of our study was that the patient and diabetologist survey links were initially sent to doctors from 46 LADA China Study collaborative hospitals in 25 representative major cities and spread on their WeChat contact networks. In addition to this snowball sampling method, patients with diabetes were surveyed via convenience sampling on 3 WeChat public accounts. Finally, our sample came from 30 provinces across China. Our research also had several limitations. First, the 1276 sampled patients did not sufficiently represent the large population of patients with diabetes in China. Second, our sampling was not stratified by geographic region, urban or rural location, socioeconomic status, age, or diabetes type. Some selection bias was unavoidable. Although the mean age of our patient sample was comparable with that of the national survey concerning diabetes prevalence in China [47], the proportion of patients with T1DM in our sample was higher than the actual disease proportion of total patients with diabetes [48]. Finally, our sampling was based on the WeChat network. Although WeChat has 1.04 billion monthly active users [49], some people do not use WeChat or surf the internet. Thus, actual app usage might be lower, particularly among elderly patients.

Our study was a cross-sectional survey. Although patients' perspectives and attitudes are very important when developing a mobile app for their use [50,51], people's attitudes toward what is useful and what might work are heavily anchored in their present experience regarding the development of technology and how it is implemented. Therefore, these findings must be updated over time as technology develops and people's perceptions change. Furthermore, many factors influence app use. Although we adjusted for some factors in the multivariate analysis, other potential confounding factors remain.

Conclusions

Using an exploratory approach, we found that awareness and use of diabetes apps among the Chinese diabetic population and the proportion of diabetologists using diabetes apps to manage patients are low. There are a few differences in the needs for and choice of diabetes apps between patients with T1DM and patients with T2DM. Therefore, designing apps targeted for different patients' needs and conducting high-quality RCTs will improve the effectiveness of apps, provide evidence for patients to choose suitable apps, and be conducive to the promotion of diabetes apps. China should increase public awareness of diabetes apps, and relevant policies and regulations are needed to support doctors' use of apps to manage patients. Diabetes app use in China has good potential. Diabetes apps are potentially effective supplements that can be used in traditional outpatient clinics to improve glycemic control in Chinese patients with diabetes.

Acknowledgments

The authors thank all the participants for sharing their perspectives in the course of this study. The authors also thank American Journal Experts for their assistance in language editing.

This work was supported by the National Key R&D Program of China (2017YFC1309604), (2018YFC1315603) and the Natural Science Foundation of Hunan Province, China (2018JJ2573).

Conflicts of Interest

None declared.

References

1. Cho NH, Shaw JE, Karuranga S, Huang Y, da Rocha JD, Ohlrogge AW, et al. IDF Diabetes Atlas: global estimates of diabetes prevalence for 2017 and projections for 2045. Diabetes Res Clin Pract 2018 Apr;138:271-281. [doi: 10.1016/j.diabres.2018.02.023] [Medline: 29496507]
2. Xu Y, Wang L, He J, Bi Y, Li M, Wang T, 2010 China Noncommunicable Disease Surveillance Group. Prevalence and control of diabetes in Chinese adults. J Am Med Assoc 2013 Sep 4;310(9):948-959. [doi: 10.1001/jama.2013.168118] [Medline: 24002281]
3. Nathan DM. Long-term complications of diabetes mellitus. N Engl J Med 1993 Jun 10;328(23):1676-1685. [doi: 10.1056/NEJM199306103282306] [Medline: 8487827]
4. Bommer C, Heesemann E, Sagalova V, Manne-Goehler J, Atun R, Bärnighausen T, et al. The global economic burden of diabetes in adults aged 20-79 years: a cost-of-illness study. Lancet Diabetes Endocrinol 2017 Dec;5(6):423-430. [doi: 10.1016/S2213-8587(17)30097-9] [Medline: 28456416]
5. Beck J, Greenwood DA, Blanton L, Bollinger ST, Butcher MK, Condon JE, 2017 Standards Revision Task Force. 2017 National Standards for Diabetes Self-Management Education and Support. Diabetes Care 2017 Dec;40(10):1409-1419. [doi: 10.2337/dci17-0025] [Medline: 28754780]
6. American Diabetes Association. 6. Glycemic targets: Standards of Medical Care in Diabetes-2018. Diabetes Care 2018 Dec;41(Suppl 1):S55-S64. [doi: 10.2337/dc18-S006] [Medline: 29222377]
7. Chen Y, Yin Z, Xie Q. Suggestions to ameliorate the inequity in urban/rural allocation of healthcare resources in China. Int J Equity Health 2014;13:34 [FREE Full text] [doi: 10.1186/1475-9276-13-34] [Medline: 24884614]

8. Li Q, Xie P. Outpatient workload in China. Lancet 2013 Jun 8;381(9882):1983-1984. [doi: 10.1016/S0140-6736(13)61198-8] [Medline: 23746895]
9. Hu Y, Zhang Z. Skilled doctors in tertiary hospitals are already overworked in China. Lancet Glob Health 2015 Dec;3(12):e737 [FREE Full text] [doi: 10.1016/S2214-109X(15)00192-8] [Medline: 26566744]
10. Chyxx. [Analysis of development and prospect of smart phone industry in China in 2018 and its development prospects] URL: http://www.chyxx.com/industry/201803/620359.html [accessed 2018-10-22] [WebCite Cache ID 73MRoAWBg]
11. Hsu J, Liu D, Yu YM, Zhao HT, Chen ZR, Li J, et al. The top Chinese mobile health apps: a systematic investigation. J Med Internet Res 2016 Aug 29;18(8):e222 [FREE Full text] [doi: 10.2196/jmir.5955] [Medline: 27573724]
12. Kleinman NJ, Shah A, Shah S, Phatak S, Viswanathan V. Impact of the Gather mHealth system on A1C: primary results of a multisite randomized clinical trial among people with type 2 diabetes in India. Diabetes Care 2016 Dec;39(10):e169-e170. [doi: 10.2337/dc16-0869] [Medline: 27493133]
13. Zhou W, Chen M, Yuan J, Sun Y. Welltang-a smart phone-based diabetes management application - improves blood glucose control in Chinese people with diabetes. Diabetes Res Clin Pract 2016 Jun;116:105-110. [doi: 10.1016/j.diabres.2016.03.018] [Medline: 27321324]
14. Waki K, Fujita H, Uchimura Y, Omae K, Aramaki E, Kato S, et al. DialBetics: a novel smartphone-based self-management support system for type 2 diabetes patients. J Diabetes Sci Technol 2014 Mar 13;8(2):209-215 [FREE Full text] [doi: 10.1177/1932296814526495] [Medline: 24876569]
15. Garg SK, Shah VN, Akturk HK, Beatson C, Snell-Bergeon JK. Role of mobile technology to improve diabetes care in adults with type 1 diabetes: the remote-T1D Study iBGStar((R)) in type 1 diabetes managemen. Diabetes Ther 2017 Aug;8(4):811-819 [FREE Full text] [doi: 10.1007/s13300-017-0272-5] [Medline: 28555339]
16. Bonoto BC, de Araújo VE, Godói IP, de Lemos LL, Godman B, Bennie M, et al. Efficacy of mobile apps to support the care of patients with diabetes mellitus: a systematic review and meta-analysis of randomized controlled trials. JMIR Mhealth Uhealth 2017 Mar 1;5(3):e4 [FREE Full text] [doi: 10.2196/mhealth.6309] [Medline: 28249834]
17. American Diabetes Association. 4. Lifestyle management: Standards of Medical Care in Diabetes-2018. Diabetes Care 2018 Jan;41(Suppl 1):S38-S50. [doi: 10.2337/dc18-S004] [Medline: 29222375]
18. Hou C, Xu Q, Diao S, Hewitt J, Li J, Carter B. Mobile phone applications and self-management of diabetes: a systematic review with meta-analysis, meta-regression of 21 randomized trials and GRADE. Diabetes Obes Metab 2018 Dec;20(8):2009-2013. [doi: 10.1111/dom.13307] [Medline: 29582538]
19. Offringa R, Sheng T, Parks L, Clements M, Kerr D, Greenfield MS. Digital diabetes management application improves glycemic outcomes in people with type 1 and type 2 diabetes. J Diabetes Sci Technol 2018 May;12(3):701-708 [FREE Full text] [doi: 10.1177/1932296817747291] [Medline: 29277103]
20. Husted GR, Weis J, Teilmann G, Castensøe-Seidenfaden P. Exploring the influence of a smartphone app (Young with Diabetes) on young people's self-management: qualitative study. JMIR Mhealth Uhealth 2018 Feb 28;6(2):e43 [FREE Full text] [doi: 10.2196/mhealth.8876] [Medline: 29490897]
21. Williams JP, Schroeder D. Popular glucose tracking apps and use of mhealth by Latinos with diabetes: review. JMIR Mhealth Uhealth 2015;3(3):e84 [FREE Full text] [doi: 10.2196/mhealth.3986] [Medline: 26307533]
22. Boyle L, Grainger R, Hall RM, Krebs JD. Use of and beliefs about mobile phone apps for diabetes self-management: surveys of people in a hospital diabetes clinic and diabetes health professionals in New Zealand. JMIR Mhealth Uhealth 2017 Jun 30;5(6):e85 [FREE Full text] [doi: 10.2196/mhealth.7263] [Medline: 28666975]
23. Trawley S, Baptista S, Browne JL, Pouwer F, Speight J. The use of mobile applications among adults with type 1 and type 2 diabetes: results from the second MILES-Australia (MILES-2) study. Diabetes Technol Ther 2017 Dec;19(12):730-738. [doi: 10.1089/dia.2017.0235] [Medline: 29028442]
24. Trawley S, Browne JL, Hagger VL, Hendrieckx C, Holmes-Truscott E, Pouwer F, et al. The use of mobile applications among adolescents with type 1 diabetes: results from Diabetes MILES Youth-Australia. Diabetes Technol Ther 2016 Dec;18(12):813-819. [doi: 10.1089/dia.2016.0233] [Medline: 27788032]
25. Conway N, Campbell I, Forbes P, Cunningham S, Wake D. mHealth applications for diabetes: user preference and implications for app development. Health Informatics J 2016 Dec;22(4):1111-1120. [doi: 10.1177/1460458215616265] [Medline: 26635324]
26. Zhang Y, Li X, Luo S, Liu C, Liu F, Zhou Z. Exploration of users' perspectives and needs and design of a type 1 diabetes management mobile app: mixed-methods study. JMIR Mhealth Uhealth 2018 Sep 21;6(9):e11400 [FREE Full text] [doi: 10.2196/11400] [Medline: 30249580]
27. Chavez S, Fedele D, Guo Y, Bernier A, Smith M, Warnick J, et al. Mobile apps for the management of diabetes. Diabetes Care 2017 Oct;40(10):e145-e146. [doi: 10.2337/dc17-0853] [Medline: 28774944]
28. Wu Y, Yao X, Vespasiani G, Nicolucci A, Dong Y, Kwong J, et al. Mobile app-based interventions to support diabetes self-management: a systematic review of randomized controlled trials to identify functions associated with glycemic efficacy. JMIR Mhealth Uhealth 2017 Mar 14;5(3):e35 [FREE Full text] [doi: 10.2196/mhealth.6522] [Medline: 28292740]
29. Chomutare T, Fernandez-Luque L, Arsand E, Hartvigsen G. Features of mobile diabetes applications: review of the literature and analysis of current applications compared against evidence-based guidelines. J Med Internet Res 2011;13(3):e65 [FREE Full text] [doi: 10.2196/jmir.1874] [Medline: 21979293]

30. Nie L, Xie B, Yang Y, Shan YM. Characteristics of Chinese m-Health applications for diabetes self-management. Telemed J E Health 2016 Jul;22(7):614-619. [doi: 10.1089/tmj.2015.0184] [Medline: 27171016]
31. Drincic A, Prahalad P, Greenwood D, Klonoff DC. Evidence-based mobile medical applications in diabetes. Endocrinol Metab Clin North Am 2016 Dec;45(4):943-965 [FREE Full text] [doi: 10.1016/j.ecl.2016.06.001] [Medline: 27823614]
32. American Diabetes Association. Standards of medical care in diabetes--2014. Diabetes Care 2014 Jan;37(Suppl 1):S14-S80. [doi: 10.2337/dc14-S014] [Medline: 24357209]
33. Wikipedia. WeChat URL: https://en.wikipedia.org/wiki/WeChat [accessed 2018-10-22] [WebCite Cache ID 73MQIKI5b]
34. Zhang X, Wen D, Liang J, Lei J. How the public uses social media WeChat to obtain health information in China: a survey study. BMC Med Inform Decis Mak 2017 Jul 5;17(Suppl 2):66 [FREE Full text] [doi: 10.1186/s12911-017-0470-0] [Medline: 28699549]
35. Statista: The Stastics Portal. Average number of contacts of WeChat users in China as of 2016 URL: https://www.statista.com/statistics/387665/wechat-china-contact-list-size/ [accessed 2018-10-22] [WebCite Cache ID 73MPvY2mm]
36. [Changsha Haoxing Information Technology Co]. [More than a questionnaire/online exam] URL: https://www.wjx.cn/ [accessed 2018-12-05] [WebCite Cache ID 74QZMJNBa]
37. Zhou Z, Xiang Y, Ji L, Jia W, Ning G, Huang G, LADA China Study Group. Frequency, immunogenetics, and clinical characteristics of latent autoimmune diabetes in China (LADA China study): a nationwide, multicenter, clinic-based cross-sectional study. Diabetes 2013 Feb;62(2):543-550 [FREE Full text] [doi: 10.2337/db12-0207] [Medline: 23086039]
38. Chyxx. [Analysis of China's per capita GDP ranking and China's per capita GDP ranking in the world in 2016] URL: https://www.chyxx.com/industry/201712/591596.html [accessed 2018-10-22] [WebCite Cache ID 73MPYR6gt]
39. McGuire HC, Ji L, Kissimova-Skarbek K, Whiting D, Aguirre F, Zhang P, et al. Type 1 diabetes mellitus care and education in China: the 3C study of coverage, cost, and care in Beijing and Shantou. Diabetes Res Clin Pract 2017 Jul;129:32-42. [doi: 10.1016/j.diabres.2017.02.027] [Medline: 28500868]
40. Karduck J, Chapman-Novakofski K. Results of the Clinician Apps Survey, how clinicians working with patients with diabetes and obesity use mobile health apps. J Nutr Educ Behav 2018 Jan;50(1):62-69.e1. [doi: 10.1016/j.jneb.2017.06.004] [Medline: 29325664]
41. Istepanian RS. Mobile applications for diabetes management: efficacy issues and regulatory challenges. Lancet Diabetes Endocrinol 2015 Dec;3(12):921-923. [doi: 10.1016/S2213-8587(15)00414-3] [Medline: 26545304]
42. Holmen H, Wahl AK, Cvancarova SM, Ribu L. Tailored communication within mobile apps for diabetes self-management: a systematic review. J Med Internet Res 2017 Jun 23;19(6):e227 [FREE Full text] [doi: 10.2196/jmir.7045] [Medline: 28645890]
43. [National Health and Wellness Commission]. [Notice on Printing and Distributing 3 Documents Such as the Internet Management and Management Measures (Trial)] URL: http://www.nhfpc.gov.cn/yzygj/s3594q/201809/c6c9dab0b00c4902a5e0561bbf0581f1.shtml [accessed 2018-10-22] [WebCite Cache ID 73MT7S8am]
44. 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]
45. Huang Z, Soljak M, Boehm BO, Car J. Clinical relevance of smartphone apps for diabetes management: a global overview. Diabetes Metab Res Rev 2018 May;34(4):e2990. [doi: 10.1002/dmrr.2990] [Medline: 29431916]
46. Pal K, Dack C, Ross J, Michie S, May C, Stevenson F, et al. Digital health interventions for adults with type 2 diabetes: qualitative study of patient perspectives on diabetes self-management education and support. J Med Internet Res 2018 Jan 29;20(2):e40 [FREE Full text] [doi: 10.2196/jmir.8439] [Medline: 29463488]
47. Wang L, Gao P, Zhang M, Huang Z, Zhang D, Deng Q, et al. Prevalence and ethnic pattern of diabetes and prediabetes in China in 2013. J Am Med Soc 2017 Dec 27;317(24):2515-2523. [doi: 10.1001/jama.2017.7596] [Medline: 28655017]
48. Bullard KM, Cowie CC, Lessem SE, Saydah SH, Menke A, Geiss LS, et al. Prevalence of diagnosed diabetes in adults by diabetes type - United States, 2016. MMWR Morb Mortal Wkly Rep 2018 Mar 30;67(12):359-361 [FREE Full text] [doi: 10.15585/mmwr.mm6712a2] [Medline: 29596402]
49. Statista: The Stastics Portal. Number of monthly active WeChat users from 2nd quarter 2010 to 1st quarter 2018 (in millions) URL: https://www.statista.com/statistics/255778/number-of-active-wechat-messenger-accounts/ [accessed 2018-10-22] [WebCite Cache ID 73MS4A7Hx]
50. Pulman A, Taylor J, Galvin K, Masding M. Ideas and enhancements related to mobile applications to support type 1 diabetes. JMIR Mhealth Uhealth 2013;1(2):e12 [FREE Full text] [doi: 10.2196/mhealth.2567] [Medline: 25100684]
51. Peng W, Yuan S, Holtz BE. Exploring the challenges and opportunities of health mobile apps for individuals with type 2 diabetes living in rural communities. Telemed J E Health 2016 Sep;22(9):733-738. [doi: 10.1089/tmj.2015.0180] [Medline: 26982017]

Abbreviations

DSMES: diabetes self-management education and support
GDP: gross domestic product
HbA_{1c}: glycosylated hemoglobin

HCP: health care professional
IQR: interquartile range
LADA: Latent Autoimmune Diabetes of Adults
mHealth: mobile health
OR: odds ratio
RCT: randomized controlled trial
T1DM: type 1 diabetes mellitus
T2DM: type 2 diabetes mellitus

Edited by G Eysenbach; submitted 31.10.18; peer-reviewed by A Martinez-Millana, J Lei, A Farmer; comments to author 29.11.18; revised version received 01.01.19; accepted 20.01.19; published 08.02.19.

Please cite as:
Zhang Y, Li X, Luo S, Liu C, Xie Y, Guo J, Liu F, Zhou Z
Use, Perspectives, and Attitudes Regarding Diabetes Management Mobile Apps Among Diabetes Patients and Diabetologists in China: National Web-Based Survey
JMIR Mhealth Uhealth 2019;7(2):e12658
URL: http://mhealth.jmir.org/2019/2/e12658/
doi:10.2196/12658
PMID:30735147

Original Paper

A Mobile Phone–Based Self-Monitoring Tool for Perioperative Gastric Cancer Patients With Incentive Spirometer: Randomized Controlled Trial

Ji Yeong Soh[1*], RN; Se Uk Lee[2*], MD; Inpyo Lee[3], MS; Ki Sang Yoon[4], BD; Changho Song[3], BS; Nam Hun Kim[5], MS; Tae Sung Sohn[6], MD; Jae Moon Bae[6], MD; Dong Kyung Chang[1,7,8], MD; Won Chul Cha[8,9,10], MD

[1]Department of Digital Health, Samsung Advanced Institute for Health Sciences & Technology, Sungkyunkwan University, Gangnam-gu, Republic of Korea

[2]Department of Emergency Medicine, Seoul National University Hospital, Seoul, Republic of Korea

[3]Product Development Division, BreaThings Co, Ltd, Seoul, Republic of Korea

[4]R&D Division, BreaThings Co, Ltd, Seoul, Republic of Korea

[5]Camera R&D Group, Samsung Electronics, Seoul, Republic of Korea

[6]Department of Surgery, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Republic of Korea

[7]Department of Gastroenterology, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Republic of Korea

[8]Health Information and Strategy Center, Samsung Medical Center, Seoul, Republic of Korea

[9]Department of Digital Health, Samsung Advanced Institute for Health Sciences & Technology, Sungkyunkwan University, Seoul, Republic of Korea

[10]Department of Emergency Medicine, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Republic of Korea

[*]these authors contributed equally

Corresponding Author:
Won Chul Cha, MD
Department of Digital Health
Samsung Advanced Institute for Health Sciences & Technology
Sungkyunkwan University
81, Irwon-ro
Gangnam-gu
Seoul, 06351
Republic of Korea
Phone: 82 10 5386 6597
Email: docchaster@gmail.com

Abstract

Background: An incentive spirometer (IS) is a medical device used to help patients improve the functioning of their lungs. It is provided to patients who have had any surgery that might jeopardize respiratory function. An incentive spirometer plays a key role in the prevention of postoperative complications, and the appropriate use of an IS is especially well known for the prevention of respiratory complications. However, IS utilization depends on the patient's engagement, and information and communication technology (ICT) can help in this area.

Objective: This study aimed to determine the effect of mobile ICT on the usage of an IS (Go-breath) app by postoperative patients after general anesthesia.

Methods: For this study, we recruited patients from April to May 2018, who used the Go-breath app at a single tertiary hospital in South Korea. The patients were randomly classified into either a test or control group. The main function of the Go-breath app was to allow for self-reporting and frequency monitoring of IS use, deep breathing, and active coughing in real time. The Go-breath app was identical for both the test and control groups, except for the presence of the alarm function. The test group heard an alarm every 60 min from 9 am to 9 pm for 2 days. For the test group alone, a dashboard was established in the nurse's station through which a nurse could rapidly assess the performance of multiple patients. To evaluate the number of performances per group, we constructed an incentive spirometer index (ISI).

Results: A total of 44 patients were recruited, and 42 of them completed the study protocol. ISI in the test group was 20.2 points higher than that in the control group (113.5 points in the test group and 93.2 points in the control group, *P*=.22). The system

usability scale generally showed almost the same score in the 2 groups (79.3 points in the test group and 79.4 points in the control group, *P*=.94). We observed that the performance rates of IS count, active coughing, and deep breathing were also higher in the test group but with no statistically significant difference between the groups. For the usefulness "yes or no" question, over 90% (38/42) of patients answered "yes" and wanted more functional options and information.

Conclusions: The use of the Go-breath app resulted in considerable differences between the test group and control group but with no statistically significant differences.

Trial Registration: ClinicalTrials.gov NCT03569332; https://clinicaltrials.gov/ct2/show/NCT03569332 (Archived by WebCite at http://www.webcitation.org/74ihKmQIX).

(JMIR Mhealth Uhealth 2019;7(2):e12204) doi:10.2196/12204

KEYWORDS

incentive spirometer; mobile health; postoperative care; gastric cancer; motivation

Introduction

Background

Vigorous and multidisciplinary efforts are required to prevent postoperative complications following prolonged general anesthesia [1], and an incentive spirometer (IS) plays a key role, especially for patients with limited mobility [2]. An incentive spirometer is a medical device used to help patients improve the functioning of their lungs. It is provided to patients who have had any surgery that might jeopardize respiratory function. The IS is used for the prevention of atelectasis, hypoxemia, pneumonia, respiratory dysfunction, and pleural effusion [3,4]. However, IS utilization has remained the same since its initial development in the 1960s, and the usage protocol has not been standardized [5-8].

Although many mobile phone apps have been developed for self-management in various clinical settings, only a few cases showing their use in postoperative care were reported [9], even though subjective activity by patients can influence their clinical hospital outcomes. Therefore, with help from mobile information and communication technology (ICT), patients can be more engaged in behaviors that can lead to better clinical results [3,10,11].

A mobile ICT for hospitalized postoperative patients can improve clinical outcomes by several pathways: it can encourage patients by informing them of important performance indices such as exercise duration [12-14], and it can connect the patient and provider with a real-time dashboard that can support prompt provider reactions. In a previous study, a postoperative group of patients using mobile ICT recovered significantly better than a control group [9].

Objective

Therefore, the goal of this study was to determine the effect of mobile ICT on the performance of postoperative patients using an IS (Go-breath) app after general anesthesia.

Methods

Study Design

This was a single-center, randomized controlled trial (RCT) assessing the effectiveness of an IS self-reporting app (Go-breath) to improve IS performance and effectiveness. This trial was registered with the US National Institutes of Health Clinical Trials Registry (NCT03569332).

Study Setting

This study was conducted at an academic tertiary center located in Seoul, Korea. The hospital's expertise is cancer care. It hosts approximately 2000 inpatient beds and registers approximately 8500 outpatients per day.

Gastric cancer is one of the major diseases treated in the institution. The hospital's Critical Pathway defines the routine care process for gastric cancer surgery as follows: The patient is admitted the evening before the surgery and discharged approximately 9 to 10 days after surgery—the discharge period of laparotomy is postoperative day (POD) 8 and for laparoscopy is POD 7—without complications; the patient receives instructions on the routine use of an IS before the surgery, and its use is initiated the morning after the surgery (POD 1).

Study Participants

Participants were enrolled from April to May 2018. We recruited patients from the surgical ward immediately after the admission process. Our eligibility criteria stipulated that patients were older than 18 years and planned to undergo robotic or laparoscopic surgery or a laparotomy. The enrollment was performed the day before the surgery, and the intervention was initiated the day after the surgery (POD 1). Following recruitment, participants were provided written informed consent and randomized using a Web-based random number generator and sealed enveloped in a 1:1 ratio. If the patient was in the intensive care unit (ICU) or in an emergency condition requiring nonelective surgery, he or she was excluded from the study. The patients were dropped from the study if the postoperative state of the patients was not difficult to predict in advance and if their postsurgical condition (and before the initiation of intervention) was consistent with the exclusion criteria such as ICU admission. Dropout was only permitted on the day of surgery (POD 0).

Study Protocol

Patients were allocated to the test and control groups after giving their consent for the study. Patients received a tablet with the Go-breath app installed, and they received information about the app and use of the IS for approximately 30 min. The IS awareness session was carried out using either video clips

(Multimedia Appendix 1) or nurses as part of the routine clinical process.

Generally, the nurses were encouraged to check the performance of the IS on the patients over 10 times every hour after the surgery, and some ward nurses checked the IS performance rate more than 10 times per hour from 9 am to 9 pm.

The Go-breath app was identical for both the test and control groups except for the alarm function. The test group heard an alarm every 60 min from 9 am to 9 pm during POD1-POD2 (a total of 24 hours). For the test group only, a dashboard was established in the nurse's station through which a nurse could rapidly assess the performance of multiple patients. On the other hand, the control group was not provided an alarm function or nurse dashboard identification. The control group only entered the number of IS performances using the app.

The measurement of IS use was unlimited, but the record of the Go-breath app and dashboard would show up to a maximum of 10 times per hour. We named this the incentive spirometer index (ISI), and a description of this index is provided in the major outcomes section. Termination of the study occurred in the morning of POD 3. Study assistants visited each patient, collected devices, conducted a short survey about usability among the patients and interviewed the patients for subjective opinions. Even though the study intervention lasted for only 2 days, the patients' outcomes were tracked until the day of discharge. An overview of the study protocol is provided in Figure 1.

Device and System

Hardware and System Architecture

The system was constructed based on the opinions of health providers and comprised a mobile phone app for patients and a server for medical staff. Patients could self-input their performance of IS, active coughing, and deep breathing every time. This information was transmitted to a nurse dashboard for a real-time check. All patients were given the same hardware (Galaxy tablet A 8.0, Samsung, purchased 2018) running Android 8.0, and the mobile Go-breath app was tested on Android 8.0 utilizing an Amazon Web service server (Figure 2).

Figure 1. Study protocol. Superscripted "a" indicates alarm every 50 min for insufficient during 9am-9pm (alarm off during 9pm-9am); "b" indicates calculation of performance score applies only to incentive spirometer (full score 240, 10 out of 10 points). V/S: blood pressure, fever, body temperature, respiration rate; CRP: C-reactive protein; IS: incentive spirometer; PA: posterior to anterior; WBC: white blood cell.

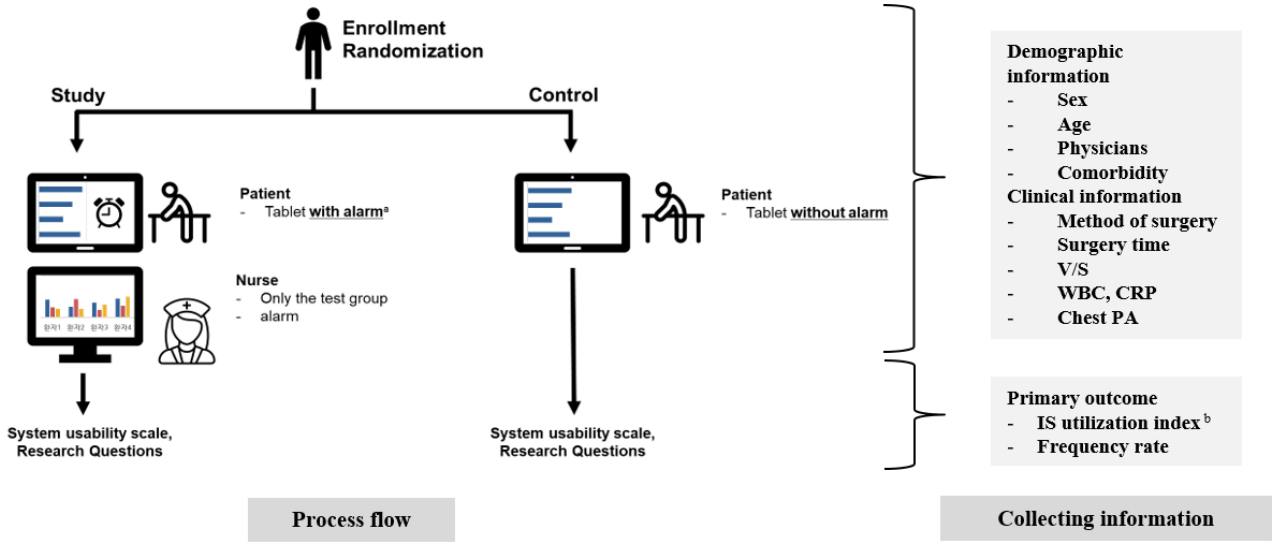
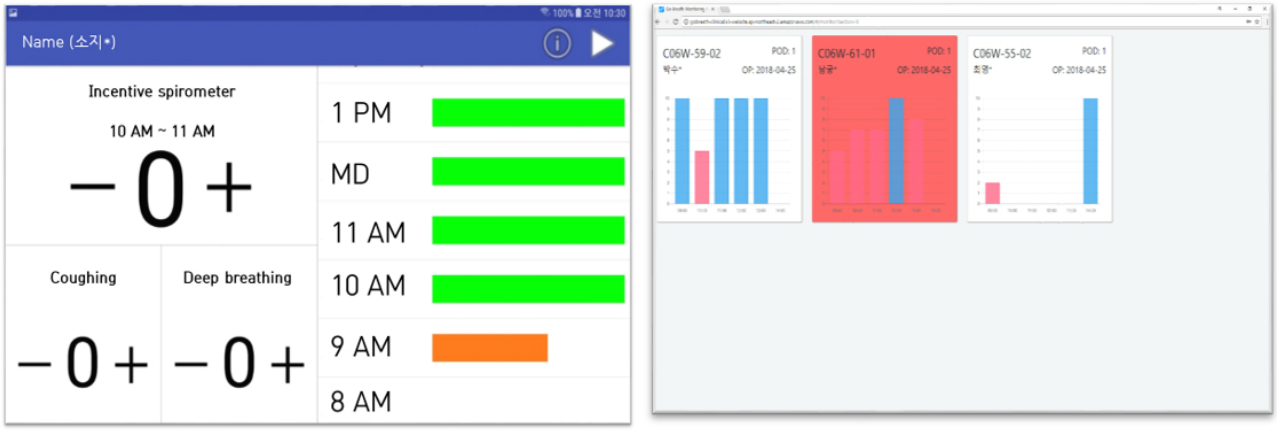


Figure 2. Screenshot of the Go-breath app (left) and dashboard (right). The original version was in Korean, but it was modified to display English.



Go-breath App

The main function of the Go-breath app was to allow for self-reporting and frequency monitoring of IS, deep breathing, and active coughing in real-time. In principle, patients were instructed to use the IS 10 times every hour and simultaneously report it to the app. The optimal number of times to use the IS was derived from the literature [2,6]. Patients could also check on their current state and edit previous indices they had input (Figure 3).

The Go-breath app was set to alarm when the number of IS uses did not reach 10 in an hour. The app was set to ring 10 min before the end of each hour to give the patients time to complete the expected number of hourly uses. When the patients fulfilled their hourly IS use, the alarm would silence automatically. However, patients could silence the alarm. When the mute icon was pressed, the alarm would be deactivated for 2 hours and would reactivate by default. This functionality was provided because some of the patients could be away from their bed for various reasons, and the alarm could disturb other patients. The system remained frozen during the study period. We had preserved the anonymity of the patients by an unidentified number.

Go-breath Dashboard

Only information from the test group was displayed on the Go-breath dashboard. The dashboard showed IS performance that did not include other measures, such as deep breathing and active coughing. Each patient was displayed as a card with his or her location and POD information. For each card, the patient's ISI during the most recent 6 hours was displayed as a bar graph. If the most recent hourly ISI was not optimal (below 10), the card turned red. The timing and alarm rules were identical to that of the Go-breath app (Figure 3).

Major Outcomes

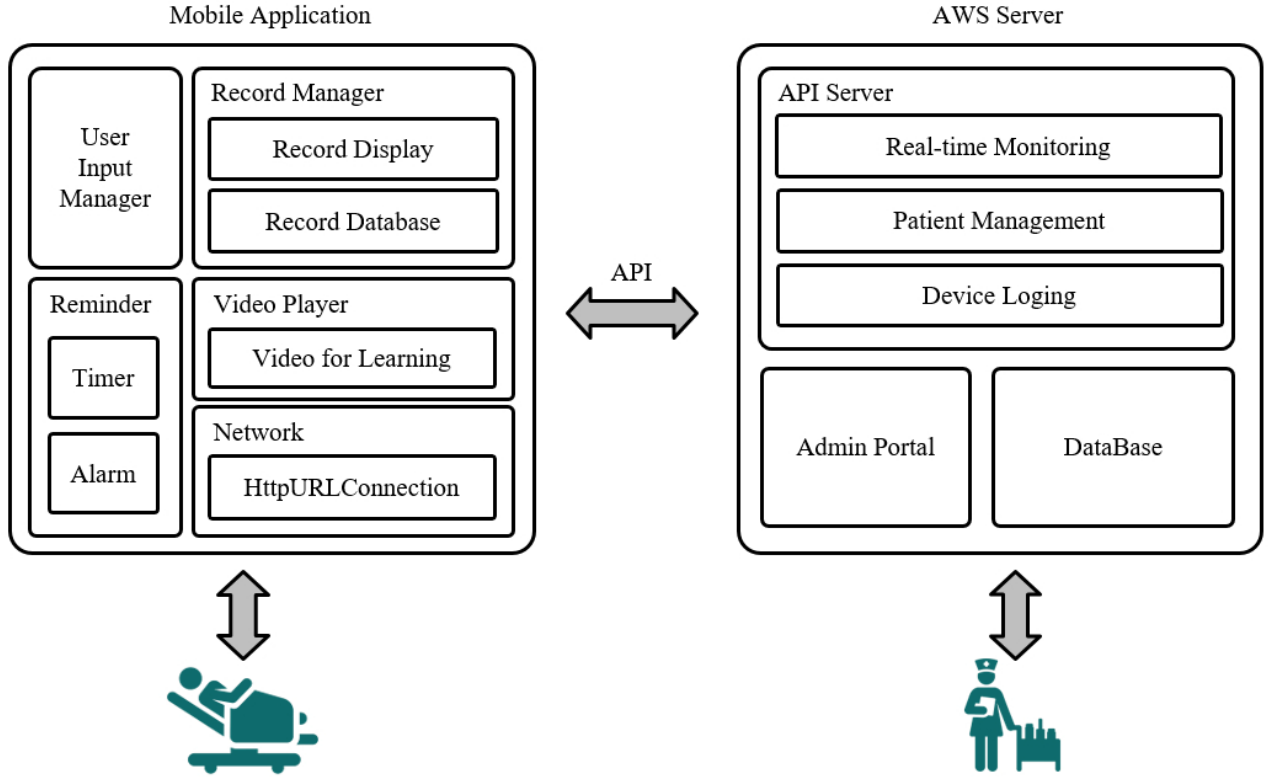
Primary Outcome

We developed the ISI, which ranged from 0 to 120 per day. The maximum score was 240 for a 2-day observation period. We developed an outcome index using clinical experts and supporting literature [2,6-8]. The ISI was defined as 10 points per each hour, the frequency of IS use was 10 times or more for that hour. If the frequency was below 10 for that hour, the score was rated as 0. ISI was only measured from 9:00 am to 9:00 pm at PODs 1 and 2. We compared the ISI of the test group and control group.

Secondary Outcome

The secondary outcome was the System Usability Scale (SUS). At the end of the study, patients completed an offline questionnaire to evaluate the system using the SUS, which is widely used in health information technology research [15-20]. The questionnaire measures the usability of hardware and software products, and it comprises only 10 items that respondents score on a 5-point Likert scale. Answers were converted to a final 0- to 100-point score, and the higher the score, the more usable the product [17,19,21].

Figure 3. System architecture. API: application programming interface; AWS: Amazon Web Service.



Other Outcomes

In addition to the major outcomes, we measured IS use frequency, active coughing frequency, deep breathing, and clinical results (as length of stay). Clinical data, such as surgery, laboratory data including white blood cell, C-reactive protein, vital signs (fever), and chest posterior to anterior, and length of stay were also collected.

Interviews

After the study, a brief face-to-face interview with every patient was performed by a research assistant. Patients answered whether the system was helpful or not and rated the score. Patients' subjective opinions of the system were also gathered during the process.

Data Collection

We collected the patients' baseline demographic information along with variables associated with surgery including surgeon, type of surgery, and duration of the operation. In the trial, all participants were asked to record log when they started and finished Go-breath app every day. Missing entries were interpreted as a missed value.

Sample Size Calculation

We calculated the sample size of this study with the G-power 3.1 program (Statistical power analyses using G*Power 3.1 [22]). Several assumptions were made to set the sample population. The power was set to 0.90, significance level was set to .05, and effect size was 1.00. The minimum number of samples for the Wilcoxon rank-sum test was 19 in each group. We assumed that the ISI would be different by 60% (12/20; test group) versus 40% (8/20; control group), with SDs for each group. This assumption was made based on expert opinions because there was no evidence that described IS frequency. Under these conditions, the sample size was 19 patients, and finally, 22 patients were selected for each group considering a dropout rate of 10%.

Results

Study Participants

We recruited 44 patients from April to May 2018, with 22 patients allocated to the test group and 22 patients allocated to the control group from patients who used the Go-breath app at a single tertiary hospital in South Korea. A total of 2 patients were excluded because of clinical conditions. The overall average age was 55.9 (SD 12.9) years. The proportion of female patients was 34.1%. All patients had gastric cancer, and the overall average surgery time was 169 (SD 46.2) min. The demographic information of both groups is shown in Table 1.

Major Outcomes

The primary ISI outcome and the secondary SUS outcome are shown in Table 2, and no statistically significant differences were noted in these outcomes between the test and control groups. SUS of both groups showed good to excellent grade [21].

For the performance measures, both the IS rate count and deep breathing showed a higher performance without a statistical significance. However, active coughing showed a significantly higher performance in the test group compared with the control group. For the clinical results, including inflammatory markers and pulmonary complications, there were no statistically significant differences between the test and control groups (Table 3).

Interviews

For the usefulness questionnaire, 90.0% of the test group and 90.9% of the control group answered yes. There were 3 domains of opinions: function addition, function alteration, and etc. The 2 most wanted functions were exercise tracking and peer patients' information, and the most frequent complaint was about Wi-Fi connection issues (Table 4).

Table 1. Demographic characteristics of patients in the test and control groups.

Category	Test group (N=22)	Control group (N=22)
Sex, n (%)		
Female	6 (27)	9 (40)
Male	16 (72)	13 (59)
Dropout, n (%)	2 (9)	0 (0)
Age (years), n (%)		
30-40	1 (4)	4 (18)
50-60	12 (54)	11 (50)
70-80	8 (36)	7 (31)
Over 80	1 (4)	0 (0)
Method of surgery, n (%)		
Laparotomy	10 (45)	8 (36)
Laparoscopy, robotics	12 (54)	14 (63)
Physicians, n (%)		
Dr K	9 (40)	7 (31)
Dr B	3 (13)	3 (13)
Dr S	5 (22)	5 (22)
Dr L	5 (22)	7 (31)
Comorbidity, n (%)		
Tuberculosis	0 (0)	1 (4)
Diabetes mellitus	2 (8)	2 (9)
Hypertension	4 (20)	8 (36)
Others	2 (10)	2 (9)
Time of surgery[a] (min), mean (SD)	168 (54)	170 (36)

[a]Time of surgery was defined as the time from anesthesia to extubation time.

Table 2. Major outcomes.

Index	Test group (N=20), mean (SD)	Control group (N=22), mean (SD)	*P* value
ISI[a]	113.5 (0.8)	93.2 (71.2)	.22
SUS[b]	79.25 (20.59)	79.43 (18.83)	.94

[a]ISI (incentive spirometer utilization index) is a score of 10 points per hour if over 10 times from 9 am to 9 pm for 2 days, with a total maximum score of 240.

[b]SUS: System Usability Scale.

Table 3. Other outcomes.

Category	Test group (N=20)	Control group (N=22)	*P* value
Performance rate, mean (SD)			
Incentive spirometer, count	139.5 (61.2)	119.0 (84.5)	.27
Active coughing	79.0 (52.78)	44.1 (43.62)	.04
Deep breathing	107.8 (66.8)	94.8 (88.33)	.49
Laboratory results, mean (SD)			
WBC[a], k/mL	6.2 (2.8)	5.9 (1.8)	.93
CRP[b], mg/dL	8.1 (4.0)	8.2 (2.9)	.07
Length of stay, mean (SD)	10.8 (1.6)	10.40 (1.00)	.57
Chest PA[c] , n (%)			
Atelectasis	2 (10)	3 (15)	.64
Fever[d], n (%)	6 (30)	10 (45)	.79

[a]WBC: white blood cell.

[b]CRP: C-reactive protein.

[c]PA: posterior to anterior.

[d]Fever was defined as 37.5ºC (99.5ºF), and it means the number of people who have been checked for fever at least once.

Table 4. Components of the qualitative interview.

Part	Contents	Responses (n)
Function addition (N=12)	"I would like to have the ability to check the number of steps when walking after surgery."	3
	"I would like a report function on the results of my performance compared with other patients and checking the number of balls up."	3
	"I wish I could hang it on a portable pole and small size device."	2
	"I would like to be provided with the training methods for coughing and deep breathing in addition."	1
	"I wish that the software has more fun-factors."	1
	"I wish my IS use data gets entered automatically."	1
	"I would like to have a medication schedule on the device."	1
Alert function	"I wish there was a notice function when the goal is achieved."	2
	"I would like to be able to customize the alarm settings myself."	1
	"I was very worried that the alarm would be annoying to other patients."	1
Other components	"I did not want to use it because of Wi-Fi connection issues."	3
	"The device and solution were somewhat difficult for the elderly."	3
	"I expected the nurses to inform me when my performance was not good, only to be disappointed when they did not."	2

Discussion

Principal Findings

The purpose of this study was to assess the effectiveness of a mobile phone app on improving the performance of the IS by postoperative patients. The major intervention of this study was to provide an interactive feedback system for patients to improve their breathing exercise behaviors. To our knowledge, this is the first RCT that involves the IS and a mobile phone app even though multiple research studies had previously shown their feasibility [9,23]. We enrolled 42 patients, split between a test and control group, and compared their ISI, which showed an approximate 20% difference between the 2 groups (113.5, SD 50.8 vs 93.2, SD 71.2), although the difference was not statistically significant.

Strength and Limitations

The negative outcome could have been influenced by several factors. First, the SDs were very high in both groups. This means the study population was heterogeneous with respect to IS use, which was not considered when designing the clinical trial. To the authors' knowledge, there was no literature with specific measures of the IS. The SD was higher than other in-hospital

indices, which was hard to foresee before the trial. In general, the evidence on the effectiveness of this app is limited and needs further investigation before implementation in mobile health care. Recent studies have shown a substantial variation in the effectiveness of mobile health apps. Previous studies have not reported a significant difference in clinical outcomes in mobile health research because of the difficulties in RCT settings, such as small sample sizes and low levels of research design. However, an effective clinical assessment of mobile health is hard to achieve, and there have been various attempts to determine health engagement and behavioral factors [24,25].

The second possible reason was that the feedback mechanism could have influenced the control group more than intended. Patients behave differently because of several factors, and many patients come from a nonmedical background. However, studies involving the behaviors of hospitalized patients still need more exploration. For example, the Diabetes Prevention Program was introduced to Mobile App America by insurance companies. Mobile App America is a company with proven effective mobile phone app development, and it focuses on participant engagement and app value [26-29]. Individuals involved in technological development, engineering, and theory tend to focus on the technology itself [30]. However, health care user engagement in a postoperation setting includes unpredictable factors such as caregiving, word setting, Wi-Fi connections, and nurse feedback.

The third possible reason could be the effectiveness of the alarms. Many patients were sharing rooms with other patients and were very concerned about the sound of the alarm bothering others. To make this kind of trial more successful, patient feedback should be carefully considered.

Nevertheless, the effect of the mobile phone app was demonstrated along with the details of the ISI and performance rate. All results were based on the difference between the test and control groups. Although the effectiveness of the app was not proven, the possible utility of a mobile phone app in the postoperation setting was shown. In addition, this was the first trial to collect real-time evidence of IS device use over a long duration. Although we could not find a standard for how patients should perform the IS, this research is significant by itself.

The negative outcome of this study is because of the diversity of the patients' behavior than the effectiveness of the intervention. Both the test and control groups showed a wide SD, which diminished the statistical power. This could have been prevented by appropriate sample size calculations, but to the authors' knowledge, no other study had revealed such a wide difference in hospitalized patients' behavior with postoperative conditions.

The reason we had negative outcomes in the SUS could vary. Though we wanted to make an interactive system and evaluate its utility, a majority of the device's functions were very similar in both groups. From the users' aspect, reporting on the system could have been the main task. The alarm function was a major intervention for recognition, but it could be a minor task for patients.

This study did not detect a statistically significant difference with Go-breath app usage between the test group and the control group. However, considerable differences were still observed between the 2 groups, and such differences may demonstrate the possible effects of ICT on patient IS engagement. The surgical procedures, comorbidities (Multimedia Appendix 2), and length of anesthesia could have also influenced the outcome, which must have been minimized by randomization.

This study had some limitations. First, we could not measure the frequency of the use of the Go-breath dashboard because of the difficulty of implementing it without substantially disturbing the clinical routines of the acting nurses. Second, because our system relied on the input from a number of patients (nonblinded), the results could have been biased by the patients' subjective cognition, willingness, or adherence to digital devices. The existence of family members could have also influenced the input. However, even with these potential biases, inputting the data itself could have positively influenced the use of IS, as with other mobile solutions [22]. The third limitation is related to the second one; the control group could have been substantially influenced by the device when inputting IS use. Even without its alarm function, individuals in the test group could see their current IS usage status, which could have encouraged them to use it more frequently, aggravating the Hawthorne effect. Finally, there was no confirmed function for the performance rate. Although we made it possible to modify the number of previous performances, it did not guarantee the accuracy of the number of IS uses. To improve this, a sensor that automatically measures the number of IS uses could provide more accurate data.

Conclusions

The use of the Go-breath app resulted in considerable differences between the test group and control group but with no statistically significant differences.

Acknowledgments

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF), funded by the Ministry of Education (NRF-2017R1D1A1B03029526), and processed as one of the creative lab projects by Samsung electronics.

Conflicts of Interest

IP, GS, NH, and CH are employees of the company that developed the Go-breath app.

Editorial notice: This randomized study was only retrospectively registered. The editor granted an exception of ICMJE rules for prospective registration of randomized trials because the risk of bias appears low and the study was considered formative. However, readers are advised to carefully assess the validity of any potential explicit or implicit claims related to primary outcomes or effectiveness, as retrospective registration does not prevent authors from changing their outcome measures retrospectively.

Multimedia Appendix 1

Video clips for incentive spirometer after surgery.

[DOCX File, 2MB - mhealth_v7i2e12204_app1.docx]

Multimedia Appendix 2

Co-morbidities in patients undergoing open surgery and robotic or laparoscopic procedures.

[DOCX File, 14KB - mhealth_v7i2e12204_app2.docx]

Multimedia Appendix 3

CONSORT‐EHEALTH checklist (V 1.6.1).

[PDF File (Adobe PDF File), 3MB - mhealth_v7i2e12204_app3.pdf]

References

1. García-Miguel FJ, Serrano-Aguilar PG, López-Bastida J. Preoperative assessment. Lancet 2003 Nov 22;362(9397):1749-1757. [doi: 10.1016/S0140-6736(03)14857-X] [Medline: 14643127]
2. Restrepo RD, Wettstein R, Wittnebel L, Tracy M. Incentive spirometry: 2011. Respir Care 2011 Oct;56(10):1600-1604 [FREE Full text] [doi: 10.4187/respcare.01471] [Medline: 22008401]
3. Sebio R, Yáñez-Brage MI, Giménez-Moolhuyzen E, Valenza MC, Reychler G, Cahalin LP. Impact of a pre-operative pulmonary rehabilitation program on functional performance in patients undergoing video-assisted thoracic surgery for lung cancer. Arch Bronconeumol 2016 May;52(5):231-232 [FREE Full text] [doi: 10.1016/j.arbres.2015.10.013] [Medline: 26747731]
4. Kumar AS, Alaparthi GK, Augustine AJ, Pazhyaottayil ZC, Ramakrishna A, Krishnakumar SK. Comparison of flow and volume incentive spirometry on pulmonary function and exercise tolerance in open abdominal surgery: a randomized clinical trial. J Clin Diagn Res 2016 Jan 1;10(1):KC01-KC06 [FREE Full text] [doi: 10.7860/JCDR/2016/16164.7064] [Medline: 26894090]
5. Malik PR, Fahim C, Vernon J, Thomas P, Schieman C, Finley CJ, et al. Incentive spirometry after lung resection: a randomized controlled trial. Ann Thorac Surg 2018 Aug;106(2):340-345. [doi: 10.1016/j.athoracsur.2018.03.051] [Medline: 29702071]
6. Hall JC, Tarala RA, Tapper J, Hall JL. Prevention of respiratory complications after abdominal surgery: a randomised clinical trial. Br Med J 1996 Jan 20;312(7024):148-152. [doi: 10.1136/bmj.312.7024.148] [Medline: 8563533]
7. Westwood K, Griffin M, Roberts K, Williams M, Yoong K, Digger T. Incentive spirometry decreases respiratory complications following major abdominal surgery. Surgeon 2007 Dec;5(6):339-342. [doi: 10.1016/S1479-666X(07)80086-2] [Medline: 18080608]
8. Agostini P, Singh S. Incentive spirometry following thoracic surgery: what should we be doing? Physiotherapy 2009 Jun;95(2):76-82. [doi: 10.1016/j.physio.2008.11.003] [Medline: 19627688]
9. Jaensson M, Dahlberg K, Eriksson M, Nilsson U. Evaluation of postoperative recovery in day surgery patients using a mobile phone application: a multicentre randomized trial. Br J Anaesth 2017 Nov 1;119(5):1030-1038 [FREE Full text] [doi: 10.1093/bja/aex331] [Medline: 29077818]
10. Hung WW, Ross JS, Farber J, Siu AL. Evaluation of the Mobile Acute Care of the Elderly (MACE) service. JAMA Intern Med 2013 Jun 10;173(11):990-996 [FREE Full text] [doi: 10.1001/jamainternmed.2013.478] [Medline: 23608775]
11. Semple J, Sharpe S, Murnaghan M, Theodoropoulos J, Metcalfe K. Using a mobile app for monitoring post-operative quality of recovery of patients at home: a feasibility study. JMIR Mhealth Uhealth 2015 Feb 12;3(1):e18 [FREE Full text] [doi: 10.2196/mhealth.3929] [Medline: 25679749]
12. Williams A, Bhatti U, Alam H, Nikolian V. The role of telemedicine in postoperative care. Mhealth 2018;4:11 [FREE Full text] [doi: 10.21037/mhealth.2018.04.03] [Medline: 29963556]
13. Dadlani R, Mani S, Mohan D, Rajgopalan N, Thakar S, Aryan S, et al. The impact of telemedicine in the postoperative care of the neurosurgery patient in an outpatient clinic: a unique perspective of this valuable resource in the developing world--an experience of more than 3000 teleconsultations. World Neurosurg 2014;82(3-4):270-283. [doi: 10.1016/j.wneu.2014.05.027] [Medline: 24887533]

14. Siddle J, Pang P, Weaver C, Weinstein E, O'Donnell D, Arkins T, et al. Mobile integrated health to reduce post-discharge acute care visits: a pilot study. Am J Emerg Med 2018 May;36(5):843-845. [doi: 10.1016/j.ajem.2017.12.064] [Medline: 29317154]
15. Borsci S, Federici S, Lauriola M. On the dimensionality of the System Usability Scale: a test of alternative measurement models. Cogn Process 2009 Aug;10(3):193-197. [doi: 10.1007/s10339-009-0268-9] [Medline: 19565283]
16. Borsci S, Federici S, Bacci S, Gnaldi M, Bartolucci F. Assessing user satisfaction in the era of user experience: comparison of the SUS, UMUX, and UMUX-LITE as a function of product experience. Int J Hum Comput Stud 2015 Jun 24;31(8):484-495. [doi: 10.1080/10447318.2015.1064648]
17. Kortum P, Bangor A. Usability ratings for everyday products measured with the System Usability Scale. Int J Hum Comput Stud 2012 Apr 11;29(2):67-76. [doi: 10.1080/10447318.2012.681221]
18. Sauro J, Lewis JR. Quantifying the User Experience: Practical Statistics for User Research. USA: Morgan Kaufmann; 2012.
19. Brooke J. SUS-a quick and dirty usability scale. In: Usability Evaluation in Industry. UK: CRC Press; 1996:4-7.
20. Bangor A, Kortum P, Miller J. Journ Usability Stud. 2009 May. Determining what individual SUS scores mean: adding an adjective rating scale URL: http://uxpajournal.org/determining-what-individual-sus-scores-mean-adding-an-adjective-rating-scale/
21. Lehmann M, Monte K, Barach P, Kindler CH. Postoperative patient complaints: a prospective interview study of 12,276 patients. J Clin Anesth 2010 Feb;22(1):13-21. [doi: 10.1016/j.jclinane.2009.02.015] [Medline: 20206846]
22. Faul F, Erdfelder E, Buchner A, Lang AG. Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. Behav Res Methods 2009 Nov;41(4):1149-1160. [doi: 10.3758/BRM.41.4.1149] [Medline: 19897823]
23. Overdijkink SB, Velu AV, Rosman AN, van Beukering MD, Kok M, Steegers-Theunissen RP. The usability and effectiveness of mobile health technology-based lifestyle and medical intervention apps supporting health care during pregnancy: systematic review. JMIR Mhealth Uhealth 2018 Apr 24;6(4):e109 [FREE Full text] [doi: 10.2196/mhealth.8834] [Medline: 29691216]
24. Robinson MD, Branham AR, Locklear A, Robertson S, Gridley T. Measuring satisfaction and usability of FaceTime for virtual visits in patients with uncontrolled diabetes. Telemed J E Health 2015 Aug 21. [doi: 10.1089/tmj.2014.0238] [Medline: 26295592]
25. Ackermann RT, Finch EA, Brizendine E, Zhou H, Marrero DG. Translating the diabetes prevention program into the community. The DEPLOY pilot study. Am J Prev Med 2008 Oct;35(4):357-363 [FREE Full text] [doi: 10.1016/j.amepre.2008.06.035] [Medline: 18779029]
26. Diabetes Prevention Program Research Group. The Diabetes Prevention Program (DPP): description of lifestyle intervention. Diabetes Care 2002 Dec 1;25(12):2165-2171 [FREE Full text] [doi: 10.2337/diacare.25.12.2165] [Medline: 12453955]
27. Yu X, Chau JP, Huo L. The effectiveness of traditional Chinese medicine-based lifestyle interventions on biomedical, psychosocial, and behavioral outcomes in individuals with type 2 diabetes: a systematic review with meta-analysis. Int J Nurs Stud 2018 Apr;80:165-180. [doi: 10.1016/j.ijnurstu.2018.01.009] [Medline: 29471267]
28. Xiao L, Yank V, Wilson SR, Lavori PW, Ma J. Two-year weight-loss maintenance in primary care-based Diabetes Prevention Program lifestyle interventions. Nutr Diabetes 2013 Jun;3:e76 [FREE Full text] [doi: 10.1038/nutd.2013.17] [Medline: 23797383]
29. Richards KA, Jones E. Customer relationship management: finding value drivers. Indust Mark Manag 2008 Apr;37(2):120-130. [doi: 10.1016/j.indmarman.2006.08.005]
30. Soh JY, Cha WC, Chang DK, Hwang JH, Kim K, Rha M, et al. Development and validation of a multidisciplinary mobile care system for patients with advanced gastrointestinal cancer: interventional observation study. JMIR Mhealth Uhealth 2018 May 7;6(5):e115 [FREE Full text] [doi: 10.2196/mhealth.9363] [Medline: 29735478]

Abbreviations

ICT: information and communication technology
ICU: intensive care unit
IS: incentive spirometer
ISI: incentive spirometer index
NRF: National Research Foundation of Korea
POD: postoperative day
RCT: randomized controlled trial
SUS: System Usability Scale

Edited by G Eysenbach; submitted 13.09.18; peer-reviewed by Y Carter, S Borsci; comments to author 08.10.18; revised version received 26.11.18; accepted 31.12.18; published 19.02.19.

Please cite as:
Soh JY, Lee SU, Lee I, Yoon KS, Song C, Kim NH, Sohn TS, Bae JM, Chang DK, Cha WC
A Mobile Phone–Based Self-Monitoring Tool for Perioperative Gastric Cancer Patients With Incentive Spirometer: Randomized Controlled Trial
JMIR Mhealth Uhealth 2019;7(2):e12204
URL: http://mhealth.jmir.org/2019/2/e12204/
doi:10.2196/12204
PMID:30777844

Original Paper

Patient Adherence to a Mobile Phone–Based Heart Failure Telemonitoring Program: A Longitudinal Mixed-Methods Study

Patrick Ware[1,2], MPH; Mala Dorai[2], MPH; Heather J Ross[3,4,5], MHS, MD, FRCPC; Joseph A Cafazzo[1,2,6], PEng, PhD; Audrey Laporte[1,7], MA; Chris Boodoo[1], BSc; Emily Seto[1,2], PEng, PhD

[1]Institute of Health Policy, Management and Evaluation, Dalla Lana School of Public Health, Toronto, Toronto, ON, Canada

[2]Centre for Global eHealth Innovation, Techna Institute, University Health Network, Toronto, ON, Canada

[3]Ted Rogers Centre for Heart Research, University Health Network, Toronto, ON, Canada

[4]Department of Medicine, University of Toronto, Toronto, ON, Canada

[5]Peter Munk Cardiac Centre, University Health Network, Toronto, ON, Canada

[6]Institute of Biomaterials and Biomedical Engineering, University of Toronto, Toronto, ON, Canada

[7]Canadian Centre for Health Economics, Toronto, ON, Canada

Corresponding Author:
Patrick Ware, MPH
Institute of Health Policy, Management and Evaluation
Dalla Lana School of Public Health
Toronto
155 College Street
Toronto, ON, M5T 3M6
Canada
Phone: 1 647 227 6015
Email: patrick.ware@mail.utoronto.ca

Abstract

Background: Telemonitoring (TM) can improve heart failure (HF) outcomes by facilitating patient self-care and clinical decision support. However, these outcomes are only possible if patients consistently adhere to taking prescribed home readings.

Objective: The objectives of this study were to (1) quantify the degree to which patients adhered to taking prescribed home readings in the context of a mobile phone–based TM program and (2) explain longitudinal adherence rates based on the duration of program enrollment, patient characteristics, and patient perceptions of the TM program.

Methods: A mixed-methods explanatory sequential design was used to meet the 2 research objectives, and all explanatory methods were guided by the unified theory of acceptance and use of technology 2 (UTAUT2). Overall adherence rates were calculated as the proportion of days patients took weight, blood pressure, heart rate, and symptom readings over the total number of days they were enrolled in the program up to 1 year. Monthly adherence rates were also calculated as the proportion of days patients took the same 4 readings over each 30-day period following program enrollment. Next, simple and multivariate regressions were performed to determine the influence of time, age, sex, and disease severity on adherence rates. Additional explanatory methods included questionnaires at 6 and 12 months probing patients on the perceived benefits and ease of use of the TM program, an analysis of reasons for patients leaving the program, and semistructured interviews conducted with a purposeful sampling of patients (n=24) with a range of adherence rates and demographics.

Results: Overall average adherence was 73.6% (SD 25.0) with average adherence rates declining over time at a rate of 1.4% per month (P<.001). The multivariate regressions found no significant effect of sex and disease severity on adherence rates. When grouping patients' ages by decade, age was a significant predictor (P=.04) whereby older patients had higher adherence rates over time. Adherence rates were further explained by patients' perceptions with regard to the themes of (1) performance expectancy (improvements in HF management and peace of mind), (2) effort expectancy (ease of use and technical issues), (3) facilitating conditions (availability of technical support and automated adherence calls), (4) social influence (support from family, friends, and trusted clinicians), and (5) habit (degree to which taking readings became automatic).

Conclusions: The decline in adherence rates over time is consistent with findings from other studies. However, this study also found adherence to be the highest and most consistent over time in older age groups and progressively lower over time for younger

age groups. These findings can inform the design and implementation of TM interventions that maximize patient adherence, which will enable a more accurate evaluation of impact and optimization of resources.

International Registered Report Identifier (IRRID): RR2-10.2196/resprot.9911

(JMIR Mhealth Uhealth 2019;7(2):e13259) doi:10.2196/13259

KEYWORDS

telemonitoring; mHealth; adherence; heart failure

Introduction

Background

Heart failure (HF) telemonitoring (TM) interventions are designed to transform traditional HF management from one of episodic care (during periods of symptom exacerbation or scheduled follow-up visits) to one of continuous management, extending into patients' daily lives. TM systems enable patients to take home readings (eg, weight, blood pressure, pulse rate, oxygen saturation, and symptoms) [1] which then get transmitted to clinicians at a remote location [2]. The main outputs of this data transfer are threefold. First, the act of taking regular measurements instills in patients a sense of active participation in their care while providing information required to engage in self-care [3,4]. Second, timely data transmission enables clinicians to catch symptom exacerbations early and allow for remote intervention [5]. Finally, even in the periods of patient stability, longitudinal data collected by TM systems provide a more holistic picture of patients' condition which can improve the quality of clinical decisions [6]. According to several meta-analyses, these mechanisms work together to improve quality of life and reduce mortality and health care utilization compared with the standard of care without TM [2,7-10]. However, large and well-designed randomized controlled trials (RCTs) have reported null or mixed results which cannot be ignored [11-14]. We have previously made the case that inconsistencies in the evidence can be explained, in part, by varying fidelity with which interventions are implemented in trials, including the degree to which patients adhere to taking prescribed home readings [15].

Despite the importance of ensuring consistent patient adherence over the course of a TM intervention, there is a dearth in the literature on this topic [16,17] and existing knowledge is difficult to generalize. First, although systematic reviews describe general trends of adherence as starting high in the early months and dropping off over time, there is significant heterogeneity with overall rates between 40% and 90% being reported across studies [3,16]. Second, much of the remote monitoring literature on adherence relates to interactive voice response (IVR)–based interventions with much fewer studies related to newer forms of TM that leverage devices already familiar to patients (eg, mobile phones) [3]. Third, adherence is defined and measured inconsistently across studies with many simply reporting engagement with the technology (eg, taking a single measure) which does not always encompass the full set of patient behaviors needed to optimize the intervention's mechanisms of action [18,19]. Finally, the phenomenon of patient adherence is typically measured in the context of RCTs, limiting the understanding of patient adherence within real-world TM contexts which may be less likely to limit patient use of an intervention to a predefined study period.

Study Objectives

In August 2016, an HF TM program was deployed as part of the standard of care in a specialty heart function clinic in Toronto, Canada. A previously published case study of this program's implementation described and explained clinician adoption as well as the degree of integration within the clinic [20]. This study aimed to describe the patient perspective within the case study. The objectives were to (1) quantify the degree to which patients adhered to taking prescribed home readings and (2) explain longitudinal adherence rates based on the duration of program enrollment, patient characteristics, and patient perceptions of the TM program.

Methods

Study Design

The study used a mixed-methods explanatory sequential design whereby overall and monthly patient adherence rates were first analyzed over a 1-year period. These adherence rates subsequently informed the sampling strategy for semistructured interviews with the objective of explaining overall adherence rates. Additional explanatory data were collected including questionnaires and reasons for leaving the program, all of which were triangulated with the interview findings to explain adherence rates.

The methods used to explain adherence were guided by the unified theory of acceptance and use of technology 2 (UTAUT2) which outlines how 7 constructs influence consumers' intention to use a technology [21]. These constructs, whose definitions have been adapted to facilitate their operationalization within this study, include the following: (1) *performance expectancy* (the degree to which using the TM system is perceived to provide benefits for patients and is analogous to *relative advantage* in the diffusion of innovation literature and *perceived usefulness* in the technology acceptance model), (2) *effort expectancy* (the degree of ease associated with patients using the TM system), (3) *facilitating conditions* (patients' perception that there are resources available to support their use of the TM system), (4) *social influence* (the extent to which others important to the patients [eg, family and friends] support the use of the TM system), (5) *hedonic motivation* (the fun or pleasure derived from using the TM system), (6) *price value* (patients' cognitive trade-off between the benefit of using the TM system and the monetary and time costs to them), and (7) *habit* (the extent to which patients take their required readings automatically because of learning). The UTAUT2 also proposes

that the influence of these 7 constructs on behavioral intention to use technology is modified by age, gender, and experience with using the technology [21].

Key methods for this study have been published in a protocol for a larger quality improvement program evaluation [22] which has been approved by the University Health Network (UHN) Research Ethics Board (16-5789). This approval included the analysis of all data collected as part of the standard of care (ie, TM usage data). Informed consent was obtained by patients who completed questionnaires and interviews.

The Intervention

Telemonitoring Technology

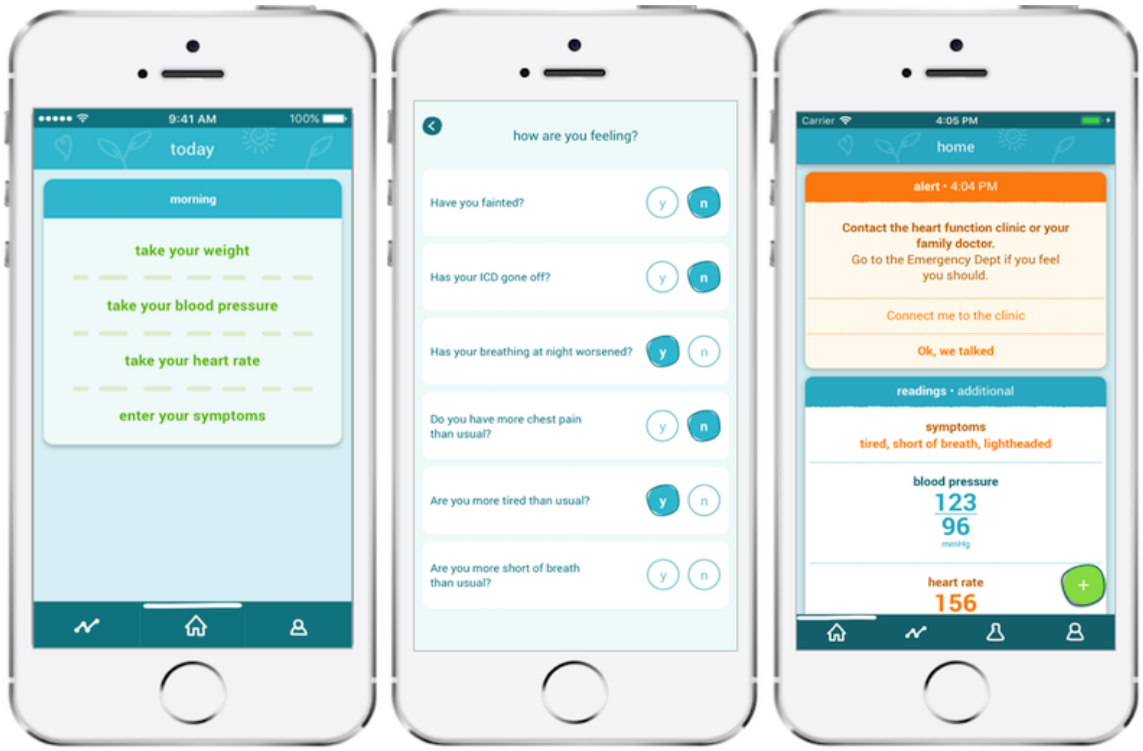
The central component of the *Medly* TM program is the *Medly* smartphone app that patients use to take weight, blood pressure, and heart rate readings as well as record their symptoms using a "yes/no" questionnaire composed of 5 to 11 "yes/no" questions. Patients are instructed to take these 4 readings daily within 30 min of each other before 12 pm. If this is done, the recorded data get processed by a clinically validated algorithm embedded in the app [23], which has been contextualized according to each patient's target thresholds for each of the 4 readings. If the algorithm identifies that key readings are within the acceptable range, the reading values are presented on the screen with a message telling patients that their readings are fine on that day. However, if the algorithm identifies that key readings are out of range or that there is a worrying trend in weight gain, the app generates and displays self-care feedback messages which are highlighted in a different color depending on the determined urgency (Figure 1). Types of self-care feedback messages include the following: informing patients when they are outside their normal range, instructing them to take their prescribed diuretic medication, and suggesting when to contact their care providers or go to the emergency department. Similarly, clinicians also receive alerts when readings are out of range which they can receive via email or view within the clinician-facing *Medly* dashboard.

Other features of the *Medly* app include the ability to view graphical trends of each reading's values and to assist with adherence, an automated phone call to their primary phone line (personal mobile phone or home landline) to remind patients if they have not yet taken morning readings by 10 am. This feature can be disabled at the patient's request. The development of *Medly* features aimed at promoting patient self-care was guided by the Connelly Framework for Self-Care in Chronic Illness [24] (a derivation of the health belief model [25]) using an iterative user-centered design process which included a formal needs assessment [26] and multiple rounds of usability testing. Further description of the *Medly* program can be found elsewhere [20,22,27].

Program Enrollment and Onboarding

As the *Medly* program is offered as part of the standard of care, enrollment is decided jointly between patients and their treating cardiologist during a follow-up appointment or after an inpatient hospital stay. After the treating cardiologist explains the *Medly* program and the patient agrees to participate, they are immediately escorted to a private room where they receive training on how to use the technology. A part of this training includes highlighting the importance of taking daily readings.

Figure 1. Screens of the Medly app showing the incomplete morning card with required readings, the symptoms questionnaire, and personalized self-care feedback after all 4 readings were taken and processed by the algorithm.



Ongoing Monitoring

Throughout their participation in the *Medly* program, patients are expected to complete the 4 daily morning readings and to follow the self-care feedback displayed in the app. If clinical alerts are triggered, a designated clinician at the heart function clinic reviews these alerts as soon as possible. Most clinical alerts result in a call being made to the patient to obtain more information or to provide relevant clinical guidance.

Offboarding

Offboarding refers to the process of ending a patient's participation in the *Medly* program. Unlike many TM interventions, the *Medly* program does not have a predefined end date. Thus, as with most medical interventions, patients remain enrolled for as long as there is a perceived clinical benefit. Patients or clinicians can, at any time, initiate a conversation about the appropriateness of the *Medly* program as part of a patient's treatment plan. Once the joint decision to offboard a patient is made, patients return any equipment they borrowed (mobile phone or peripheral devices) which get recycled and used for future participants.

Adaptations to the Program Since Its Launch

When the program first launched in August 2016, patients were provided with a *Medly* kit which included a smartphone with a data plan and with the *Medly* app already downloaded along with a Bluetooth-enabled weight scale and blood pressure cuff. This enabled data from the peripheral devices to be transmitted directly and automatically to the *Medly* app. Training and ongoing technical support was provided by an analyst from the hospital's telehealth department, and the triage of clinical alerts was done by nurse practitioners on staff at the clinic. Since program launch, 2 key changes have been implemented to enable the sustainability and scalability of the program. First, in January 2018, it became possible for patients with iPhones to download the *Medly* app on their own smartphones and to use their personal weight scales and blood pressure cuffs (the app is now also available for Android users; however, this option was not available at the time of data analysis). Patients without Bluetooth-enabled peripheral devices manually entered readings directly into the app. Second, as of May 2018, a *Medly* coordinator role was created whereby a registered nurse took over the role of triaging clinical alerts in addition to providing frontline technical support. Details and rational for these changes have been published elsewhere [27].

Measuring Patient Adherence

As patients are instructed to take weight, blood pressure, heart rate, and symptom readings every morning (these 4 readings are required for the *Medly* algorithm to generate self-care instructions for patients and alerts for clinicians), adherence was defined as the proportion of days patients took all 4 morning readings over the total number of days they were enrolled up to 1 year. Owing to ongoing enrollment, not all patients had completed 1 year at the time of analysis. Thus, varying durations were accounted for in the proportion denominator. Similarly, monthly adherence rates were calculated as the number of completed morning readings over each 30-day period following the date of their enrollment up to 1 year. Proportions were multiplied by 100 to get monthly adherence rates expressed as a percentage of prescribed completed readings. To further understand patient engagement, we also calculated incomplete adherence rates which were defined as the percentage of days patients took at least 1 reading but not all 4 (such that, although data are transmitted, no clinical alerts or patient feedback is generated). Usage data required to determine adherence rate were collected between August 2016 and October 2018 and extracted from the *Medly* program server.

Explaining Patient Adherence

Quantitative Data and Analyses

Explanatory Variables

Patient demographic variables were collected to characterize the patient population using a questionnaire administered immediately after program enrollment to patients who provided informed consent (n=174). Simple linear regression for full and incomplete adherence over time was performed. In addition, because adherence was collected using repeated measures over a 12-month period, a panel multivariate regression approach was used to determine the impact of time on adherence when controlling for key variables. Preliminary diagnostics included the Hausman and Lagrange multiplier to choose between pooled ordinary least squares, fixed effects, or random effects models and the Breusch-Pagan test to detect the presence of heteroscedasticity [28]. Ultimately, random effects models with cluster-robust standard errors (to adjust for the presence of heteroscedasticity) [29] proved best suited for the dataset. Selected explanatory variables for the multivariate regression included age categorized by decade and sex (both moderating variables in the UTAUT2) [21], and New York Heart Association functional classification (NYHA class), a subjective measure of HF symptom severity based on the hypothesis that sicker patients may benefit more. Data for baseline NYHA class (sometimes documented as a range), age, and sex were extracted from patients' chart in the hospital electronic medical record.

Patient Questionnaire

As part of a larger questionnaire used in an impact evaluation [22], consenting patients responded to questions about their satisfaction with the *Medly* program at 6 and 12 months, offering an opportunity to triangulate these quantitative findings with results from patient interviews (described below). Items in the satisfaction questionnaire could be classified according to the key UTAUT2 constructs of *performance expectancy* (3 items) and *effort expectancy* (4 items); no questionnaire items could be classified within the remaining UTAUT2 constructs and thus were not quantitatively assessed.

Descriptive statistics for the questionnaire responses, adherence rates, and linear regressions were performed using SPSS version 24 (IBM Corporation). Multivariate regression analyses were conducted in RStudio v.1.0.153 (RStudio Inc) using the "plm" package [30]. For all statistical tests, a *P* value of less than .05 was used to indicate statistical significance. Temporal trends were graphically represented using Microsoft Excel (Microsoft Corporation).

Qualitative Data and Analyses

Reasons for Offboarding

Reasons leading to patients being offboarded were recorded in the *Medly* coordinator's records as part of the standard offboarding procedures. These reasons were qualitatively analyzed and classified into themes before being transformed into a count for each category.

Patient Interviews

Semistructured interview guides, which were developed to understand the patients' experience in the *Medly* program, included probes based on the constructs in the UTAUT2. Participants were identified using a purposeful sampling approach to ensure a variety of opinions and to reach information saturation [31]. Variables considered in this sampling approach were age, sex (age and gender are both moderators in the UTAUT2), overall adherence rates, and time since enrollment. The latter involved selecting patients who had been enrolled for different durations, including baseline (to understand initial perceptions without actual experience), approximately 1 month (the intervention was still fresh but patients had been using it long enough to experience benefits and barriers to use), and approximately 6 and 12 months (to align with questionnaire administration and assess patients' perceptions after longer term use). Interviews were recorded and took place in a private room during a scheduled clinic visit or over the telephone.

Interview transcripts were analyzed by 2 independent investigators (PW and MD) using the framework method [32]. This approach involved a first round of largely deductive thematic analysis using an initial coding framework based on the UTAUT2 constructs. PW and MD met to discuss results of the first round and to agree upon subthemes within those constructs. Next, a second round of independent coding was done using the updated coding framework, which was followed by a meeting to discuss contradictory codes and passages. The management of source documents and coding was accomplished with the help of NVivo version 11 (QSR International).

Results

Characteristics of Study Participants

Participants of the *Medly* program were predominantly male (184/232, 79.3%) and had an average age of 57.6 (SD 16.0) years. Other demographics presented in Table 1 are representative of the patient characteristics typically followed in this urban heart function clinic. With regard to HF severity, approximately half experienced relatively mild HF symptoms daily with 48.5% (109/225) having an NYHA class of 2 or less at the time of program enrollment and the average left ventricular ejection fraction of patients was 32.1 (SD 13.2). Most patients included in this analysis (201/231, 87%) used the full kit version of the *Medly* system, 8.2% (19/231) used their personal smartphone but were given peripheral devices by the clinic, and the remaining 4.8% (11/231) used their personal smartphone and either purchased or used their own weight scales and blood pressure cuffs. The option for patients to use their own equipment started approximately 1.5 years after the launch of the program [27].

Overall and Longitudinal Adherence Rates

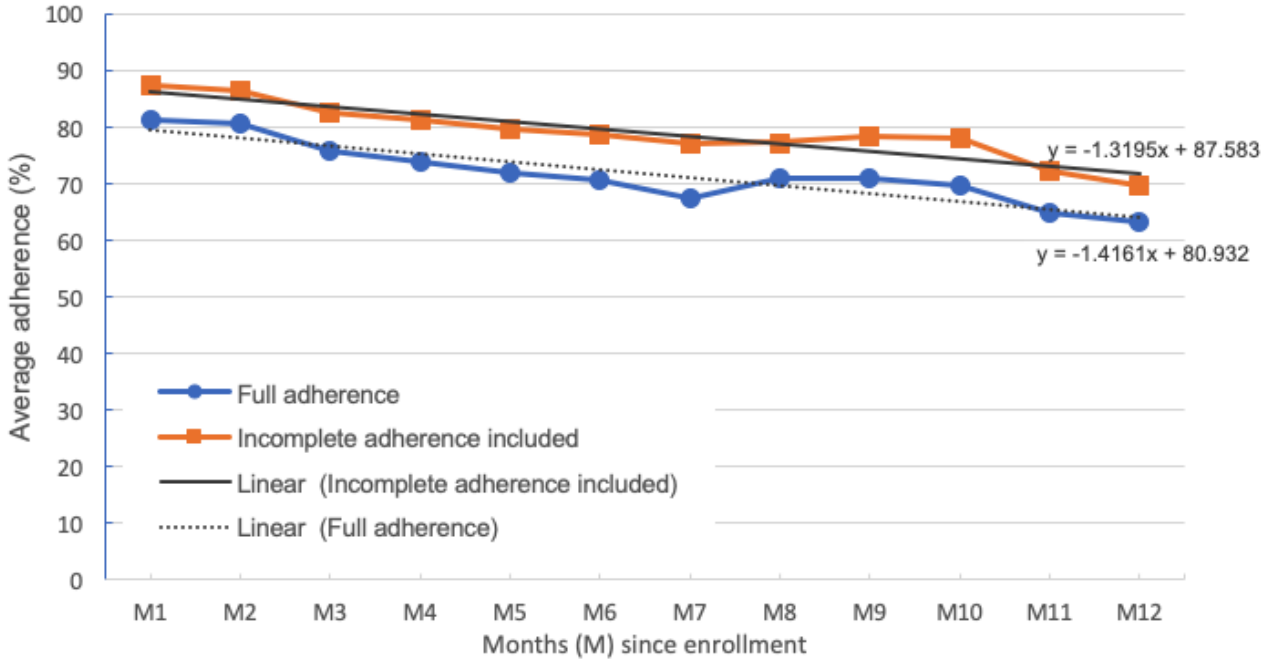
The average overall adherence rate for the 231 patients included in the analysis was 73.6% (SD 25.0), indicating that the average patient completed their prescribed morning readings 5 days per week over the course of their enrollment in the program. When considering days where patients took at least 1 but fewer than all 4 morning readings (ie, including incomplete adherence), the average rate was 80.0% (SD 21.7). Longitudinal examination of monthly adherence rates shows a relatively high average adherence in the first month of 81.2% (SD 23.0) with a gradual decline to 63.1% (SD 37.0) after 12 months of enrollment (see Figure 2). Outputs of the simple linear regression indicates that time is a significant predictor of adherence (beta=−1.42, $P<.001$) with each month since enrollment accounting for a 1.4% decrease in adherence.

Table 1. Characteristics of patients included in the quantitative analysis of overall and longitudinal adherence.

Characteristic	Statistics
Age (years), mean (SD)	57.6 (16.0)
Age (years; categorical), n (%)	
70 or more	60 (25.9)
60-69	56 (24.1)
50-59	50 (21.6)
40-49	34 (14.7)
39 or less	32 (13.8)
Sex, n (%)	
Male	184 (79.3)
Female	48 (20.7)
Ethnicity, n (%)	
White	115 (66.0)
Black	14 (8.0)
Asian	21 (12.1)
Other	24 (13.8)
Rurality, n (%)	
Urban	100 (58.1)
Suburban	49 (28.5)
Rural	23 (13.4)
Place of birth, n (%)	
Canada	85 (48.9)
Elsewhere	89 (51.1)
Highest education achieved, n (%)	
Less than high school	13 (7.5)
High school	34 (19.5)
College or university	127 (73.0)
Income in Can $, n (%)	
<$15,000	26 (15.1)
$15,000-$49,999	57 (33.1)
>$50,000	58 (33.7)
Preferred not to answer	31 (18.0)
Work, n (%)	
Working full time	35 (20.2)
Working part time	17 (9.8)
Retired	87 (50.3)
Unemployed/homemaker	14 (8.1)
Other	20 (11.6)
Supplementary health insurance, n (%)	
Yes	104 (60.8)
No	67 (39.2)

Characteristic	Statistics
New York Heart Association functional classification, n (%)	
2 or less	109 (48.5)
2-3	48 (21.3)
3 or more	68 (30.3)
Left ventricular ejection fraction, mean (SD)	32.1 (13.2)
Have a smartphone, n (%)	
Yes	119 (70.4)
No	50 (29.6)
Comfort with smartphone, n (%)	
Not comfortable	5 (4.0)
Somewhat comfortable	24 (19.2)
Comfortable	47 (37.6)
Very comfortable	49 (39.2)
Equipment used by patients, n (%)	
Full *Medly* kit	201 (87.0)
Patients used personal phone and were provided with peripherals	19 (8.2)
Patients used all personal equipment	11 (4.8)

Figure 2. Average full adherence rates compared with adherence rates which include incomplete adherence over time.



Quantitative Results Explaining Adherence

Multivariate Regression

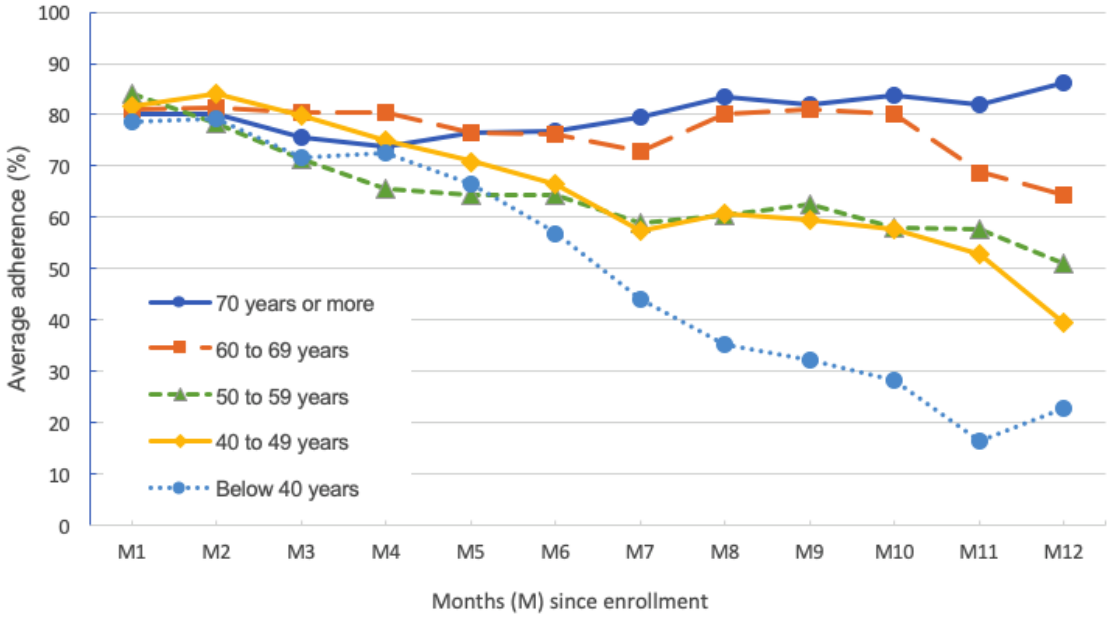
Random effects multivariate regression with cluster-robust standard errors was performed as described in the Methods section. The results, presented in Table 2, confirm a significant effect of time on adherence with each passing month starting after the second month of program enrollment. Patient age was a significant predictor of adherence (P=.04); the positive coefficient indicates that adherence rates were higher with each increasing age category such that older patients maintained higher adherence over time. Figure 3 shows adherence rates over time with regard to the age groups included in the regression model. Disease severity (NYHA class) and sex were not significant predictors of adherence.

Table 2. Random effects multivariate regression with cluster-robust standard errors (SE) showing the effect of time, sex, New York Heart Association (NYHA) class, and age on average adherence.

Variables	Coefficient (beta)	SE	*P* value
Intercept	87.57	4.03	<.001
Month 1	Ref[a]	—[b]	—
Month 2	-1.27	1.98	.52
Month 3	-5.63	2.36	.02
Month 4	-8.21	2.80	.004
Month 5	-9.84	2.85	<.001
Month 6	-12.65	3.05	<.001
Month 7	-15.87	3.33	<.001
Month 8	-12.45	3.32	<.001
Month 9	-13.71	3.63	<.001
Month 10	-15.11	4.20	<.001
Month 11	-19.55	4.84	<.001
Month 12	-20.98	5.13	<.001
Sex	-2.33	5.61	.68
NYHA class	-0.34	2.60	.90
Age	3.49	1.68	.04

[a]Month 1 is the reference category to which all other levels of the time variable (months 2 to 12) are compared in the multivariate regression model.
[b]Not applicable.

Figure 3. Average adherence rates over time by age group showing higher adherence over time for older age groups.



Patient Questionnaire

Results from the patient questionnaires show that a clear majority of patients perceived value in using the *Medly* system after 6 months with 90.6% (87/96) agreeing or strongly agreeing with the statement that the TM system is important for managing their HF and 87.4% (83/95) agreeing with the statement that it would be useful for them to continue using the system (Table 3). The percentage of patients who agree with these same statements increased to 95.8% (46/48) and 93.9% (46/49) after 12 months, respectively. Responses related to *effort expectancy* at 6 months similarly show a high level of agreement with 92.7% (89/96) of patients agreeing with the statements that the TM system was easy to use and to learn how to use it. Perceptions of ease of use remained consistent with 89.4% (42/47) and 91.8% (45/49) agreeing with these same statements at 12 months.

Table 3. Patient perceptions of the benefits and effort of using the *Medly* TM system at 6 and 12 months postenrollment.

Item in questionnaire	Agree or Strongly agree, n (%)	
	6 months	12 months
Performance expectancy		
The monitoring system is important for managing my heart failure	87 (90.6)	46 (95.8)
I think using the monitoring system improved my health	65 (70.7)	36 (75.0)
It would be useful for me to keep using the monitoring system	83 (87.4)	46 (93.9)
Effort expectancy		
Learning to operate the monitoring system was easy for me	89 (92.7)	45 (91.8)
I found the monitoring system to be easy to use	85 (92.4)	42 (89.4)
Taking my blood pressure at home was easy	93 (96.9)	47 (95.9)
Taking my weight was easy	93 (96.9)	47 (95.9)

Table 4. Classification of reasons for patient offboarding.

Reason for offboarding	Statistics, n (%)
Clinician–initiated offboarding	
Received heart transplant or surgical repair of the heart	14 (22.9)
Switched to more invasive form of remote monitoring (eg, CardioMEMS)	5 (8.1)
Patient recovered ventricular function	4 (6.5)
Significant change in health status (eg, shift to palliative care)	6 (9.8)
Patient was not compliant with taking readings or with following clinician instructions	3 (4.9)
Patient–initiated offboarding	
Not interested in participating or a belief that the benefits are not worth the effort	5 (8.1)
Stress caused by taking daily readings	4 (6.5)
Life circumstances (eg, shift work and sick relatives)	2 (3.2)
Poor eyesight	1 (1.6)
Other (eg, unknown, moved provinces)	5 (8.1)
Mortality	12 (19.6)

Qualitative Results Explaining Adherence

Reasons for Offboarding

Of the 61 patients who left the *Medly* program during the study period, 52% (32/61) were offboarded because a change in their HF condition made it such that the *Medly* program would no longer be a beneficial part of their care plan. A total of 3 patients were offboarded because they were not adhering to taking measures or following clinician instructions (Table 4). A further 28% (17/61) of patients chose to leave the program because of a lack of interest or a feeling that the benefits of enrollment were not worth the effort, that daily monitoring was causing stress, and for other unknown reasons. Finally, 20% (12/61) of the offboarding were because of patient death. These deaths were attributed to the severity and natural progression of HF.

Interview Findings

Interview Participant Characteristics

The interviewed participants (n=24) largely matched the distribution of age and sex of the larger patient sample as shown in Table 5. The patients interviewed had overall adherence rates ranging between 22.2% and 98.6% and were interviewed at various times since program enrollment. This included 17% of patients (4/24) being interviewed the day they were onboarded and 2 patients who agreed to participate after deciding they wanted to leave the program.

Table 5. Participant characteristics for semistructured interviews.

Participant ID	Sex	Age at enrollment (years)	Time of interview since enrollment (month)	Average adherence rate (%)
HFpro009	M	76	12	92.2
HFpro011	M	72	12	80.0
HFpro018	M	60	6[a]	90.0
HFpro019	M	46	3	30.3
HFpro027	M	59	6	82.3
HFpro028	M	67	6	93.1
HFpro037	F	62	0	90.3
HFpro038	M	63	6	97.5
HFpro048	M	44	1[a]	96.7
HFpro052	M	83	6	70.3
HFpro059	M	76	6	54.2
HFpro060	F	81	6	96.3
HFpro061	M	62	6	45.0
HFpro064	M	45	6	67.9
HFpro089	M	57	9	72.2
HFpro091	F	61	12	94.4
HFpro107	M	54	6	62.8
HFpro109	M	41	1	68.3
HFpro129	M	22	1	22.2
HFpro131	F	71	0	87.0
HFpro154	M	50	1	96.1
HFpro157	F	45	0	98.6
HFpro158	F	52	0	82.4
HFpro168	F	65	6	90.6

[a]Interview conducted after offboarding.

Interview Themes

Interview themes were classified according to UTAUT2 constructs of *performance expectancy*,*effort expectancy*, *facilitating conditions*,*social influence*, and *habit*; no statements related to *hedonic motivation* or *price value* were identified by the coders. No overarching patterns emerged in the themes based on patients' age, sex, or time since enrollment. Therefore, the themes and representative quotes discussed in the following sections predominantly help distinguish between high and low adherers.

Performance Expectancy

This theme refers to the perceived benefits, both expected and experienced, of being part of the *Medly* program. Subthemes included (1) self-management support, (2) peace of mind, (3) relationship with care team, and (4) lack of context.

Self-Management Support

The most commonly mentioned benefit of the *Medly* program is that it supports patients in their ability to self-manage their HF. Participants discuss how the system, by enabling them to take daily readings, keeps them accountable and provides guidance in their self-care tasks:

> *I rarely ever took my weight which was a big issue getting admitted into the hospital because I retain so much water. So yeah, it's been helpful for monitoring things that I normally wouldn't…It keeps me on track and lets me know if I need to take medication that I normally don't have to take it...I had the expectation that it would be helpful for me because it keeps me on a routine and it's lived up to those expectations.* [HFpro154]

Underpinning this self-management support is the immediateness of the patient self-care feedback which allows patients to plan for their day around the results of readings they have just taken:

> *I can start my day off with knowing that I've got to be extra careful…I'm going to plan my day from what Medley is telling me. That's how it helps me every morning, I know what to do and what not to do for the rest of the day.* [HFpro089]

Peace of Mind

The automated self-care feedback works alongside clinician monitoring to provide many patients with peace of mind. For some, this peace of mind brings a heightened sense of confidence when they are trying to decide if their symptoms are bad enough to warrant a trip to the hospital.

> *I'm very diligent...I basically rely on it. I just love the peace of mind that it represents. When you're as sick as I was, it's good to have a big brother or big sister out there.* [HFpro107]

> *It gives me comfort [knowing that] somebody's watching over me, I don't have to go to the hospital all the time.* [HFpro052]

Relationship With Care Team

Patients who had been in the program for longer periods said that it improved their relationship with clinical members of their care team:

> *At first it didn't bother me [that I didn't have a lot of interaction] but now actually you gain trust, you know, a relationship with the person on the other end...When [the Medly nurse] calls sometimes, we're talking for 20 minutes and she's really getting a full history of things because you just can't get a full history on a minute conversation.* [HFpro089]

However, for some, this closer relationship helps explain lower adherence in patients who did not like the idea of clinicians being able to see transgressions of daily life:

> *There is a feeling, and it sort of upsets me or disturbs me that they know [everything I do]. If I want to go out on a binge and watch a soccer game or a hockey game and eat lots …they're going to know because my weights going to go up and so there's a fear of, "Oh God, I'm going to be told off." Will it stop me [from] doing that? No, what will stop me doing it is the fact that it's bad for my health… I forget to take my weight [laughs] because there's a feeling that they're looking.* [HFpro061]

Lack of Context

Patients with lower overall adherence rates expressed the opinion that the readings, particularly the symptoms, do not accurately capture the full context of their health status. Consequently, some felt that the self-care feedback messages did not always reflect how they were feeling and they eventually learned that they should not immediately act on some alerts:

> *[The feedback messages] are a little bit too alarming... because sometimes they say "have somebody take you immediately to emergency," and usually it turns out it's okay, so I got used to that.* [HFpro009]

Effort Expectancy

Patient uptake and adherence can also be explained by the perceived efforts involved in using the system. Subthemes included (1) usability and (2) technical difficulties.

Usability

The qualitative findings mirror those from the questionnaire insofar as most patients found the *Medly* system easy to learn and use. Furthermore, some patients were frequent travelers and described how the portability of the system allowed them to continue taking their readings wherever they were:

> *I'm of the age where I'm not as computer literate with cellphones... But it was fine, it's easy. If I can learn it, it's pretty easy to learn.* [HFpro027]

> *It's been all over Canada with me… we just throw it in the car. I even took my weight at Tim Horton's first thing in the morning and that was in Edmonton. We just left the hotel and it was in the car and…before we left [on] the road I said I didn't take my weight. And my wife went out and got it, hooked it up to the Wi-Fi at the Tim Hortons and bam, [I] took it right there.* [HFpro089]

A minority of patients described some difficulty using the peripheral devices and the cognitive effort in trying to decide how to accurately answer the "yes/no" symptom questions:

> *The equipment is highly sensitive. For me I have dizziness constantly because of my medications and my low blood pressure and my heart condition. So, if I sway or move on the scale, the scale has different readings…The scale is very narrow... and I'm a very wide guy and I need to have my feet spread apart in order to be stable on the scale... The same with my arm, if I move my arm a little bit or anything, (it) will make the blood pressure monitor go into error mode and it's frustrating* [HFpro064]

> *With the [symptoms] questionnaire sometimes I'm sort of on the edge because it says "Are they worse?" Well no, they're not any worse but sometimes I am like a little bit short of breath.* [HFpro091]

Technical Issues

The quality of the system was perceived as high across all patients and time points. However, several patients, from various adherence levels, recounted experiencing technical issues, particularly related to Bluetooth connectivity between the peripheral devices and the smartphone.

> *There are times when I'm not impressed, because I weigh myself but it doesn't record. And, then I get this call saying that I didn't do it and it throws me off. But, generally, it's okay. I've had, I think about three times where it's misfired sort to speak...I think, "what did I do wrong?"* [HFpro060]

Facilitating Conditions

Facilitating conditions are the resources and support available to facilitate the use of *Medly*. Subthemes included (1) technical support, (2) automated adherence calls, and (3) informal caregivers.

Technical Support

The technical issues described did not seem to have severely impacted adherence rates because of the easy access to technical

support services. In addition, the presence of comprehensive onboarding process helped with initial uptake:

> *I can't see [a reason to not use Medly]. I mean, as far as running into technical difficulties, you've given me all the numbers, there's people I can contact so I don't foresee there being that big of a problem that I wouldn't be able to work through it.* [HFpro027]

> *[The training was] a piece of cake...it was private, we were in a closed room, the information was face to face, the equipment was right there, it had hands-on... it was great.* [HFpro018]

Automated Adherence Calls

The automated call that is sent to patients if they have not taken their readings before 10 am, although sometimes described as annoying, was a facilitating condition expressed by many:

> *[Taking my readings] is what I do first thing in the morning before I get the phone call with the annoying ringing… I do [appreciate the call but]...I'm also a single dad of two kids so any opportunity that I possibly can to rest and sleep I take it... A text would be better than a phone call.* [HFpro064]

Informal Caregivers

Although the technology is designed to support the HF patient experience, some patients receive support from family members who help with all aspects of HF management, including reminding them to take their *Medly* readings:

> *I have a built-in monitor at home [laugh]...which is very, very beneficial because if [my wife] wasn't there at all, you know, I'd probably be worse than I am, as far as habits are concerned. So, having that extra person, she polices me pretty good.* [HFpro028]

Social Influence

Social influence is the degree to which individuals in patients' lives were supportive of the use of TM. Patient responses revealed that their family and friends overwhelmingly support their participation in the *Medly* program:

> *Everybody knows [I use Medly]. I write about it all the time...everybody is very envious that I'm on this type of program and envious that I have doctors that care about me this much.* [HFpro064]

Some patients had family members who raised privacy concerns, but this does not seem to have impacted the willingness to participate:

> *Well, [my friends and family] think it's great. There's a few that think it's kind of Big Brother [saying] "wow, they know a lot of information on you." But their fear of Big Brother is kind of secondary to my doctor need[ing] to know what's going on.* [HFpro019]

Finally, the fact that the *Medly* program was endorsed by a trusted clinician also appears to have been a motivating factor for patients:

> *[My cardiologist] is always very supportive saying, "[Medly]'s really doing a good job for you keeping you out of emergency."* [HFpro009]

Habit

Evidence of the formation of a habit was more prominent in patients with higher adherence rates with many describing how taking measures eventually became a part of their delay routine. Once a habit was established, events or conditions breaking that routine explained why some readings were missed:

> *It's part of a habit now. I don't forget it. ...[It's] automatic.* [HFpro038]

> *[When I forget] I think I can smell my wife making the coffee... The sense of smell is very strong …and it just beats the other senses out of my head that say "go weigh yourself first." The smell of fresh coffee coming down the hallway, honestly that's the only time I would miss it.* [HFpro089]

Although the formation of a habit helps, it is not essential to ensure high adherence. One patient described the hassle of taking daily readings yet still maintained a high level of adherence (90.0%) throughout their enrollment. In this example, factors such as guilt and recognizing the importance of a behavior may be enough to motivate daily readings even if this behavior did not become automatic:

> *I don't like doing it every day, it's a bit of drag, because sometimes I want to sleep in and I feel kind of guilty because I haven't got it done...I'm just sort of getting old and lazy and don't want to do anything but that's part of my regimen, I never miss.* [HFpro018]

Discussion

Principal Findings

This study has presented the findings of a mixed-methods study seeking to describe and explain patient adherence rates to taking daily prescribed home readings over a 1-year period in the context of a mobile phone–based HF TM program offered as part of the standard of care. Results found an average overall adherence rate of 73.6% and an average 1.4% drop in adherence with each passing month. The random effects model, which enabled repeated measures of monthly adherence (effect of time) to be included in the same regression as other demographic variables, found a significant effect of age on monthly adherence rates. Specifically, adherence rates were highest (and more consistent over time) for the older age group (70 years or more) and were progressively lower for each younger decade.

Additional methods employed could not fully explain the temporal decline in adherence, but they did provide evidence that patients' perceptions of the program and other contextual factors contribute to explaining higher and lower adherence rates. Factors explaining patients' motivation to adhere include the following: (1) perceived benefits of the program (self-management support, peace of mind, and improvement in clinical care), (2) ease of use, (3) a positive opinion of the program from family and friends, (4) supporting services

(training and technical support), and (5) the ability to form a habit. Themes explaining low and imperfect adherence included the following: (1) technical issues, (2) life circumstances that interfered with a formed habit, and (3) a perception that the benefits of the program were suboptimal because of the system's inability to adequately capture and communicate the full context of patients' health state. These explanatory findings fit within the constructs of the UTAUT2 of *performance expectancy*, *effort expectancy*, *facilitation conditions*, *social influence*, and *habit*.

There were no findings related to the UTAUT2 constructs of *hedonic motivation* and *price value*; however, this is likely because of the context in which this study was conducted. First, although patients expressed numerous benefits, it remains that the use of TM systems occurs in the context of disease management and therefore is unlikely to be described as fun or enjoyable. Second, Canada has a public payer health system which means that patients did not have to pay out-of-pocket to use the technology. In addition, those who used their own smartphones and peripheral devices either already had that equipment or were assessed for their ability to pay or cover the costs through supplementation health insurance. Thus, patients were not put in a position of having to weigh the supplemental personal costs and benefits of being part of the *Medly* program.

Finally, although the principal aim of this study was to explain adherence using a definition based on the prescribed patient behavior needed to optimize program benefits, the finding that the incomplete adherence rate was 6.4% higher than full adherence should not be discounted. A certain percentage of these incomplete morning readings are likely due to the Bluetooth connectivity issues expressed by patients which would have prevented taking a weight or blood pressure reading until the issue could be fixed. Other possible explanations may include patients not recognizing or remembering the importance of taking the full set of readings. Alternatively, patients may make the decision to take measures that are most relevant based on how they are feeling (because of a high sense of self-efficacy for self-management) and may not necessarily lead to poorer outcomes. These hypotheses cannot be confirmed by the explanatory data generated in this study and should be empirically tested in future studies. The impact of adherence rates and health outcomes will be explored in a subgroup analysis of the upcoming impact evaluation of the *Medly* program [22].

Comparison With Previous Work

Measuring Adherence

The findings from this study are in line with the literature review by Maeder et al, which found that adherence rates in home–based telehealth projects ranged from 40% to 90% and tended to be higher in earlier months before dropping off over time [16]. A recent and similar study looked at adherence to taking vital signs in TM interventions addressing various conditions and found an average adherence rate of 64.1% to scheduled daily readings. However, this study also found a trend toward increasing adherence after a steep initial drop off which is difficult to interpret alongside our results. The authors did not fully explain this initial drop off but hypothesized that patients may be encouraged to adhere only after longitudinal values could be generated and that they have had enough time to experience the value of the intervention [33]. In another study, adherence to completing IVR calls was 90% in HF patients, [34] but calls were only scheduled once per week, again, limiting comparison to a regimen that asks patients to take readings daily.

Explaining Adherence

The finding that older patients maintained a high level of adherence is seen in another remote monitoring study [34] but without explanation. The UTAUT2 proposes that the moderating impact of age is such that the effect of *effort expectancy* and *facilitating conditions* is strongest in older people [21]. In addition, it has been found that after a habit has been formed through repeated use, it becomes more difficult for changes in one's external environment to override that behavior in older compared with younger people [21]. In other words, the ease of use of the *Medly* system and the availability of supporting services likely led to higher use in older patients which would have led to the formation of a habit. A habit which, once formed, would be more difficult to disrupt compared with younger patients. Although this may explain some of the differences between age groups, it is also possible that younger patients experience more potential distractors (work and dependent children) than older patients. Further research is required to understand the effect of age when it comes to adherence to TM interventions.

Factors explaining higher and lower adherence in this study were similar to the barriers and facilitators to TM, mobile health, and telehealth use described in the literature. With regard to *performance expectancy*, the literature cites similar perceived benefits including the degree to which the intervention improves health management (including both self- and clinician-directed management), peace of mind, and enhanced relationship between patients and clinicians [3,4,19,35-41].

Often-cited factors related to *effort expectancy* include user-friendliness of the equipment, technical barriers, health literacy or language barriers, and limited answer options [3,4,19,36-40,42]. The latter was seen in this study with patients who struggled with the "yes/no" format of symptoms.

In terms of *facilitating conditions*, studies support the availability of technical support services and features to help with remembering as important factors in technology use [3,19,37,38].

Similar factors related to *social influence* are discussed in a systematic review by O'Connor, which cites the lack of clinical endorsement as a barrier to patient uptake [40]. A survey study found that the construct of social influence contributed to explaining patients' intention to use electronic health systems beyond what could already be explained by performance and effort expectancy [36]. In this study, we found overwhelming support for the use of *Medly* by family, friends, and the patient's treating cardiologist. Thus, although social influence was likely not a strong enough factor in isolation, it probably contributed to higher adherence in the initial stages of enrollment.

Barriers such as failure of daily readings to become automatic and integrated into patients' everyday tasks are cited in the literature [4,40] and were categorized in the construct of *habit* in our study. If an automatic habit is not created, the added energy of taking daily readings likely contributes to user fatigue over time.

This discussion is intended to highlight the prevalence of UTAUT2 themes in the literature. However, it is important to recognize that each TM intervention is different and may yield different experiences for patients. For example, a study by Fairbrother et al concluded that, although patients experienced peace of mind, the TM intervention did not increase the sense of ownership over their condition [43]. This contrasts with many of the patients in our study who described *Medly* as facilitating self-management. This is likely explained by the automated self-care feedback messages not part of the TM system in the Fairbrother study. This is an example of how different patient experiences can explain some of the heterogeneity of results across adherence studies. This study focused on patient perceptions because it is individuals' experiences that are most likely to influence the degree to which they adhere to a TM program. The evaluation of the program's outcomes, including quantitative measures for quality of life and self-care, is outside of the scope of this study and will be discussed in an upcoming publication [22].

Limitations

Several limitations related to this study's pragmatic design should be considered when interpreting the results. First, unlike TM interventions in RCTs, the *Medly* program was adapted over the 2-year period in which data collection occurred. As described elsewhere [27], these changes were made such that the essence of the intervention was maintained but it remains that some patients may have had different experiences.

Second, because enrollment in the intervention was not contingent on patients being part of a study, we were conscious of not overburdening patients with interviews at multiple time points. This decision meant that we did not collect qualitative data of individuals as they progressed through the program.

Third, reasons for offboarding were limited to the administrative data collected by clinicians and not all patients in the program had consented to being approached for an interview.

Fourth, the lack of strict inclusion criteria means that, by experimental standards, selection bias likely occurred such that enrolled patients were more likely to be engaged and not face language barriers. Furthermore, limiting the generalizability of findings are the demographics of the sample and the overrepresentation of male patients in the *Medly* program. This is consistent with other studies that find an under-representation of women enrolled in heart function clinics and clinical research despite a similar prevalence of HF among both sexes [44,45]. Although exploring the reasons for why fewer women have access to HF management interventions is outside the scope of this study, there is clearly a need for research into sex- or gender–based differences, including as it relates to uptake, use, and adherence to TM interventions. For instance, it is possible that the reason sex was not correlated with adherence rates in the regression analysis is because of a relatively small number of female participants in this study.

Fifth, the sample size of available data got progressively smaller with each passing month, and although this was accounted for in the regression analyses, it is likely that patients with strong negative opinions of the program left the program before the 12-month point (and thus did not complete the questionnaires).

Sixth, a previously published protocol [22] describes the use of quantitative methods for measuring patient adherence and interviews guided by the UTAUT2 to explain these adherence rates. Other data were collected as part of this pragmatic evaluation (ie, patient satisfaction questionnaire and reasons for offboarding) and were reported because they offer an opportunity to further explain patient adherence through triangulation with patient interviews. However, because the satisfaction questionnaire was not initially developed to explain adherence, it only contained items related to 2 of the 7 UTAUT2 constructs. Researchers conducting questionnaire–based work related to the UTAUT2 should consider using tools which include the validated items for that framework [21].

Finally, we did not have data allowing us to account for periods when patients were unable to take readings for legitimate reasons (eg, traveling, admitted in the hospital, and system down time), which would ideally be accounted for when measuring adherence. This limitation likely underestimates true adherence rates in the *Medly* program.

Recommendations

On the basis of the findings from this study, we agree with recommendations from other studies that patients should receive comprehensive training and may benefit from refresher sessions aimed at reminding them of the proper use of the TM system, the benefits of the TM intervention, and the process for obtaining technical support when needed [6,16]. We also advocate for the involvement of supportive family members in those discussions and as part of the onboarding process. In addition, because many HF patients receive support from informal caregivers, further research into how best to incorporate that role within the design of TM systems would be beneficial.

Reminders (such as adherence calls) were found to be important in this study. Therefore, developers of TM systems should offer a range of options (eg, phone call, text, and app notification) that users can choose from based on their preferences such that these reminders do not become so disruptive that they opt to disable the feature.

Finally, study results offer important insights into how the user-centered design of TM systems should be conducted. Although there is value for scenario–based usability testing in laboratory environments, new TM systems should also be piloted in the real world with users of all age groups before full deployment. This is needed to allow TM designers to understand how patients use (or do not use) the system in the context of their existing habits and personal lives. In addition, although self-care messages can be simulated in a single usability testing session, it is preferable to give patients the opportunity to use a TM system over a period of time to evaluate the accuracy and

appropriateness of self-care messages in response to fluctuations in their health state.

Conclusions

This study has presented the results of a mixed-methods study aimed at explaining longitudinal adherence rates of patients enrolled in a mobile phone–based HF TM program. The study found that, on average, patients took weight, blood pressure, heart rate, and symptom readings for 73.6% of the days they were enrolled in the program. Results also showed a consistent decline in adherence over the 12 months which is further influenced by patients' age such that patients in older age groups maintained higher and more consistent adherence rates throughout the study period whereas the declining rate of adherence became progressively more pronounced for younger age groups. Finally, interview findings indicated that the perceived benefits of the program, ease of use, social and technical support, and ability to form a habit around taking daily readings further explained levels of adherence. These findings can inform the design of TM interventions that maximize patient adherence. When implemented in the context of effectiveness trials, interventions with high fidelity of use will enable a more accurate evaluation of impact, and when implemented as part of the standard of care, they will ensure the optimization of resources and satisfaction among patient and clinician users.

Acknowledgments

The authors wish to thank the patients who participated in this study.

Authors' Contributions

PW led the overall design, data collection, data analysis, and write-up of this study. HR, AL, JC, and ES contributed to the design. MD contributed to the analysis and interpretation of the qualitative data, and CB contributed to the analysis of the quantitative data. All authors reviewed and edited the manuscript. All authors read and approved the final version of the manuscript. This study was conducted as part of the PhD dissertation work of PW who receives a stipend from the Institute of Health Policy, Management, and Evaluation at the University of Toronto.

Conflicts of Interest

HR, JC, and ES are considered inventors of the *Medly* system under the intellectual property policies of the UHN and may benefit from future commercialization of the technology by UHN.

References

1. Anker SD, Koehler F, Abraham WT. Telemedicine and remote management of patients with heart failure. Lancet 2011 Aug 20;378(9792):731-739. [doi: 10.1016/S0140-6736(11)61229-4] [Medline: 21856487]
2. Kitsiou S, Paré G, Jaana M. Effects of home telemonitoring interventions on patients with chronic heart failure: an overview of systematic reviews. J Med Internet Res 2015;17(3):e63 [FREE Full text] [doi: 10.2196/jmir.4174] [Medline: 25768664]
3. Gorst SL, Armitage CJ, Brownsell S, Hawley MS. Home telehealth uptake and continued use among heart failure and chronic obstructive pulmonary disease patients: a systematic review. Ann Behav Med 2014 Dec;48(3):323-336 [FREE Full text] [doi: 10.1007/s12160-014-9607-x] [Medline: 24763972]
4. Vassilev I, Rowsell A, Pope C, Kennedy A, O'Cathain A, Salisbury C, et al. Assessing the implementability of telehealth interventions for self-management support: a realist review. Implement Sci 2015;10:59 [FREE Full text] [doi: 10.1186/s13012-015-0238-9] [Medline: 25906822]
5. Desai AS, Stevenson LW. Connecting the circle from home to heart-failure disease management. N Engl J Med 2010 Dec 9;363(24):2364-2367. [doi: 10.1056/NEJMe1011769] [Medline: 21080836]
6. Hanley J, Ure J, Pagliari C, Sheikh A, McKinstry B. Experiences of patients and professionals participating in the HITS home blood pressure telemonitoring trial: a qualitative study. BMJ Open 2013;3(5) [FREE Full text] [doi: 10.1136/bmjopen-2013-002671] [Medline: 23793649]
7. Yun JE, Park J, Park H, Lee H, Park D. Comparative effectiveness of telemonitoring versus usual care for heart failure: a systematic review and meta-analysis. J Card Fail 2018 Jan;24(1):19-28. [doi: 10.1016/j.cardfail.2017.09.006] [Medline: 28939459]
8. Lin M, Yuan W, Huang T, Zhang H, Mai J, Wang J. Clinical effectiveness of telemedicine for chronic heart failure: a systematic review and meta-analysis. J Investig Med 2017 Mar 22;65(5):899-911. [doi: 10.1136/jim-2016-000199] [Medline: 28330835]
9. Nakamura N, Koga T, Iseki H. A meta-analysis of remote patient monitoring for chronic heart failure patients. J Telemed Telecare 2014 Jan;20(1):11-17. [doi: 10.1177/1357633X13517352] [Medline: 24352899]
10. Xiang R, Li L, Liu SX. Meta-analysis and meta-regression of telehealth programmes for patients with chronic heart failure. J Telemed Telecare 2013 Jul;19(5):249-259. [doi: 10.1177/1357633X13495490] [Medline: 24163234]
11. Chaudhry SI, Mattera JA, Curtis JP, Spertus JA, Herrin J, Lin Z, et al. Telemonitoring in patients with heart failure. N Engl J Med 2010 Dec 9;363(24):2301-2309 [FREE Full text] [doi: 10.1056/NEJMoa1010029] [Medline: 21080835]

12. Koehler F, Koehler K, Deckwart O, Prescher S, Wegscheider K, Kirwan B, et al. Efficacy of telemedical interventional management in patients with heart failure (TIM-HF2): a randomised, controlled, parallel-group, unmasked trial. Lancet 2018 Sep;392(10152):1047-1057 [FREE Full text] [doi: 10.1016/S0140-6736(18)31880-4]
13. Koehler F, Winkler S, Schieber M, Sechtem U, Stangl K, Böhm M, Telemedical Interventional Monitoring in Heart Failure Investigators. Impact of remote telemedical management on mortality and hospitalizations in ambulatory patients with chronic heart failure: the telemedical interventional monitoring in heart failure study. Circulation 2011 May 3;123(17):1873-1880 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.111.018473] [Medline: 21444883]
14. Ong MK, Romano PS, Edgington S, Aronow HU, Auerbach AD, Black JT, Better Effectiveness After Transition–Heart Failure (BEAT-HF) Research Group. Effectiveness of remote patient monitoring after discharge of hospitalized patients with heart failure: the Better Effectiveness After Transition -- Heart Failure (BEAT-HF) randomized clinical trial. JAMA Intern Med 2016 Mar;176(3):310-318 [FREE Full text] [doi: 10.1001/jamainternmed.2015.7712] [Medline: 26857383]
15. Ware P, Seto E, Ross HJ. Accounting for complexity in home telemonitoring: a need for context-centred evidence. Can J Cardiol 2018 Jul;34(7):897-904. [doi: 10.1016/j.cjca.2018.01.022] [Medline: 29861204]
16. Maeder A, Poultney N, Morgan G, Lippiatt R. Patient compliance in home-based self-care telehealth projects. J Telemed Telecare 2015 Dec;21(8):439-442. [doi: 10.1177/1357633X15612382] [Medline: 26556057]
17. Or CK, Karsh B. A systematic review of patient acceptance of consumer health information technology. J Am Med Inform Assoc 2009 Aug;16(4):550-560 [FREE Full text] [doi: 10.1197/jamia.M2888] [Medline: 19390112]
18. Sieverink F, Kelders SM, van Gemert-Pijnen JE. Clarifying the concept of adherence to eHealth technology: systematic review on when usage becomes adherence. J Med Internet Res 2017 Dec 06;19(12):e402 [FREE Full text] [doi: 10.2196/jmir.8578] [Medline: 29212630]
19. Simblett S, Greer B, Matcham F, Curtis H, Polhemus A, Ferrão J, et al. Barriers to and facilitators of engagement with remote measurement technology for managing health: systematic review and content analysis of findings. J Med Internet Res 2018 Jul 12;20(7):e10480 [FREE Full text] [doi: 10.2196/10480] [Medline: 30001997]
20. Ware P, Ross HJ, Cafazzo JA, Laporte A, Gordon K, Seto E. Evaluating the implementation of a mobile phone-based telemonitoring program: longitudinal study guided by the consolidated framework for implementation research. JMIR Mhealth Uhealth 2018 Jul 31;6(7):e10768 [FREE Full text] [doi: 10.2196/10768] [Medline: 30064970]
21. Venkatesh V, Thong JY, Xu X. Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. MIS Q 2012;36(1):157. [doi: 10.2307/41410412]
22. Ware P, Ross HJ, Cafazzo JA, Laporte A, Seto E. Implementation and evaluation of a smartphone-based telemonitoring program for patients with heart failure: mixed-methods study protocol. JMIR Res Protoc 2018 May 03;7(5):e121 [FREE Full text] [doi: 10.2196/resprot.9911] [Medline: 29724704]
23. Seto E, Leonard KJ, Cafazzo JA, Barnsley J, Masino C, Ross HJ. Developing healthcare rule-based expert systems: case study of a heart failure telemonitoring system. Int J Med Inform 2012 Aug;81(8):556-565. [doi: 10.1016/j.ijmedinf.2012.03.001] [Medline: 22465288]
24. Connelly CE. An empirical study of a model of self-care in chronic illness. Clin Nurse Spec 1993 Sep;7(5):247-253. [Medline: 8221532]
25. Janz NK, Becker MH. The Health Belief Model: a decade later. Health Educ Q 1984;11(1):1-47. [doi: 10.1177/109019818401100101] [Medline: 6392204]
26. Seto E, Leonard KJ, Masino C, Cafazzo JA, Barnsley J, Ross HJ. Attitudes of heart failure patients and health care providers towards mobile phone-based remote monitoring. J Med Internet Res 2010 Nov 29;12(4):e55 [FREE Full text] [doi: 10.2196/jmir.1627] [Medline: 21115435]
27. Ware P, Ross H, Cafazzo J, Laporte A, Gordon K, Seto E. User-centered adaptation of an existing heart failure telemonitoring program to ensure sustainability and scalability: qualitative study. JMIR Cardio 2018 Dec 06;2(2):e11466. [doi: 10.2196/11466]
28. Hill R, Griffiths W, Lim G. Principles of econometrics. Hoboken, NJ: Wiley; 2008.
29. Stock JH, Watson MW. Heteroskedasticity-robust standard errors for fixed effects panel data regression. Econometrica 2008 Jan;76(1):155-174. [doi: 10.1111/j.0012-9682.2008.00821.x]
30. Croissant Y, Millo G. Panel data econometrics in R: the plm package. J Stat Soft 2008;27(2). [doi: 10.18637/jss.v027.i02]
31. Francis JJ, Johnston M, Robertson C, Glidewell L, Entwistle V, Eccles MP, et al. What is an adequate sample size? Operationalising data saturation for theory-based interview studies. Psychol Health 2010 Dec;25(10):1229-1245. [doi: 10.1080/08870440903194015] [Medline: 20204937]
32. Gale NK, Heath G, Cameron E, Rashid S, Redwood S. Using the framework method for the analysis of qualitative data in multi-disciplinary health research. BMC Med Res Methodol 2013;13:117 [FREE Full text] [doi: 10.1186/1471-2288-13-117] [Medline: 24047204]
33. Celler B, Argha A, Varnfield M, Jayasena R. Patient adherence to scheduled vital sign measurements during home telemonitoring: analysis of the intervention arm in a before and after trial. JMIR Med Inform 2018 Apr 9;6(2):e15. [doi: 10.2196/medinform.9200] [Medline: 29631991]

34. Piette JD, Rosland A, Marinec NS, Striplin D, Bernstein SJ, Silveira MJ. Engagement with automated patient monitoring and self-management support calls: experience with a thousand chronically ill patients. Med Care 2013 Mar;51(3):216-223 [FREE Full text] [doi: 10.1097/MLR.0b013e318277ebf8] [Medline: 23222527]
35. Greenhalgh T, A'Court C, Shaw S. Understanding heart failure; explaining telehealth - a hermeneutic systematic review. BMC Cardiovasc Disord 2017 Dec 14;17(1):156 [FREE Full text] [doi: 10.1186/s12872-017-0594-2] [Medline: 28615004]
36. de Veer AJ, Peeters JM, Brabers AE, Schellevis FG, Rademakers JJ, Francke AL. Determinants of the intention to use e-Health by community dwelling older people. BMC Health Serv Res 2015 Mar 15;15(1). [doi: 10.1186/s12913-015-0765-8] [Medline: 25889884]
37. Cimperman M, Makovec BM, Trkman P. Analyzing older users' home telehealth services acceptance behavior-applying an Extended UTAUT model. Int J Med Inform 2016 Jun;90:22-31. [doi: 10.1016/j.ijmedinf.2016.03.002] [Medline: 27103194]
38. Kraai IH, Luttik ML, de Jong RM, Jaarsma T, Hillege HL. Heart failure patients monitored with telemedicine: patient satisfaction, a review of the literature. J Card Fail 2011 Aug;17(8):684-690. [doi: 10.1016/j.cardfail.2011.03.009] [Medline: 21807331]
39. Matthew-Maich N, Harris L, Ploeg J, Markle-Reid M, Valaitis R, Ibrahim S, et al. Designing, implementing, and evaluating mobile health technologies for managing chronic conditions in older adults: a scoping review. JMIR Mhealth Uhealth 2016 Jun 09;4(2):e29 [FREE Full text] [doi: 10.2196/mhealth.5127] [Medline: 27282195]
40. O'Connor S, Hanlon P, O'Donnell CA, Garcia S, Glanville J, Mair FS. Understanding factors affecting patient and public engagement and recruitment to digital health interventions: a systematic review of qualitative studies. BMC Med Inform Decis Mak 2016 Sep 15;16(1):120 [FREE Full text] [doi: 10.1186/s12911-016-0359-3] [Medline: 27630020]
41. Granja C, Janssen W, Johansen MA. Factors determining the success and failure of eHealth interventions: systematic review of the literature. J Med Internet Res 2018 May 01;20(5):e10235 [FREE Full text] [doi: 10.2196/10235] [Medline: 29716883]
42. Guzman-Clark JR, van Servellen G, Chang B, Mentes J, Hahn TJ. Predictors and outcomes of early adherence to the use of a home telehealth device by older veterans with heart failure. Telemed J E Health 2013 Mar;19(3):217-223. [doi: 10.1089/tmj.2012.0096] [Medline: 23268695]
43. Fairbrother P, Ure J, Hanley J, McCloughan L, Denvir M, Sheikh A, et al. Telemonitoring for chronic heart failure: the views of patients and healthcare professionals - a qualitative study. J Clin Nurs 2014 Jan;23(1-2):132-144. [doi: 10.1111/jocn.12137] [Medline: 23451899]
44. Pressler S. Women with heart failure are disproportionately studied as Compared with prevalence: a review of published studies from 2013. J Cardiovasc Nurs 2016;31(1):84-88. [doi: 10.1097/JCN.0000000000000212] [Medline: 25419948]
45. McSweeney J, Pettey C, Lefler L, Heo S. Disparities in heart failure and other cardiovascular diseases among women. Womens Health (Lond) Jul 2012;8(4):473-485. [doi: 10.2217/whe.12.22] [Medline: 22757737]

Abbreviations

HF: heart failure
IVR: interactive voice response
NYHA class: New York Heart Association functional classification
RCT: randomized controlled trial
TM: telemonitoring
UHN: University Health Network
UTAUT2: unified theory of acceptance and use of technology 2

Edited by G Eysenbach; submitted 29.12.18; peer-reviewed by L de Witte, S Kitsiou, A Miranda; comments to author 24.01.19; revised version received 25.01.19; accepted 11.02.19; published 26.02.19.

Please cite as:
Ware P, Dorai M, Ross HJ, Cafazzo JA, Laporte A, Boodoo C, Seto E
Patient Adherence to a Mobile Phone–Based Heart Failure Telemonitoring Program: A Longitudinal Mixed-Methods Study
JMIR Mhealth Uhealth 2019;7(2):e13259
URL: http://mhealth.jmir.org/2019/2/e13259/
doi:10.2196/13259
PMID:30806625

Original Paper

Initial Assessment of a Comprehensive Digital Smoking Cessation Program That Incorporates a Mobile App, Breath Sensor, and Coaching: Cohort Study

Jennifer D Marler[1], MD; Craig A Fujii[1], BS; David S Utley[1], MD; Lydia J Tesfamariam[1], BS; Joseph A Galanko[2], PhD; Heather Patrick[1], PhD

[1]Carrot Inc, Redwood City, CA, United States

[2]Biostatistics Core for the Center for Gastrointestinal Biology and Disease and the biostatistician for the Clinical Nutrition Research Center, Department of Medicine, Division of Gastroenterology and Hepatology, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

Corresponding Author:
Jennifer D Marler, MD
Carrot Inc
1600 Seaport Blvd
Suite 150
Redwood City, CA, 94063
United States
Phone: 1 4157577696
Email: marler@carrot.co

Abstract

Background: Cigarette smoking is the leading cause of preventable morbidity and mortality, excess health care expenditure, and lost work productivity. Otherwise effective evidence-based treatments have had limited success owing to challenges with access, engagement, and scale. Pivot is a comprehensive digital smoking cessation program that incorporates a Food and Drug Administration–cleared carbon monoxide breath sensor, smartphone app, and text-based human coaching.

Objective: This initial evaluation of Pivot aimed to assess participant engagement, changes in attitudes toward quitting, and changes in smoking behavior.

Methods: US cigarette smokers aged 18 to 65 years who smoked ≥5 cigarettes per day (CPD) were recruited online. Participants completed a screening call, electronic informed consent, registration, and onboarding before beginning Pivot. Pivot includes 5 sequential stages (Explore, Build, Mobilize, Quit, and Secure), taking 14.5 to 18.5 weeks to complete. Data were collected via app and online questionnaires. Outcomes included engagement and retention (ie, weeks of active engagement and Pivot stage progression); attitudes toward quitting (ie, quit readiness, quit confidence, and expected difficulty maintaining quit); and smoking behavior (ie, quit attempts, cigarette reduction, and abstinence (7- and 30-day point prevalence abstinence [PPA]).

Results: A total of 319 participants completed onboarding (intention-to-treat [ITT] sample); 272/319 participants (85.3%) completed the end-of-Pivot questionnaire (study completer sample). Most (212/319, 66.5%) were not ready to quit in the next 30 days at baseline. On average, participants actively engaged in the program for a mean 12.4 (SD 7.1) weeks. Pivot stage completion rates were Explore: 88.7% (283/319), Build: 57.4% (183/319), Mobilize: 43.6% (139/319), Quit: 41.1% (131/319), and Secure: 39.5% (126/319). Repeated measures linear mixed model analyses demonstrated positive changes in attitudes from baseline to Mobilize (pre-Quit): increased confidence to quit (4.2 to 7.4, $P<.001$) and decreased expected difficulty maintaining quit (3.1 to 6.8, $P<.001$). The quit attempt rate (ie, those making ≥1 quit attempt lasting ≥1 day) was 79.4% (216/272, completer). At the end of Pivot, 7-day PPA rates were 32.0% (102/319, ITT) and 37.5% (102/272, completer); 30-day PPA rates were 27.6% (88/319, ITT) and 32.4% (88/272, completer). Moreover, 30-day PPA rates were comparable among those ready and not ready to quit in the next 30 days at baseline. Of those not achieving abstinence, 25.9% (44/170, completer) achieved ≥50% reduction in CPD by study end.

Conclusions: This study evaluated Pivot's initial performance with comparable quit rates among those ready and not ready to quit in the next 30 days at entry. The present data, considered with the program's accessibility, innovation, evidence-based foundation, and design for all smokers, suggest Pivot has the potential to address limitations of reach and scale and thereby advance smoking cessation efforts.

Trial Registration: ClinicalTrials.gov NCT03295643; https://clinicaltrials.gov/ct2/show/NCT03295643 (Archived by WebCite at http://www.webcitation.org/75TiNe6BE).

(JMIR Mhealth Uhealth 2019;7(2):e12609) doi:10.2196/12609

KEYWORDS

smoking cessation; mobile applications; health promotion; cell phone

Introduction

Background

In the United States, cigarette smoking is responsible for 480,000 deaths annually, and more than 16 million individuals live with a smoking-related illness [1]. Meaningful gains have been made in smoking cessation over the last decade, with the prevalence of cigarette smoking among US adults decreasing from 20.9% in 2005 to 14.8% in 2017 [2,3]. Despite this progress, smoking remains the leading cause of preventable morbidity and mortality.

Approximately 70% of adult smokers want to quit, and a little over half attempt to quit each year. Yet, quit rates are low, with less than 1 in 10 quit attempts being successful. This is related, in part, to underutilization of proven treatments, which include behavioral counseling and pharmacotherapy [4]. Use of behavioral counseling or cessation medications approximately doubles quit rates with combined use being more effective than either alone [4-6]. Even so, fewer than a third of those making a quit attempt use any evidence-based cessation methods, and less than 5% use the optimal approach of combination counseling and pharmacotherapy [7,8]. There are multiple causes of this underutilization, including health care access, transportation challenges, time constraints, convenience, awareness of service availability, desire to quit unassisted, and costs. State quitlines overcome many of these barriers but, on average, reach only about 1% of smokers annually [9]. Overall, the end result is that otherwise effective treatments suffer limited population reach, which ultimately undermines their impact on smoking cessation.

Acknowledging the limitations of current treatment methods, particularly relating to reach and scalability, the rise of smartphone ownership presents an appealing medium for smoking cessation treatment. As of 2017, 77% of Americans owned a smartphone, with majority-level ownership rates across various demographic and socioeconomic characteristics (eg, ≥75% ownership rates for varying ethnic and racial groups and ≥67% for all income groups) [10]. Research suggests smokers are actively using smartphones in quit attempts, with 1 study reporting an average of 779,400 downloads of Android platform–based smoking cessation apps per month worldwide [11]. Another study reported that a little over half of the smokers had downloaded smoking cessation apps in the past and, of these, three-quarters had made quit attempts using an app [12].

Despite this interest and use, quality among smoking cessation apps is disparate. Haskins et al evaluated 158 peer-reviewed articles addressing mobile apps for smoking cessation and identified 177 unique apps relevant to smoking cessation in the App Store for iPhone and 139 in Google Play for Android. They ultimately identified only 6 apps with some level of scientific support in the peer-reviewed literature, only 3 of which were available in the app stores [13]. Although, to date, there has been limited availability of evidence-based mobile smoking cessation programs, recent assessments of such programs have yielded encouraging results. Bricker et al evaluated a smoking cessation app focused on Acceptance and Commitment Therapy (ACT) in 99 adult smokers. At 2-month follow-up, 84% of study completers were satisfied with the program, 21% had achieved 7-day point prevalence abstinence (PPA), and 11% had achieved 30-day PPA [14]. In 416 smokers, Iacoviello et al assessed a smoking cessation app designed to deliver the essential features of the United States Clinical Practice Guideline (USCPG) for treating tobacco use and dependence [4]. They reported a 7-day PPA of 45.2% and 30-day PPA of 26.2% in an intent-to-treat (ITT) analysis [15].

Building on these efforts, Pivot is a comprehensive digital smoking cessation program that brings together multiple evidence-based components into a seamless solution. In addition to delivering the USCPG through a multiphase mobile app, Pivot includes the first Food and Drug Administration (FDA)-cleared personal carbon monoxide (CO) breath sensor, dedicated human coaching delivered through in-app text messaging, and a program designed for individuals with varying levels of readiness to quit. In line with wearable devices, the CO breath sensor provides real-time personal biometric data to users. This leverages the findings of several published studies [16-19] as well as expert opinion [20,21], which suggest that personal CO breath sample data can be educational and motivational and may lead to changes in attitudes toward quitting and smoking behavior. To that end, the CO breath sensor is incorporated in the Pivot program as an engagement tool, with the intention that users will find their expired CO values informative and motivational. Pivot coaching incorporates evidence-based principles of smoking cessation treatment combined with innovation through digital delivery. Continuity of care is achieved through a dedicated human coach who partners with the participant for the entire Pivot journey. Communication with one's coach is conducted through asynchronous text messaging within the Pivot app, allowing the participant to initiate outreach or respond to their coach using a modality and timeline that fits into their life. To increase reach, Pivot has been designed to support users along the entire spectrum of readiness to quit, from being unsure or ambivalent to being highly motivated. Pivot does not assume participants are ready to quit smoking at entry and, consequently, does not start with choosing a quit date or developing a quit plan but rather begins with self-exploration and awareness building.

A feasibility study of the first stage of Pivot (ie, Explore) examined program engagement, changes in attitudes toward

quitting, self-reported changes in smoking behavior, and program acceptability [22]. Engagement rates over the 9 days of Explore were as follows: ≥80% of participants (34 to 39 of 41) used the CO sensor ≥1 time per day and over 55% (23 to 27 of 41) used it ≥5 times per day; all 9 in-app activities had completion rates of ≥80% (33 to 40 of 41). Furthermore, coach-initiated contacts received response rates of ≥73% (30 to 39 of 41). Results also yielded significant positive changes in attitudes toward quitting smoking from baseline to study exit, including increased readiness to quit, lower perceived difficulty quitting, and greater expectations of success. At study exit, 78% of participants (32/41) reported that they had decreased the number of cigarettes smoked per day since beginning the study. They also rated program quality and satisfaction as very high. Although these results are promising, there were some limitations. First, this was an initial feasibility study that examined only the first stage of Pivot: Explore. Additional research is needed to examine longer-term engagement throughout the Pivot program. Finally, although participants reported decreasing cigarette smoking, this feasibility study did not include more robust indicators of changes in smoking such as number of quit attempts and 7- and 30-day PPA.

Objectives

This pragmatic, exploratory study is the first evaluation of the Pivot program in the hands of its intended users. The purpose of this study was to assess participant engagement, changes in attitudes toward quitting, and changes in smoking behavior.

Methods

Study Design

This was a prospective open-label, single-arm study of Pivot (Carrot Inc). The study was conducted as a pragmatic, exploratory study to assess the initial performance of the Pivot program. The study examined engagement (ie, weeks active in the program and stage progression), changes in attitudes toward quitting smoking (ie, readiness to quit, confidence about quitting, and anticipated difficulty quitting), and changes in smoking behavior (ie, quit attempts, reduction in cigarettes smoked per day, and 7- and 30-day PPA).

Consent and Ethical Approval

All participants provided electronic informed consent before participation. The study was reviewed and approved by the Solutions institutional review board (Little Rock, AR, USA) protocol number 2017/09/22 and registered with Clinicaltrials.gov NCT03295643.

Eligibility and Recruitment

Participants were enrolled from October 2017 to March 2018. Potential participants were identified via advertisements on Web media (ie, Facebook, Instagram, Twitter, Google Ads, Reddit, and smokefree.gov), with a link to an online screening form. Contact information and data on age, sex, smartphone ownership, employment status, and smoking behavior were collected. To be eligible for participation, individuals had to meet all of the following eligibility criteria: aged 18 to 65 years, English speaking, smoke ≥5 cigarettes per day (CPD), own and use a compatible smartphone (iPhone 5 and above, operating system iOS 9.0 and above, or Android 4.4 and above, operating system Android 4.4 and above), be employed for ≥20 hours a week, and live in the United States. Although we aim for broad availability of Pivot through multiple channels such as private and public insurers, direct-to-consumer, and not-for-profit foundations, Pivot will initially be available to individuals through their employers (self-insured employers or employee wellness programs). As such, we applied the employment requirement to assess Pivot in individuals closely aligned with Pivot's initial user population. This approach was informed by recent reports, which indicate that 75% of US employers offer wellness resources and/or a general wellness program with representation across all occupational and wage groups [23-25]. In addition, in most US states, being employed 20 hours per week or more is a requirement for benefits eligibility, including access to wellness offerings.

Nonproportional quota sampling was employed with percentage limits applied to age, CPD, stage of change, and sex (Table 1). The eligibility criteria and nonproportional quota sampling were used to achieve a study population that reflects the initial targeted commercialized population (ie, employed adult smokers likely to have access to a wellness benefit offering, representing the spectrum of readiness to quit). Individuals were called on a first-come-first-serve basis with nonproportional quota sampling enrollment guidelines applied. Initially, Pivot was released for iPhone only. Toward the end of enrollment, Pivot became available for use on the Android platform, and enrollment shifted accordingly.

Those who met the eligibility criteria and met the nonproportional quota sampling requirements were contacted by phone to describe the study and allow potential participants to ask questions. Those expressing interest in study participation were emailed a Web address where they could register and complete the electronic informed consent. Participants did not have to indicate intent to quit smoking as a condition of study participation.

Onboarding

Upon providing electronic informed consent and completing the online registration, participants received an email with a link to the baseline questionnaire, which assessed demographics, smoking history, and attitudes toward quitting (ie, readiness to quit, confidence to quit, and perceived difficulty of quitting). They were also mailed the CO breath sensor and instructions to download the Pivot program onto their smartphone. Getting started in Pivot was a self-guided process. Participants used the provided instructions for downloading the Pivot program on their smartphone and pairing the CO sensor to the Pivot app. They also had access to customer service as needed. Once this process was complete, the participant was considered to have onboarded. After successful onboarding, they were paired with a live, dedicated coach who provided one-on-one support over the course of the study.

Table 1. Nonproportional quota sampling enrollment: target and actual values (n=319).

Category	Targeted enrollment values (%)	Actual enrollment values, n (%)
Age (years)		
18-29	<20	24 (7.5)
30-60	≥70	281 (88.1)
≥61	<10	14 (4.4)
Cigarettes per day		
5-10	<25	68 (21.3)[a]
11-30	≥65	236 (74.0)
>30	<10	15 (4.7)
Stage of change		
Intend to quit within 30 days	≥20	107 (33.5)
Intend to quit within 6 months	≥20	201 (63.0)
Not thinking of quitting	<20	11 (3.5)
Sex		
Female	40-60	184 (57.7)

[a]One participant reported smoking 5 cigarettes per day at screening, then entered 4 cigarettes per day on the baseline questionnaire.

Figure 1. Pivot carbon monoxide breath sensor.

Pivot Program

Pivot is a comprehensive digital solution that includes (1) the first FDA-cleared, over-the-counter CO breath sensor (Figure 1), which communicates with a smartphone and app via Bluetooth; (2) the multiphase Pivot mobile app; and (3) dedicated human coaching delivered one-on-one through in-app text messaging. The program has been developed for commercialization and designed for delivery in the context of employee wellness programs and health plans. Pivot leverages evidence-based principles and clinical best practices. Specifically, Pivot uses the USCPG-recommended 5 As (Ask, Advise, Assess, Assist, and Arrange), tailors on readiness to quit [4], encourages the use of FDA-approved pharmacotherapy [4,26-28], capitalizes on effective methods for smoking cessation (eg, motivational interviewing, cognitive behavioral therapy, and self-determination theory) [4,29-31], and provides behavioral counseling through a live, dedicated coach [4,28,32,33].

The Pivot journey consists of 5 sequential stages (Explore, Build, Mobilize, Quit, and Secure). To advance to the next Pivot stage, the preceding stage must be completed. Pivot begins with *Explore* (9 days), which is designed for anyone who smokes,

to raise awareness and interest in moving forward. In Explore, users take samples with the Pivot Breath Sensor, log cigarettes, get to know their coach, and complete daily activities to understand their smoking patterns and explore how smoking affects their lives. The second phase of Pivot is *Build* (1 day to 4 weeks), which is tailored to users' readiness, motivation, and confidence. Build culminates with users setting a quit date and building a quit plan. Next is *Mobilize* (7 days), which provides opportunities for users to put into practice individual elements of their quit plan, one at a time, in preparation for quit day. The fourth phase of Pivot is *Quit* (7 days), which begins on the user's selected quit day and continues through the first week of living smoke-free. Quit incorporates a daily check-in feature to allow users to track their progress and set daily goals to reinforce the idea of quitting as a process. *Secure* (11 weeks) is a natural extension of Quit and focuses on supporting users in developing internal, sustainable motivation to stay smoke-free for good. With continued coaching support, self-monitoring, and practice, Secure is designed to help Pivot's newly smoke-free users learn to navigate the challenges that come in the first few months after quitting. Sample screenshots from the Pivot app from different stages are shown in Figure 2.

Throughout the Pivot program, participants could provide breath samples, log cigarettes, complete in-app activities, and interact with their coach via in-app chat. During the study, the suggested breath sampling frequency was once per hour while awake during Explore through Mobilize, decreasing to a few breath samples per day thereafter. This was discussed during the screening call with acknowledgment that breath sampling should ultimately be undertaken in a way that works with each participant's schedule. Participants were also encouraged to log all cigarettes they smoked using the in-app logging feature, preferably as soon as possible after smoking. A daily push notification alerted participants to their daily in-app activity (Explore through Quit) or check-in (Quit and Secure). Finally, participants received behavioral counseling via coaching throughout the entire program. Coach-initiated contact included outreach 3 times a week from Explore through the first 30 days in Secure, once per week for the next 30 days in Secure, and every other week for the last 30 days in Secure. Participants could initiate contact with their coach as frequently as desired. Coaching was undertaken through asynchronous in-app text messaging, thus allowing participants to respond to coach-initiated contact or to initiate contact with their coach whenever it was convenient for them.

Procedure

The Pivot program was designed to be 14.5 to 18.5 weeks in duration, depending on participant navigation of the program stages (Build lasts 1 to 28 days depending on participant readiness and desired pace). Participants were asked to complete electronic questionnaires in-app and online through Survey Monkey at baseline, upon completion of each stage, and at the end of Pivot. Participants were notified of the online questionnaires via email with a link to the questionnaire. Automated email reminders were sent daily until online questionnaire completion, the emailing of the next study questionnaire, or in the case of the final questionnaire, until the end of the study period. Participants were compensated US $10 to US $50 per completed study questionnaire and US $50 for returning the CO breath sensor for up to a total of US $265, using Visa gift cards. Compensation was not associated with use of the various components of Pivot, level of engagement, or smoking/quitting status.

Data Collection

Data were collected electronically through participant input in the Pivot online registration form, Pivot app, and online questionnaires. Study data were imported directly into a secure database (PostgreSQL, PostgreSQL Global Development Group).

Outcome Variables and Measurement

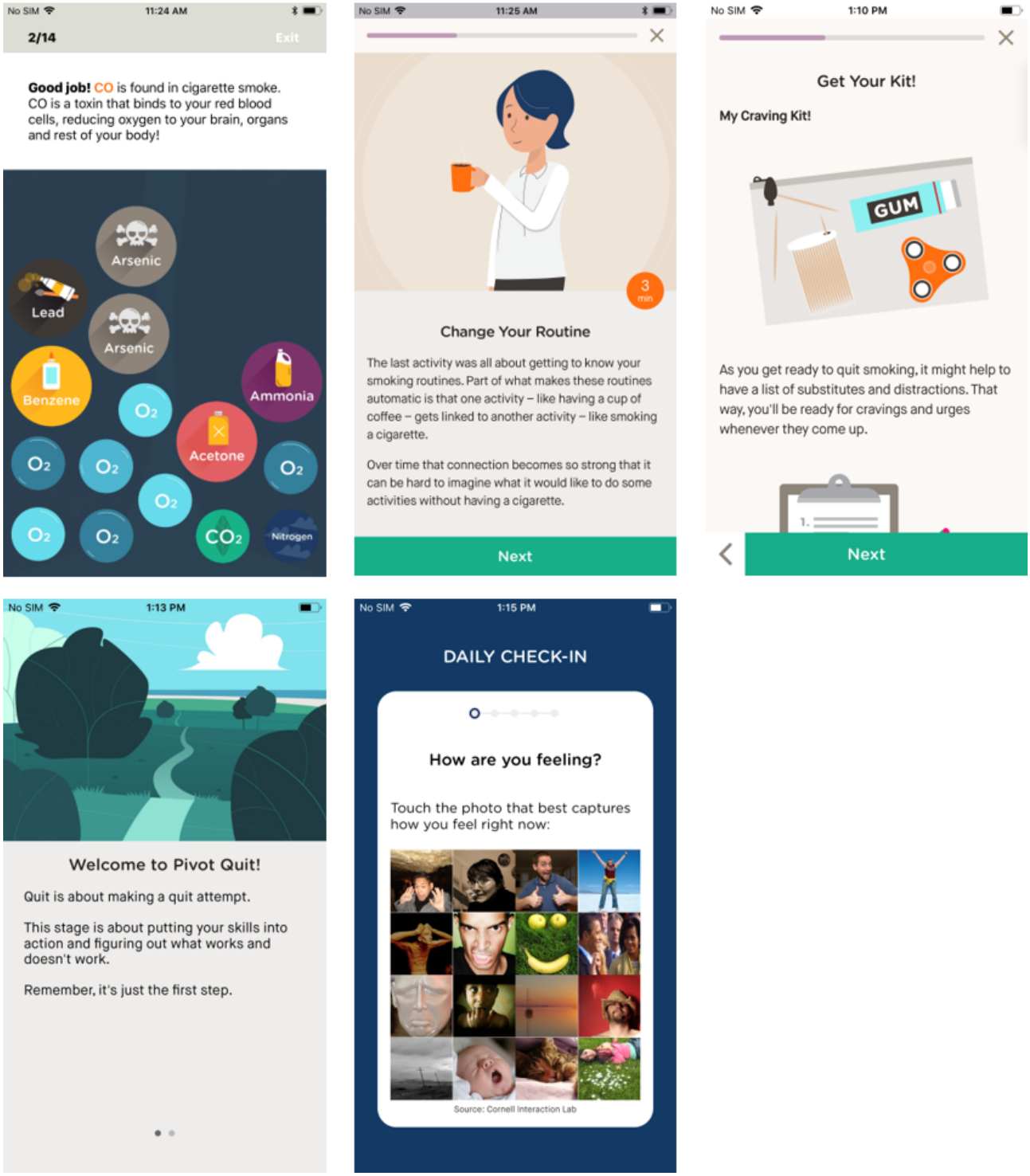
Baseline Characteristics

Baseline characteristics included demographic information (age, sex, race or ethnicity, household income, and education) and smoking behavior data (cigarettes smoked per day, duration of smoking, and number of quit attempts over the past 12 months).

Engagement

A total of 3 metrics of engagement were assessed: Pivot stage completion, weeks of active engagement, and program completion. Pivot stage completion is defined as the percent of participants who completed each stage based on the number of those who onboarded. Weeks of active engagement was defined as doing at least 1 of the following: doing a breath sample, logging a cigarette, starting or completing a daily activity, challenge or check-in, or messaging one's coach. Opening the app without doing one of the aforementioned actions and receipt of a message from one's coach did not count as engagement. Finally, program completion was defined as those who completed onboarding and completed Explore through Secure. Completing Secure was defined as (1) having been in Secure for 11 weeks, (2) having engaged with the program at least once during the 11 weeks of Secure as defined above, and (3) providing smoking status (CPD and 7- and 30-day point prevalence) in the final questionnaire.

Figure 2. Sample screenshots from the Pivot app, different stages.



Attitudes Toward Quitting

Measurements assessing attitudes toward quitting included readiness to quit/stage of change [34,35], confidence to quit, and anticipated difficulty maintaining quit status [36,37]. These measures were chosen to reflect a range of psychosocial indicators that have been shown to influence the likelihood of attempting to quit smoking and/or cessation (ie, motivation and self-efficacy). Participants completed the readiness to quit/stage of change item at baseline and the end of Explore. To complete the second phase of Pivot, Build, participants had to have selected a quit date within the next 14 days. Therefore, it did not make sense to ask participants the readiness to quit/stage of change item beyond the end-of-Explore survey. Confidence to quit and anticipated difficulty maintaining quit status were assessed at baseline and at the end of Explore, Build, and Mobilize. The specific items used to assess participant attitudes are detailed in Table 2.

Table 2. Measurements assessing attitudes toward quitting.

Question	Answer options/scale
Are you seriously thinking of quitting smoking? (stage of change)	Yes, within the next 30 days
	Yes, within the next 6 months
	No, not thinking of quitting
If you were to quit smoking right now, how successful would you be?	Scale 1 to 10 (1=not at all successful, 10=completely successful)
If you were to quit smoking right now, how difficult do you think it would be to stay smoke-free?	Scale 1 to 10 (1=really hard to stay quit, 10=really easy to stay quit)

Changes in Smoking Behavior

A total of 4 metrics were used to assess changes in smoking behavior: smoking reduction, quit attempts, and 7-day and 30-day PPA. At the end of each stage, monthly in Secure, and on the final questionnaire, participants were asked the number of cigarettes smoked per day. *Smoking reduction* was evaluated in all participants before Quit and in the subset of participants who never achieved 7-day PPA using the following metrics: change in CPD, percentage change in CPD, and percentage of participants who achieved ≥50% reduction in CPD. *Quit attempt* was defined as going at least 1 day without smoking cigarettes, even a single puff. *Point-prevalence abstinence (7-day and 30-day)* was assessed as a primary outcome on the final study questionnaire. We also asked PPA for secondary outcomes at the end of Quit (7-day PPA) and monthly in Secure. Participants were considered to have achieved 7-day (30-day) PPA if they answered "no" to the following question: "In the last 7 (30) days have you smoked any cigarettes, even a single puff?" As the Pivot program has no face-to-face contact, and data collection is achieved through remote means using the app and electronic questionnaires, biochemical verification of smoking status was not pursued in accordance with previous recommendations [38].

Sample Size

Previous evaluation has shown that changes in attitudes toward quitting are meaningful predictors of quit attempts [39]. On the basis of a previous assessment of 41 individuals using the first stage of Pivot (Explore), we estimated that the mean (SD) change in ratings assessing attitudes toward quitting (confidence to quit and expected difficulty maintaining quit) would be ≥1 (4) just before reaching Quit [22]. On the basis of these estimates, there was 80% power to detect a significant difference in these ratings with a sample size of 101. As this was an initial study of the complete Pivot program, and in the context of known high attrition rates with mobile health apps [40,41], we applied conservative retention estimates drawn from other similar studies. Specifically, the target enrollment of 310 was estimated to yield at least 100 participants still engaged at the end of Secure.

Statistical Analyses

Statistical analyses were conducted using all available data. For engagement, data were collected through the Pivot app to capture stage completion, weeks of active engagement, and program completion. Changes in attitudes toward quitting were assessed from baseline to the end of Explore, Build, and Mobilize (ie, pre-Quit); changes in cigarettes smoked per day were assessed from baseline to the end of each Pivot stage and the final questionnaire. Participants served as their own controls, and comparisons were made to no change. To evaluate changes in attitudes or cigarettes smoked per day over time, repeated measures linear mixed model analyses were performed using a compound symmetric correlation matrix to model the repeated measures within subjects. As these measurements were taken at the same Pivot stage (not necessarily the same time), stage was used as a surrogate for time. To make specific comparisons across time, F statistics were computed using the results from the model. To determine whether the effects over time were the same among subgroups, subgroup and time by subgroup interaction terms were added to the models with time. Analyses were conducted to calculate mean (SD) for normally distributed variables for actual data or mean (SE) for modeled data and median (interquartile range) values in instances of non-normally distributed variables. Furthermore, one-sample t tests were used for numerical data. Fisher exact or chi-square tests were used for categorical data. McNemar test was used for 2-category match-paired data. Cohen kappa statistic was used for 3-category match-paired data. Analyses were conducted using SAS Version 9.4 (SAS Institute). Statistical significance was set at $P<.05$.

In the assessment of cessation (PPA), 2 sets of analyses were performed. In the first, an ITT analysis, individuals who did not respond to PPA questions were assumed to be smoking. The ITT analysis was performed as a standard assessment with long-standing use in traditional clinical studies. However, this method is subject to biases in effect size estimates and may lead to errors [42,43]. Consequently, a study completer analysis was also performed, which only included individuals who completed the final questionnaire. Participants were sent the final questionnaire regardless of whether or not they completed the Pivot program. For additional assessments performed at the end of the study (quit attempts, CPD, and smoking reduction), a study completer analysis was performed. This analysis approach comports with previous reports assessing app-based digital cessation programs [14,15].

Results

Enrollment

From October 2017 to March 2018, 7382 online screening forms were received; 3436 met the screening eligibility criteria, and 1435 received outbound phone calls from study staff. Most phone calls were not answered or returned. After the phone calls, registration links were emailed to 417 potential

participants; 362 completed the online registration and informed consent. A total of 319 participants completed onboarding and comprised the ITT sample. Moreover, 85.3% (272/319) of the participants completed the final questionnaire and comprised the study completer sample. A total of 47 participants were lost to follow-up or withdrew consent. Study enrollment and attrition are depicted in the participant flow diagram in Figure 3.

Baseline Characteristics

The study sample consisted of 57.7% women (184/319), had a mean age of 42.8 (SD 10.2) years, smoked a mean of 17.7 (SD 7.6) CPD at baseline, and had been smoking for an average of 26.4 (SD 10.7) years. Most study participants (294/319, 92.2%) used iPhones.

At baseline, one-third of participants (107/319, 33.5%) indicated they were seriously thinking of quitting smoking in the next 30 days. Two-thirds of participants (212/319, 66.5%) did not intend to quit smoking within the next 30 days. Specifically, 63.0% (201/319) indicated they were thinking of quitting in the next 6 months, and 3.5% (11/319) indicated they were not seriously thinking of quitting smoking. On average, participants had made 2.1 (SD 3.3) quit attempts over the past 12 months. Study demographic details are provided in Table 3.

Figure 3. Consolidated Standards of Reporting Trials flow diagram of participants.

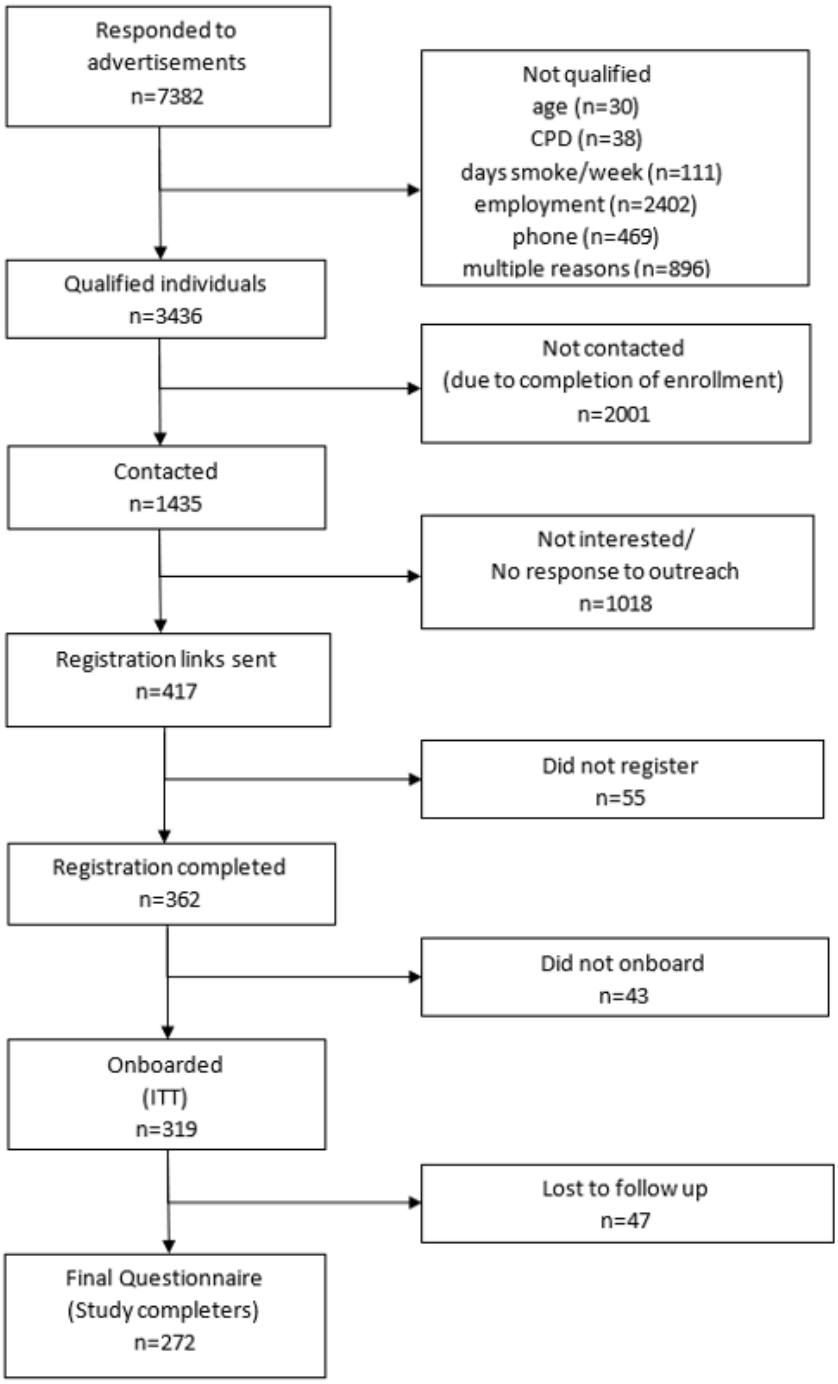


Table 3. Baseline demographics.

Demographics	Statistics
Completed onboarding, n	319
Female, n (%)	184 (57.7)
Age (years), mean (SD)	42.8 (10.2)
20-29, n (%)	24 (7.5)
30-39, n (%)	108 (33.9)
40-49, n (%)	103 (32.3)
50-59, n (%)	63 (19.7)
60-69, n (%)	21 (6.6)
Cigarettes smoked per day, mean (SD)	17.7 (7.6)
Years smoking, mean (SD)	26.4 (10.7)
Quit attempts in last 12 months, mean (SD)	2.1 (3.3)
Smartphone, n (%)	
iPhone	294 (92.2)
Android	25 (7.8)
Ethnicity, n (%)	
White	264 (82.8)
African American	22 (6.9)
Hispanic	15 (4.7)
Asian	5 (1.6)
American Indian or Alaska Native	4 (1.3)
Native Hawaiian or other Pacific Islander	2 (0.6)
Other	7 (2.2)
US region, n (%)	
South	123 (38.6)
Midwest	73 (22.9)
West	68 (21.3)
Northeast	55 (17.2)
Highest level of education or degree attained, n (%)	
Professional or doctorate degree	8 (2.5)
Master's degree	18 (5.6)
Bachelor's (4-year) degree	70 (21.9)
Associate's (2-year) degree	52 (16.3)
Some college	112 (35.1)
High school/GED[a]	56 (17.6)
Some high school	3 (0.9)
Annual household income (US dollars), n (%)	
Less than $25,000	30 (9.4)
$25,000-$34,999	48 (15.0)
$35,000-$49,999	58 (18.2)
$50,000-$74,999	60 (18.8)
$75,000-$99,999	56 (17.6)
$100,000-$149,999	40 (12.5)

Demographics	Statistics
$150,000 or more	17 (5.3)
Prefer not to answer	10 (3.1)
Are you seriously thinking of quitting smoking? n (%)	
Yes, within the next 30 days	107 (33.5)
Yes, within the next 6 months	201 (63.0)
No, not thinking of quitting	11 (3.5)
Attitude toward smoking ratings, mean (SD)	
If you were to quit smoking now how successful would you be? (1=not successful at all, 10=completely successful)	4.2 (2.7)
If you were to quit smoking right now how difficult do you think it would be to stay smoke-free? (1=really hard; 10=really easy)	3.1 (2.5)

[a]GED: General Education Diploma.

Engagement

The proportion of participants who completed each Pivot stage is detailed in Figure 4. As shown, after some drop-off in Build, retention rates remained relatively stable for the remainder of program participation. Overall, 39.5% (126/319) of participants who onboarded completed the Pivot program.

Participants were active in the program for a mean of 12.4 (SD 7.1) weeks. Time to complete each stage varied for some from original projections with participants taking longer in Mobilize, Quit, and Secure (Table 4).

Beyond active engagement by week, as a secondary analysis, we explored patterns of daily engagement to gain a more detailed understanding of how participants interacted with Pivot (Figure 5). Overall, participants consistently used all Pivot components throughout their engagement with the Pivot program. Notable use pattern variation included performance of the highest average number of daily breath samples in Explore, completion of the highest average number of daily Pivot app activities in Build, and an overall decrease in daily logging of cigarettes from one stage to the next. These use patterns are expected. Specifically, breath sampling is particularly salient to Explore's focus on self-awareness and exploration of how smoking affects one's life. Build is a time for learning and skill building, which are primarily accomplished through completion of daily app activities. Cigarette logging would decrease both as a result of actual decreases in smoking and decreased compliance with logging during advancement in the program. In addition, there was continued but decreased use of all program components in Secure, consistent with an expected decrease in need for program intensity as participants achieve cessation.

Figure 4. Proportion of participants who completed each Pivot stage, by those who onboarded (n=319).

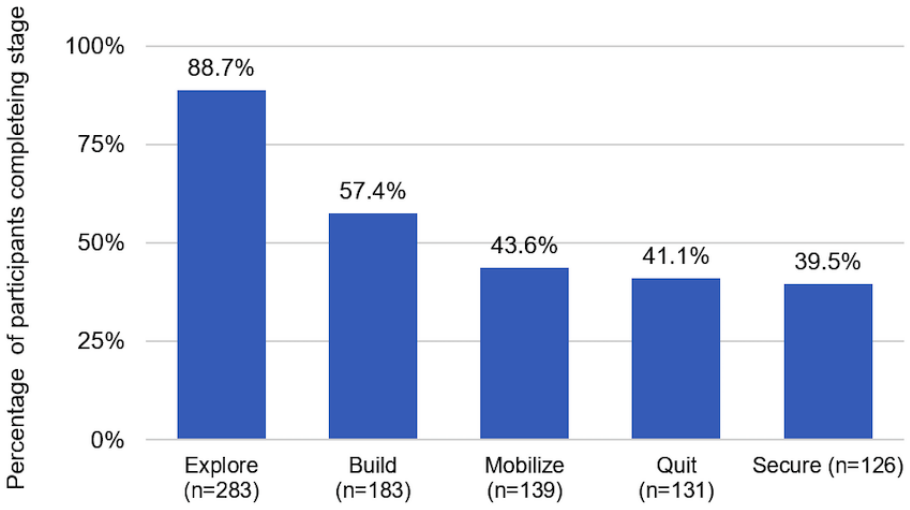
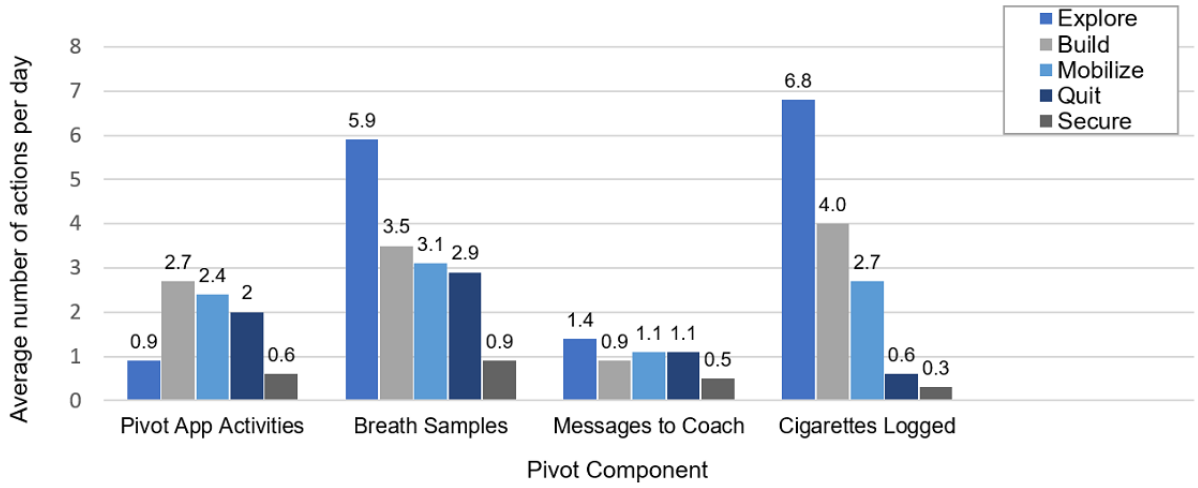


Table 4. Projected and actual days spent in each Pivot stage.

Stage	Projected days	Actual days, mean (SD)
Explore	9	9.3 (4.1)
Build	1-28	19.8 (26.2)
Mobilize	7	18.0 (23.0)
Quit	7	9.1 (10.0)
Secure	77	87.3 (7.7)

Figure 5. Participant engagement with the different components of Pivot; average actions per day, by stage, for all participants who entered stage.



Change in Attitudes Toward Quitting

Change in readiness to quit (stage of change) was assessed at baseline and at the end of Explore (Table 5). Readiness to quit improved with 55.9% (157/281) of respondents indicating they were thinking of quitting in the next 30 days at the end of Explore compared with 33.5% (94/281) at baseline (P<.001). At the end of Explore, 28.1% (79/281) increased their readiness to quit, 66.2% (186/281) stayed the same, and 5.7% (16/281) decreased (P<.001).

Participants' confidence to quit and perceived difficulty of quitting also improved (Figures 6 and 7). In repeated measures linear mixed model analyses, end-of-stage results were compared with baseline and each other; all changes were statistically significant (P<.001, except for end of Build vs end of Explore, where P=0.01 for confidence to quit).

Table 5. Change in readiness to quit from baseline to the end of Explore.

Readiness to Quit: Baseline	Readiness to Quit[a]: End of Explore, n (%)			
	Yes, within the next 30 days	Yes, within the next 6 months	No, not thinking of quitting	Total
Yes, within the next 30 days	82 (29.2)	12 (4.3)	0 (0)	94 (33.5)
Yes, within the next 6 months	74 (26.3)	98 (34.9)	4 (1.4)	176 (62.6)
No, not thinking of quitting	1 (0.4)	4 (1.4)	6 (2.1)	11 (3.9)
Total	157 (55.9)	114 (40.6)	10 (3.6)	281 (100.0)[b]

[a]Readiness to Quit assessed via stage of change question: Are you seriously thinking of quitting smoking? A. Yes, within the next 30 days; B. Yes, within the next 6 months; C. No, not thinking of quitting.

[b]283 participants completed Explore; however, 2 of these participants did not answer this question on the end of Explore questionnaire.

Figure 6. Changes in confidence to quit. Estimate of means and standard errors based on linear mixed model.

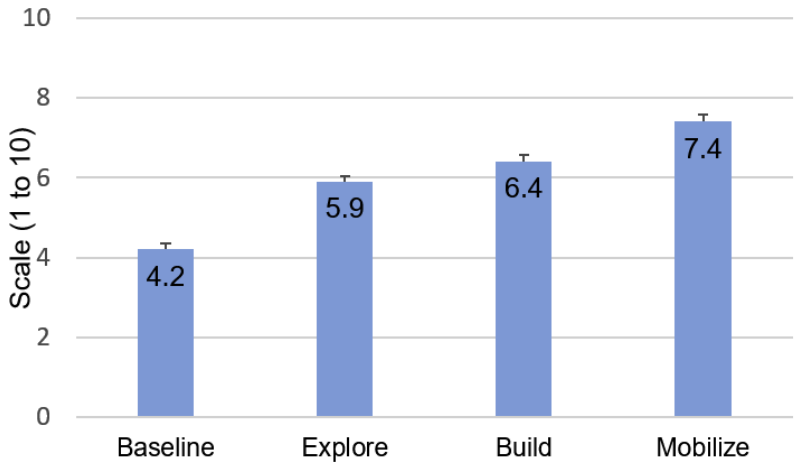
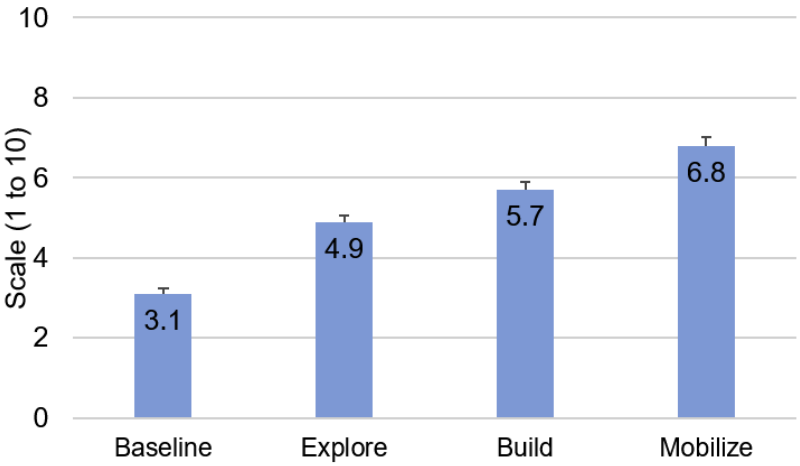


Figure 7. Changes in perceived difficulty of staying quit. Estimate of means and standard errors based on linear mixed model.



Changes in Smoking Behaviors

Change in Cigarettes Per Day by Stage

Repeated measures linear mixed model analysis was performed, with projected end-of-stage CPD values compared with baseline and each other (Table 6). End-of-stage percent changes in CPD were statistically significant (all $P<.001$ with $P=.003$ for Mobilize vs Build), except at the end of Explore versus baseline in those ready to quit in the next 30 days (percentage change in CPD=−10.2%, $P=.13$). Among those who completed Mobilize and provided data (n=139), CPD were reduced by 50.7% (SD 32.6) from baseline to immediately before entering Quit.

Using repeated measures linear mixed model analysis, 31.8% of study participants were projected to reduce their CPD by ≥50% before entering Quit (Table 7). Among those who completed Mobilize and provided data, 53.2% (74/139) reduced their CPD by ≥50% before entering Quit. At the end of the study, among the study completers who did not achieve at least 7-day PPA (n=170), CPD were reduced by 29.1% (SD 34.5), and 25.9% (44/170) reduced their CPD by ≥50%. Pre-Quit (Explore, Build, and Mobilize) there were no statistically significant differences in CPD reduction rates or in the proportion of participants reducing CPD by ≥50% between those ready and not ready to quit in the next 30 days at entry.

Table 6. Average percentage change in cigarettes per day (CPD) by pre-Quit stage, repeated measures linear mixed model analysis.

Group	Pivot stage		
	Explore	Build	Mobilize
All, mean % (SE)	−12.6 (2.5)	−30.7 (2.9)	−44.5 (3.0)
Baseline Readiness to Quit: within the next 30 days, mean % (SE)	−10.2 (6.6)	−27.7 (7.1)	−45.1 (7.4)
Baseline Readiness to Quit: *not* within the next 30 days, mean % (SE)	−13.8 (2.0)	−32.1 (2.4)	−44.0 (2.5)

Table 7. Proportion of participants who reduced cigarettes per day (CPD) by ≥50% by pre-Quit stage, repeated measures linear mixed model analysis.

Group	Pivot stage		
	Explore	Build	Mobilize
All	11.3 (8.1-15.6)	21.2 (16.8-26.4)	31.8 (26.6 -37.5)
Baseline Readiness to Quit: within the next 30 days, estimated percentage (95% CI)	14.9 (9.0-23.6)	22.3 (15.0-31.9)	37.2 (28.1-47.4)
Baseline Readiness to Quit: *not* within the next 30 days, estimated percentage (95% CI)	9.5 (6.1-14.6)	20.6 (15.5-27.0)	29.1 (23.1-36.0)

Table 8. Among participants with increased motivation to stop smoking in Pivot, what was the top reason?

Reason	Statistics, n (%)
Using the sensor to measure carbon monoxide in my breath	80 (34.3)
Trying a variety of strategies to reduce or quit	61 (26.2)
Learning about the effects of smoking through daily activities	38 (16.3)
Coaching	33 (14.2)
Other	16 (6.9)
Receiving support from friends, family, and others	5 (2.1)
Total	233 (100)

Quit Attempts

Overall, 79.4% (216/272, study completers) reported making at least 1 quit attempt during their participation in the Pivot program. On average, study completers made 2.4 (SD 3.1) quit attempts during Pivot. Among the participants who, at baseline, were planning to quit smoking within the next 30 days, 89.1% (82/92, study completers) made a quit attempt during the study. Among the participants who, at baseline, were *not* planning to quit smoking within the next 30 days, 74.4% (134/180, study completers) made a quit attempt during the study.

Point Prevalence Abstinence (7-Day and 30-Day)

At the end of the study, 32.0% (102/319) achieved 7-day PPA and 27.6% (88/319) achieved 30-day PPA using an ITT analysis. Using data from study completers, 37.5% (102/272) achieved 7-day PPA and 32.4% (88/272) achieved 30-day PPA.

Subset analyses of those who were ready to quit within the next 30 days at baseline versus those who were not ready to quit demonstrated similar cessation rates between the 2 groups. Using an ITT analysis, among those who, at baseline, were ready to quit smoking in the next 30 days, 32.7% (35/107) achieved 7-day PPA and 27.1% (29/107) achieved 30-day PPA; for those, at baseline, who were *not* ready to quit smoking in the next 30 days, end-of-study 7-day and 30-day PPA were 31.6% (67/212) and 27.8% (59/212), respectively. A similar pattern emerged for the study completer analysis. Among those who, at baseline, were ready to quit smoking in the next 30 days, 38.0% (35/92) achieved 7-day PPA and 31.5% (29/92) achieved 30-day PPA; for those, at baseline, who were not ready to quit smoking in the next 30 days, end-of-study 7-day and 30-day PPA were 37.2% (67/180) and 32.8% (59/180), respectively. None of these differences in PPA between subsets were statistically significant.

In addition to the primary endpoint with the end-of-study questionnaire, 7-day PPA and 30-day PPA data were also collected as secondary outcomes periodically throughout the Pivot program: at the end of Quit (7-day PPA only) and in Secure at 30, 60, and 90 days (7-day and 30-day PPA). Using all available data, 36.1% (115/319) achieved 7-day PPA and 31.0% (99/319) achieved 30-day PPA at some point during the study.

Participant Feedback

Participants were asked to consider the various components of Pivot. Among the participants who reported Pivot increased their motivation to stop smoking (85.7%, 233/272), using the breath sensor was the most common reason for the increased motivation (Table 8). Trying a variety of strategies to reduce or quit and learning about the effects of smoking through daily activities were the second and third most common reasons, respectively.

Discussion

Principal Findings

This was the initial study of the Pivot program; a comprehensive, multiphase digital smoking cessation program that includes a mobile CO breath sensor, the Pivot app, and dedicated human coaching. The majority (212/319, 66.5%) of participants were not ready to quit in the next 30 days at study entry. Overall, 39.5% (126/319) of participants completed the Pivot program. Participants were active in the program for an average of 12.4 weeks.

There was a positive, statistically significant shift in attitudes toward quitting, including readiness to quit/stage of change, confidence in quitting, and perceived difficulty of staying quit. With regard to changes in smoking behavior, 53.2% of participants who completed Mobilize decreased CPD by at least half just before entering Quit. Moreover, 30-day PPA was achieved by 27.6% (88/319) and 32.4% (88/272) of participants for ITT and study completer analyses, respectively. Notably,

30-day abstinence rates were comparable among those who were and were *not* ready to quit in the next 30 days at baseline: 27.1% (29/107) versus 27.8% (59/212; ITT), respectively. Among the study completers who did not achieve abstinence, 25.9% (44/170) reduced their CPD by ≥50% by the end of the study.

Readiness to Quit

This study was unique in its inclusion of individuals who were not ready to quit smoking in the next 30 days. This study design element reflected an overarching goal of Pivot to engage individuals along the entire spectrum of readiness to quit. Prochaska et al estimated that approximately 20% of smokers are thinking of quitting smoking in the next 30 days, 35% to 40% are thinking of quitting in the next 6 months, and 40% to 45% are not seriously thinking of quitting. They noted that cessation professionals, "approaching patients and settings only with action-oriented programs are likely to under-serve or misserve the majority of their target population" [44]. Despite the fact that the USCPG provides clear guidance on how to work with those who are ambivalent or otherwise not ready to quit in the next 30 days [4], few programs actively recruit or even include these individuals. Pivot had high enrollment of individuals not ready to quit in the next 30 days and comparable program retention and cessation rates between these individuals and those who were ready to quit in the next 30 days. This suggests Pivot addresses this historical shortcoming of focusing solely on those smokers who are ready to quit in the short term and has the potential to increase reach by effective inclusion of a larger proportion of the smoking population.

Comparison With Prior Work

In addition to inclusion of individuals who were not ready to quit smoking in the next 30 days, differentiating aspects of Pivot include the mobile CO breath sensor and dedicated in-app human coaching. These differences, along with those in program duration and study population characteristics such as employment status, age, and baseline CPD, limit comparability of these programs. Acknowledging these differences, Table 9 is included to provide context for current and previous abstinence rates from app-based cessation programs. Overall, Pivot abstinence rates are favorable when assessed among app-based cessation programs.

Recently, Krishnan et al assessed the use of a personal CO breath sensor (iCO monitor, Bedfont Scientific Ltd) with the COach2Quit app. Participants were randomized to receive brief advice (control, n=52) or brief advice plus the iCO monitor with the COach2Quit app (intervention, n=50). Participants in the intervention group were sent reminders from the app to use the breath sensor twice per day. On the basis of the user's CO result after a breath sample, the app sent messages from a predefined text library and provided a graphical display of the CO readings. At 30-day follow-up, 1 participant in each study arm had quit smoking. There were no significant differences between study groups in changes in CO levels or CPD between baseline and 30-day follow-up [45]. The use of a personal CO breath sensor was incorporated in this study and our study. However, there were several notable differences in the programs and study populations, including the intensity of and delivery mechanism of counseling, suggested breath sampling frequency, the app-delivered Pivot journey, and study population demographics (age, employment, and education levels). These differences may have contributed to the different outcomes.

Adherence and Attrition

Poor adherence and attrition are known problems with technology-based treatment platforms such as apps. Ubhi et al evaluated SmokeFree28, a 28-day app-based cessation program. Among the 1170 participants, 470 (40.2%) used the app for 7 days or more and only 226 (19.3%) used it for 28 days or more [40]. Iacoviello et al reported on the use of an app-based cessation program by 416 smokers. In that program, users chose a quit date 7 to 21 days from enrollment with study endpoint data collected at 8 weeks. Study participants used the program for an average of 5.3 weeks [15]. In Pivot, participants were actively engaged with the program for an average of 12.4 weeks, indicating durable use among participants.

Of 99 participants in a study assessing a digital cessation program based on ACT, Bricker et al reported that at 2-month follow-up, 24% of participants had completed the program [14]. In this study, 39.5% (126/319) of participants completed the Pivot program. The greatest attrition occurred during the Build stage; 64.7% (183/283) of those who started Build completed it, in contrast to ≥76.0% for the other Pivot stages. To complete Build and advance in Pivot, one must choose a quit date and create a quit plan. This requirement reflects the overall linear program flow, which requires completion of one stage before moving on to the next. This design may deter further advancement in those not ready to quit, and it may also hinder engagement and data collection from individuals on the other end of the spectrum, those who make a quit attempt before reaching Pivot's Quit phase. Although the overall Pivot program completion rate is favorable in the context of published rates, learnings from this study suggest the potential for app flow modifications that allow more flexibility in navigation, such that early quitters can access Quit content and those not ready to quit can maintain a sense of progress by accessing additional educational and motivational materials.

Table 9. 30-day point prevalence abstinence (PPA) in app-based cessation programs.

Author	Program components	Study characteristics	n	ITT[a] 30-day PPA[b], n (%)	Study completer 30-day PPA, n (%)
Bricker JB et al [14]	App: applies Acceptance and Commitment Therapy; users create a quit plan, complete 8 core modules, use "Urge Pass" tracker and "Anytime Coaching" (other ACT[c]-based activities)	Single arm pilot trial (assessment of program updates following previous RCT[d]); Mean age: 38 years; Female: 78%; Employed: 70%; CPD[e]: not reported, 32% smoked ≥1 pack per day; Seriously thinking of quitting in the next 30 days: 100%; Follow-up: 2 months	99	9 (9.1)	9 (11.0)
Iacoviello BM et al [15]	App: tailored quit plan of missions and personalized messages that adhere to USCPG[f]. Includes controlled breathing, personalized messaging, social engagement, encouragement of pharmacotherapy and medication adherence, and digital diversions	Single arm trial; Mean age: 36 years; Female: 59%; Employment not reported; CPD: 16.7; Seriously thinking of quitting in the next 30 days: 100%; Follow-up: 2 months	416	109 (26.2)	109 (29.9)
Marler JD et al (this study)	CO[g] breath sensor—user tracks effect of smoking on breath CO; App: applies USCPG; includes daily activities, challenges, cigarette logging, encouragement of pharmacotherapy, quit plan, practice quits, and check-ins; Coaching: asynchronous, dedicated human coaching via in-app text messaging, based on Cognitive Behavioral Therapy and Self-Determination Theory	Single arm trial; Mean age: 43 years; Female: 58%; Employed: 100% (≥20 hours a week); CPD: 17.7; Seriously thinking of quitting in the next 30 days: 33.5%; Follow-up: 14.5 to 18.5 weeks	319	88 (27.6)	88 (32.4)

[a]ITT: intention-to-treat.

[b]PPA: point prevalence abstinence.

[c]ACT: Acceptance and Commitment Therapy.

[d]RCT: randomized controlled trial.

[e]CPD: cigarettes per day.

[f]USCPG: United States Clinical Practice Guideline.

[g]CO: carbon monoxide.

Participant Experience: Carbon Monoxide Breath Sensing

Most study participants (233/272, 85.7%) indicated their motivation to quit smoking increased while using Pivot, reporting use of the breath sensor as the most common reason. Breath sensor use behavior during the study supports this preference and generally aligns with previous findings. Participants gave an average of 5.9 breath samples per day during Explore compared with 5.9 to 8.1 samples per day during Explore in a previous study [22]; comparison beyond Explore was not possible due to the shorter duration of the previous study. Assessment by Krishnan et al of a smoking cessation program that also utilized a personal CO breath sensor and accompanying app reported similar findings regarding reception of the technology; 91% of participants liked having the breath sensor and app to help them quit smoking, and 86% reported that using both motivated them to quit smoking [45]. Although the use of a personalized CO breath sensor for biometric feedback is a relatively new approach in smoking cessation programs, the data of this study are encouraging and engender additional questions for future study regarding the optimal use of this tool.

Indicators of Future Outcomes of Behavior

Abstinence rates are paramount in smoking cessation research, and the abstinence rates in this study were favorable in the context of similar programs. Nonetheless, nicotine addiction is a disease state that rarely, if ever, exceeds a 50% cure rate per treatment attempt with smokers making an average of 6 to 29 quit attempts before quitting successfully [46]. Accordingly, indicators of future outcomes or behavior are of particular interest. In this study, 28.1% (79/281) of participants progressed from one stage of readiness to the next at the end of Explore; this stage advancement is important because it translates to a doubling of one's chance of taking action in the next 6 months [44]. Among the study completers who did not achieve abstinence during the study, the average reduction in CPD at the end of the study was 29.1%, and 25.9% participants decreased their CPD by ≥50%. This behavior change is meaningful as the rates of quit attempts or cessation itself

significantly increase among those who reduce CPD by ≥50% [47].

Limitations

This study has a few limitations. First, the majority of participants (294/319, 92.2%) were iOS users. Ubhi et al assessed 1386 users of the SmokeFree28 app and reported differences in iOS and Android users. iOS users were more likely to have made a quit attempt within the last 12 months and set their quit date on the day of registration and were less likely to have used cessation medication to support their quit attempt compared with Android users [48]. Further study of Pivot in Android users will inform the understanding of if and how their outcomes differ from iOS users. Second, the eligibility criterion of employment ≥20 hours per week may limit the applicability of the results among the general public, particularly in individuals with lower socioeconomic status or serious mental health issues. Notably, despite the ≥20 hours per week employment eligibility requirement, approximately 25% of study participants had an annual household income less than US $35,000. As we did not collect information on mental health, we are unable to comment on the representation of individuals with mental health conditions in the study population. In addition, although we used the eligibility criterion of employment ≥20 hours a week as a proxy for individuals more likely to have access to an employee wellness program or employer-provided health care, study participants were not directly recruited from such employer-provided programs, thereby limiting direct conclusions in these populations. Third, although most study participants were not ready to quit in the next 30 days, only 3.5% (11/319) of all study participants fell into the specific category of "not thinking of quitting," limiting assessment in this group. Overall, the study results should be considered in the context of the aforementioned participant smartphone, employment, and readiness to quit characteristics, with additional study of Pivot needed in a broader population of smokers.

In addition, participants were compensated for their participation. Although efforts were made to minimize the influence of payments, such as keeping individual payments under US $50, incorporating a several week delay between questionnaire completion and payment receipt, and not linking payment to use of program components or smoking outcomes, we cannot exclude some influence of study payment on participant behavior. Finally, the study was conducted in a pragmatic fashion to assess the Pivot program as it is used in real-world contexts. Pivot is characterized by a sequential multistage design. These stages were created with certain time frame estimates in mind, specifically, that it would take 9 days, 1 to 28 days, 7 days, 7 days, and 11 weeks to complete Explore, Build, Mobilize, Quit, and Secure, respectively. However, actual participant behavior demonstrated longer durations spent in some stages (Mobilize through Secure), and the study was designed to allow for this type of participant behavior. This flexibility is valuable in that it mirrors how Pivot will be available to and used by its initial intended users (employees of self-insured employers or employers with wellness programs). Nonetheless, this approach is different from the more traditional and rigid 30- or 60-day assessments that are linked directly to enrollment date, and this difference is worth acknowledging.

Conclusions

Pivot is a comprehensive digital smoking program, which combines proven cessation principles derived from the USCPG with the innovation of a mobile CO breath sensor and in-app text-based human coaching. In this initial assessment, Pivot was shown to be engaging and quit rates were aligned with those in the peer-reviewed literature. This was true in both individuals who were and were not ready to quit in the next 30 days, with comparable retention and cessation outcomes between the 2 groups. Pivot leverages accessible and nearly ubiquitous smartphone technology and engages individuals along the entire spectrum of readiness to quit—something few programs have done to date. The hope is this combination will translate to increased reach; these initial results are encouraging. Looking forward, this study informs the future development of Pivot. Next steps include implementing and evaluating refinements to the program based on present learnings and further evaluation via a randomized controlled trial.

Acknowledgments

The authors would like to thank the scientific advisors, Drs Judith (Jodi) Prochaska, Danielle E Ramo, and Sharon M Hall for their contributions to the development of Pivot. The authors would also like to thank Susanna Walsh and Laura Dotson for assistance with recruiting, technical support, and collaborative tracking of participants. In addition, the authors would like to thank Matt Day and Corrie Mills for assistance with recruiting and all the coaches for the clinical expertise and compassion they bring to all Pivot participants. Joseph A Galanko's work is supported by National Institutes of Health (NIH) Grant P30 DK 034987.

Authors' Contributions

JDM, HP, DSU, and CAF designed the study. LJT, JDM, and CAF recruited participants. JDM ran the study. LJT performed administrative study functions and assisted with database management and analysis. CAF managed the database and conducted data analyses. JAG conducted statistical data analyses. JDM and HP prepared the original draft of the manuscript. HP, DSU, CAF, JAG, and LJT reviewed and provided comments on the manuscript before submission.

Conflicts of Interest

JDM, HP, CAF, LJT, and DSU are employees of Carrot Inc, the developer of the app and devices used in this study. DSU is the President and CEO of Carrot Inc and an investor in the company.

References

1. Centers for Disease Control and Prevention. 2018. Current cigarette smoking among adults in the United States URL: https://www.cdc.gov/tobacco/data_statistics/fact_sheets/adult_data/cig_smoking/index.htm [accessed 2018-10-15] [WebCite Cache ID 73BsFDCm2]
2. Jamal A, King BA, Neff LJ, Whitmill J, Babb SD, Graffunder CM. Current cigarette smoking among adults - United States, 2005-2015. MMWR Morb Mortal Wkly Rep 2016 Nov 11;65(44):1205-1211 [FREE Full text] [doi: 10.15585/mmwr.mm6544a2] [Medline: 27832052]
3. Centers for Disease Control and Prevention. Early Release of Selected Estimates Based on Data From the National Health Interview Survey, January–March 2017 URL: https://www.cdc.gov/nchs/data/nhis/earlyrelease/Earlyrelease201709_08.pdf [accessed 2018-10-14] [WebCite Cache ID 73BvP5wOr]
4. US Department of Health and Human Services, Public Health Service. Treating Tobacco Use and Dependence: 2008 Update - Clinical Practice Guideline. Scotts Valley, CA, USA: Createspace Independent Publishing Platform; 2019.
5. Siu AL, US Preventive Services Task Force. Behavioral and pharmacotherapy interventions for tobacco smoking cessation in adults, including pregnant women: US Preventive Services Task Force recommendation statement. Ann Intern Med 2015 Oct 20;163(8):622-634. [doi: 10.7326/M15-2023] [Medline: 26389730]
6. Patnode CD, Henderson JT, Thompson JH, Senger CA, Fortmann SP, Whitlock EP. Behavioral counseling and pharmacotherapy interventions for tobacco cessation in adults, including pregnant women: a review of reviews for the US Preventive Services Task Force. Ann Intern Med 2015 Oct 20;163(8):608-621. [doi: 10.7326/M15-0171] [Medline: 26389650]
7. US Surgeon General. 2018. The Health Consequences of Smoking—50 Years of Progress: A Report of the Surgeon General URL: https://www.surgeongeneral.gov/library/reports/50-years-of-progress/full-report.pdf [accessed 2018-10-11] [WebCite Cache ID 73BvjiHl3]
8. Babb S, Malarcher A, Schauer G, Asman K, Jamal A. Quitting smoking among adults - United States, 2000-2015. MMWR Morb Mortal Wkly Rep 2017 Jan 6;65(52):1457-1464 [FREE Full text] [doi: 10.15585/mmwr.mm6552a1] [Medline: 28056007]
9. Centers for Disease Control and Prevention. 2014. Best Practices for Comprehensive Tobacco Control Programs 2014 URL: https://www.cdc.gov/tobacco/stateandcommunity/best_practices/pdfs/2014/comprehensive.pdf [WebCite Cache ID 73DCGYVLG]
10. Pew Research Center. 2018. Mobile fact sheet URL: http://www.pewinternet.org/fact-sheet/mobile/ [WebCite Cache ID 73DCVRPYU]
11. Abroms LC, Westmaas JL, Bontemps-Jones J, Ramani R, Mellerson J. A content analysis of popular smartphone apps for smoking cessation. Am J Prev Med 2013 Dec;45(6):732-736 [FREE Full text] [doi: 10.1016/j.amepre.2013.07.008] [Medline: 24237915]
12. BinDhim NF, McGeechan K, Trevena L. Who uses smoking cessation apps? A feasibility study across three countries via smartphones. JMIR Mhealth Uhealth 2014 Feb 6;2(1):e4 [FREE Full text] [doi: 10.2196/mhealth.2841] [Medline: 25098439]
13. Haskins BL, Lesperance D, Gibbons P, Boudreaux ED. A systematic review of smartphone applications for smoking cessation. Transl Behav Med 2017 Dec;7(2):292-299 [FREE Full text] [doi: 10.1007/s13142-017-0492-2] [Medline: 28527027]
14. Bricker JB, Copeland W, Mull KE, Zeng EY, Watson NL, Akioka KJ, et al. Single-arm trial of the second version of an acceptance & commitment therapy smartphone application for smoking cessation. Drug Alcohol Depend 2017 Dec 1;170:37-42 [FREE Full text] [doi: 10.1016/j.drugalcdep.2016.10.029] [Medline: 27870987]
15. Iacoviello BM, Steinerman JR, Klein DB, Silver TL, Berger AG, Luo SX, et al. Clickotine, a personalized smartphone app for smoking cessation: initial evaluation. JMIR Mhealth Uhealth 2017 Apr 25;5(4):e56 [FREE Full text] [doi: 10.2196/mhealth.7226] [Medline: 28442453]
16. Beard E, West R. Pilot study of the use of personal carbon monoxide monitoring to achieve radical smoking reduction. J. Smok Cessat 2012 Jul 13;7(01):12-17 [FREE Full text] [doi: 10.1017/jsc.2012.1]
17. Choi WY, Kim CH, Lee OG. Effects of brief smoking cessation education with expiratory carbon monoxide measurement on level of motivation to quit smoking. Korean J Fam Med 2013 May;34(3):190-198 [FREE Full text] [doi: 10.4082/kjfm.2013.34.3.190] [Medline: 23730486]
18. Risser NL, Belcher DW. Adding spirometry, carbon monoxide, and pulmonary symptom results to smoking cessation counseling: a randomized trial. J Gen Intern Med 1990;5(1):16-22. [doi: 10.1007/BF02602303] [Medline: 2405112]
19. Jamrozik K, Vessey M, Fowler G, Wald N, Parker G, van Vunakis H. Controlled trial of three different antismoking interventions in general practice. Br Med J (Clin Res Ed) 1984 May 19;288(6429):1499-1503 [FREE Full text] [doi: 10.1136/bmj.288.6429.1499] [Medline: 6426618]

20. Fagerström KO. Assessment of the smoker who wants to quit. Monaldi Arch Chest Dis 2001 Apr;56(2):124-127. [Medline: 11499300]
21. Bittoun R. Carbon monoxide meter: the essential clinical tool—the 'Stethoscope'—of smoking cessation. J Smok Cessat 2012 Feb 21;3(02):69-70 [FREE Full text] [doi: 10.1375/jsc.3.2.69]
22. Patrick H, Fujii CA, Glaser DB, Utley DS, Marler JD. A comprehensive digital program for smoking cessation: assessing feasibility in a single-group cohort study. JMIR Mhealth Uhealth 2018 Dec 18;6(12):e11708 [FREE Full text] [doi: 10.2196/11708] [Medline: 30563807]
23. Society for Human Resource Management. 2018 Employee Benefits: The Evolution of Benefits URL: https://www.shrm.org/hr-today/trends-and-forecasting/research-and-surveys/Documents/2018%20Employee%20Benefits%20Report.pdf [accessed 2018-10-15] [WebCite Cache ID 73DD1Wodz]
24. Bureau of Labor Statistics. Employee access to wellness programs in 2017 URL: https://www.bls.gov/opub/ted/2018/employee-access-to-wellness-programs-in-2017.htm [accessed 2018-10-16] [WebCite Cache ID 73DD9RpWK]
25. Bureau of Labor Statistics. Median age of the labor force, by sex, race and ethnicity URL: https://www.bls.gov/emp/tables/median-age-labor-force.htm [accessed 2018-10-16] [WebCite Cache ID 73DDH9dM4]
26. Centers for Disease Control and Prevention. 2014. Coverage for tobacco use cessation treatments URL: https://www.cdc.gov/tobacco/quit_smoking/cessation/pdfs/coverage.pdf [accessed 2018-06-29] [WebCite Cache ID 73DDWzpo2]
27. Stead LF, Perera R, Bullen C, Mant D, Hartmann-Boyce J, Cahill K, et al. Nicotine replacement therapy for smoking cessation. Cochrane Database Syst Rev 2012 Nov 14;11:CD000146. [doi: 10.1002/14651858.CD000146.pub4] [Medline: 23152200]
28. US Preventive Services Task Force. 2018. Final Recommendation Statement - Tobacco Smoking Cessation in Adults, Including Pregnant Women: Behavioral and Pharmacotherapy Interventions URL: https://www.uspreventiveservicestaskforce.org/Page/Document/RecommendationStatementFinal/tobacco-use-in-adults-and-pregnant-women-counseling-and-interventions1 [accessed 2018-10-16] [WebCite Cache ID 73DE8FUi2]
29. Miller WR, Rollnick S. Motivational Interviewing: Preparing People For Change Second Edition. New York City: The Guilford Press; 2002.
30. Lindson-Hawley N, Thompson TP, Begh R. Motivational interviewing for smoking cessation. Cochrane Database Syst Rev 2015 Mar 2(3):CD006936. [doi: 10.1002/14651858.CD006936.pub3] [Medline: 25726920]
31. Williams GC, Niemiec CP, Patrick H, Ryan RM, Deci EL. The importance of supporting autonomy and perceived competence in facilitating long-term tobacco abstinence. Ann Behav Med 2009 Jun;37(3):315-324 [FREE Full text] [doi: 10.1007/s12160-009-9090-y] [Medline: 19373517]
32. Abrams DB, Niaura R, Brown RA, Emmons KM, Goldstein MG, Monti PM. The Tobacco Dependence Treatment Handbook: A Guide To Best Practices. New York City: The Guilford Press; 2003.
33. Lancaster T, Stead LF. Individual behavioural counselling for smoking cessation. Cochrane Database Syst Rev 2017 Dec 31;3:CD001292. [doi: 10.1002/14651858.CD001292.pub3] [Medline: 28361496]
34. DiClemente CC, Prochaska JO, Fairhurst SK, Velicer WF, Velasquez MM, Rossi JS. The process of smoking cessation: an analysis of precontemplation, contemplation, and preparation stages of change. J Consult Clin Psychol 1991 Apr;59(2):295-304. [doi: 10.1037/0022-006X.59.2.295] [Medline: 2030191]
35. Spencer L, Pagell F, Hallion ME, Adams TB. Applying the transtheoretical model to tobacco cessation and prevention: a review of literature. Am J Health Promot 2002;17(1):7-71. [doi: 10.4278/0890-1171-17.1.7] [Medline: 12271754]
36. Hall SM, Havassy BE, Wasserman DA. Commitment to abstinence and acute stress in relapse to alcohol, opiates, and nicotine. J Consult Clin Psychol 1990 Apr;58(2):175-181. [doi: 10.1037/0022-006X.58.2.175] [Medline: 2335634]
37. Hall SM, Havassy BE, Wasserman DA. Effects of commitment to abstinence, positive moods, stress, and coping on relapse to cocaine use. J Consult Clin Psychol 1991 Aug;59(4):526-532. [doi: 10.1037/0022-006X.59.4.526] [Medline: 1918556]
38. SRNT Subcommittee on Biochemical Verification. Biochemical verification of tobacco use and cessation. Nicotine Tob Res 2002 May;4(2):149-159. [doi: 10.1080/14622200210123581] [Medline: 12028847]
39. Vangeli E, Stapleton J, Smit ES, Borland R, West R. Predictors of attempts to stop smoking and their success in adult general population samples: a systematic review. Addiction 2011 Dec;106(12):2110-2121. [doi: 10.1111/j.1360-0443.2011.03565.x] [Medline: 21752135]
40. Ubhi HK, Michie S, Kotz D, Wong WC, West R. A mobile app to aid smoking cessation: preliminary evaluation of SmokeFree28. J Med Internet Res 2015 Jan 16;17(1):e17 [FREE Full text] [doi: 10.2196/jmir.3479] [Medline: 25596170]
41. Consumer Health Information Corporation. 2018. Motivating patients to use smartphone apps URL: http://www.consumer-health.com/motivating-patients-to-use-smartphone-health-apps/ [accessed 2018-10-16] [WebCite Cache ID 73DEsl6SE]
42. Hallgren KA, Witkiewitz K. Missing data in alcohol clinical trials: a comparison of methods. Alcohol Clin Exp Res 2013 Dec;37(12):2152-2160 [FREE Full text] [doi: 10.1111/acer.12205] [Medline: 23889334]
43. Nelson DB, Partin MR, Fu SS, Joseph AM, An LC. Why assigning ongoing tobacco use is not necessarily a conservative approach to handling missing tobacco cessation outcomes. Nicotine Tob Res 2009 Jan;11(1):77-83. [doi: 10.1093/ntr/ntn013] [Medline: 19246444]

44. Prochaska JO, Norcross JC, Diclemente C. Applying the stages of change. Psychother Aust 2013;19:15. [doi: 10.1093/med:psych/9780199845491.003.0034]
45. Krishnan N, Elf JL, Chon S, Golub JE. COach2Quit: a pilot randomized controlled trial of a personal carbon monoxide monitor for smoking cessation. Nicotine Tob Res 2018 Aug 29. [doi: 10.1093/ntr/nty182] [Medline: 30169740]
46. Chaiton M, Diemert L, Cohen JE, Bondy SJ, Selby P, Philipneri A, et al. Estimating the number of quit attempts it takes to quit smoking successfully in a longitudinal cohort of smokers. BMJ Open 2016 Dec 9;6(6):e011045 [FREE Full text] [doi: 10.1136/bmjopen-2016-011045] [Medline: 27288378]
47. Hughes JR. Reduced smoking: an introduction and review of the evidence. Addiction 2000 Jan;95 Suppl 1:S3-S7. [doi: 10.1046/j.1360-0443.95.1s1.5.x] [Medline: 10723815]
48. Ubhi HK, Kotz D, Michie S, van Schayck OC, West R. A comparison of the characteristics of iOS and Android users of a smoking cessation app. Transl Behav Med 2017 Dec;7(2):166-171 [FREE Full text] [doi: 10.1007/s13142-016-0455-z] [Medline: 28168609]

Abbreviations

ACT: Acceptance and Commitment Therapy
CO: carbon monoxide
CPD: cigarettes per day
FDA: Food and Drug Administration
ITT: intention-to-treat
PPA: point prevalence abstinence
USCPG: United States Clinical Practice Guideline

Edited by G Eysenbach; submitted 25.10.18; peer-reviewed by A Munger, B Iacoviello, S Luo, K Hallgren; comments to author 12.12.18; revised version received 18.01.19; accepted 20.01.19; published 04.02.19.

Please cite as:
Marler JD, Fujii CA, Utley DS, Tesfamariam LJ, Galanko JA, Patrick H
Initial Assessment of a Comprehensive Digital Smoking Cessation Program That Incorporates a Mobile App, Breath Sensor, and Coaching: Cohort Study
JMIR Mhealth Uhealth 2019;7(2):e12609
URL: http://mhealth.jmir.org/2019/2/e12609/
doi:10.2196/12609
PMID:30670372

Original Paper

Evaluating Mobile Health Apps for Customized Dietary Recording for Young Adults and Seniors: Randomized Controlled Trial

Ying-Chieh Liu[1,2], PhD; Chien-Hung Chen[3], MSc; Ya-Chi Tsou[1], MSc, RD; Yu-Sheng Lin[2], MD; Hsin-Yun Chen[4], MSc; Jou-Yin Yeh[5], BS; Sherry Yueh-Hsia Chiu[5,6], PhD

[1]Department of Industrial Design, College of Management, Chang Gung University, Taoyuan, Taiwan

[2]Health Promotion Center, Department of Internal Medicine, Chang Gung Memorial Hospital, Taoyuan, Taiwan

[3]Cyber Security Technology Institute, Institute for Information Industry, Taipei, Taiwan

[4]Department of Nutrition Therapy, Chang Gung Memorial Hospital, Taoyuan, Taiwan

[5]Department of Health Care Management and Healthy Aging Research Center, College of Management, Chang Gung University, Taoyuan, Taiwan

[6]Division of Hepatogastroenterology, Department of Internal Medicine, Kaohsiung Chang Gung Memorial Hospital, Kaohsiung, Taiwan

Corresponding Author:
Sherry Yueh-Hsia Chiu, PhD
Department of Health Care Management and Healthy Aging Research Center
College of Management
Chang Gung University
259, Wen-Hwa 1st Road
Kwei-Shan
Taoyuan, 333
Taiwan
Phone: 886 3 2118800 ext 5250
Fax: 886 3 2118345
Email: sherrychiu@mail.cgu.edu.tw

Abstract

Background: The role of individual-tailored dietary recording in mobile phone health apps has become increasingly important in management of self-health care and population-based preventive service. The development of such mobile apps for user-centered designing is still challengeable and requires further scientific evidence.

Objective: This study aims to conduct a randomized trial to assess the accuracy and time efficiency of two prototypes for dietary recoding utilization related to the input method of food intake.

Methods: We first present an innovative combinatorial concept for dietary recording to account for dish variation. One prototype was a self-chosen tab app that featured choosing each food ingredient to synthesize an individual dish, whereas the other was an autonomous exhaustive list app that provided one selection from a comprehensive list of dish items. The concept included commercially available choices that allowed users to more accurately account for their individual food selection. The two mobile apps were compared in a head-to-head parallel randomized trial evaluation. Young adults (n=70, aged 18-29) and older adults (n=35, aged 55-73) were recruited and randomized into two groups for accuracy and response time evaluation based on 12 types of food items in use of the developed self-chosen tab and autonomous exhaustive list apps, respectively.

Results: For the trials based on the self-chosen tab (53 participants) and autonomous exhaustive list groups (52 participants), the two prototypes were found to be highly accurate (>98%). The self-chosen tab app was found to be more efficient, requiring significantly less time for input of 11 of 12 items (P<.05). The self-chosen tab users occasionally neglected to select food attributes, an issue which did not occur in the autonomous exhaustive list group.

Conclusions: Our study contributes through the scientific evaluation of the transformation step into prototype development to demonstrate that a self-chosen tab app has potentially better opportunity in effectiveness and efficiency. The combinatorial concept offers potential for dietary recording and planning which can account for high food item variability. Our findings on prototype development of diversified dietary recordings provide design consideration and user interaction for related further app development and improvement.

Trial Registration: ISRCTN Registry ISRCTN86142301; http://www.isrctn.com/ ISRCTN86142301 (Archived by WebCite at http://www.webcitation.org/74YLEPYnS)

(JMIR Mhealth Uhealth 2019;7(2):e10931) doi:10.2196/10931

KEYWORDS

customized dietary recording; prototypes; user-centered design; utilization; mobile health; mHealth; randomized trial

Introduction

As part of a new direction in food service marketing [1], many fast food chains, restaurants, and coffee and tea shops now encourage customers to select food alternatives to meet their special or individual needs. Programs in such places at Wendy's [2] and Starbucks [3] allow customers to select alternative cooking methods and ingredient types and portions, thus helping customers meet their individual taste and dietary preferences.

However, the diversity of foods in real-world contexts poses a significant challenge to such services accurately accounting for actual food intake. Recent advances in information and communication technology have led to the development of innovative methods for reporting dietary intake using mobile phones in a domain referred to as mHealth [4,5]. Accurate dietary reporting is fundamental to counting calories and calculating nutrient intake. Examples of innovation in dietary intake include sized photography for food portion measurement [6], expert direct observation of photographs [7,8], using digital images for the assessment of food intake [9], and using mobile phone cameras to capture food images [10]. However, new methods are still lacking to prescribe individualized food alternatives that not only provide valid intake data, but also are suited for different types of users. Recently, the emphasis has trended toward providing more innovative and effective applications in digital health interventions. Essential components of the framework in development of mHealth solutions are utilized to leverage the potential outcomes (ie, theory of user-centered innovation) [11,12], mobile health user-centered design [13-15], and a theory-driven and user-centered approach [13]. Previous studies [14-16] have emphasized integrating these components in a comprehensive framework.

In this study, we investigate the prototyping step of a mobile health app. Prototypes commonly evolve from original concept in more than one design variant (ie, divergent steps) [12]. The usability of these design variants is subjected to systematic evaluation and requires comparative evidence from user interaction [17]. This research conducts a scientific evaluation to explicate the differences of the two design variants. The consolidated statistical results provide a better understanding of prototype suitability. Several pilot studies have attempted to develop and evaluate a specific design [18-20]. However, such attempts have largely failed to consider design variants or conduct randomized trials to evaluate efficacy.

This paper presents an innovative concept for selecting or creating individual meals for a mobile health app. The proposed concept is used to develop two mobile apps to help users select a wide variety of food alternatives. The first app is a self-chosen tab app, which allows users to choose and click each food ingredient to synthesize a food. The second app is an autonomous exhaustive list app, in which users scroll through and select from a comprehensive list of combined food ingredients. The concept aims to help users specify the desired food item with detailed food information (eg, sugar content, method of preparation) and capture various food ingredient types. The efficacies of the two methods were assessed experimentally.

Methods

App Design

We developed our apps based on user-centered design approaches [18,20-22]. Key development steps included a review of the relevant literature and commercial mobile apps. Innovative design ideas were brainstormed and then reviewed to develop concepts for helping users of different age groups select individual-tailored food items. A physician and two dieticians evaluated the initial prototypes in terms of usability, and the completeness and accuracy of dietary information. Two alternatives were finally identified. Further, these two prototypes were evaluated from a usability perspective in randomized trials to compare usability in accuracy and response time to reflect actual food items in predefined meals.

Combinatorial Concept

The combinatorial concept for customized food designation was addressed. Foods were represented in terms of one or more main food ingredients as well as their related choices of side ingredients that affect the caloric and macronutrient content of each food item. Side ingredients are named "food attributes" in this paper. Each main food ingredient was culturally recognized as representing a major food group. For example, the main food ingredient of coffee mocha was the coffee part without milk, sugar, or other additives. The food attributes of coffee mocha included all commercially available choices, such as various types of milk, quantities of sugar, and flavorings (eg, cinnamon power). In this paper, mocha has up to three types of food attributes: toppings (tapioca pearls, hsian-tsao herbal tea jelly, ice cream, coconut konjac jelly), type of milk (whole, low fat, fat free), and sugar quantity (regular, less, half, quarter, and none). Other food items might involve different cooking methods (steam, boil, stir fry, deep fry, pan fry, and salad). In another example, a "stir-fried egg with tomatoes" consists of two main ingredients (egg and tomatoes). Food attributes include the method of preparation (stir frying) which determines the quantity of oil used.

Main food ingredients are organized in subgroups in a tree-like structure. In summary, the overall structure is presented in terms of group, subgroup, and main ingredient. The prototypes have 14 food groups, containing 49 subgroups, which, in turn, contain more than 1000 main food ingredients. Group and subgroup classifications are designed with input from a senior dietitian, and reflect widely used classification schemes. Food groups and corresponding food attribute choices are subject to variation depending on culture and habit. Finally, "mixed foods" in our

study incorporate two or more main food ingredients to account for new food items not included in the other food groups.

Design of Prototypes

Based on the concept, two prototypes were implemented in the Android operating system for use in mobile devices. The two prototypes share a common interface and procedures for selecting food groups and subgroups. The two prototypes then differ in terms of operations used to determine corresponding food attributes. The self-chosen tab allowed users to choose required food attribute(s) to compose a food item. In the autonomous exhaustive list, users scrolled through and selected from a comprehensive list of food items including different food attributes.

The first screen in each prototype included a scrollable list of food groups, showing five groups at a time. In the self-chosen tab app (see Figure 1 a), each field in the list featured a colorful icon on the left and a textual description on the right. However, in the autonomous exhaustive list app (see Figure 2 a), it only provided a textual description. The background of each field featured alternating white and gray lines for additional clarity. The contrast of the text and the background was also considered for better readability. The text font used was Microsoft JhengHei. In a 7-inch display, individual characters measured approximately 1.5 × 1.0 cm.

The second and third screens (Figure 1 b,c and Figure 2 b,c) respectively present the subgroup and main food ingredients. The layout design, mode of user interaction, and text features are identical to those of the first screen.

In the self-chosen tab app (see Figure 1 d-f), the user first selects the main food ingredient and food attributes. These choices are then displayed in individual tabs. The user then taps a food attribute (eg, “sugar”) and selects a descriptive characteristic (eg, “half sugar”; Figure 1 e). Once the required food attribute(s) is specified, the user clicks the “confirm” button (Figure 1 f) to complete the food item (Figure 1 g). The entire operation sequence is shown in Multimedia Appendix 1.

For the autonomous exhaustive list app, after selecting the food group and subgroup, the user then selects the main food ingredient (Figure 2 c) and is then presented with a scrolling list (Figure 2 d). Each item in the list provides a textual description of the main food ingredient and corresponding combinations of food attributes separated by commas (eg, “mocha, 5 points sugar, tapioca, zero fat”). This list provides all possible combinations of side ingredients. Using coffee mocha as an example, there are six toppings choices, five milk choices, and five sugar quantity choices for a permutation of 150 variants (6*5*5=150). Users scroll through the list to select the desired combination. The entire operation sequence is shown in Multimedia Appendix 2.

Another interface is used to select mixed foods (see Combinatorial Concept in Methods). Beginning with the “mixed food” option on the first screen, the user is presented with two empty frames in the upper half of the screen (see Multimedia Appendix 3 and Multimedia Appendix 4), which allow users to select two of five subgroups from icons in the lower section of the screen. This design allows users to combine two or three main food ingredients from the following subgroups: meat, eggs, mushrooms, vegetables, and beans. Having selected the main food ingredients to be combined, the user then selects the corresponding food attributes, including the method of preparation.

In the self-chosen tab app, the user selects two subgroups that are presented with two corresponding scrolling lists (see Multimedia Appendix 3) from which the user then selects one main food ingredient. The users then select the proper associated attribute (see the lower section in Multimedia Appendix 3). The operation of this sequence is shown in Multimedia Appendix 5.

In the autonomous exhaustive list app, after selecting two subgroups, the user is then presented with a scrollable list containing all possible combinations of the main food ingredients and food attributes (Multimedia Appendix 4). The operation sequence is shown in Multimedia Appendix 6.

Study Design and Participant Recruitment

A parallel two-group randomized trial was designed to evaluate and compare the effectiveness of the self-chosen tab and autonomous exhaustive list apps in terms of correct food input accuracy and end user response time. The study protocol was reviewed by the Ethics Committee of Chang Gung Memorial Hospital and received approval from the Institutional Review Board (103-2745B, ISRCTN 86142301). Participants aged 18 to 29 years and 55 to 73 years were recruited through local colleges and hospitals, respectively. Those who had severe diseases were excluded.

For the young adults’ recruitment, we announced and introduced the study flowchart and contents before course start, which was carried out by one master’s student. The participant’s contact information was noted at the bottom of the flowchart, so they could be reached for participation registry. Conversely, the elder participants were introduced to this study by a research assistant in Chang Gung Memorial Hospital. Those who were willing to participate were noted on the list. Baseline data and informed consent were acquired following registration.

Randomization

To ensure an even age distribution, two random number lists were generated by SAS software [23]. Our recruitment and implementation were performed based on the order of randomization lists with a 1:1 ratio. Individual appointments were then made for evaluation. Overall, there were 53 (36 young adults and 17 older adults) and 52 (34 young adults and 18 older adults) participants assigned to the self-chosen tab and autonomous exhaustive list groups, respectively.

Figure 1. The self-chosen tab interface design and operation.

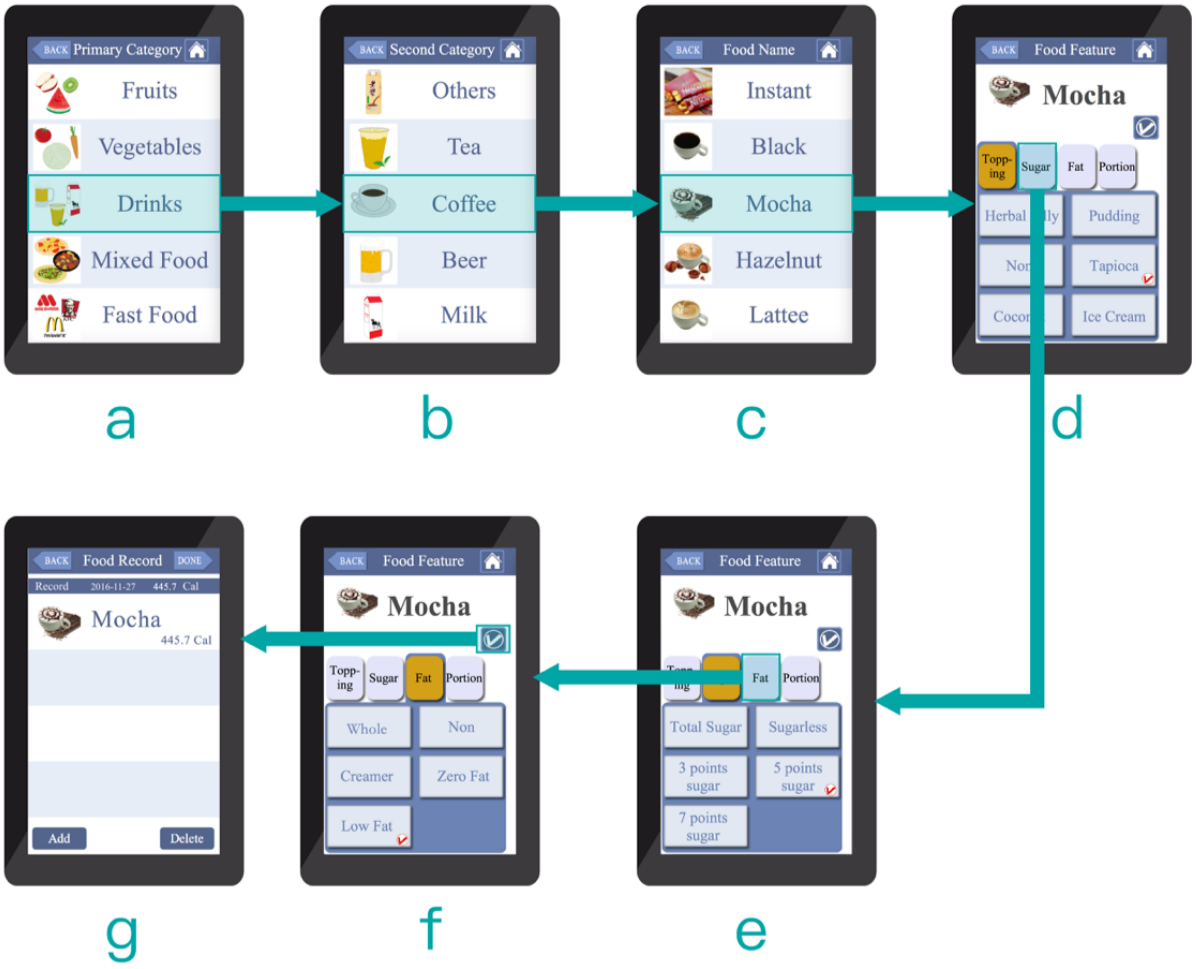
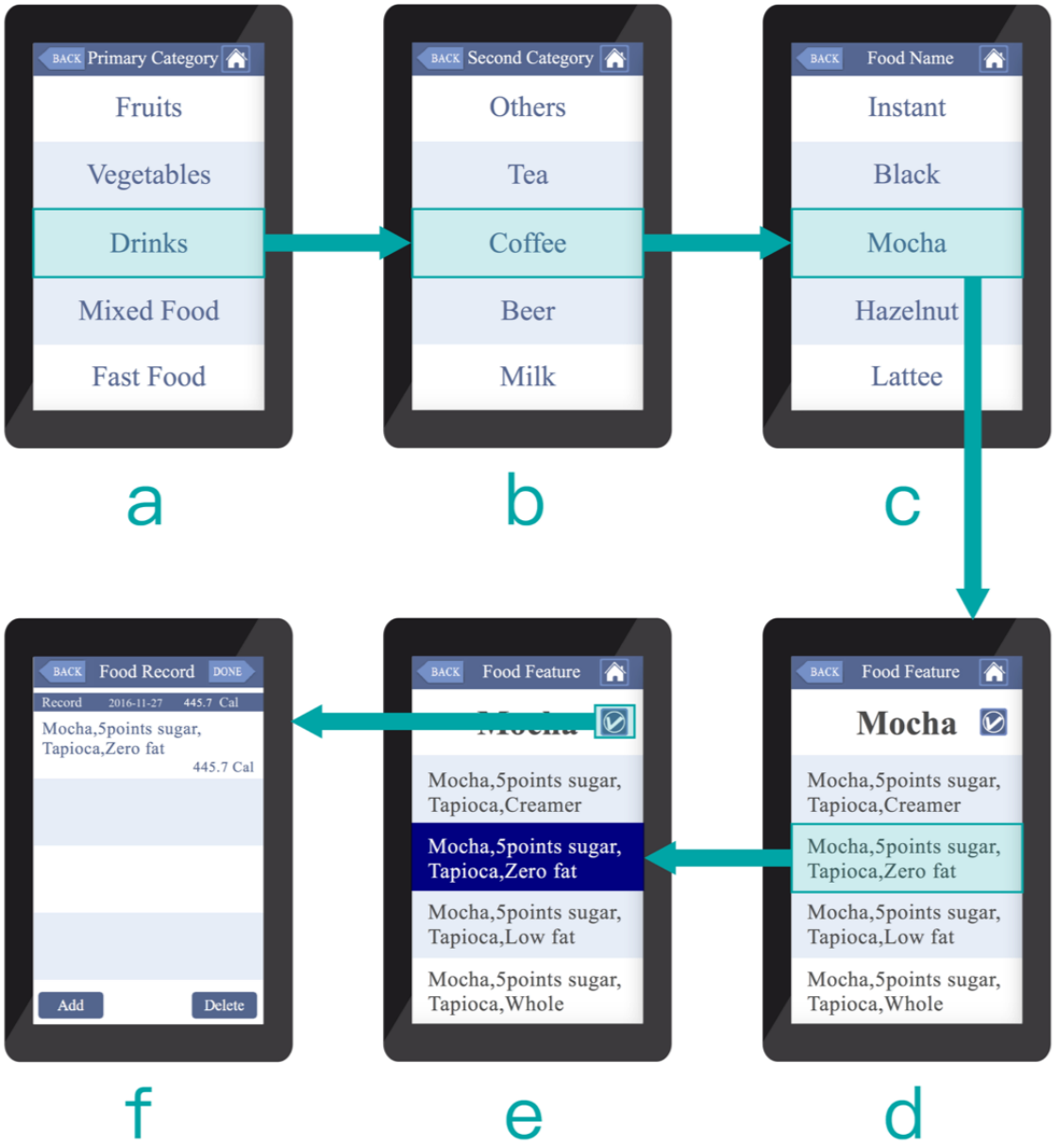


Figure 2. The autonomous exhaustive list interface design and operation.



Evaluation Outcomes

App usability was assessed in terms of accuracy and response time as the primary endpoint in the task of reporting the food items. Accuracy was defined as the number of correct counts divided by the overall counts. “Correct” was defined as the participant selecting the correct main food ingredient(s) as well as the correct food attribute(s) for each of the 12 items, given unlimited switching among groups and/or subgroups. Response time was recorded and embedded in apps in milliseconds for the time elapsed from a user’s selection (clicking) of a certain main food ingredient (see Figure 1 c and Figure 2 c) to complete in food attribute(s) selection and to click at the “confirm” button (see Figure 1 f and Figure 2 e).

Assessment Procedures

The assessment was performed by a research assistant who first administered a basic background questionnaire to analyze the distribution of relevant experience. The questionnaire collected participants’ self-reported baseline information, including gender, age, body mass index (BMI), department/unit, and experience with nutrition-related courses, health education program, and cooking. The research assistant then demonstrated the use of both apps through one meal with four food items (steamed sweet potato, boiled goose meat, stir-fried mushroom, fried tofu, and apple juice) to familiarize participants with app operation. After the demonstration, each user was allowed to practice app operation for 3 minutes to warm up. Each participant was asked to observe two actual meals and to record each item in one prototype. Each meal represented a typical lunch or dinner meal. The meals (see Multimedia Appendix 7) were prepared with real food in appropriate portions and

presented consistently on a plate throughout the experiment. All food items were accompanied by clearly visible and comprehensible labels to prevent errors due to misidentification. All the measures, including onscreen responses and time durations, were automatically collected within the mobile app. The assessment period was from June 2014 to January 2015. No harm or unintended effects were seen in either group.

The first meal featured Chinese steamed bread, a fried chicken drumstick, cold tofu with preserved eggs, stir-fried eggs with shredded carrot, stir-fried napa cabbage with bacon, and green tea. The second meal featured boiled rice, fried pork chop, stir-fried shredded pork with green pepper, stir-fried egg with tomatoes, stir-fried bitter melon with salted duck eggs, and milk tea.

Participants were given 3 minutes to complete each task. Having completed the first meal, the participant continued to label the second meal without taking a rest. All participants completed the assessment.

In the self-chosen tab group, each participant first observed the meal on the table. They then scrolled through first screen (ie, food group list) and selected the appropriate entry for each group. This then launched the second screen (ie, the subgroup list), followed by the third screen (ie, main food ingredient list). If the participant was unable to find the desired main food ingredient, he/she could click the “back” button at the top left of the screen to return to the previous screen for reselection.

After choosing the desired main food ingredient, the participant then selected the corresponding food attributes by switching to the other tabs. After selection was complete, the participant clicked the “confirm” button to complete the selection.

Having selected the staple food, main course, and beverage dishes, the participant then clicked the “mixed food” item on the first screen. On the mixed food screen (see Multimedia Appendix 3), participants were allowed to select two subgroups. In the main food ingredient screen (see Multimedia Appendix 3), the participant scrolled through the menu to find and select the desired main food ingredient, and then confirmed the selection. Once the desired main food ingredients had been selected, the participant clicked the desired method of preparation.

In the autonomous exhaustive list group, participants followed the same operations as self-chosen tab for the first and second screens. However, in the third (mixed food) screen, the app presented a comprehensive list of foods from which the participant would select the desired food item.

Analysis

For the baseline questionnaire information, the dichotomous variables (eg, gender, unit, and experience) were demonstrated by count for both the self-chosen tab and autonomous exhaustive list groups. The continuous variables were also collected by numeric value with individual level and analyzed by each group, such as age and BMI for baseline, and time consumed in seconds. We conducted a chi square test and *t* test for dichotomous and continuous variables for proportion and mean examination of two-group comparisons, respectively. When the difference in duration of task completion (measured in seconds) was used as a continuous variable, we used the independent *t* test to compare the two groups based on the intention-to-treat principle. All statistical tests were two-tailed, and *P* values less than .05 were considered statistically significant. All statistical analyses were performed using SAS software version 9.4.

As for the determination of sample size, our primary outcomes were both based on the accuracy and time consumed of reporting food items. Estimating the required sample size was based on a previous study [6]. Given a statistical power of 80% and a two-tailed alpha level of 5%, between the self-chosen tab and autonomous exhaustive list approaches, the sample size requirement for each arm was determined to be 50 participants under 10% accuracy difference and 37 participants assumed 4 seconds mean difference with standard deviation of 6 seconds. Therefore, a minimal sample size with 50 participants was required for each arm.

Results

Characteristics of the Participants

A total of 105 participants completed the tests, including 70 university students recruited from Chang Gung University (18-29 years) and 35 older adults recruited among volunteers at Chang Gung Memorial Hospital (55-73 years). The participants were assigned at random to the self-chosen tab (53 participants) and autonomous exhaustive list groups (52 participants), with source populations in each group represented in proportion to the overall population (see Figure 3) for an overall mean age of 35 (SD 19.50) years. Experience and expertise in nutrition, general health education, and cooking did not differ significantly between the two groups. The baseline information distributions did not reveal significant differences among the two groups, thus confirming the random allocation by chance (see Table 1).

Completion Accuracy and Trial and Error

Table 2 summarizes the response accuracy for all 12 food items. The self-chosen tab and autonomous exhaustive list groups achieved respective overall accuracy levels of 97.77% (1228/1256) and 98.53% (1214/1232). The most frequently mislabeled food items in self-chosen tab group were green tea and stir-fried shredded pork and green pepper; green tea was the most frequently mislabeled item in autonomous exhaustive list. In the self-chosen tab group, 10 of 12 items were mislabeled zero or one time. In the autonomous exhaustive list group, 11 of 12 items were mislabeled zero or one time.

Figure 3. App evaluation flow using randomized design.

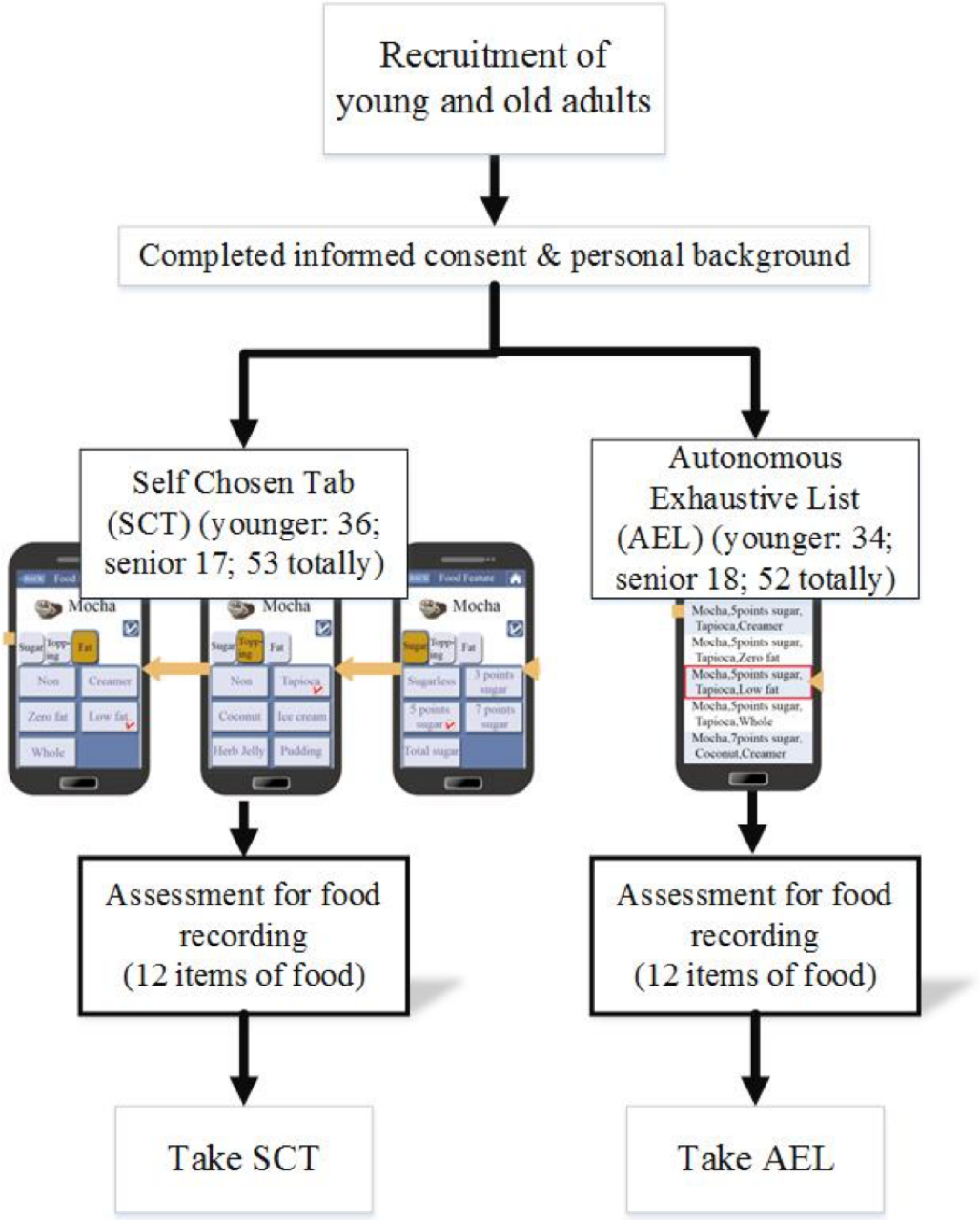


Table 1. Distribution of characteristics among the self-chosen tab and autonomous exhaustive list groups.

Variable and classification	Total (N=105)	Self-chosen tab (n=53)	Autonomous exhaustive list (n=52)	*P* value
Participants, n (%)				.78
Young	70 (67)	36 (68)	34 (65)	
Senior	35 (33)	17 (328)	18 (35)	
Gender, n (%)				.78
Male	35 (33)	17 (32)	18 (35)	
Female	70 (67)	36 (68)	34 (65)	
Age (years), mean (SD)				
Overall	34.70 (19.50)	33.68 (18.94)	35.75 (20.18)	.59
Young	21.19 (2.09)	20.92 (1.90)	21.47 (2.26)	.27
Senior	61.74 (5.01)	60.71 (3.85)	62.72 (5.85)	.24
Body mass index (BMI)	22.00 (3.57)	21.61 (3.31)	22.39 (3.80)	.27
Unit, n (%)				.81
Hospital	35 (33)	17 (32)	18 (35)	
Medical and others	20 (19)	12 (23)	8 (15)	
Industrial design	28 (27)	13 (24)	15 (29)	
Information management	22 (21)	11 (21)	11 (21)	
Experiences of nutrition-related courses, n (%)				.85
Yes	21 (20)	11 (21)	10 (19)	
No	84 (80)	42 (79)	42 (81)	
Experience of health education program, n (%)				.18
Yes	15 (14)	10 (19)	5 (10)	
No	90 (86)	43 (81)	47 (90)	
Experience in cooking, n (%)				.92
Yes	53 (50)	27 (51)	26 (50)	
No	52 (50)	26 (49)	26 (50)	

Table 2. Accuracy comparison among self-chosen tab and autonomous exhaustive list groups.

Types of food	Self-chosen tab		Autonomous exhaustive list	
	Correct/Incorrect	Error description	Correct/Incorrect	Error description
Single				
Chinese steamed bread				
Overall	53/0		51/0	
Young	36/0		33[a]/0	
Senior	17/0		18/0	
Fried chicken drumstick				
Overall	53/0		51/1	
Young	36/0		34/0	
Senior	17/0		17/1	Selected “deep fried” instead of “stir fried”
Green tea				
Overall	48/5		46/6	
Young	33/3	Did not select “nondairy creamer” (n=2); did not select “no topping”	32/2	Selected “pudding” instead of “no topping”; selected incorrect sugar quantity
Senior	15/2	Did not select “nondairy creamer”; did not select “no topping”	14/4	Selected incorrect sugar quantity; selected incorrect milk type (n=3)
Boiled rice				
Overall	52/1		52/0	
Young	36/0		34/0	
Senior	16/1	Selected “boiled” instead of “steamed”	18/0	
Fried pork chop				
Overall	50/1		50/0	
Young	35[a]/0		32[b]/0	
Senior	15[a]/1	Did not select “stir fried”	18/0	
Tea with milk				
Overall	50/1		51/1	
Young	34[b]/0		34/0	
Senior	16/1	Did not two food attributes (ie, “pudding” and “low-fat milk”)	17/1	Selected incorrect milk type
Mixed food				
Cold tofu with preserved eggs				
Overall	53/0		51/0	
Young	36/0		33[a]/0	
Senior	17/0		18/0	
Stir-fried eggs with shredded carrot				
Overall	53/0		49/1	
Young	36/0		31[b]/1	Selected “deep fried” instead of “stir fried”
Senior	17/0		18/0	
Stir-fried Napa cabbage with bacon				

Types of food	Self-chosen tab		Autonomous exhaustive list	
	Correct/Incorrect	Error description	Correct/Incorrect	Error description
Overall	52/1		52/0	
Young	36/0		34/0	
Senior	16/1	Did not select “stir fried”	18/0	
Stir-fried shredded pork and green pepper				
Overall	47/4		52/0	
Young	34/2	Did not select “stir fried”; selected “stir fried” instead of “deep fried”	34/0	
Senior	13[a,c]/2	Did not select “stir fried” (n=2)	18/0	
Stir-fried egg with tomatoes				
Overall	53/0		51/0	
Young	36/0		33[a]/0	
Senior	17/0		18/0	
Stir-fried bitter melon with salted duck eggs				
Overall	50/1		51/0	
Young	34[a]/1	Selected “stir fried” instead of “deep fried”	33[a]/0	
Senior	16[a]/0		18/0	

[a]One participant did not pass the entry criteria and was not counted.

[b]Two participants did not pass the entry criteria and were not counted.

[c]One participant was not counted due to log data error.

In the self-chosen tab group, 11 of 13 incorrect answers were that the participant did not select “no toppings” (n=2), did not select method of preparation (n=5), did not select “nondairy creamer” (n=3), and did not select two food attributes (n=1). The rest of the incorrect answers were selecting the wrong method of preparation (n=2). In the autonomous exhaustive list group, all nine incorrect answers were due to incorrect selection of food attributes; none resulted from the user neglecting to select an attribute.

Time Needed to Complete the Operation for Efficiency Evaluation

Table 3 summarizes the time participants needed to select attributes for the 12 food items. The self-chosen tab group significantly outperformed the autonomous exhaustive list group on all but one food item (boiled rice). The young participants in the self-chosen tab group had a significant time advantage over their counterparts in the autonomous exhaustive list for all but three items (Chinese steamed bread, fried chicken drumstick, and boiled rice). Senior participants in the self-chosen tab group had a significant time advantage over their autonomous exhaustive list counterparts for all but fried chicken drumstick and boiled rice.

Table 3. Time duration for operating assessment.

Type, food, and age group	Response time (seconds), mean (SD)		*t* value for difference	*P* value
	Self-chosen tab	Autonomous exhaustive list		
Single				
Chinese steamed bread				
All	3.01 (1.23)	4.84 (4.29)	–2.49	.02
Young	2.48 (0.75)	3.19 (2.64)	–1.30	.20
Senior	4.67 (0.90)	8.28 (5.09)	–2.41	.03
Fried chicken drumstick				
All	3.40 (2.48)	6.04 (6.12)	–2.43	.02
Young	2.60 (1.77)	4.05 (3.91)	–1.71	.10
Senior	5.90 (2.81)	10.20 (7.83)	–1.75	.10
Green tea				
All	12.40 (8.13)	26.15 (14.90)	–4.93	<.001
Young	8.92 (2.53)	17.78 (6.82)	–6.14	<.001
Senior	23.22 (10.08)	43.58 (11.60)	–4.20	<.001
Boiled rice				
All	2.71 (1.90)	3.63 (2.94)	–1.61	.11
Young	1.91 (0.56)	2.47 (1.32)	–1.98	.06
Senior	5.19 (2.47)	6.06 (3.87)	–0.59	.56
Fried pork chop				
All	3.09 (1.37)	4.71 (2.67)	–3.29	.002
Young	2.60 (0.75)	3.34 (1.11)	–2.89	.006
Senior	4.63 (1.73)	7.57 (2.73)	–2.82	.01
Tea with milk				
All	11.76 (4.99)	26.04 (17.27)	–4.83	<.001
Young	9.77 (3.47)	17.26 (7.70)	–4.47	<.001
Senior	17.95 (3.79)	44.34 (17.50)	–5.07	<.001
Mixed food				
Cold tofu with preserved eggs				
All	6.15 (2.83)	31.36 (25.78)	–5.91	<.001
Young	5.02 (1.55)	20.38 (13.00)	–5.87	<.001
Senior	9.67 (3.09)	54.24 (31.01)	–4.95	<.001
Stir-fried eggs with shredded carrot				
All	5.46 (2.13)	32.74 (27.59)	–6.00	<.001
Young	4.63 (1.24)	20.56 (16.85)	–4.71	<.001
Senior	8.05 (2.29)	58.13 (28.82)	–5.99	<.001
Stir-fried napa cabbage with bacon				
All	5.90 (2.19)	27.49 (18.34)	–7.11	<.001
Young	4.89 (1.08)	19.15 (11.54)	–6.16	<.001
Senior	9.03 (1.75)	44.87 (17.98)	–6.86	<.001
Stir-fried shredded pork and green pepper				
All	5.09 (1.52)	38.20 (25.03)	–8.03	<.001

Type, food, and age group	Response time (seconds), mean (SD)		*t* value for difference	*P* value
	Self-chosen tab	Autonomous exhaustive list		
Young	4.58 (1.07)	26.76 (13.24)	–8.35	<.001
Senior	6.70 (1.67)	62.02 (27.40)	–6.98	<.001
Stir-fried egg with tomatoes				
All	5.01 (1.93)	24.31 (15.12)	–7.70	<.001
Young	4.12 (0.89)	16.56 (7.90)	–7.83	<.001
Senior	7.77 (1.64)	40.47 (13.79)	–8.14	<.001
Stir-fried bitter melon with salted duck eggs				
All	5.73 (1.90)	32.66 (21.47)	–7.60	<.001
Young	4.96 (0.91)	21.52 (6.74)	–12.20	<.001
Senior	8.13 (2.20)	55.86 (23.25)	–7.07	<.001

Discussion

Based on a combinatorial concept in dietary recording, we developed two prototypes and assessed and compared their efficacies through randomized trials including both young adults and seniors. Assessment focused on accuracy, time efficiency, and the method's potential.

Accuracy

Both the self-chosen tab and autonomous exhaustive list groups demonstrated high accuracy results (Table 2) for reporting food items in two different meals regardless of age group.

Errors in the self-chosen tab group occurred because the user did not select the appropriate attribute in a certain category or because the user selected the wrong attribute within the category. These two error types would lead to the incorrect calorie nutrient intake calculation. Senior participants had a relatively higher rate of incorrect responses, possibly due to reduced vision or cognitive ability. Among the wrong answers, 11 occurred because the participant did not select the appreciate food attributes. These errors could possibly be due to interface or training issues and they could be addressed by future improvements to the user interface and training protocol.

The autonomous exhaustive list group was more likely to make attribute selection errors than the self-chosen tab group, possibly because of the large number of lists with similar food descriptions. However, the autonomous exhaustive list group did not produce any instance of failing to select the appropriate attribute because the autonomous exhaustive list design automatically presented a comprehensive list of all possible food items.

Time Efficiency

The self-chosen tab app was found to be more efficient than the autonomous exhaustive list app, with 11 of 12 food items requiring less time to complete (boiled rice being the only exception). During input, the self-chosen tab users spent time selecting each appropriate attribute from the tab menu and then switched to the next tab to select a certain attribute. The overall mean time spent for each selection ranged from 2.7 seconds (SD 1.90) for boiled rice to 12.4 seconds (SD 8.13) for green tea. However, the autonomous exhaustive list required participants to browse multiple lists to locate specific combinations of food items, taking between 3.6 seconds (SD 2.94) for boiled rice and 26.2 seconds (SD 14.90) for green tea.

However, input tasks for mixed foods in the self-chosen tab app took between 5 to 6 seconds per attribute, with beverages being the most time-consuming items to classify because the participants needed to select three attributes from each tab menu, whereas the other food items only required selecting a single attribute. The input operations for mixed foods are inefficient in the autonomous exhaustive list app. In the autonomous exhaustive list app, boiled rice takes the least time to input because the option list only includes five items.

Additionally, inputting the second meal required less time than the first meal for both groups and apps, likely because the first meal familiarized participants with the system operation.

Combinatorial Concept Potential

To accurately describe actual food intake, the mobile app allows users to select meal customization to correctly describe food alternatives when the current database is unable to describe the actual food item. Food composition databases can be time-consuming and difficult to maintain given constant updates and new recipes [24]. Incomplete databases can negatively impact the dietary recording accuracy. The proposed concept provides an advantage in that it can derive a broad range of food alternatives from a small number of known elements, ensuring that users can supplement incomplete food databases to account for their actual food intake. The food list is generated by combinatorial algorithms of elements (both main ingredients and food attributes) in the database. For example, a mixed food with three main food ingredients, each of which has 10 choices, will create a permutation of 1000 (10*10*10=1000) combinations. One potential issue is that the system will produce many combinations which are highly unlikely to be consumed in real life. This list will be displayed in the autonomous exhaustive list interface, which will significantly slow searching and selecting. Furthermore, the user might have difficulty in selecting from among similar food variants that differ only in

terms of a single ingredient or attribute. This is not an issue in self-chosen tab app because it does not provide an exhaustive permutation of food item attributes.

Limitations and Future Research

The experiments described here were conducted under laboratory conditions using a predetermined list of food items, the participants were recruited from the college and the hospital, and the data collected does not explain the reasons for incorrect selections. As for participant recruitment in this study, only participants aged 18 to 29 years and 55 to 73 years were included; therefore, the results cannot be applied for external validity. A large cohort with a full age range would be helpful for further explanation.

Future research should focus on improving both prototypes, and the development of new apps designed for use in actual dining contexts. Further work also needs to consider additional aspects and variables (including food portion sizes and combinations of more than two food items). The list of foods and the combinations require further evaluation in other cultures. The results of such efforts will help determine whether this mixed food idea can be used to effectively replace a food database including all food items.

Conclusion

A wide range of mHealth apps have been developed with diverse innovative designs, but the effectiveness in user interaction of such developed prototypes can be difficult to evaluate rigorously. This research demonstrates the application of design innovation in implementing a concept in individual-tailored dietary recording and using randomized trials with two target groups. Experimental results show that the two developed apps achieve a high degree of accuracy in describing a wide variety of food items among target users in two distinct age groups. The self-chosen tab app performs better both in terms of accuracy and time response. Furthermore, the concept has potential to account for a broader diversity of foods in customized dietary recording. The concept and the results in user interaction provide a scientific evidence for the continued development of related dietary recording apps.

Acknowledgments

The authors appreciate the efforts and cooperation of all study participants at the Food Interaction Design Lab of the Chang Gung University Industrial Design Department, and the Health Informatics Lab of the Health Care Management Department in the College of Management. This work was supported by the Ministry of Science and Technology, Taiwan (NSC-105-2221-E-182-044; NSC-104-2221-E-182-049), and Research Fund of Chang Gung Memorial Hospital and Chang Gung University (BMRP-B81; BMRPD 67).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Mobile app based on self-chosen tab with one main food ingredient.

[ZIP File (Zip Archive), 9MB - mhealth_v7i2e10931_app1.zip]

Multimedia Appendix 2

Mobile app based on autonomous exhaustive list with one main food ingredient.

[ZIP File (Zip Archive), 27MB - mhealth_v7i2e10931_app2.zip]

Multimedia Appendix 3

Mixed food selection process based on self-chosen tab app.

[PNG File, 294KB - mhealth_v7i2e10931_app3.png]

Multimedia Appendix 4

Mixed food selection process based on autonomous exhaustive list app.

[PNG File, 270KB - mhealth_v7i2e10931_app4.png]

Multimedia Appendix 5

Mobile app based on self-chosen tab with mixed (two or three) main food ingredients.

[ZIP File (Zip Archive), 14MB - mhealth_v7i2e10931_app5.zip]

Multimedia Appendix 6

Mobile app based on autonomous exhaustive list with mixed foods (two or three) main food ingredients.

[ZIP File (Zip Archive), 12MB - mhealth_v7i2e10931_app6.zip]

Multimedia Appendix 7

Experimental setup.

[ZIP File (Zip Archive), 1MB - mhealth_v7i2e10931_app7.zip]

Multimedia Appendix 8

CONSORT‐EHEALTH checklist (V 1.6.1).

[PDF File (Adobe PDF File), 448KB - mhealth_v7i2e10931_app8.pdf]

References

1. Adolphs C, Winkelmann A. Personalization research in e-commerce: a state of the art review (2000 - 2008). J Electron Com Re 2010 Aug 4;11(4):326-341 [FREE Full text]
2. Wendy's. Fresh, healthy fast food salads URL: https://menu.wendys.com/en_US/category/freshmade-salads/ [accessed 2018-09-06] [WebCite Cache ID 72Dr54Lbn]
3. Starbucks. App for iPhone and Android URL: https://www.starbucks.com/coffeehouse/mobile-apps [accessed 2018-09-06] [WebCite Cache ID 72DsIGDph]
4. WHO Global Observatory for eHealth. mHealth: New Horizons for Health through Mobile Technologies (Global Observatory for Ehealth). World Health Organization: Geneva; 2011.
5. Istepanian R, Laxminarayan S, Pattichis C. M-Health: Emerging Mobile Health Systems. New York: Springer; 2005.
6. Liu Y, Chen C, Lee C, Lin Y, Chen H, Yeh J, et al. Design and usability evaluation of user-centered and visual-based aids for dietary food measurement on mobile devices in a randomized controlled trial. J Biomed Inform 2016 Dec;64:122-130 [FREE Full text] [doi: 10.1016/j.jbi.2016.10.001] [Medline: 27720984]
7. Sullivan S, Bopp M, Weaver D, Sullivan D. Innovations in calculating precise nutrient intake of hospitalized patients. Nutrients 2016 Jul 04;8(7):E412 [FREE Full text] [doi: 10.3390/nu8070412] [Medline: 27384584]
8. Timon CM, Astell AJ, Hwang F, Adlam TD, Smith T, Maclean L, et al. The validation of a computer-based food record for older adults: the Novel Assessment of Nutrition and Ageing (NANA) method. Br J Nutr 2015 Feb 28;113(4):654-664. [doi: 10.1017/S0007114514003808] [Medline: 25630436]
9. Matthiessen TB, Steinberg FM, Kaiser LL. Convergent validity of a digital image-based food record to assess food group intake in youth. J Am Diet Assoc 2011 May;111(5):756-761. [doi: 10.1016/j.jada.2011.02.004] [Medline: 21515125]
10. Chen YS, Wong JE, Ayob AF, Othman NE, Poh BK. Can Malaysian young adults report dietary intake using a rood diary mobile application? A pilot study on acceptability and compliance. Nutrients 2017 Jan 13;9(1) [FREE Full text] [doi: 10.3390/nu9010062] [Medline: 28098770]
11. Azar KM, Lesser LI, Laing BY, Stephens J, Aurora MS, Burke LE, et al. Mobile applications for weight management: theory-based content analysis. Am J Prev Med 2013 Nov;45(5):583-589. [doi: 10.1016/j.amepre.2013.07.005] [Medline: 24139771]
12. Brown T. Change by Design: How Design Thinking Can Transform Organizations and Inspire Innovation. New York: HarperCollins; 2009.
13. van der Weegen S, Verwey R, Spreeuwenberg M, Tange H, van der Weijden T, de Witte L. The development of a mobile monitoring and feedback tool to stimulate physical activity of people with a chronic disease in primary care: a user-centered design. JMIR Mhealth Uhealth 2013 Jul 02;1(2):e8 [FREE Full text] [doi: 10.2196/mhealth.2526] [Medline: 25099556]
14. Dicianno BE, Henderson G, Parmanto B. Design of mobile health tools to promote goal achievement in self-management tasks. JMIR Mhealth Uhealth 2017 Jul 24;5(7):e103 [FREE Full text] [doi: 10.2196/mhealth.7335] [Medline: 28739558]
15. Giunti G, Mylonopoulou V, Rivera Romero O. More Stamina, a gamified mHealth solution for persons with multiple sclerosis: research through design. JMIR Mhealth Uhealth 2018 Mar 02;6(3):e51 [FREE Full text] [doi: 10.2196/mhealth.9437] [Medline: 29500159]
16. Curtis KE, Lahiri S, Brown KE. Targeting parents for childhood weight management: development of a theory-driven and user-centered healthy eating app. JMIR Mhealth Uhealth 2015 Jun 18;3(2):e69 [FREE Full text] [doi: 10.2196/mhealth.3857] [Medline: 26088692]
17. 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]
18. Mummah S, Robinson T, King A, Gardner C, Sutton S. IDEAS (Integrate, Design, Assess, and Share): a framework and toolkit of strategies for the development of more effective digital interventions to change health behavior. J Med Internet Res 2016 Dec 16;18(12):e317 [FREE Full text] [doi: 10.2196/jmir.5927] [Medline: 27986647]

19. Matheson GO, Klügl M, Engebretsen L, Bendiksen F, Blair SN, Börjesson M, et al. Prevention and management of non-communicable disease: the IOC consensus statement, Lausanne 2013. Br J Sports Med 2013 Nov;47(16):1003-1011. [doi: 10.1136/bjsports-2013-093034] [Medline: 24115479]
20. Witteman HO, Dansokho SC, Colquhoun H, Coulter A, Dugas M, Fagerlin A, et al. User-centered design and the development of patient decision aids: protocol for a systematic review. Syst Rev 2015 Jan 26;4:11 [FREE Full text] [doi: 10.1186/2046-4053-4-11] [Medline: 25623074]
21. Norman D, Draper S. User Centered System Design: New Perspectives on Human-Computer Interaction. Hillsdale, NJ: Lawrence Earlbaum Associates; 1986.
22. Hermawati S, Lawson G. Managing obesity through mobile phone applications: a state-of-the-art review from a user-centred design perspective. Pers Ubiquit Comput 2014 Jan 20;18(8):2003-2023. [doi: 10.1007/s00779-014-0757-4]
23. Deng C, Graz J. Generating randomization schedules using SAS programming. 2002 Presented at: 27th Annual SAS Users Group International Conference; Apr 14-17, 2002; Orlando, FL URL: http://www2.sas.com/proceedings/sugi27/p267-27.pdf
24. Carter M, Hancock N, Albar S, Brown H, Greenwood D, Hardie L, et al. Development of a new branded UK food composition database for an online dietary assessment tool. Nutrients 2016 Aug 05;8(8) [FREE Full text] [doi: 10.3390/nu8080480] [Medline: 27527214]

Abbreviations

BMI: body mass index

Edited by G Eysenbach; submitted 03.05.18; peer-reviewed by M Boaz, V Mylonopoulou; comments to author 16.08.18; revised version received 06.10.18; accepted 26.10.18; published 15.02.19.

Please cite as:
Liu YC, Chen CH, Tsou YC, Lin YS, Chen HY, Yeh JY, Chiu SYH
Evaluating Mobile Health Apps for Customized Dietary Recording for Young Adults and Seniors: Randomized Controlled Trial
JMIR Mhealth Uhealth 2019;7(2):e10931
URL: http://mhealth.jmir.org/2019/2/e10931/
doi:10.2196/10931
PMID:30767906

Original Paper

Exploring Users' Experiences of the Uptake and Adoption of Physical Activity Apps: Longitudinal Qualitative Study

Dario Baretta[1], PhD; Olga Perski[2,3], PhD; Patrizia Steca[1], PhD

[1]Department of Psychology, University of Milano-Bicocca, Milano, Italy

[2]Department of Clinical, Educational & Health Psychology, University College London, London, United Kingdom

[3]Institute of Digital Health, University College London, London, United Kingdom

Corresponding Author:
Dario Baretta, PhD
Department of Psychology
University of Milano-Bicocca
U6 Building, 3rd Floor
Piazza dell'Ateneo Nuovo,1
Milano, 20126
Italy
Phone: 39 3405946987
Email: d.baretta@campus.unimib.it

Abstract

Background: Although smartphone apps might support physical activity (PA), engagement with them tends to be low.

Objective: This study aimed to examine potential users' needs and preferences regarding their engagement with PA apps during a first exposure to a never-used PA app and after 2 weeks' usage.

Methods: A longitudinal, one-arm qualitative study was conducted with potential PA app users. At baseline, participants (N=20) were asked to explore 1 of 3 randomly allocated PA apps while thinking aloud. Semistructured interview techniques allowed participants to elaborate on their statements. After 2 weeks, follow-up interviews explored participants' (n=17) lived experiences of real-world app use. Verbal reports from both time points were analyzed using inductive thematic analysis.

Results: Features that promote a fair and simple user experience, support users' self-regulation skills, and address users' exercise motives were considered important for engagement both during a first exposure and after a 2-week use of PA apps. Features that support users' need for relatedness as well as those that facilitate users to implement their intentions were expected to be important for engagement mainly during a first exposure to PA apps. Proactive and tailored features that integrate behavioral, psychological, and contextual information to provide adaptive exercise plans and just-in-time support were considered relevant to sustain engagement over time.

Conclusions: App features that address users' exercise motives, promote self-regulation, and fulfill users' need for relatedness might promote engagement with PA apps. Tailored and proactive features were expected to promote sustained engagement.

(JMIR Mhealth Uhealth 2019;7(2):e11636) doi:10.2196/11636

KEYWORDS

physical activity; smartphone apps; engagement; person-based approach; longitudinal qualitative study; thematic analysis

Introduction

Background

Regular physical activity (PA) is associated with a reduced risk of mortality caused by cardiovascular disease, diabetes, obesity, and some forms of cancer [1,2]. Nevertheless, worldwide, a third of adults do not meet the recommended PA guidelines [3]. Hence, the development and promotion of effective behavior change interventions that target PA is a public health priority.

A range of different behavior change interventions to increase PA have been found to be effective [4] and there is evidence supporting their cost-effectiveness [5]. However, most of these have been delivered in face-to-face settings and are hence unsuitable for targeting large, geographically diverse populations and for supporting behavior change in real time. Mobile phone technology has the potential to deliver effective PA interventions [6,7] and constitutes an economically viable tool to reach large populations [8]. Thus far, evidence suggests that Web- and

mobile phone–based interventions have moderate effects on PA [8,9]; however, effect sizes are heterogeneous and have been found to vary depending on study design, mode of delivery, intervention components, outcome measures, and study length. It has, therefore, been argued that we should err on the side of caution when interpreting and generalizing early evidence from digital PA interventions [8].

It is assumed that some level of engagement with digital behavior change interventions is a precondition for their effectiveness [10]. Engagement with digital behavior change interventions has been conceptualized both as a behavior and as a subjective experience, involving dimensions of interest, attention, and affect [11].

Engagement as a behavior refers to the extent (eg, amount, frequency, and duration) of usage of a digital intervention over time. Evidence shows that digital behavior change interventions typically suffer from low usage and high abandonment rates. For example, it has been estimated that 26% of commercially available health and fitness apps are downloaded and used only once by each user and that 74% of such apps are abandoned after their tenth use [12]. The most frequently reported reasons why users abandon health apps include a lack of desired features and having abandoned one's health goal [13]. As results from a meta-analysis of PA internet-delivered interventions found a negative relationship between early attrition and intervention effectiveness, researchers have argued that it is important to promote sustained engagement with digital PA interventions [9]. In regard to usage, a PA mobile health study has shown that although the number of user log-ins declined throughout the intervention period, it was not associated with PA behavior [14]. It has been, therefore, proposed that focusing just on broad summative usage metrics does not enable to adequately characterize and evaluate engagement with digital interventions [11,15].

Engagement as a subjective experience refers to the interaction with a digital intervention and involves feelings of interest, enjoyment, and attention [11]. By exploring this particular aspect of engagement, researchers may better understand how and why users interact with specific intervention content and design features and help explain different usage patterns [10]. For example, a review of health and fitness apps found a positive association between engagement and the inclusion of interactive features such as behavioral tracking and semi-automated options [16]. However, little attention has been paid to potential PA app users' subjective experiences of such features. Furthermore, given that psychological factors, such as motivation, have been found or hypothesized to influence both engagement with digital behavior change interventions and the effect of specific behavior change techniques (BCTs) on engagement [11], it appears vital to explore users' psychological needs and preferences with regards to PA app engagement.

According to the person-based approach to intervention development, it is important to elicit and address the needs, perspectives, and experiences of potential users during the design process to develop apps that promote individual engagement [17]. To this aim, the use of mixed-methods research designs with a specific focus on qualitative methodologies (eg, focus groups, interview with open-ended questions, and think-aloud studies) has thus been recommended [17,18]. Previous research targeting different health-related behaviors in diverse populations has used existing commercial apps as stimuli to elicit users' thoughts about factors that are expected to be important for engagement in real life [19-21]. Indeed, through the exploration and use of existing apps, potential users are prompted to reflect on their attitudes toward particular app features and how these might act as facilitators or barriers to uptake and adoption of the app.

Objectives

The aim of this study was to guide the selection of design features to implement in PA apps for nonclinical, adult populations. Through a combination of think-aloud methodology and in-depth interview techniques, this study examined (1) what features potential users expect to be important for engagement with PA apps during first exposure to never-used, randomly allocated, and commercially available PA app and (2) what features are judged to be important for supporting engagement and satisfactory experiences after 2 weeks' usage of the same PA app.

Methods

Study Design

This study used a longitudinal, single-arm design with a 2-week follow-up using think-aloud methodology and semistructured interview techniques. Since 74% of health apps are abandoned after their tenth use [12] and their usage tends to vary between a few times per week to twice per day [22], a 2-week period was considered to be sufficient to examine relevant changes in engagement with PA apps. Think aloud is a method that requires participants to verbalize their thoughts and impressions while performing a task [23]. The benefit of using think-aloud methodology is that it directly captures participants' ongoing thought processes during a specific experience and, when applied to digital interventions, it allows researchers to observe and analyze users' reactions to every element of the intervention [17]. In this study, participants were asked to verbalize their impressions, thoughts, and feelings while downloading and exploring a specific, never-used PA app randomly allocated from a pool of 3 apps identified by the researchers according to specific criteria (for more details about these criteria see the section "Mobile Apps"). Overall, 3 different apps were included as this allowed us to (1) expose participants to a greater number of intervention components (ie, BCTs) as shown in Table 1, and (2) obtain more valid conclusions and to be more confident in generalizing current results to other PA apps. Semistructured interviews were subsequently used to retrospectively investigate (1) statements made during the think-aloud task at baseline and (2) users' thoughts and impressions after a 2-week period of *ad libitum* use of 1 of the 3 randomly allocated apps. The ethical permission for the study was granted before data collection by the ethics committee of the University of Milan-Bicocca.

Table 1. Behavior change techniques implemented in the selected physical activity (PA) apps, coded using the behavior change technique taxonomy (v1).

Behavior change techniques	Runtastic Running & Fitness Tracker	Endomondo - Running & Walking	Runkeeper - GPS Track Run Walk
1.1. Goal setting (behavior)	✔[a]	✔	✔
1.3. Goal setting (outcome)	—[b]	—	✔
1.4. Action planning	—	—	✔
1.9. Commitment	—	✔	—
2.2. Feedback on behavior	✔	✔	✔
2.3. Self-monitoring of behavior	✔	✔	✔
2.4. Self-monitoring outcomes of behavior	✔	✔	✔
2.7. Feedback on outcomes of behavior	✔	✔	✔
3.3. Social support	✔	✔	✔
4.1. Instruction on how to perform the behavior	✔	—	—
5.1. Information about health consequences	✔	—	—
5.4. Monitoring of emotional consequences	✔	—	✔
6.2. Social comparison	✔	✔	✔
9.1. Credible source	✔	—	—
10.5. Social reward	✔	✔	✔

[a]✔ indicates BCT is present.

[b]— indicates BCT is absent.

Participants

Participants were recruited to represent a subsample of the Italian, adult, nonclinical population not meeting the PA guidelines and interested in using an app to increase PA. Participants were recruited in the Italian county of Lombardy and were eligible to take part in the study if they (1) were aged between 30 and 50 years (a more narrow age range was preferred to a wider one to prevent collecting data from a heterogeneous sample characterized by too varying motivations and subjective experiences); (2) did not have any pre-existing health conditions that would impede on the ability to engage in PA (eg, cardiovascular diseases, heart failure, and pulmonary conditions); (3) did not have a diagnosis of a clinical condition that would benefit from doing PA (eg, hypertension and diabetes) as clinical populations were likely to have different motivations compared with the intended users (ie, healthy adults); (4) reported PA behavior not in line with the recommended PA guidelines (ie, 150 min of moderate PA or 75 min of vigorous PA per week); (5) were willing to increase PA and interested in doing so using an app; and (6) owned an Android or iOS smartphone.

Sampling

Participants were recruited through social media (eg, Facebook), snowball sampling methods (ie, each participant was asked to find another participant), and posters placed in a large university campus. The aim of the study and the eligibility criteria were specified in the recruitment materials. Recruitment into the study was discontinued when no new insights and novel themes were identified in the interviews (ie, when theoretical saturation was achieved) [24].

Procedure

Interested participants filled out a Web-based screening questionnaire to verify study eligibility. Eligible participants were contacted to schedule the baseline session and were asked to provide written informed consent before taking part in the study. Of them, 1 of 3 PA apps was randomly allocated to each participant before the baseline session. The baseline session included a pretask interview, a think-aloud task, and a semistructured interview (see Multimedia Appendix 1). After taking part in the baseline session, participants were asked to use the app for 2 weeks. No specific instructions were provided in terms of app use (eg, frequency or duration of use) to avoid influencing users' engagement with and experiences of the app. Participants were, however, encouraged to use the app to increase their PA according to their own goals. Participants were contacted after 1 week to verify if they were using the app and to ask if they were willing to continue using the app for another week. The participants took part in a telephone interview 2 weeks after the baseline session using an interview schedule as guide (see Multimedia Appendix 1).

During the recruitment, each participant was given a unique code. A separate file in which codes were linked to personal identifiers (eg, name and contacts) was created, and any personal identifiers in the data file were removed and replaced with the codes. The de-identified data files and the file with personal identifiers were stored separately in secure locations, with access restricted to the study personnel.

Measures

In the online screening questionnaire, participants were asked to report information about (1) age; (2) gender; (3) physical conditions that prevent participants from being physically active (eg, chronical back pain and acute cardiovascular disease); (4) physical conditions associated with a lack of PA behavior (eg, hypertension and obesity); (5) PA during leisure time, measured using the Global Physical Activity Questionnaire [25], with consequent assessment of whether PA was in line with the recommended guidelines; (6) intention to increase PA assessed through a 6-point Likert scale ranging from 1 (I do not intend to increase my PA level) to 6 (I intend to increase my PA level by the next month); (7) interest in increasing PA with the aid of a mobile app (yes vs no); and (8) ever use of a PA app. The screening questionnaire was hosted by Qualtrics software (Qualtrics, Provo, UT) [26].

Mobile Apps

The PA apps identified for the study were *Runtastic Running & Fitness Tracker*, *Endomondo - Running & Walking*, and *Runkeeper - GPS Track Run Walk* (henceforth, *Runtastic*, *Endomondo*, and *Runkeeper*). The researchers have no association with the development or marketing of these apps. The inclusion criteria for the app selection were as follows: (1) their main goal was to increase PA among healthy adults, (2) they were freely available both on iTunes and Google Play, (3) they incorporated many BCTs (see Table 1 for a description of the BCTs incorporated in the selected apps, coded using behavior change technique taxonomy (v1) [27]) that are common among top-ranked publicly available PA apps [28], and (4) they had received high user ratings (ie, 4.5 stars), which is considered to be an indicator of positive user experiences with PA apps. Finally, it should be noted that all the selected apps contained advertisements and in-app purchases.

Data Analysis

Quantitative data from the online screening questionnaire were analyzed using descriptive statistics with SPSS version 22 [29]. Qualitative data were audio-recorded and transcribed verbatim. Transcripts from the baseline and follow-up activities were analyzed using inductive thematic analysis [30]. Thematic analysis is characterized by 6 phases: (1) familiarizing with the data, (2) generating initial codes, (3) searching for themes, (4) reviewing themes, (5) defining and naming themes, and (6) producing the report. Baseline and follow-up transcripts were considered as 2 different datasets and were analyzed separately. Data and repeated patterns that were considered pertinent to the aims of the study were coded by the first author. New inductive codes were labeled as they were identified during the coding process and the results of the coding were iteratively discussed with 2 researchers (SN and SB). The next stage involved searching for themes; the first author reviewed the codes one-by-one, organizing the findings to combine different codes that focus on similar aspects. The ordered data were reviewed and revised in discussion with 2 researchers (SN and SB) and were subsequently organized into themes. Recruitment stopped when theoretical saturation was achieved (ie, no new themes were identified). Resolution of disagreements and agreement on the final themes was reached through discussion among all coauthors. After having defined and named themes, examples of relevant transcripts were selected to illustrate themes. Data were analyzed in their original language to preserve original meanings, although coding and themes were formulated in English only. Illustrative quotes were translated by the first and second authors.

Results

Descriptive Statistics

A total of 20 participants (55% male; mean age 39.8 years; SD 7.0) took part in the study. Overall, 35% of participants had made an attempt to increase their PA in the past 6 months and 35% had previously used a PA app to increase their PA. Among the 20 participants who completed the baseline session, 17 participants (17/20, 85%) completed the follow-up interview. In total, 2 participants were excluded as they did not use the app during the 2-week study period, and 1 participant could not be successfully recontacted. Table 2 presents information about app allocation, presence at follow-up, and self-reported app use (ie, frequency and duration of use). It should be noted that use is defined in terms of how frequently and for how long the apps were used to track PA; however, this might not necessarily be related to the overall duration of time spent doing PA during the study and it might not reflect the intensity of use (eg, the amount or depth of use per log-in).

Thematic Analysis

In total, 6 themes were developed in relation to the first research question and were labeled: “fair and simple user experience,” “sense of autonomy and self-regulation of behavior,” “features that address users’ exercise motives,” “need for relatedness,” and “tailored support and action planning.” Overall, 1 subtheme was developed in relation to the second theme: “efficient and reliable monitoring and feedback of PA.” Overall, 3 subthemes were developed in relation to the fourth theme: “peer support,” “coaching support,” and “social comparison.”

A total of 5 themes were developed in relation to the second research question, 4 of which overlapped with themes from the baseline interviews. Hence, the labels were retained: “fair and simple user experience,” “sense of autonomy and self-regulation of behavior,” “features that address users’ exercise motives,” and “need for relatedness.” A new theme was developed, labeled: “proactive motivational features” (see Table 3 for a description of the themes and Multimedia Appendix 2 for additional extracts illustrating each theme).

A Fair and Simple User Experience

Baseline

Participants emphasized the importance of PA apps being able to promote a simple, easy, and fair interaction with its users. In particular, participants wanted the app to be simple to use, not cognitively demanding, or time consuming. Furthermore, participants expressed a strong dislike for obtrusive interactions such as inappropriate reminders or unexpected advertisements that redirect users to external websites:

> *It [user interface] should be simple and intuitive because this app is very complicated and I feel a bit*

lost. [P16 characterizing desired user interface during post-task interview]

Participants were particularly annoyed when they were asked to manually enter information about themselves or their activity levels into the apps. Rather, they reported a preference for apps that automatically register their information:

I hate having to [manually] register everything I do! [P2 commenting on the user interface of the given app during think-aloud task]

Follow-Up

After 2 weeks of *ad libitum* use, participants still maintained that simplicity and ease of use are essential prerequisites for app engagement. The perceived quality of the interactions participants had with their allocated app appeared to influence their overall satisfaction and affect their willingness to fully explore the app's functionality:

I immediately understood how to use it and how to get it up and running. [P7 commenting on the user interface of the given app]

It's not so easy to find information or understand how it works. You need to spend some time playing around with it [...] it should be organised in an easier way. [P9 commenting on the user interface of the given app]

Table 2. Information about app allocation to participants, presence at follow-up and self-reported usage of the app during physical activity session. Average use during the 2-week study was 4.3 times and average time duration per use was 73.6 min.

Participant	App	Follow-up[a]	Self-reported usage of the app during physical activity	
			Frequency of use during the 2-week study	Average time duration per use (min)
P1	Endomondo	✔	7 times	105
P2	Runtastic	✔	2 times	31
P3	Runkeeper	—	N/A[b]	N/A
P4	Endomondo	✔	6 times	65
P5	Runtastic	✔	6 times	45
P6	Runkeeper	✔	4 times	28
P7	Endomondo	✔	5 times	60
P8	Runtastic	✔	4 times	115
P9	Runkeeper	✔	7 times	75
P10	Endomondo	—	N/A	N/A
P11	Runtastic	✔	2 times	18
P12	Runkeeper	✔	2 times	30
P13	Endomondo	—	N/A	N/A
P14	Runkeeper	✔	1 time	50
P15	Endomondo	✔	5 times	35
P16	Runtastic	✔	7 times	N/A
P17	Runkeeper	✔	3 times	240
P18	Endomondo	✔	6 times	120
P19	Endomondo	✔	3 times	70
P20	Runkeeper	✔	3 times	90

[a]Presence at follow-up is denoted by ✔ while absence is denoted by —.

[b]N/A: not assessed.

Table 3. Summary of themes and subthemes identified in (1) the baseline interviews and (2) the follow-up interviews.

Theme and subthemes[a]	Description	Baseline	Follow-up
A fair and simple user experience	Features that enhance an easy, simple, and fair app interaction (ie, not cognitively demanding or time consuming)	✔[b]	✔
Features that promote autonomy and self-regulation of behavior	Features that support users' autonomous behavior regulation toward personal goals (eg, self-monitoring, feedback, and goal setting)	✔	✔
Efficient and reliable monitoring and feedback of PA	Features that support the efficient monitoring of PA (physical activity) and related parameters (eg, calories and heart rate)	✔	✔
Features that address users' exercise motives	Features that focus on health, fitness, and weight loss are expected to be motivating and engaging as they highlight the effects of PA on users' individual exercise motives	✔	✔
Need for relatedness	Features that leverage the motivational component of social support and facilitate the interaction with human elements	✔	✔
Peer support	Features that leverage peer support to overcome laziness, increase commitment, and as a prompt to be more active	✔	—[c]
Coaching support	Coaching features that recommend what to do to achieve personal goals and facilitate the receipt of emotional support	✔	—
Social comparison	Features that allow users to share PA performance with others and to compete against one another	✔	✔
Tailored action planning	Features that help users to implement their exercise intentions and achieve goals through tailored plans	✔	—
Proactive motivational features	Features that prompt and stimulate users by sparking their interest and increasing motivation (eg, reminders and suggestions)	—	✔

[a]Subthemes are denoted by italics.

[b]✔ indicates theme is present.

[c]— indicates theme is absent.

Features That Promote a Sense of Autonomy and Self-Regulation of Behavior

Baseline

During baseline assessment, participants perceived monitoring, feedback, and goal-setting features as fundamental self-regulation components to build the entire app around, as these features were expected to support users' autonomy and help regulate behavior toward their personal goals. Participants expected PA apps to include features that highlight their progress and failures; thus, clearly illustrating discrepancies between their current behavior and goals:

> *Looking at the percentage of achieved goals and getting a notification with relevant feedback saying: "You're below average, above average, you are achieving your goal or not..." Well, it might be a way to remind me of my commitment and verify the percentage of accomplishments.* [P15 commenting on existing and desired features during post-task interview; discrepancy between current behavior and goal and feedback on PA]

Follow-Up

Participants maintained that self-regulation features such as monitoring, feedback, and goal setting had a positive effect on their engagement with the app. In particular, participants reported appreciating the possibility to monitor their progress over time and, consequently, regulate their behavior and effort:

> *I have to admit that it's stimulating to be able to log all of my walks, how many calories I've burnt, the hydration I have lost, the average pace [...]. To me, [these features] are quite engaging, they make me more prone to open [the app] and log what I've done.* [P1 commenting on existing features; self-monitoring of PA, feedback on PA, self-monitoring of outcome of PA, and feedback on outcome of PA]

Participants also reported the need for autonomy and having an active role in the regulation of their behavior:

> *The goal shouldn't be imposed [by the app] because in that case I won't achieve it. The goal needs to come from inside - I don't want to do X km because the app says so, but because I want to.* [P8 commenting on existing features; goal setting]

Subtheme: Efficient and Reliable Monitoring and Feedback of Physical Activity

Baseline

Participants expressed the desire for an app that efficiently monitors PA and related variables (eg, calories and heart rate). It was, therefore, considered important that the app could distinguish between specific types of PA and that it would provide tailored and reliable monitoring and classification of PA:

> *The app should be accurate, particularly the GPS! In order for me to use an app, it has to provide me with precise and not generic information, otherwise*

it is not worth using. [P8 characterizing desired features during post-task interview; self-monitoring of PA]

Follow-Up

In line with the baseline interviews, monitoring features were considered to be important as they allow the collection of reliable information about PA and health-related parameters, hence enabling users to regulate their behavior:

> *It looks like there's no possibility of assessing heart rate...in some apps, similar to this one, you can assess your blood pressure on your own with a blood-pressure monitor and verify your minimum and maximum blood pressure on a daily basis.* [P1 characterizing desired features; self-monitoring of behavior and self-monitoring of outcome of PA]

Features That Address Users' Exercise Motives

Baseline

Participants reported an interest in apps that address their specific exercise motives. Health- and fitness-related participation motives were repeatedly reported as the main goals for being physically active. Specifically, participants expressed a preference for features that highlight fitness and weight loss. Hence, apps that allowed users to set calorie- and energy expenditure-related goals and monitor progress toward such goals were highly appreciated:

> *The fact that you can specify a training plan and, at the same time, control your heart rate and nutrition and figure out how these things interact...Well, in my opinion, these are the kinds of things that make an app engaging.* [P1 characterizing existing and desired features during post-task interview; action planning, self-monitoring of PA, and nutrition]

> *[The app] could, for example, be complemented by nutritional aspects...calories burnt, calorie intake.* [P15 characterizing desired features during pretask interview; self-monitoring of PA and nutrition]

Follow-Up

After 2 weeks of use, participants reaffirmed their positive attitudes toward features that take account of their personal motives for doing PA. In particular, nutrition suggestions, feedback about calories intake and consumption, and weight loss were thought to support engagement with PA apps:

> *If it was complemented and integrated with other things, such as calories, a diet, suggestions about nutrition – things that aren't present in this app or that are paid – it would probably be more engaging. Otherwise people get bored after a while.* [P15 characterizing desired features; feedback on outcome of PA and instruction on how to perform a behavior—nutrition]

Need for Relatedness

Baseline

Participants highlighted the motivational aspect of *human elements*. According to most participants, PA apps should provide opportunities for social interactions, mostly conceived of as a way to obtain human, empathic support, thus fostering a higher level of commitment to one's personal PA goal:

> *For me, it should be...I don't know how to explain...a bit more human.* [P14 characterizing desired features during pretask interview; social support—emotional]

Follow-Up

After having used the app for 2 weeks, participants still highlighted the importance of app features that foster a sense of relatedness by creating a digitally enabled social environment that is supportive and nonjudgmental:

> *In my opinion, the app starts being enjoyable/entertaining when you have "group use."* [P16 characterizing desired features [social support—unspecified]

However, different from the baseline interviews, the need for relatedness was mostly mentioned in relation to features that support social comparison. Hence, this constitutes the only subtheme.

Subtheme 1: Peer Support

Baseline

Opportunities for interacting with and doing PAs with peers were perceived as a way to boost one another's motivation and to leverage social commitment through the organization of group activities. Indeed, users described peer support as a way to overcome laziness. They also expected that peer support would make exercise more enjoyable and that it would function as a prompt to be more active:

> *I believe that [doing PA with friends] is stimulating, it's like when you go to the gym alone or with a friend, it's the same idea, you can support each other.* [P16 commenting on existing features during post-task interview; social support—emotional]

Subtheme 2: Coaching Support

Baseline

Participants held positive attitudes toward the possibility of interacting with coaching features embedded into the app, as these were expected to provide practical exercise suggestions and to facilitate the receipt of emotional support:

> *The best way to attain a goal is to have a person that helps you...the app should be a substitute for a personal trainer.* [P2 characterizing desired features during post-task interview; social support—practical]

Subtheme 3: Social Comparison

Baseline

Most participants believed that sharing information about PA with one's wider social network through an app is inappropriate

and expressed the desire to avoid social comparison and maintain focus on their own goals:

> *I don't like sharing...because I would come across as trying to brag, wouldn't I?...when I do a specific thing it's just for me, I don't like to involve too many people.* [P11 commenting on existing features during post-task interview; social comparison]

However, some participants expressed a desire for more competitive features, contingent on competition being framed as a friendly way to support each other rather than competing against other users:

> *The opportunity to compete is great - everyone can set their goal and support each other!* [P3 commenting on existing features during post-task interview; social comparison, goal setting, and social support—emotional]

Follow-Up

Most participants had not shared personal results with other users when such features were available, as social comparison features were thought to expose users to others' judgmental evaluations:

> *I'm not interested in sharing my runs and my results. I'm not a "social" person, so why would I start sharing the embarrassing results about my runs.* [P2 commenting on existing features; social comparison]

Although generally perceived as inappropriate, some participants mentioned that social incentives and challenges might make PA apps more engaging. Specifically, features that promote healthy competition were considered as a potential trigger to motivate users and make their exercise more enjoyable:

> *It would be interesting to get a group of friends together and say: "Let's compete and see who's training more, who's burning more calories." This would make [the app] more interesting.* [P16 characterizing desired interaction with features; social comparison]

Tailored Action Planning

Baseline

Features that help users to implement their PA intentions and achieve goals were judged to be important for engagement with PA apps. Participants expected PA apps to provide action plans to achieve individual goals:

> *They give you a training plan and you achieve your goals, that's great!* [P4 commenting on existing features during think-aloud task; action planning]

When developing action plans, participants wanted the PA app to provide a step-by-step guide and provide users with suggestions and advice about how to exercise:

> *Something that gradually guides you towards your goals, step-by-step, perhaps also suggesting what kind of physical activity to do and providing advice.* [P15 characterizing desired features during pretask interview; action planning and instruction about how to perform PA]

Proactive Motivational Features

Follow-Up

Some participants expressed skepticism toward the app's ability to boost motivation and to provide novel features that would spark their interest. According to these participants, the app was considered to be a functional tool rather than a trigger for initiating exercise:

> *I wasn't particularly engaged, perhaps because it didn't have additional functionalities to the other one that I had already used. This app didn't stimulate me to search for information, look at my stats or to train more.* [P8 commenting on the given app; general]

Other participants who had enjoyed their app experience expressed positive attitudes toward proactive features such as reminders and suggestions that served as triggers both for doing PA and for supporting engagement with the app:

> *It provides suggestions about how much activity to do per week, how to increase it, etc. That's what I liked a lot.* [P1 commenting on existing features; instruction on how to perform PA and action planning]

Moreover, some participants experienced a lack of proactive features and expressed their desire for a more interactive and prompting app:

> *I would have preferred it to remind me to exercise, like: "You are lazy, you should go out for a walk" [...] This app is too quiet, it's not stimulating.* [P11 characterizing desired features; prompts or cues]

Discussion

Summary of Main Findings

This study examined what PA app features potential users expect to influence app engagement during first exposure to never-used, commercially available PA apps and after 2 weeks of app use. Participants' thoughts and impressions were elicited through observing direct experiences of PA app use in the laboratory and through asking participants to report on their experiences of app use in their everyday lives to obtain ecologically valid insights. Thanks to the longitudinal study design, it was possible to compare what PA app features were considered important for engagement at different time points. A key advantage of the longitudinal qualitative study design was the ability to gain a more in-depth understanding of factors influencing users' engagement with PA apps both initially and over a period.

At baseline and follow-up, participants expressed the desire for PA apps that are simple to use, intuitive, and not cognitively demanding. Our findings suggest that PA apps should be easy to use, avoiding confusing and inadequate user experiences, as these constitute barriers to engagement. This finding supports previous research, which highlights the importance of observing how users interact with apps in real time to assess their ease of use and how they fit within users' everyday lives [18].

Participants in this study also highlighted the need for features that support the self-regulation of behavior to achieve personal exercise goals (eg goal setting, monitoring, and feedback), which supports previous findings [19]. The positive effect of personal goal attainment on individual well-being has been largely recognized [31] and the goal regulation process has been specifically described by Carver and Scheier [32] as a discrepancy-reducing feedback loop in which individuals are motivated to self-regulate their actions according to feedback to achieve a specific goal. Our findings suggest that the availability of self-regulation features might allow users to self-organize their experiences and behavior and hence, as proposed by self-determination theory [33], foster experiences of autonomy. Our findings, corroborated with results from recent reviews and qualitative studies, indicate that app features that support self-regulation are expected to increase engagement [11,16], are appreciated by users [34], and are associated with PA intervention effectiveness [4,35]. This suggests that self-regulation features should be considered as core components of PA apps.

Addressing users' PA participation motives emerged as a further important aspect of engagement with PA apps. In line with previous research [13], most participants preferred features that focus on fitness, nutrition, and weight loss as these were the main reasons for engaging in PA. For instance, some participants were strictly interested in suggestions about how to lose weight and wanted to track relevant parameters (eg, weight loss and calorie intake). Although intrinsic exercise motives, relative to more controlled ones (eg, body-related motives), have been shown to be positively associated with PA [36], some forms of less autonomous regulation (eg, identification with the outcomes of PA) are expected to regulate short-term behavior [37] and constitute an important motivational component of exercise [38]. Hence, app features that address users' exercise motives might be leveraged as a trigger for engagement with PA apps as well as for boosting motivation in the early stages of behavior change.

Users expressed varying types and levels of needs relating to the connectedness with other users. Some participants believed that opportunities for connecting with peers might help to develop a climate of social commitment that increases motivation and makes exercising a more enjoyable activity. However, the subtheme labeled *peer support* was not identified during the follow-up interviews. This might be because of the fact that participants, in spite of their initial inclination to seek support from peers, were unwilling to connect with strangers or because they were worried by social comparison.

The motivational aspect of human support was also mirrored by users' preferences for coaching features directly embedded into PA apps. These results support those from a focus group study about preferred PA app features [19], which found that coaching features were thought to be as an advantage in PA apps. Moreover, these findings are consistent with previous research suggesting that the desire to continue working with a digital behavior change intervention is higher when supported by the presence of human relational skills (eg, empathy and social dialog) designed into a computer interface, as this fosters the therapeutic *working alliance* [39]. Moreover, findings from a scoping review of Web-based interventions highlighted the potential for supportive virtual coaches to counteract low adherence in digital behavior change interventions [40]. It is worth noting that this subtheme was not identified during the follow-up interviews. A plausible explanation for this is that coaching features, which were mentioned as desirable by participants during the baseline interviews, were completely lacking from the selected PA apps.

Consistent with previous research into different behaviors [20,34], potential users of PA apps were reluctant to share PA information with their social networks. In our study, most participants thought that social comparison features would hinder their motivation to exercise because of exposure to others' negative judgments. However, some participants wanted to take part in social challenges, contingent on these being *friendly*, thus facilitating the receipt of mutual social support.

An additional theme labeled *tailored action planning* emerged only during the baseline interviews and relates to the role of the app in supporting users to implement their intentions and to achieve their goals. Users expressed their desire to be provided with tailored PA plans and suggestions based on their goals and current progress. This highlights the crucial role of action planning features in PA apps to facilitate goal attainment [41]. This theme was not identified as important during the follow-up interviews. A possible explanation is that action planning is a volitional rather than a motivational variable [41]. Hence, although action planning may help individuals to implement their intentions, it does not motivate and nudge PA app users to engage with the digital behavior change intervention itself over time.

After having used the app for 2 weeks, participants expressed the desire for PA apps with proactive features that boost their motivation. Participants who enjoyed their allocated app mentioned their liking of proactive features (eg, suggestions and reminders), and these were also mentioned as a point for improvement by participants who were not satisfied with their allocated apps. Hence, the development of a proactive app that is able to tailor the intervention content and timing according to individual differences is hypothesized to be a key element for supporting engagement with PA apps. These results are in line with recent advances in intervention design aimed at the development of adaptive interventions (eg, just-in-time adaptive interventions) that integrate behavioral, psychological, and contextual information to deliver more tailored and potentially effective strategies to increase PA [42,43]. Proactive features may actively support users' self-regulation skills, reducing users' responsibility for the behavior change process. Although this theme was not identified during the baseline interviews, it represented a key issue in the follow-up interviews when participants were prompted to reflect on their actual app use. This finding provides support for the argument that sustained engagement with PA apps cannot simply rely on users' motivation. Instead, motivation to continue engaging might be the result of moment-to-moment engagement that is triggered by motivational and proactive app features.

Implications

Thanks to the longitudinal design, this study facilitated the development of a comprehensive picture of potential users' needs and preferences regarding PA apps, with clear implications for the development of future, or modification of existing PA apps for physically inactive adults.

The findings that potential users are interested in self-regulatory features supported current design practice in PA apps; however, some improvements can be introduced. First, users should be able to monitor and assess the PA they perform in a more reliable and efficient way. Second, features that enable users to draw their attention to how their PA progresses and to any discrepancy between PA behavior and goals may help to guide their self-regulation and efforts.

As potential users are likely to practice PA to improve fitness and lose weight, PA apps should (1) deliver suggestions about what exercises to do and how to exercise to achieve specific health-related goals, (2) provide nutritional advice and opportunities for monitoring food intake, (3) monitor and provide feedback on health-related progress (eg, weight loss, waistline reduction, and blood pressure), and (4) permit users to estimate the daily and weekly difference between calorie intake and consumption.

While interacting with PA apps, simplicity and usability influenced users' perception of the app. Indeed, simple apps were seen as more easy to use and, thus, more immediate and useful. In particular, users expected that apps should not be cognitive demanding or time consuming. Therefore, developers should focus on designing PA apps that assure immediate and efficient use, avoiding poor user experiences that constitute a barrier to engagement and behavior change. To this end, design and implementation phases should be characterized by a constant and iterative evaluation of the user experience in real-world settings.

Further findings suggested that the design of PA apps may be improved by providing opportunities to connect users to one another and to easily arrange for social PA events. As pointed out, there may also be merit in designing PA apps that interact with users in the same way as a personal coach. Hence, developers may consider the opportunity to characterize PA apps by the presence of human relational skills or components (eg, virtual coaching, empathy, and social dialog).

Finally, proactive features (human support, tailored PA plans and suggestions, and context aware prompts) constitute one of the main gaps of commercial PA apps. This is a crucial point because, as emerged in this study, the development of a proactive app that is able to tailor the intervention content according to individual differences and diverse circumstances is hypothesized to be a key element for supporting effective and sustained engagement with digital interventions. To this end, a close collaboration between behavioral and computer scientists is required for developing machine learning techniques that can help understanding, predicting, and meaningfully addressing users' needs.

Limitations

This study also has a few limitations. First, participants' heterogeneous levels of digital literacy might have influenced the quality of the experience with the apps. For example, it is possible that some participants had difficulties with their allocated apps because of low digital literacy. We tried to minimize such limitations as far as possible by only including top-ranked apps, thus preventing participants from interacting with low-quality user interfaces. Second, we asked participants to use the app for 2 weeks to preserve the ecological validity of the study findings but did not specify amount and frequency of use. However, it is possible that the accuracy and depth of participants' answers in the follow-up interviews was influenced by the amount and frequency of app use during the 2 weeks. Moreover, 3 freely available apps were chosen for this study as stimuli to elicit participants' needs and preferences. Although the choice of these apps was based on criteria aiming to provide participants with high-quality apps, it is possible that a greater number of apps might have elicited different and even more generalizable insights. However, we believe that our findings may generalize to other commercial apps as the selected ones were characterized by most of the BCTs generally implemented in PA apps [28]. Finally, the 2-week longitudinal study design was aimed at understanding what and why particular app features sustained engagement over time. The second time point was defined on the basis of statistics suggesting that 74% of health apps are abandoned after their tenth use [12] and their usage mainly varies from a few times each week to 2 times each day [22] and, consequently, assuming that abandonment may occur around the end of the second week of use. Our results confirmed previous statistics [22]; however, none of the participants used the app for more than 7 times. Therefore, findings from this study cannot answer the question what and why design features influence engagement with PA apps over a longer period.

Conclusions

Our findings suggest that designers of PA apps may benefit from taking into account users' exercise motives (ie, fitness and weight loss) and supporting users with self-regulation features such as monitoring, feedback, and goal setting. As participants expressed a desire for features that foster the sense of relatedness, features that leverage the motivational aspects of relational and emphatic support in the behavior change process should be considered in the design of PA apps. Similarly, preferences for tailored action planning and proactive features drew attention to the need for the active role of the app as a complement to other features that support users' self-regulation skills.

Acknowledgments

The authors would like to acknowledge and thank the master students who supported data collection and first stage of data analysis: Sara Bergamini and Sophia Nicolosi.

Authors' Contributions

DB and PS conceived the study and defined the study protocol. DB, with support from two master students (SB and SN), recruited participants and analyzed the data. DB drafted the first version of the manuscript. OP and PS contributed to the data analysis and interpretation. All authors reviewed and edited the manuscript and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Think-aloud and semistructured interview protocol for baseline and follow-up.

[PDF File (Adobe PDF File), 39KB - mhealth_v7i2e11636_app1.pdf]

Multimedia Appendix 2

Additional excerpts from the think-aloud and interview sessions illustrating each theme.

[DOCX File, 23KB - mhealth_v7i2e11636_app2.docx]

References

1. Arem H, Moore SC, Patel A, Hartge P, de Berrington GA, Visvanathan K, et al. Leisure time physical activity and mortality: a detailed pooled analysis of the dose-response relationship. JAMA Intern Med 2015 Jun;175(6):959-967 [FREE Full text] [doi: 10.1001/jamainternmed.2015.0533] [Medline: 25844730]
2. Fogelholm M. Physical activity, fitness and fatness: relations to mortality, morbidity and disease risk factors. A systematic review. Obes Rev 2010 Mar;11(3):202-221. [doi: 10.1111/j.1467-789X.2009.00653.x] [Medline: 19744231]
3. Hallal PC, Andersen LB, Bull FC, Guthold R, Haskell W, Ekelund U, Lancet Physical Activity Series Working Group. Global physical activity levels: surveillance progress, pitfalls, and prospects. Lancet 2012 Jul 21;380(9838):247-257. [doi: 10.1016/S0140-6736(12)60646-1] [Medline: 22818937]
4. Greaves CJ, Sheppard KE, Abraham C, Hardeman W, Roden M, Evans PH, IMAGE Study Group. Systematic review of reviews of intervention components associated with increased effectiveness in dietary and physical activity interventions. BMC Public Health 2011;11:119 [FREE Full text] [doi: 10.1186/1471-2458-11-119] [Medline: 21333011]
5. Cobiac LJ, Vos T, Barendregt JJ. Cost-effectiveness of interventions to promote physical activity: a modelling study. PLoS Med 2009 Jul 14;6(7):e1000110 [FREE Full text] [doi: 10.1371/journal.pmed.1000110] [Medline: 19597537]
6. 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]
7. Patrick K, Griswold WG, Raab F, Intille SS. Health and the mobile phone. Am J Prev Med 2008 Aug;35(2):177-181 [FREE Full text] [doi: 10.1016/j.amepre.2008.05.001] [Medline: 18550322]
8. Müller AM, Alley S, Schoeppe S, Vandelanotte C. The effectiveness of e-& mHealth interventions to promote physical activity and healthy diets in developing countries: a systematic review. Int J Behav Nutr Phys Act 2016 Oct 10;13(1):109 [FREE Full text] [doi: 10.1186/s12966-016-0434-2] [Medline: 27724911]
9. Davies CA, Spence JC, Vandelanotte C, Caperchione CM, Mummery WK. Meta-analysis of internet-delivered interventions to increase physical activity levels. Int J Behav Nutr Phys Act 2012;9:52 [FREE Full text] [doi: 10.1186/1479-5868-9-52] [Medline: 22546283]
10. Yardley L, Spring BJ, Riper H, Morrison LG, Crane DH, Curtis K, et al. Understanding and promoting effective engagement with digital behavior change interventions. Am J Prev Med 2016 Nov;51(5):833-842. [doi: 10.1016/j.amepre.2016.06.015] [Medline: 27745683]
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;7(2):254-267. [doi: 10.1007/s13142-016-0453-1] [Medline: 27966189]
12. Cision PRWeb. Motivating Patients to Use Smartphone Health Apps URL: http://www.prweb.com/releases/2011/04/prweb5268884.htm [WebCite Cache ID 6zpLYAEmz]
13. Murnane E, Huffaker D, Kossinets G. Mobile health apps: adoption, adherence, and abandonment. In: Adjunct Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015

ACM International Symposium on Wearable Computers. 2015 Presented at: UbiComp/ISWC'15 Adjunct; September 7 - 11, 2015; Osaka, Japan p. 261-264. [doi: 10.1145/2800835.2800943]
14. Duncan M, Vandelanotte C, Kolt GS, Rosenkranz RR, Caperchione CM, George ES, et al. Effectiveness of a web- and mobile phone-based intervention to promote physical activity and healthy eating in middle-aged males: randomized controlled trial of the ManUp study. J Med Internet Res 2014;16(6):e136 [FREE Full text] [doi: 10.2196/jmir.3107] [Medline: 24927299]
15. Morrison LG, Geraghty AW, Lloyd S, Goodman N, Michaelides DT, Hargood C, et al. Comparing usage of a web and app stress management intervention: an observational study. Internet Interv 2018 Apr;12:74-82 [FREE Full text] [doi: 10.1016/j.invent.2018.03.006] [Medline: 30135771]
16. Bardus M, van Beurden SB, Smith JR, Abraham C. A review and content analysis of engagement, functionality, aesthetics, information quality, and change techniques in the most popular commercial apps for weight management. Int J Behav Nutr Phys Act 2016;13(1):35 [FREE Full text] [doi: 10.1186/s12966-016-0359-9] [Medline: 26964880]
17. Yardley L, Morrison L, Bradbury K, Muller I. The person-based approach to intervention development: application to digital health-related behavior change interventions. J Med Internet Res 2015;17(1):e30 [FREE Full text] [doi: 10.2196/jmir.4055] [Medline: 25639757]
18. van Gemert-Pijnen JE, Peters O, Ossebaard HC. Improving eHealth. The Hague, the Netherlands: Eleven International Publishing; 2013.
19. Middelweerd A, van der Laan DM, van Stralen MM, Mollee JS, Stuij M, te Velde SJ, et al. What features do Dutch university students prefer in a smartphone application for promotion of physical activity? A qualitative approach. Int J Behav Nutr Phys Act 2015 Mar 01;12:31 [FREE Full text] [doi: 10.1186/s12966-015-0189-1] [Medline: 25889577]
20. Perski O, Blandford A, Ubhi HK, West R, Michie S. Smokers' and drinkers' choice of smartphone applications and expectations of engagement: a think aloud and interview study. BMC Med Inform Decis Mak 2017 Feb 28;17(1):25 [FREE Full text] [doi: 10.1186/s12911-017-0422-8] [Medline: 28241759]
21. Puszkiewicz P, Roberts AL, Smith L, Wardle J, Fisher A. Assessment of cancer survivors' experiences of using a publicly available physical activity mobile application. JMIR Cancer 2016 May 31;2(1):e7 [FREE Full text] [doi: 10.2196/cancer.5380] [Medline: 28410168]
22. Robbins R, Krebs P, Jagannathan R, Jean-Louis G, Duncan DT. Health app use among US mobile phone users: analysis of trends by chronic disease status. JMIR Mhealth Uhealth 2017 Dec 19;5(12):e197 [FREE Full text] [doi: 10.2196/mhealth.7832] [Medline: 29258981]
23. Jaspers MW, Steen T, van den Bos C, Geenen M. The think aloud method: a guide to user interface design. Int J Med Inform 2004 Nov;73(11-12):781-795. [doi: 10.1016/j.ijmedinf.2004.08.003] [Medline: 15491929]
24. Morse J. Theoretical Saturation. In: Lewis-Beck MS, Bryman AE, Futing Liao T, editors. The SAGE Encyclopedia of Social Science Research Methods. Thousand Oaks: SAGE Publications; 2004.
25. Armstrong T, Bull F. Development of the World Health Organization Global Physical Activity Questionnaire (GPAQ). J Public Health 2006 Mar 2;14(2):66-70. [doi: 10.1007/s10389-006-0024-x]
26. Qualtrics. 2017. URL: https://www.qualtrics.com/ [accessed 2018-11-30] [WebCite Cache ID 74K8lDAMZ]
27. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. Ann Behav Med 2013 Aug;46(1):81-95. [doi: 10.1007/s12160-013-9486-6] [Medline: 23512568]
28. Bondaronek P, Alkhaldi G, Slee A, Hamilton FL, Murray E. Quality of publicly available physical activity apps: review and content analysis. JMIR Mhealth Uhealth 2018 Mar 21;6(3):e53 [FREE Full text] [doi: 10.2196/mhealth.9069] [Medline: 29563080]
29. IBM. 2014. IBM SPSS software URL: https://www.ibm.com/analytics/spss-statistics-software [accessed 2018-11-30] [WebCite Cache ID 74K8anLqY]
30. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
31. Steca P, Monzani D, Greco A, D'Addario M, Cappelletti E, Pancani L. The effects of short-term personal goals on subjective well-being. J Happiness Stud 2015 Jun 18;17(4):1435-1450. [doi: 10.1007/s10902-015-9651-7]
32. Carver CS, Scheier MF. Themes and issues in the self-regulation of behavior. In: Wyer Jr RS, editor. Perspectives on behavioural self-regulation: Advances in social cognition. 1st edition. New York: Psychology Press; 1999:1-105.
33. Deci EL, Ryan RM. The "what" and "why" of goal pursuits: human needs and the self-determination of behavior. Psychol Inq 2000 Oct;11(4):227-268. [doi: 10.1207/s15327965pli1104_01]
34. Dennison L, Morrison L, Conway G, Yardley L. Opportunities and challenges for smartphone applications in supporting health behavior change: qualitative study. J Med Internet Res 2013;15(4):e86 [FREE Full text] [doi: 10.2196/jmir.2583] [Medline: 23598614]
35. Knittle K, Nurmi J, Crutzen R, Hankonen N, Beattie M, Dombrowski SU. How can interventions increase motivation for physical activity? A systematic review and meta-analysis. Health Psychol Rev 2018 Sep;12(3):211-230. [doi: 10.1080/17437199.2018.1435299] [Medline: 29385950]

36. Teixeira PJ, Carraça EV, Markland D, Silva MN, Ryan RM. Exercise, physical activity, and self-determination theory: a systematic review. Int J Behav Nutr Phys Act 2012;9:78 [FREE Full text] [doi: 10.1186/1479-5868-9-78] [Medline: 22726453]
37. Deci E, Ryan RM. Intrinsic motivation and self-determination in human behavior. New York: Springer; 1985.
38. Edmunds J, Ntoumanis N, Duda JL. A test of Self-Determination Theory in the exercise domain. J Appl Social Pyschol 2006 Sep;36(9):2240-2265. [doi: 10.1111/j.0021-9029.2006.00102.x]
39. Bickmore T, Gruber A, Picard R. Establishing the computer-patient working alliance in automated health behavior change interventions. Patient Educ Couns 2005 Oct;59(1):21-30. [doi: 10.1016/j.pec.2004.09.008] [Medline: 16198215]
40. Scholten M, Kelders SM, van Gemert-Pijnen JE. A Scoped Review of the Potential for Supportive Virtual Coaches as Adjuncts to Self-guided Web-Based Interventions. Cham: Springer; 2017 Presented at: 12th International Conference, PERSUASIVE 2017; April 4–6, 2017; Amsterdam, The Netherlands p. 43-54. [doi: 10.1007/978-3-319-55134-0_4]
41. Sniehotta FF, Scholz U, Schwarzer R. Bridging the intention–behaviour gap: planning, self-efficacy, and action control in the adoption and maintenance of physical exercise. Psychol Health 2005 Apr;20(2):143-160. [doi: 10.1080/08870440512331317670]
42. Baretta D, Sartori F, Greco A, Melen R, Stella F, Bollini L, et al. Wearable devices and AI techniques integration to promote physical activity. In: Human-Computer Interaction with Mobile Devices and Services Adjunct. New York: ACM; 2016 Presented at: 18th International Conference on Human-Computer Interaction with Mobile Devices and Services; September 06-09, 2016; Florence, Italy. [doi: 10.1145/2957265.2965011]
43. Korinek EV, Phatak SS, Martin CA, Freigoun MT, Rivera DE, Adams MA, et al. Adaptive step goals and rewards: a longitudinal growth model of daily steps for a smartphone-based walking intervention. J Behav Med 2018 Feb;41(1):74-86. [doi: 10.1007/s10865-017-9878-3] [Medline: 28918547]

Abbreviations

BCT: behavior change technique
PA: physical activity

Edited by G Eysenbach; submitted 19.07.18; peer-reviewed by K Knittle, S Miller, P Wark; comments to author 30.08.18; revised version received 17.10.18; accepted 29.10.18; published 08.02.19.

Please cite as:
Baretta D, Perski O, Steca P
Exploring Users' Experiences of the Uptake and Adoption of Physical Activity Apps: Longitudinal Qualitative Study
JMIR Mhealth Uhealth 2019;7(2):e11636
URL: http://mhealth.jmir.org/2019/2/e11636/
doi:10.2196/11636
PMID:30735143

Original Paper

Can Brief, Daily Training Using a Mobile App Help Change Maladaptive Beliefs? Crossover Randomized Controlled Trial

María Roncero[1*], PhD; Amparo Belloch[1*], PhD; Guy Doron[2*], PhD

[1]Research and Treatment Unit for Obsessive-Compulsive and Related Disorders, Department of Personality, Faculty of Psychology, University of Valencia, Valencia, Spain

[2]Baruch Ivcher School of Psychology, Interdisciplinary Center Herzliya, Herzliya, Israel

[*]all authors contributed equally

Corresponding Author:
Guy Doron, PhD
Baruch Ivcher School of Psychology
Interdisciplinary Center Herzliya
PO Box 167
Herzliya, 46150
Israel
Phone: 972 99522850
Email: Gdoron@idc.ac.il

Abstract

Background: Obsessive-compulsive disorder (OCD) is a disabling condition with a wide variety of clinical presentations including contamination fears, fear of harm, and relationship-related obsessions. Cognitive behavioral models of OCD suggest that OC symptoms result from catastrophic misinterpretations of commonly occurring intrusive experiences and associated dysfunctional strategies used to manage them. OCD-related maladaptive beliefs including inflated responsibility, importance and control of thoughts, perfectionism, and intolerance for uncertainty increase the likelihood of such misinterpretations.

Objective: Considering accumulating evidence suggesting that mobile health (mHealth) apps based on cognitive-behavioral principles may lead to significant reductions in psychopathological symptoms, we assessed the effectiveness of a novel cognitive training app (GGRO) designed to challenge OCD-related beliefs.

Methods: A total of 97 students were randomized to groups undertaking immediate-use (iApp) or delayed use (dApp) of GGRO. All participants were requested to complete Web-based assessments, with questionnaires relating to maladaptive beliefs, mood, and OC symptoms at baseline (T1), 15 days from baseline (T2), and 30 days from baseline (T3). Participants in iApp group started using the app at baseline and continued using the app for 15 consecutive days. They were then requested to stop using the app until T3. Participants in the dApp group were requested to wait for 15 days and only then start using the app (crossover) for 15 consecutive days.

Results: All participants used the app for a mean of 14.07 (SD 1.41) days with 2.94 levels per day. Consistent with previous findings, app use was associated with medium-large effect size reductions in both iApp (n=51) and dApp (n=46) groups. In the iApp group, all effects remained significant during the 15 days of follow-up. Analyses focusing on the first two assessment occasions revealed significant treatment × repeated measures interactions on maladaptive beliefs, several OC symptom measures, and self-esteem.

Conclusions: This study provides further evidence for the efficacy of GGRO as a mobile-delivered training exercise that is useful for reducing OCD-related beliefs and symptoms.

Trial Registration: ClinicalTrials.gov NCT03571464; https://clinicaltrials.gov/ct2/show/NCT03571464 (Archived by WebCite at http://www.webcitation.org/7675sYPsH)

(JMIR Mhealth Uhealth 2019;7(2):e11443) doi:10.2196/11443

KEYWORDS

obsessive compulsive disorder; cognitive therapy; maladaptive beliefs; mobile apps; relationships

Introduction

Obsessive-compulsive disorder (OCD) is a disabling disorder that causes impairment in multiple areas of patients' lives [1,2]. OCD is characterized by the presence of repetitive unwanted and disturbing intrusive thoughts, images or urges (obsessions), or ritualistic and repetitive acts (compulsions) [3]. The content of OCD is heterogenic, comprising themes as scrupulosity [4], repugnant obsessions [5], moral and physical contamination fears [6], cleaning compulsions, obsessional doubts, and relationship-related obsessions [7].

Cognitive behavioral therapy (CBT) combined with exposure and ritual prevention is the first choice of psychological treatment recognized by the National Institute for Clinical Excellence [8]. CBT models of OCD postulate that catastrophic misinterpretation of intrusive thoughts or images and urges and the use of counterproductive cognitive and behavioral strategies to manage them lead to their escalation into chronic obsessions [9-11]. A number of maladaptive beliefs have been found to be associated with this catastrophic misinterpretation: inflated responsibility, overimportance of thoughts, desire to control one's thoughts, overestimation of threat, and need for certainty and perfectionism [12,13].

Many individuals, however, have great difficulties in accessing CBT therapy, either because of their high cost, the stigma associated with treatment, or the lack of available trained professionals [14,15]. Information and Communication Technologies including mobile apps and internet-based interventions have been suggested to increase accessibility and availability of CBT-based interventions [16-18]. Such alternative CBT-delivery systems are consistent with the stepped-care approach for OCD [8]. Clients with OCD may begin with low-intensity interventions (eg, self-help materials) and, if needed, gradually receive more intense and expert interventions [19].

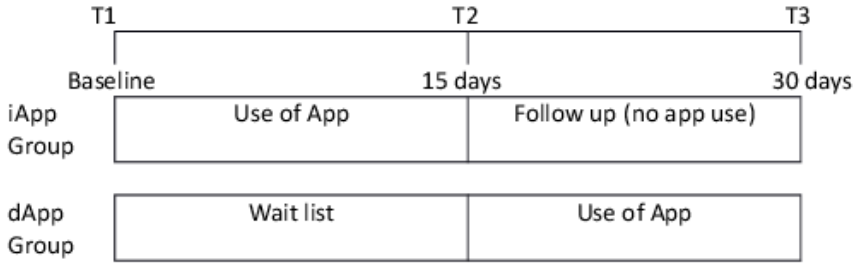
Information and Communication Technologies have been implemented to a significantly smaller extent in the treatment or prevention of OCD symptoms than in that of other mental disorders. Most studies assessed the efficacy of video-conference or telephone therapy used in exposure and response prevention [20-22]. In addition, most existing CBT-based mobile apps translate internet-delivered desktop treatment programs into mobile apps without exploiting the special advantages of the mobile app platform. These programs often have a long duration (eg, more than 20 min per interaction) and involve tasks requiring high internal motivation, a long attention span, and high persistence from users (eg, enter a significant amount of free text) [23].

Recently, an exploratory study evaluated a brief, game-like training exercise for challenging OCD beliefs delivered via a mobile app platform named "GGRO - GG relationship doubt and obsessions V1.1" [24]. GGRO is one of various mobile apps designed by GGApps (Herzliya, Israel) to challenge beliefs associated with a range of psychological difficulties (eg, depression, body image distress, and low self-esteem). GGRO was specifically designed to challenge maladaptive beliefs that underlie common OCD symptoms (eg, contamination and repugnant thoughts) as well as relationship obsessions (eg, obsessive preoccupations regarding the suitability of the relationship or the relationship partners) [7,12,25]. The platform was designed to help users learn to respond to statements that challenge OCD-related beliefs by embracing them (ie, pulling them down toward themselves) and rejecting statements that are consistent with beliefs underlying OCD symptoms and low self-esteem (ie, throwing them upward, away from themselves; see Methods section). Following CBT principles, statements challenging OCD-related beliefs include alternative, more adaptive interpretations of thoughts, emotions, and events as well as statements encouraging approach behavioral strategies (eg, tolerance of negative feelings and acceptance of thoughts). Increasing accessibility to such statements is expected to reduce adherence to OCD-related beliefs and associated symptoms.

The results of this study, which included 20 participants from a nonclinical population, suggested that training for 3 min a day for a period of 15 days was associated with significant, large effect-size reductions in the levels of OCD-related beliefs measured by the Obsessive Beliefs Questionnaire (OBQ) - Short form [24]. Participants also showed significant pre-post training decreases in the levels of OCD symptoms, measured by the Obsessive-Compulsive Inventory - Reduced (OCI-R) [26], including relationship-related OCD symptoms measured using the Relationship Obsessive-Compulsive Inventory (ROCI) [27] and the Partner-Related Obsessive-Compulsive Symptoms Inventory (PROCSI; R Moulding, GD, unpublished data, 2019). Moreover, pre-post changes in the levels of OCD-beliefs were associated with a reduction in OCD symptom levels.

The aim of this study was to further evaluate the efficacy of GGRO in reducing OCD-related maladaptive beliefs and OCD symptoms. Specifically, a randomized controlled trial with crossover design was carried out in a nonclinical student population to assess pre-post changes in the levels of OCD-related maladaptive beliefs and OCD symptoms, including relationship OCD (ROCD) symptoms, self-esteem, and depression symptoms following 15 days of GGRO use. Our main hypothesis was that students using GGRO immediately following baseline assessment (immediate-use App group, iApp) would exhibit greater declines in obsessive compulsive–related beliefs than students who did not use GGRO in this phase of the study (delayed-use App group, dApp; Figure 1). Consistent with previous research showing an association between OCD symptoms and self-esteem [28,29], we expected a decrease in OCD and ROCD symptoms and an increase in self-esteem in students in the iApp group relative to the students in the dApp control group. Following crossover (T2), we expected that user gains in the iApp group would be maintained in T3. In this phase, we anticipated that students starting to use GGRO (dApp) would show statistically significant reductions in OCD-related maladaptive beliefs and symptoms and an increase in self-esteem from T2 to T3 assessments. Consistent with a previous study using GGRO in a student population [24], we did not expect a significant reduction in depression symptoms.

Figure 1. Study design with both groups. iApp: immediate-use App; dApp: delayed-use App.



Methods

Participants

Participants were recruited from the University of Valencia during the first semester of the 2016/2017 course from nine classes at the Psychology Faculty. The students were invited to voluntarily participate in a study about beliefs, self-talk, mood, and relationships. Participants interested in participating were informed of their rights and provided online informed consent in accordance with university Institutional Review Board standards. The study received the approval of the University of Valencia ethics committee (H-1488382719361). Inclusion criteria included native Spanish speaking, experience of at least one stable romantic relationship, and possession of a mobile device capable of installing GGRO (available for download via Google Play or the App Store).

Consistent with common practice in OCD-related research, the sample used in the present study comprised nonclinical participants [30]. Like individuals who are clinically diagnosed with OCD, nonclinical participants tend to engage in compulsive behaviors to alleviate distress [31]. Furthermore, taxometric studies of OCD [32] have found that OCD symptoms and obsessive compulsive–related beliefs are best conceptualized as continuous dimensional rather than categorical.

All participants volunteered and were included in a draw for a prize of a dinner for two (valued at 30€). A total of 98 students attended a recruitment seminar wherein they were explained the general procedure of the research. They were then asked to download that app and complete the pretreatment evaluation (Time 1, T1) on Qualtrics [33], which is a secure online survey platform. Emails with the corresponding survey links were sent to participants at Time 2 (T2) and Time 3 (T3). From the 98 students who wanted to take part in the study, one was excluded because he did not have a stable partner in the present or past. The final 97 participants (79 women, 81.4%) were second- and third-year students in the Bachelor of Arts program, with ages ranging from 18 to 65 years (mean 21.56; SD 7.07). The majority (61.9%) reported having a medium socioeconomic status (28.9% below average and 9.3% above average). More than half of the participants (56.3%) were in a romantic relationship at the time of the study (median length, 33 months).

Study Design

The study was a randomized controlled trial with a crossover design (Figure 1). The intervention was a mobile-delivered cognitive training using GGRO. Participants were randomized to an App first group (iApp, n=51) or a wait list crossover group (dApp, n=46). Participants in the iApp group started using the app immediately (T1) for a period of 15 consecutive days (until T2). They were then requested to stop using the app until the end of the trial. Participants randomized to the dApp group were requested to start using the app at T2 (15 days after the iApp group). They were then requested to use the app (crossover) for the following 15 days. In both groups, participants were instructed to complete 3 levels a day (approximate 3 min a day). The CONSORT-EHEALTH checklist is presented as Multimedia Appendix 1.

Randomization

Randomization was carried out in a 1:1 ratio and based on a prespecified computer-generated randomization list [34]. Group assignment was performed onsite using the next available number on the randomization list.

Intervention

GGRO was developed by the author GD, an expert in OCD and related disorders, in collaboration with Gur Ilany, a mobile platform developer. This app consists of training exercises intended to help people increase accessibility to functional self-statements that facilitate adaptive interpretations of thoughts, emotions, and events associated with OCD (Figure 2). Users are presented with “blocks” featuring statements such as “I take things as they come” or “Everything can end in a catastrophe.” Users then have to respond to these statements by either embracing them (ie, pulling the “blocks” downwards toward themselves) or rejecting them (ie, throwing the “blocks” upward away from themselves).

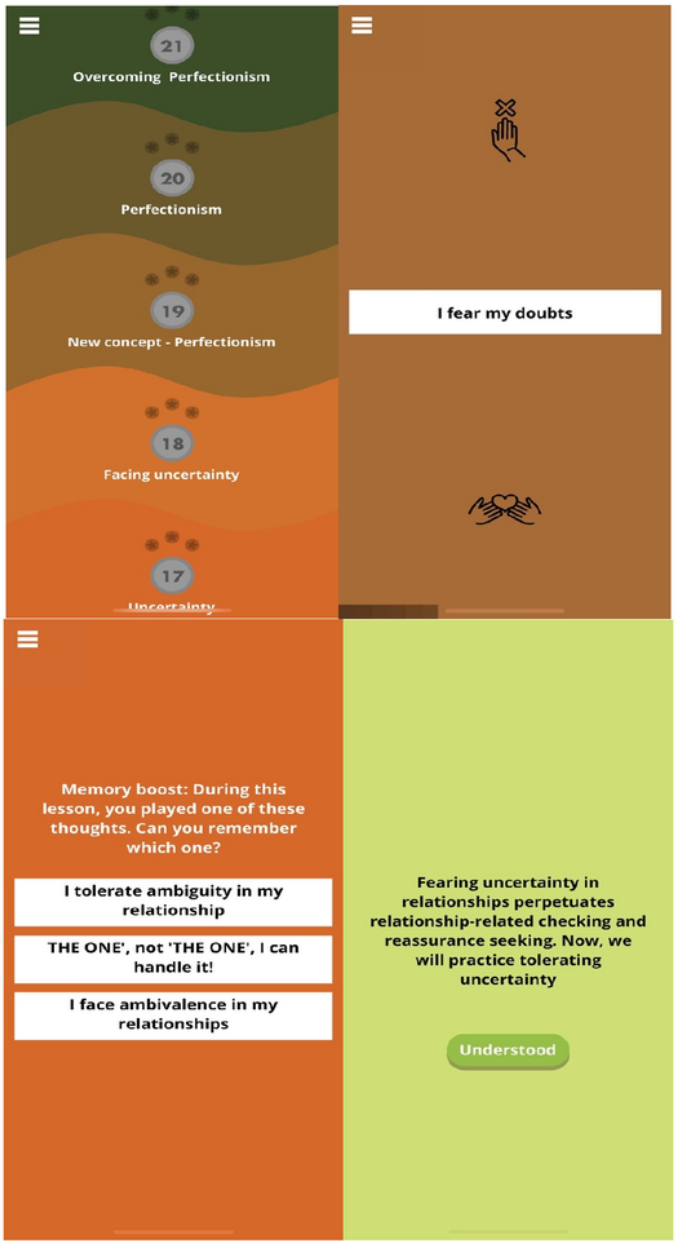
Users progressively completed 45 levels dedicated to OCD-related maladaptive beliefs (3 levels per belief) such as dealing with threat, importance of thoughts, and overcoming perfectionism. In this way, the user is exposed to alternative interpretations of the relevant maladaptive belief in each stage, increasing accessibility to functional self-statements that encourage adaptive interpretations for thoughts, emotions, and events (eg, the occurrence of distressing doubts) associated with OCD. For instance, statements challenging perfectionism may include “Mistakes teach me to overcome my fears” and

"Imperfect[ion] is human." Users are also encouraged to adopt approach behavioral strategies (rather than avoidance) including tolerance of negative emotions by responding to statements such as "I can tolerate doubts."

Following the completion of each level, the user receives feedback, depending on the length of time it took them to complete the level (0 to 3 stars). A short memory-evaluation screen (ie, memory boost) follows this feedback. In this screen, three statements are presented to the user. The user has to recall which of the statements appeared in the level he/she just completed. The correct response results in a "Correct!" message, and an incorrect response is followed by "You'll get it next time" feedback message. The two types of feedback increase attention to the training and encourage engagement.

The user then progresses to the next level. Three levels address a specific maladaptive belief. Before dealing with a new belief, a screen is presented with the rationale for challenging the specific maladaptive belief. For example, before learning to challenge overestimation of threat, users are presented the statement, "The world can be dangerous, but the tendency to look for danger all the time increases fears and anxieties. Let's learn to reduce this tendency!" Following completion of six levels pertaining to two beliefs (eg, importance of thoughts and overestimation of threat), users may see an encouraging statement such as "Excellent! Now you've learned how to better deal with your thoughts and to better recognize the way you overestimate threat." Push notifications remind users to use the app each day. Following the completion of 3 levels on a given day, a screen prompting users to stop using the app for that day appears. Users are also advised to train once a day at a preset time rather than in response to distressing thoughts or events. GGRO requires a mobile device with an operating system iOS 7 or above or android 4.2 or above.

Figure 2. GGRO screenshot.



Measures

The Obsessive-Compulsive Inventory - Reduced

The OCI-R [26,27] is a self-report inventory composed of 18 items ranked on a 5-point Likert scale, ranging from 0 (*not at all*) to 4 (*extremely*), which assess OCD symptoms. The OCI-R possesses good internal consistency for the total score (alphas ranging from .81 to .93 across samples) [26]. In our study, the internal consistency for the total scale (Cronbach alpha) was .84 at T1, .83 at T2, and .83 at T3.

The Obsessive Beliefs Questionnaire - Short Form

The OBQ-20 [12] is the abbreviated version of the 44-item OBQ - Revised [12]. The OBQ-20 is a self-report questionnaire assessing pan-situational cognitions associated with OCD. It is composed of 20 items ranked on a 7-point scale, ranging from 1 (*disagree very much*) to 7 (*agree very much*). The OBQ-20 has shown satisfactory psychometric properties [35,36]. The internal consistency of the scale as a whole in our sample (Cronbach alpha) was .88 at T1, .93 at T2, and .94 at T3.

The Relationship Obsessive-Compulsive Inventory - Short Version

The ROCI - Short version (S) is a shortened version of the ROCI [37], a 12-item measure assessing three dimensions of relationship-centered ROCD symptoms: love for the partner, the "rightness" of the relationship, and the partner's love for the participant. The ROCI-S consists of 6 items, of which 2 items assess each of the three abovementioned relationship-centered ROCD dimensions (the 2 items showing the highest average loaded on the two original ROCI validation studies; R Moulding, GD, unpublished data, 2019). In an independent sample (n=714; 302 women; mean age 38.73 years; SD 12.65 years), the mean of these six items (ROCI-S total score) showed good reliability (Cronbach alpha=.85) and correlated very highly (r=.97) with the total ROCI total scores. In the current study, the internal consistency (Cronbach alpha) of the mean of all ROCI items was .80 at T1, .83 at T2, and .79 at T3.

The Partner-Related Obsessive-Compulsive Symptoms Inventory - 6-Item Version

The PROCSI - 6-item version (Si) (R. Moulding, G.D., unpublished data, 2019) is an abbreviated version of the PROCSI [38], a 24-item measure assessing partner-focused ROCD symptoms. The PROCSI-Si consists of 6 items. The items selected showed the highest correlation of a single item with the relevant subscale in one-half of a randomly split sample (n=356; 151 women; mean age 38.58 years, SD 12.55 years). The mean of these six items (PROCSI-Si total score) showed good reliability (Cronbach alpha=.90) and correlated very highly (r=.98) with the total PROCSI total scores in this sample. The PROCSI-Si total score also showed good reliability scores (Cronbach alpha=.92) and correlated highly (r=.98) with the PROCSI total scores in the independent half of the sample (n=356; 151 females; mean age 38.88 years; SD 12.79 years). The internal consistency of PROCSI-Si (Cronbach alpha) in the current sample was .78 at T1, .83 at T2, and .77 at T3.

The Depression, Anxiety, Stress Scale - Short Version

The short version of the Depression, Anxiety, Stress Scale (DASS) [39-42] is a self-report questionnaire that evaluates negative emotional symptoms (depression, anxiety, and stress). The short version consists of 21 items rated on a 4-point scale, ranging from 0 (*did not apply to me at all*) to 3 (*applied to me very much or most of the time*). In this study, only the depression scale (7 items) was used. The DASS scales have been shown to have high internal consistency [42]. The internal consistency of the depression scale (Cronbach alpha) in the current sample was .90 at T1, .90 at T2, and .91 at T3.

The Single-Item Self-Esteem Scale

The Single-Item Self-Esteem Scale (SISE) [43] is a self-report measure that determines the extent to which the sentence "I have a high self-esteem" describes participants on a 9-point scale, ranging from 1 (*not very true for me*) to 9 (*very true for me*). The SISE has been found to have high test-retest reliability, criterion validity coefficients above .80 (median=.93 after correcting for unreliability) with the Rosenberg Self-Esteem Scale, and a similar pattern of construct validity coefficients as the Rosenberg Self-Esteem Scale with 35 different constructs [43]. Using longitudinal data, Robins et al [43] estimated the reliability of the SISE to be .75.

Statistical Analysis

Statistical analyses were performed using Statistical Package for the Social Sciences (SPSS Inc, Chicago, IL). In order to avoid overoptimistic estimates of the efficacy of the training [44], an intention-to-treat analysis using the last-observation-carried-forward method was used [45]. Descriptive statistics were used to report means, SDs, and frequencies. In addition, *t* and χ^2 tests were performed to assess differences between groups and in age, relationship duration (in months), sex, socioeconomic level, belief, and symptoms measures (OBQ-20, OCI-R, PROCSI-Si, ROCI, DASS, and SISE). A series of repeated measures analysis of variance with Bonferroni adjustments was performed to evaluate pre-post scores in both study groups. The Effect Size Determination Program [46] was used to calculate Cohen *d* values.

Results

Principal Findings

A total of 97 participants met the inclusion criteria and participated in the study. Mean scores for outcome measures and characteristics of the two groups did not differ significantly at baseline (Table 1).

Table 1. Descriptive statistics and comparisons between immediate-use App (iApp) group and delayed-use App (dApp) group in sociodemographic variables and outcome measures at baseline.

Characteristics	iApp (n=51)	dApp (n=46)	t / χ^2	*df*	*P* value	Cohen *d*
Age (years), mean (SD)	22.88 (9.23)	20.09 (2.73)	1.97	95	.05	0.4
Gender, %						
Men	25.5	10.9	3.42	1	.06	0.37
Women	74.5	89.1				
Socioeconomic status, %						
Low	3.9	2.2	3.64	3	.30	0.15
Medium-low	31.4	19.6				
Medium	52.9	71.7				
Medium-high	11.8	6.5				
Relationship duration (months), mean (SD)	45.37 (86.96)	19.30 (14.78)	2.00	95	.05	0.4
OCI-R[a] (score), mean (SD)	1.79 (0.57)	1.78 (0.40)	0.09	95	.92	0.02
OBQ-20[b] (score), mean (SD)	3.24 (0.96)	3.08 (0.84)	0.84	95	.40	0.17
ROCI-S[c] (score), mean (SD)	1.76 (0.69)	1.73 (0.61)	0.24	95	.81	0.05
PROCSI-Si[d] (score), mean (SD)	1.66 (0.57)	1.62 (0.69)	0.29	95	.77	0.06
DASS-D[e] (score), mean (SD)	1.74 (0.67)	1.53 (0.49)	1.74	95	.08	0.35
SISE[f] (score), mean (SD)	3.12 (1.09)	3.37 (0.97)	-1.20	95	.23	0.24

[a]OCI-R: Obsessive-Compulsive Inventory - Reduced.

[b]OBQ-20: Obsessive Beliefs Questionnaire-20.

[c]ROCI-S: Relationship Obsessive-Compulsive Inventory - Short version.

[d]PROCSI-Si: Partner-Related Obsessive-Compulsive Symptoms Inventory.

[e]DASS-D: Depression, Anxiety, Stress Scale - Depression.

[f]SISE: Single-Item Self-Esteem Scale.

At 15 days (T2), 79 of 97 participants (81.4%) completed the study and 62 of 97 (65.1%) completed the 15-day follow-up (T3). Participants who dropped out during the study period did not differ in age (t_{95}= 0.58, *P*=.56), gender (χ^2_1=3.3, *P*=.79), relationship duration (t_{95}= 0.04, *P*=.97), or socioeconomic status (χ^2_3=3.3, *P*=.34) compared to participants who did not drop out (Figure 3).

Tables 2 and 3 present the means and SDs for iApp and dApp participants, respectively, on all measures and testing occasions. All participants used the app for a mean of 14.07 (SD 1.41) days, with a mean of 2.94 (SD 0.37) levels per day. Additionally, the mean of the highest level completed by participants was 40.93 (SD 10.20) levels of the 45 levels. There were no significant differences between the two groups with regard to days used, mean of levels per day, and highest level achieved.

Between-Group Differences (iApp Group Versus dApp Group)

Analyses of the first two assessment occasions (T1 and T2) revealed significant treatment × repeated measures interactions in OBQ ($F_{1,95}$=17.06, *P*<.001, *d*=0.84, PROCSI-Si ($F_{1,95}$=4.28, *P*=.04, *d*=0.42), and SISE ($F_{1,95}$=4.36, *P*=.04, *d*=0.42). These results indicated that students in the iApp group exhibited fewer OCD-related beliefs, fewer partner-focused ROCD symptoms, and higher self-esteem than their waiting list counterparts on the second assessment occasion (Figure 4).

iApp Group Within-Group Effects and 15-Day Follow-Up Effects

In the iApp group, we expected pre-post reduction in OCD-related beliefs and symptoms as well as retention of these effects in the follow-up period. Thus, pre-to-final changes were specifically examined via repeated measures analysis of variance between T1 and T3 and between T2 and T3. A significant decline in pre-to-final changes was found in the OBQ, OCI-R, PROCSI-Si, ROCI-S, and SISE scores. Further, the differences found between T1 and T2 were maintained in T3. Moreover, the only statistically significant difference found was in the PROCSI-Si scores that indicated an additional significant improvement between T2 and T3 (Table 2).

Figure 3. CONSORT flow diagram of participants through the trial.

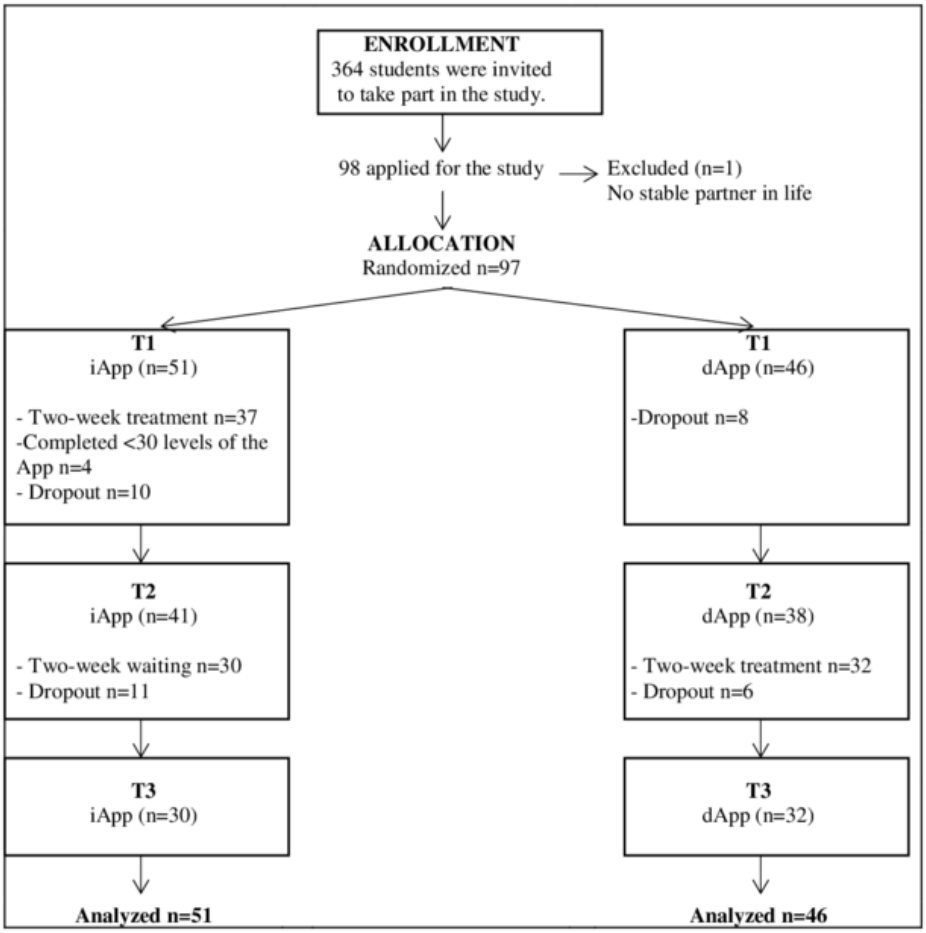


Table 2. Descriptive statistics and comparisons among periods for the immediate-use App (iApp) group.

Scale	T1 (points), mean (SD)	T2 (points), mean (SD)	T3 (points), mean (SD)	T1 vs T3			T2 vs T3		
				$F_{1,50}$	*P* value	*d*	$F_{1,50}$	*P* value	*d*
OCI-R[a]	1.79 (0.57)	1.59 (0.50)	1.59 (0.49)	10.87	.002	0.65	0.00	.99	0
OBQ-20[b]	3.24 (0.96)	2.66 (1.10)	2.57 (1.18)	51.39	<.001	1.42	1.20	.28	0.21
ROCI-S[c]	1.76 (0.69)	1.60 (0.62)	1.58 (0.63)	5.65	.02	.47	0.07	.79	0.05
PROCSI-Si[d]	1.66 (0.57)	1.49 (0.55)	1.39 (0.48)	30.00	<.001	1.109	5.98	.02	0.48
DASS-D[e]	1.74 (0.67)	1.62 (0.64)	1.69 (0.69)	0.53	.47	0.14	3.84	.06	0.39
SISE[f]	3.12 (1.09)	3.31 (1.14)	3.33 (1.21)	7.13	.01	.53	0.11	.74	0.06

[a]OCI-R: Obsessive-Compulsive Inventory.

[b]OBQ-20: Obsessive Beliefs Questionnaire - Short form.

[c]ROCI-S: Relationship Obsessive-Compulsive Inventory - Short version.

[d]PROCSI-Si: Partner-Related Obsessive-Compulsive Symptoms Inventory - Six item version.

[e]DASS-D: Depression, Anxiety, Stress Scale-Depression subscale.

[f]SISE: Single-Item Self-Esteem Scale.

Table 3. Descriptive statistics and comparisons among periods for the delayed-use App (dApp) group.

Scale	T1 (points), mean (SD)	T2 (points), mean (SD)	T3 (points), mean (SD)	T1 vs T2			T2 vs T3		
				$F_{1,45}$	*P*	*d*	$F_{1,45}$	*P*	*d*
OCI-R[a]	1.78 (0.40)	1.66 (0.36)	1.54 (0.32)	7.28	.01	0.56	9.09	.004	0.61
OBQ-20[b]	3.08 (0.84)	3.02 (1.03)	2.48 (1.04)	.41	.53	0.13	27.52	.001	1.07
ROCI-S[c]	1.73 (0.61)	1.66 (0.62)	1.43 (0.38)	1.70	.20	0.27	12.52	.001	0.72
PROCSI-Si[d]	1.62 (0.69)	1.60 (0.72)	1.38 (0.50)	.10	.75	0.06	9.41	.004	0.62
DASS-D[e]	1.53 (0.49)	1.52 (0.50)	1.44 (0.37)	.00	.96	0	2.35	.13	0.32
SISE[f]	3.37 (0.97)	3.35 (1.01)	3.61 (0.91)	.07	.78	0.05	6.75	.01	0.53

[a]OCI-R: Obsessive-Compulsive Inventory.

[b]OBQ-20: short form of the Obsessive Beliefs Questionnaire.

[c]ROCI-S: Relationship Obsessive-Compulsive Inventory - short version.

[d]PROCSI-Si: Partner-Related Obsessive-Compulsive Symptoms Inventory - 6-item version.

[e]DASS-D: Depression, Anxiety, Stress Scale-Depression subscale.

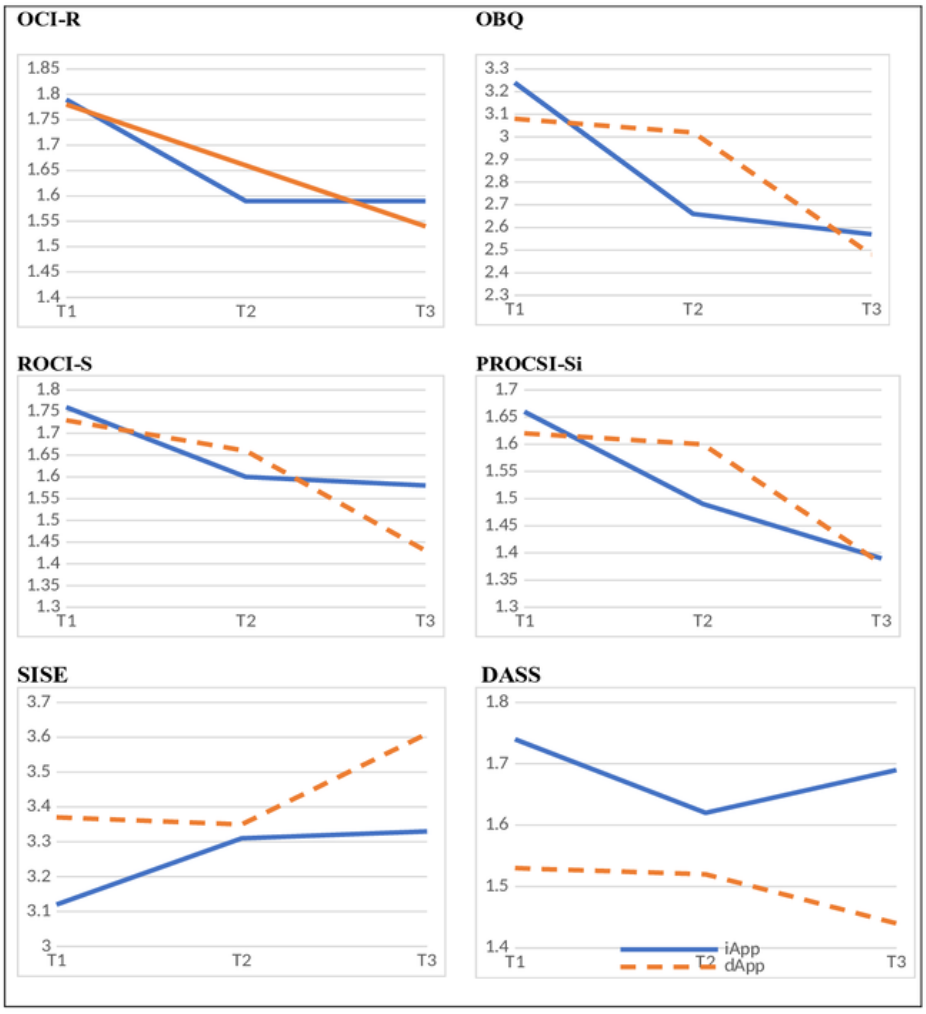
[f]SISE: Single-Item Self-Esteem Scale.

dApp Group Within-Group Effects

In the dApp group, we expected that crossover (ie, use of the app) would be associated with a significant decrease in OCD beliefs and symptom measures. Indeed, within-group differences between T2 and T3 following the crossover indicated significant reductions in the OBQ, PROCSI-Si, ROCI-S, and SISE scores. No differences were found in the DASS scores.

Unexpectedly, participants showed a significant decrease in OCD symptoms (OCI-R) between T1 and T2. Nevertheless, additional significant reduction in OCI-R scores was found between T2 and T3 (Table 3).

Figure 4. Graphs of the measures across T1, T2, and T3 for iApp and dApp groups. iApp: immediate-use App; dApp: delayed-use App; DASS: Depression, Anxiety, Stress Scale - Short version; OBQ: Obsessive Beliefs Questionnaire; OCI-R: Obsessive-Compulsive Inventory; PROCSI-Si: Partner-Related Obsessive-Compulsive Symptoms Inventory - 6-item version; ROCI-S: The Relationship Obsessive-Compulsive Inventory - Short version; SISE: The Single-Item Self-Esteem Scale.



Discussion

Mobile apps based on CBT principles have unique advantages including wide reach, continuous availability, appeal to young people, very low cost, and progress monitoring. Accumulating evidence shows that such apps can lead to significant reductions in psychopathological symptom and maladaptive behaviors [47-49]. The present randomized control study evaluated the efficacy of a mobile app platform named GGRO, which was designed to challenge OCD-related maladaptive beliefs. Consistent with a previous exploratory investigation [24], our results indicated that 15 days of brief daily training using GGRO led to significant reductions in OCD-related beliefs. Moreover, reductions in OCD-related beliefs were maintained for a follow-up period of 2 weeks. These results provide support for the stepped-care approach for OCD [8], suggesting that OCD-related beliefs and symptoms can be reduced using alternative, low-intensity modes of treatment delivery.

Relative to the waitlist control group in our study, individuals using GGRO for 2 weeks showed fewer OCD-related beliefs, fewer ROCD symptoms, and higher self-esteem. Moreover, the change after the training was stably maintained after 15 days (follow-up). Once our waiting list group started using GGRO (after crossover), participants in this group showed a reduction in OCD-related beliefs, OCD, and ROCD symptoms. Interestingly, an exception was partner-focused OC symptoms, which showed further improvements at follow-up, indicating a generalization of results and suggesting that the maladaptive beliefs targeted in GGRO may be particularly relevant to this ROCD presentation. Indeed, GGRO includes levels related to partner-value contingency of self (ie, self-esteem that is overly dependent on the partner's perceived value [50]), which may be particularly pertinent to partner-focused ROCD symptoms [51]

Unexpectedly, OCD symptoms declined in both the iApp and dApp groups during the initial 2 weeks. Fluctuations in the intensity of student requirements/evaluations may have coincided with the time period of our study and attenuated participants' OCD symptoms in the waitlist group. Indeed, most of the student data were collected during the mid-semester, a

period when students have fewer exams and other evaluations. Importantly, however, the use of GGRO was associated with a further reduction in OCD symptoms following the crossover, thus supporting the efficacy of GGRO in reducing OCD symptoms.

Consistent with a previous study [24], the levels of depression did not show any statistically significant change. Indeed, GGRO was specifically designed to target OCD-related beliefs and symptoms. GGRO was not designed to challenge depression-related maladaptive beliefs per se (eg, hopelessness and helplessness) and therefore does not include content designed to challenge depression-related beliefs. Considering that participants' depression levels were relatively low, this may explain why the depression symptoms did not show a significant reduction.

Our findings are consistent with those of previous research showing the efficacy of CBT-based apps in cognitive and behavioral change [52,53]. According to these models, commonly occurring intrusive experiences escalate into obsessions due to their catastrophic misinterpretations [12]. Strongly held maladaptive beliefs such as perfectionism, exaggerated importance attributed to the content of thoughts or their occurrence, and low tolerability for uncertainty increase the likelihood of such catastrophic interpretations.

The GGRO app evaluated in this study was designed to challenge maladaptive beliefs associated with OCD. By introducing users to interpretations that are inconsistent with their OCD-related beliefs, adherence to such beliefs was expected to weaken. This, in turn, was expected to reduce catastrophic interpretations of intrusions and decrease OCD symptoms. Indeed, previous findings showed that a change in OCD belief levels among users of GGRO was associated with a reduction in OCD symptoms further from the OCD symptom levels before using GGRO [24]. Consistent with this finding, the results of our study suggest that daily training, which involved active response from users to catastrophic interpretations of intrusions and their alternatives, was shown to lead to a significant reduction in maladaptive beliefs and associated symptoms.

Although the findings of our study are consistent with the expectations, our study has some important limitations. The sample used in our study comprised mainly female students from the general population. Indeed, the prevalence of OCD has been observed to be equal between men and women, or slightly higher in women [3]. Recent reviews support the utility of nonclinical participants in OCD-related research [31]. Moreover, initial evidence suggests that the use of the GG platform with individuals presenting with OCD [54] may reduce OCD-related beliefs and symptoms. Nevertheless, clinical populations may be different from nonclinical participants in symptom-related impairment; the lack of such symptom-related impairment in nonclinical participants may facilitate reduction in OCD-related beliefs and symptoms compared to a clinical population. In addition, the absence of mental disorders was not confirmed by clinical interviews. Future studies may benefit from evaluating the usefulness of GGRO in individuals with OCD.

Previous research using similar methodologies showed dropout rates comparable to ours [55,56]. We also performed intention-to-treat analysis with the last observation carried forward method [45] to prevent overestimation of treatment effects. Nevertheless, care should be taken in the interpretation of our results. Future studies may benefit from the use of additional dropout-reduction strategies (eg, monetary or course credit compensation).

A great majority of mobile apps designed for OCD are oriented toward self-applied therapy and track or guide exposure and response prevention [57,58]. However, their efficacy has not been empirically demonstrated with controlled studies [23,59]. In this regard, this randomized control study furthers our knowledge about the efficacy of alternative CBT-delivery systems for OCD.

GGRO was designed as a brief and easy training platform to challenge maladaptive beliefs and associated interpretations of thoughts and events. As such, this platform could complement traditional CBT interventions as an intersession work or relapse-prevention tool; thus, it is an instrument at the service of the therapist, and not a way to replace CBT. Moreover, this cost-effective and accessible mobile platform could be used in populations at risk of OCD and related disorders to reduce levels of maladaptive beliefs. Future studies should assess the usefulness of similar apps for other symptoms such as body image distress and depression. Indeed, reducing the levels of maladaptive beliefs in at-risk populations using cost-effective, accessible mobile platforms such as the one used in this study may increase resilience to a wide variety of psychological disorders. Furthermore, such a platform may be useful for relapse prevention following treatment.

Acknowledgments

This study has been supported by the Spanish MINECO-Grant PSI2013-44733-R and Acciones Especiales, University of Valencia Grant UV-INV-AE17-706208.

Conflicts of Interest

GD is a codeveloper of GGRO. GD is also a cofounder of GGapps.net. GGRO is the subject of this evaluation and therefore has financial interest to GGapps.net.

Multimedia Appendix 1

CONSORT - EHEALTH checklist (V 1.6.1).

[PDF File (Adobe PDF File), 603KB - mhealth_v7i2e11443_app1.pdf]

References

1. Bobes J, González MP, Bascarán MT, Arango C, Sáiz PA, Bousoño M. Quality of life and disability in patients with obsessive-compulsive disorder. Eur Psychiatry 2001 Jun;16(4):239-245. [Medline: 11418275]
2. Huppert J, Simpson H, Nissenson K, Liebowitz M, Foa E. Quality of life and functional impairment in obsessive-compulsive disorder: a comparison of patients with and without comorbidity, patients in remission, and healthy controls. Depress Anxiety 2009;26(1):39-45 [FREE Full text] [doi: 10.1002/da.20506] [Medline: 18800368]
3. American Psychiatric Association. Diagnostic and statistical manual of mental disorders: DSM-5. Washington, DC: American Psychiatric Association; 2019.
4. Abramowitz J, Jacoby R. Obsessive-compulsive and related disorders: a critical review of the new diagnostic class. Annu Rev Clin Psychol 2015;11:165-186. [doi: 10.1146/annurev-clinpsy-032813-153713] [Medline: 25581239]
5. Moulding R, Aardema F, O'Connor KP. Repugnant obsessions: A review of the phenomenology, theoretical models, and treatment of sexual and aggressive obsessional themes in OCD. Journal of Obsessive-Compulsive and Related Disorders 2014 Apr;3(2):161-168. [doi: 10.1016/j.jocrd.2013.11.006]
6. Elliott CM, Radomsky AS. Meaning and mental contamination: Focus on appraisals. Clin Psychol 2013 Mar 12;17(1):17-25. [doi: 10.1111/cp.12002]
7. Doron G, Derby D, Szepsenwol O. Relationship obsessive compulsive disorder (ROCD): A conceptual framework. Journal of Obsessive-Compulsive and Related Disorders 2014 Apr;3(2):169-180. [doi: 10.1016/j.jocrd.2013.12.005]
8. National Institute for Clinical Excellence. Obsessive-compulsive Disorder: Core Interventions in the Treatment of Obsessive-compulsive Disorder and Body Dysmorphic Disorder (Nice Guidelines). London, UK: Royal College of Psychiatrists; 2006.
9. Clark DA, Radomsky AS. Introduction: A global perspective on unwanted intrusive thoughts. Journal of Obsessive-Compulsive and Related Disorders 2014 Jul;3(3):265-268. [doi: 10.1016/j.jocrd.2014.02.001]
10. Rachman S, de Silva P. Abnormal and normal obsessions. Behaviour Research and Therapy 1978;16(4):233-248. [doi: 10.1016/0005-7967(78)90022-0]
11. Salkovskis PM. Obsessional-compulsive problems: A cognitive-behavioural analysis. Behaviour Research and Therapy 1985;23(5):571-583. [doi: 10.1016/0005-7967(85)90105-6]
12. Obsessive Compulsive Cognitions Working Group. Psychometric validation of the obsessive belief questionnaire and interpretation of intrusions inventory--Part 2: Factor analyses and testing of a brief version. Behav Res Ther 2005 Nov;43(11):1527-1542. [Medline: 16299894]
13. Obsessive Compulsive Cognitions Working Group. Cognitive assessment of obsessive-compulsive disorder. Behav Res Ther 1997 Jul;35(7):667-681. [Medline: 9193129]
14. Belloch A, Del Valle G, Morillo C, Carrió C, Cabedo E. To seek advice or not to seek advice about the problem: the help-seeking dilemma for obsessive-compulsive disorder. Soc Psychiatry Psychiatr Epidemiol 2009 Apr;44(4):257-264. [doi: 10.1007/s00127-008-0423-0] [Medline: 18668187]
15. Del Valle Del Valle G, Carrió C, Belloch A. Temporary stages and motivational variables: Two complementary perspectives in the help-seeking process for mental disorders. Rev Psiquiatr Salud Ment 2018;11(3):176-183. [doi: 10.1016/j.rpsm.2017.07.005] [Medline: 29033054]
16. Wootton BM, Dear BF, Johnston L, Terides MD, Titov N. Remote treatment of obsessive-compulsive disorder: A randomized controlled trial. Journal of Obsessive-Compulsive and Related Disorders 2013 Oct;2(4):375-384. [doi: 10.1016/j.jocrd.2013.07.002]
17. Mahoney A, Mackenzie A, Williams A, Smith J, Andrews G. Internet cognitive behavioural treatment for obsessive compulsive disorder: A randomised controlled trial. Behav Res Ther 2014 Dec;63:99-106. [doi: 10.1016/j.brat.2014.09.012] [Medline: 25461784]
18. Price M, Yuen E, Goetter E, Herbert J, Forman E, Acierno R, et al. mHealth: a mechanism to deliver more accessible, more effective mental health care. Clin Psychol Psychother 2014;21(5):427-436 [FREE Full text] [doi: 10.1002/cpp.1855] [Medline: 23918764]
19. Tolin D, Diefenbach G, Gilliam C. Stepped care versus standard cognitive-behavioral therapy for obsessive-compulsive disorder: a preliminary study of efficacy and costs. Depress Anxiety 2011 Apr;28(4):314-323 [FREE Full text] [doi: 10.1002/da.20804] [Medline: 21381157]
20. Baer L, Jacobs DG, Cukor P, O'Laughlen J, Coyle JT, Magruder K. Automated telephone screening survey for depression. JAMA 1995 Jun 28;273(24):1943-1944. [Medline: 7783305]
21. Lovell K, Cox D, Haddock G, Jones C, Raines D, Garvey R, et al. Telephone administered cognitive behaviour therapy for treatment of obsessive compulsive disorder: randomised controlled non-inferiority trial. BMJ 2006 Oct 28;333(7574):883 [FREE Full text] [doi: 10.1136/bmj.38940.355602.80] [Medline: 16935946]

22. Vogel P, Launes G, Moen E, Solem S, Hansen B, Håland AT, et al. Videoconference- and cell phone-based cognitive-behavioral therapy of obsessive-compulsive disorder: a case series. J Anxiety Disord 2012 Jan;26(1):158-164. [doi: 10.1016/j.janxdis.2011.10.009] [Medline: 22119331]
23. Van Ameringen M, Turna J, Khalesi Z, Pullia K, Patterson B. There is an app for that! The current state of mobile applications (apps) for DSM-5 obsessive-compulsive disorder, posttraumatic stress disorder, anxiety and mood disorders. Depress Anxiety 2017 Dec;34(6):526-539. [doi: 10.1002/da.22657] [Medline: 28569409]
24. Roncero M, Belloch A, Doron G. A novel approach to challenging OCD related beliefs using a mobile-app: An exploratory study. J Behav Ther Exp Psychiatry 2018 Jun;59:157-160. [doi: 10.1016/j.jbtep.2018.01.008] [Medline: 29425951]
25. Doron G, Derby D, Szepsenwol O, Nahaloni E, Moulding R. Relationship Obsessive-Compulsive Disorder: Interference, Symptoms, and Maladaptive Beliefs. Front Psychiatry 2016;7:58 [FREE Full text] [doi: 10.3389/fpsyt.2016.00058] [Medline: 27148087]
26. Foa E, Huppert J, Leiberg S, Langner R, Kichic R, Hajcak G, et al. The Obsessive-Compulsive Inventory: development and validation of a short version. Psychol Assess 2002 Dec;14(4):485-496. [Medline: 12501574]
27. Belloch A, Roncero M, García-Soriano G, Carrió C, Cabedo E, Fernández-Álvarez H. The Spanish version of the Obsessive-Compulsive Inventory-Revised (OCI-R): Reliability, validity, diagnostic accuracy, and sensitivity to treatment effects in clinical samples. Journal of Obsessive-Compulsive and Related Disorders 2013 Jul;2(3):249-256. [doi: 10.1016/j.jocrd.2013.05.001]
28. Doron G, Szepsenwol O, Karp E, Gal N. Obsessing about intimate-relationships: testing the double relationship-vulnerability hypothesis. J Behav Ther Exp Psychiatry 2013 Dec;44(4):433-440. [doi: 10.1016/j.jbtep.2013.05.003] [Medline: 23792752]
29. Ehntholt KA, Salkovskis PM, Rimes KA. Obsessive–compulsive disorder, anxiety disorders, and self-esteem: an exploratory study. Behaviour Research and Therapy 1999 Aug;37(8):771-781. [doi: 10.1016/S0005-7967(98)00177-6]
30. Abramowitz JS, Fabricant LE, Taylor S, Deacon BJ, McKay D, Storch EA. The relevance of analogue studies for understanding obsessions and compulsions. Clinical Psychology Review 2014 Apr;34(3):206-217. [doi: 10.1016/J.CPR.2014.01.004]
31. Muris P, Merckelbach H, Clavan M. Abnormal and normal compulsions. Behaviour Research and Therapy 1997 Mar;35(3):249-252. [doi: 10.1016/S0005-7967(96)00114-3]
32. Haslam N, Williams BJ, Kyrios M, McKay D, Taylor S. Subtyping obsessive-compulsive disorder: A taxometric analysis. Behavior Therapy 2005 Feb;36(4):381-391. [doi: 10.1016/S0005-7894(05)80120-0]
33. Qualtrics. URL: https://www.qualtrics.com/ [accessed 2019-01-21] [WebCite Cache ID 75bHWCo3o]
34. Research Randomizer. URL: https://www.randomizer.org/ [accessed 2019-01-03] [WebCite Cache ID 759H0G2AF]
35. Fergus TA, Carmin CN. The Validity and Specificity of the Short-Form of the Obsessive Beliefs Questionnaire (OBQ). J Psychopathol Behav Assess 2013 Nov 8;36(2):318-328. [doi: 10.1007/s10862-013-9398-6]
36. Moulding R, Anglim J, Nedeljkovic M, Doron G, Kyrios M, Ayalon A. The Obsessive Beliefs Questionnaire (OBQ): examination in nonclinical samples and development of a short version. Assessment 2011 Sep;18(3):357-374. [doi: 10.1177/1073191110376490] [Medline: 20634421]
37. Doron G, Derby DS, Szepsenwol O, Talmor D. Tainted love: Exploring relationship-centered obsessive compulsive symptoms in two non-clinical cohorts. Journal of Obsessive-Compulsive and Related Disorders 2012 Jan;1(1):16-24. [doi: 10.1016/j.jocrd.2011.11.002]
38. Doron G, Derby DS, Szepsenwol O, Talmor D. Flaws and all: Exploring partner-focused obsessive-compulsive symptoms. Journal of Obsessive-Compulsive and Related Disorders 2012 Oct;1(4):234-243. [doi: 10.1016/J.JOCRD.2012.05.004]
39. Antony MM, Bieling PJ, Cox BJ, Enns MW, Swinson RP. Psychometric properties of the 42-item and 21-item versions of the Depression Anxiety Stress Scales in clinical groups and a community sample. Psychological Assessment 1998;10(2):176-181. [doi: 10.1037//1040-3590.10.2.176]
40. Clara I, Cox B, Enns M. Confirmatory factor analysis of the Depression - Anxiety - Stress Scales in depressed and anxious patients. J Psychopathol Behav Assess 2001;23(1):61-67. [doi: 10.1023/1011095624717]
41. Henry JD, Crawford JR. The short-form version of the Depression Anxiety Stress Scales (DASS-21): construct validity and normative data in a large non-clinical sample. Br J Clin Psychol 2005 Jun;44(Pt 2):227-239. [doi: 10.1348/014466505X29657] [Medline: 16004657]
42. Lovibond P, Lovibond S. The structure of negative emotional states: Comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. Behaviour Research and Therapy 1995 Mar;33(3):335-343. [doi: 10.1016/0005-7967(94)00075-U]
43. Robins RW, Hendin HM, Trzesniewski KH. Measuring Global Self-Esteem: Construct Validation of a Single-Item Measure and the Rosenberg Self-Esteem Scale. Pers Soc Psychol Bull 2016 Jul 02;27(2):151-161. [doi: 10.1177/0146167201272002]
44. Gupta S. Intention-to-treat concept: A review. Perspect Clin Res 2011 Jul;2(3):109-112 [FREE Full text] [doi: 10.4103/2229-3485.83221] [Medline: 21897887]
45. Overall JE, Tonidandel S, Starbuck RR. Last-observation-carried-forward (LOCF) and tests for difference in mean rates of change in controlled repeated measurements designs with dropouts. Social Science Research 2009 Jun;38(2):492-503. [doi: 10.1016/j.ssresearch.2009.01.004]

46. Wilson D. Effect Size Determination Program (Version 2.0, Excel Macro Application). University of Maryland, College Park 2001.
47. Donker T, Petrie K, Proudfoot J, Clarke J, Birch M, Christensen H. Smartphones for smarter delivery of mental health programs: a systematic review. J Med Internet Res 2013 Nov 15;15(11):e247 [FREE Full text] [doi: 10.2196/jmir.2791] [Medline: 24240579]
48. Lau HM, Smit JH, Fleming TM, Riper H. Serious Games for Mental Health: Are They Accessible, Feasible, and Effective? A Systematic Review and Meta-analysis. Front Psychiatry 2016;7:209 [FREE Full text] [doi: 10.3389/fpsyt.2016.00209] [Medline: 28149281]
49. Rathbone AL, Prescott J. The Use of Mobile Apps and SMS Messaging as Physical and Mental Health Interventions: Systematic Review. J Med Internet Res 2017 Dec 24;19(8):e295 [FREE Full text] [doi: 10.2196/jmir.7740] [Medline: 28838887]
50. Doron G, Szepsenwol O. Partner-focused obsessions and self-esteem: An experimental investigation. J Behav Ther Exp Psychiatry 2015 Dec;49(Pt B):173-179. [doi: 10.1016/j.jbtep.2015.05.007] [Medline: 26092536]
51. Doron G, Derby D. Assessment treatment of relationship-related OCD symptoms (ROCD): A modular approach. In: Abramowitz JS, McKay DE, Storch A, editors. The Wiley handbook of obsessive-compulsive disorder across the lifespan. Hoboken, NJ: John Wiley & Sons; 2017.
52. Harrison V, Proudfoot J, Wee P, Parker G, Pavlovic D, Manicavasagar V. Mobile mental health: review of the emerging field and proof of concept study. J Ment Health 2011 Dec;20(6):509-524. [doi: 10.3109/09638237.2011.608746] [Medline: 21988230]
53. Watts S, Mackenzie A, Thomas C, Griskaitis A, Mewton L, Williams A, et al. CBT for depression: a pilot RCT comparing mobile phone vs. computer. BMC Psychiatry 2013 Feb 07;13:49 [FREE Full text] [doi: 10.1186/1471-244X-13-49] [Medline: 23391304]
54. Pascual-Vera B, Roncero M, Doron G, Belloch A. Assisting relapse prevention in OCD using a novel mobile app-based intervention: A case report. Bull Menninger Clin 2018;82(4):390-406. [doi: 10.1521/bumc.2018.82.4.390] [Medline: 30589573]
55. Arean PA, Hallgren KA, Jordan JT, Gazzaley A, Atkins DC, Heagerty PJ, et al. The Use and Effectiveness of Mobile Apps for Depression: Results From a Fully Remote Clinical Trial. J Med Internet Res 2016 Dec 20;18(12):e330 [FREE Full text] [doi: 10.2196/jmir.6482] [Medline: 27998876]
56. Roepke AM, Jaffee SR, Riffle OM, McGonigal J, Broome R, Maxwell B. Randomized Controlled Trial of SuperBetter, a Smartphone-Based/Internet-Based Self-Help Tool to Reduce Depressive Symptoms. Games Health J 2015 Jun;4(3):235-246. [doi: 10.1089/g4h.2014.0046] [Medline: 26182069]
57. Boisseau C, Schwartzman C, Lawton J, Mancebo M. App-guided exposure and response prevention for obsessive compulsive disorder: an open pilot trial. Cogn Behav Ther 2017 Nov;46(6):447-458. [doi: 10.1080/16506073.2017.1321683] [Medline: 28565937]
58. Olbrich H, Stengler K, Olbrich S. Smartphone based Geo-Feedback in obsessive compulsive disorder as facilitatory intervention: A case report. Journal of Obsessive-Compulsive and Related Disorders 2016 Jan;8:75-78. [doi: 10.1016/j.jocrd.2016.01.001]
59. Bakker D, Kazantzis N, Rickwood D, Rickard N. Mental Health Smartphone Apps: Review and Evidence-Based Recommendations for Future Developments. JMIR Ment Health 2016 Mar 01;3(1):e7 [FREE Full text] [doi: 10.2196/mental.4984] [Medline: 26932350]

Abbreviations

CBT: cognitive behavioral therapy
dApp: delayed-use App group
DASS: Depression, Anxiety, Stress Scale - Short version
DASS-D: Depression, Anxiety, Stress Scale - Depression subscale
iApp: immediate-use App group
OBQ-20: Obsessive Beliefs Questionnaire - Short form
OCD: obsessive-compulsive disorder
OCI-R: Obsessive-Compulsive Inventory - Reduced
PROCSI: Partner-Related Obsessive-Compulsive Symptoms Inventory
PROCSI-Si: Partner-Related Obsessive-Compulsive Symptoms Inventory - 6-item version
ROCD: Relationship OCD
ROCI: The Relationship Obsessive-Compulsive Inventory
ROCI-S: The Relationship Obsessive-Compulsive Inventory - Short version
SISE: Single-Item Self-Esteem Scale

Edited by G Eysenbach; submitted 29.06.18; peer-reviewed by M Ghisi, D Bakker; comments to author 07.10.18; revised version received 24.10.18; accepted 09.12.18; published 13.02.19.

Please cite as:
Roncero M, Belloch A, Doron G
Can Brief, Daily Training Using a Mobile App Help Change Maladaptive Beliefs? Crossover Randomized Controlled Trial
JMIR Mhealth Uhealth 2019;7(2):e11443
URL: http://mhealth.jmir.org/2019/2/e11443/
doi:10.2196/11443
PMID:30758294

Original Paper

Log2Lose: Development and Lessons Learned From a Mobile Technology Weight Loss Intervention

Ryan Shaw[1*], PhD, RN; Erica Levine[2], MPH; Martin Streicher[2], MS; Elizabeth Strawbridge[1], MS, RD; Jennifer Gierisch[2,3], PhD; Jane Pendergast[2], PhD; Sarah Hale[2], PhD; Shelby Reed[2], PhD; Megan McVay[4], PhD; Denise Simmons[2], BS; William Yancy[2,3], MD; Gary Bennett[5], PhD; Corrine Voils[6,7*], PhD

[1]School of Nursing, Duke University, Durham, NC, United States
[2]School of Medicine, Duke University, Durham, NC, United States
[3]Durham Veterans Affairs Medical Center, Durham, NC, United States
[4]College of Health and Human Performance, University of Florida, Gainesville, FL, United States
[5]Duke Global Health Institute, Duke University, Durham, NC, United States
[6]William S Middleton Memorial Veterans Hospital, Madison, WI, United States
[7]University of Wisconsin, Madison, WI, United States
[*]these authors contributed equally

Corresponding Author:
Ryan Shaw, PhD, RN
School of Nursing
Duke University
307 Trent Drive, DUMC 3322
Durham, NC, 27710
United States
Phone: 1 9196849434
Email: ryan.shaw@duke.edu

Abstract

Background: Providing financial incentives has gained popularity as a strategy to promote weight loss, but questions remain about how best to utilize them. A promising mobile health strategy provides users with near-real-time financial incentives based on both the process of weight loss (behavioral modification) and actual weight loss. To maximize the impact of this strategy, a methodology is needed to close the gap between the desired behavior and the financial incentive. Leveraging mobile health tools—such as mobile phone apps, cellular body weight scales that transmit data to physicians and researchers, and text messaging for instructions and encouragement—has the potential to close this gap.

Objective: This study aimed to describe the development of an innovative technology-based solution and lessons learned from a feasibility trial—Log2Lose—that encouraged individuals to lose weight by providing near-real-time financial incentives for weight loss and/or dietary self-monitoring.

Methods: We recruited participants (N=96) with a body mass index greater than or equal to 30 kg/m^2 for a 24-week weight loss trial. Participants received a behavioral intervention of biweekly, in-person group sessions and were instructed to log a minimum number of daily calories in MyFitnessPal and to step on the BodyTrace cellular scale at least twice per week. In a 2×2 design, participants were randomized into 4 groups to receive financial incentives for the following: (group 1) weekly weight loss and dietary self-monitoring, (group 2) dietary self-monitoring only, (group 3) weekly weight loss only, or (group 4) no financial incentives. Diet and weight data from the devices were obtained through application programming interfaces. Each week, we applied algorithms to participants' data to determine whether they qualified for a monetary incentive (groups 1-3). A text message notified these participants of whether they met weight loss and/or self-monitoring requirements to earn an incentive and the amount they earned or would have earned. The money was uploaded to a debit card.

Results: Our custom-engineered software platform analyzed data from multiple sources, collated and processed the data to send appropriate text messages automatically, and informed study staff of the appropriate incentives. We present lessons learned from the development of the software system and challenges encountered with technology, data transmission, and participants (eg, lost connections or delayed communication).

Conclusions: With consistent and constant validation checks and a robust beta test run, the process of analyzing data and determining eligibility for weekly incentives can be mostly automated. We were able to accomplish this project within an academic health system, which required significant security and privacy safeguards. Our success demonstrates how this methodology of automated feedback loops can provide health interventions via mobile technology.

Trial Registration: ClinicalTrials.gov NCT02691260; https://clinicaltrials.gov/ct2/show/NCT02691260

(JMIR Mhealth Uhealth 2019;7(2):e11972) doi:10.2196/11972

KEYWORDS

weight loss; diet; cell phone; mobile phone

Introduction

Background

More than one-third of the US population has obesity [1]. There are significant physical, psychological, and financial ramifications to this epidemic, including increased risk of type 2 diabetes mellitus, hypertension, and depression and medical cost burden [2]. Weight loss of at least 5% improves functional status, decreases blood pressure, and can reduce the incidence of diabetes by more than 50% over 3 years [3]. Although many weight loss interventions yield an average weight loss of at least 5%, many individuals still struggle to adhere to such interventions, reducing the proportion of individuals who achieve clinically significant weight loss.

Several efficacious interventions have encouraged weight loss by offering incentives. However, a key issue that remains is whether individuals should be offered financial incentives for weight loss, weight loss behaviors, or both. Few studies have examined these questions in the context of weight loss [4]. The few studies that have evaluated incentives for weight loss provided incentives at the end of the study instead of providing incentives for weight loss achieved throughout the program in near real-time when behaviors are being performed [5-9]. This has led to a considerable gap between health behaviors performed and receipt of incentives. Furthermore, previous studies that incented weight loss or dietary self-monitoring did so during in-person sessions that required individuals to turn in self-monitoring records and weigh in [4]. This is logistically feasible but ties the reward to attendance of a specific meeting [10].

Objectives

On the basis of our hypothesis that rewards need to be proximal to the reward-winning behaviors and that they should not be tied to in-person attendance, our team developed an innovative technology solution—Log2Lose. This mobile health tool encourages individuals to lose weight by providing near-real-time financial rewards for weight loss and/or dietary self-monitoring. This study describes how we approached developing Log2Lose and refined the technology over time based on lessons learned. This information can assist others who are creating similar scalable interventions.

Methods

Study Design and Overview

This study was approved by the Duke University Medical Center Institutional Review Board (ClinicalTrials.gov: NCT02691260). To evaluate feasibility of study processes [11], we recruited participants (N=96) through community advertisements with a body mass index greater than or equal to 30 kg/m^2 for a 24-week weight loss feasibility trial in North Carolina, United States. We sequentially enrolled 3 cohorts (n=34, n=31, and n=31), each consisting of 4 randomized groups: group 1 received incentives for dietary self-monitoring and weight loss, group 2 received incentives for dietary self-monitoring only, group 3 received incentives for interim weight loss, and group 4 received weekly text messages but no incentives. Each group had a mean of 8.75 participants across the 3 cohorts.

Participants received a group-based behavioral intervention focusing on either carbohydrate (cohorts 1 and 3) or calorie and fat (cohort 2) restriction. Each group met with a registered dietitian for 1.5 hours every 2 weeks over 24 weeks. All participants were instructed to record at least 1000 calories (female) or 1200 calories (male) daily via the MyFitnessPal mobile phone app 5 or more days per week, including 1 weekend day, and to weigh themselves at least twice per week on a cellular digital scale (BodyTrace) that transmitted weights to researchers. The rationale for the caloric threshold is as follows. The goal of self-monitoring is to illuminate the discrepancy between goals and actual behavior so that people can make changes to align their goals and behaviors [12]. By tracking at least 1000 or 1200 calories daily, participants will have enough information about their dietary patterns to inform dietary modifications. This threshold is slightly lower than the caloric goal for most people seeking weight loss and allows for small lapses in tracking. The rationale for weighing at least twice per week is that there needs to be enough weight data to calculate weight loss.

Participants in the 3 incentive conditions did not know how much money they could win in any given week. The maximum they could receive over the course of the study was US $300, with a range of US $0 (or $2 in cohort 3) to US $28 weekly. At the end of each week, we sent a text message notifying each participant whether she/he had earned an incentive and, if so, how much. The money was uploaded to a MasterCard debit card via ClinCard, a payment system for clinical trials [6]. Participants in the no-incentive condition also received a weekly text message, which encouraged them to self-weigh and log

their dietary intake in MyFitnessPal. During cohorts 2 and 3, we sent an additional text message twice a week to participants with words of encouragement or skill-building tips for weight loss.

Commercial Software and Hardware Utilized

It is beneficial to utilize commercially available hardware and software to maximize the potential for scaling weight loss programs as well as positively affect adoption of interventions. Using existing solutions allows researchers to capitalize on work already done by designers, front-end software engineers, and hardware manufacturers instead of having to design and develop their own systems. To that end, we utilized products already on the market to deliver the Log2Lose intervention.

To read and analyze data, we used Prompt (Figure 1) [13], an integrated digital health platform that can test and disseminate evidence-based health solutions and improve participant engagement in clinical and observational studies. Prompt is designed to send and receive data via multiple modalities (text message, interactive voice response, email, or application programming interface [API]). It operates as an interactive, automated tool to help users self-monitor behaviors; receive automated, tailored feedback based on performance; and/or communicate with interventionists. Prompt extends the ability of clinicians and researchers to design and test innovative behavioral interventions and to improve participant engagement in clinical studies. It utilizes tools on Twilio, Heroku, and Amazon S3 and a range of third-party APIs to send, receive, and store data. For this study, Prompt aggregated data via APIs from a dietary app and cellular scale. In addition, it analyzed data per Log2Lose specifications to perform validation checks and to compose Log2Lose text messages.

To retrieve near-real-time diet data, we asked all participants to download the MyFitnessPal and Fitbit dietary tracking apps on their mobile phones. We chose to use the MyFitnessPal app for participant tracking because it provided the appropriate carbohydrate counting interface needed for this study in addition to calorie tracking, whereas Fitbit did not. During the study period, MyFitnessPal did not have an open API (ie, it did not offer an open connection to other software apps). However, MyFitnessPal could share data with the Fitbit app. As a workaround to obtain MyFitnessPal data, we accessed the needed data from the Fitbit API. Study staff helped each participant link his or her MyFitnessPal and Fitbit accounts to each other (linking is a feature on each mobile phone app). Participants added their personal profiles to Prompt, and study staff connected their Prompt profiles to their Fitbit accounts using Open Authorization, a secure authorization protocol.

To retrieve body weight measurements, we provided each participant with a BodyTrace digital scale. These scales transmit readings via cellular service to researchers. We then connected the scale ID to participants' Prompt profiles. The BodyTrace scale, which connects to a cloud-based database, was chosen for use in this study because (1) it does not require a complicated setup (ie, does not need to be manually connected to a Wi-Fi network); (2) broadband access is not yet ubiquitous across semirural and rural parts of North Carolina, making it possible to use the scale throughout the state; (3) BodyTrace has an open API and provides research staff access to a back-end database, which enables direct collection of data from BodyTrace servers; and (4) it allowed staff to investigate any issues using the BodyTrace back-end database. Although participants were told to weigh themselves at least twice per week (preferably at the beginning and end of each week to maximize weight differences) to permit calculation of weekly weight loss, they were encouraged to weigh daily because individuals who weigh daily tend to lose more weight than those who weigh less frequently [14]. Prompt collected weights daily via the BodyTrace API and time stamped and stored them in a database. Of note, participants could game the system on their home BodyTrace scale by having someone else weigh in for them or by not putting all their weight on the scale. Therefore, our software algorithm detected weights in a single day with a difference greater than 10% and removed those weights to provide some assurance that the measurements came from the same person.

Study staff received access to a multi-factor authentication and password-protected user interface (UI) specifically designed for Log2Lose. This UI allowed staff to view participant information, a list of participants who met the study criteria, a list of participants who earned incentives each week, and the incentive amounts. Although the system automatically analyzed the data and determined who qualified for incentives, because of institutional protocols, study staff had to request money transfers, and a financial administrator had to approve them. This created a delay of approximately 1 day between notification of incentive via text message and receipt of the money on the debit card. Research staff instructed participants on how to use MyFitnessPal and the BodyTrace scale. Participants received tip sheets to help them resolve common error messages that might appear on the digital scale. Research staff were trained to troubleshoot the devices.

Development of Algorithms to Capture Data

The study design required customized algorithms to collect and analyze data from disparate sources. More than 100 individual algorithms were developed and run automatically, which addressed 2 streams of data (diet app and scale) used differently across 4 randomized groups and differences in payments every week for 24 weeks. Incentives were calculated each week based on the following criteria: (1) whether the participant was sufficiently active based on his or her transmission of data that week; (2) the study week, which determined the incentive amount; (3) the group to which the participant belonged; and (4) whether the participant met the applicable criterion (Table 1).

As 1 validation check, at the end of each week, Prompt reviewed all active participants who were due for an incentive calculation.

Figure 1. The custom-engineered software platform known as Prompt.

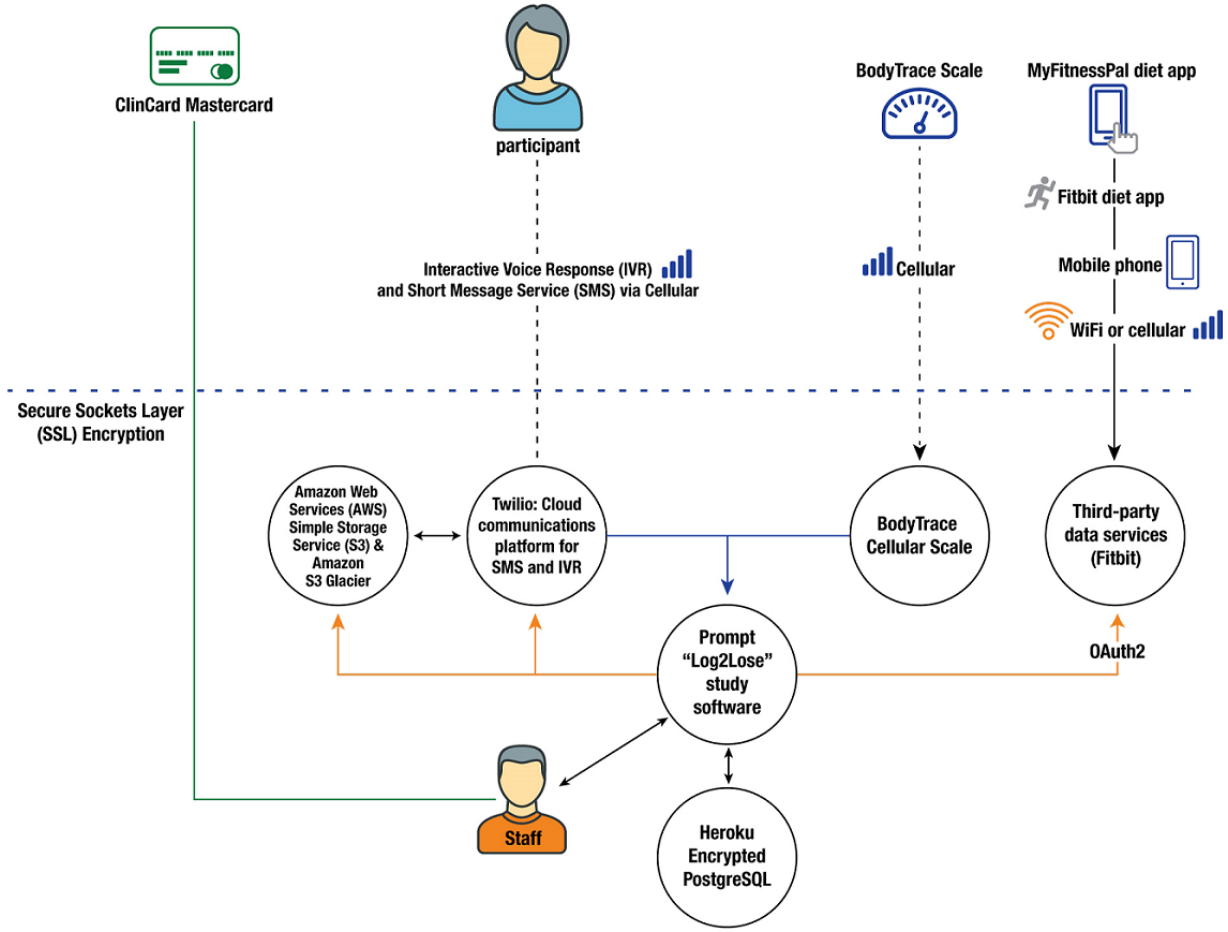


Table 1. Incentive requirements.

Data source	Requirement for incentive
MyFitnessPal	Log at least 1200 calories for men and 1000 calories for women 5 or more days per week, including 1 weekend day
BodyTrace	Any amount of weight loss, which is calculated as the difference between the first and last weight on the previous 7 days
Prompt	Participant status=active

For each participant, Prompt read the incentive schedule, the group assignment for that participant, and the diet and/or weight loss results for that week. Prompt then displayed on the UI whether each participant who was due for an incentive qualified for one and, if so, how much. As a second validation check, study staff reviewed the UI at the end of each week to ensure participants qualified or did not qualify for an incentive. They then edited, approved, or did not approve an incentive. Within 24 hours (longer over holidays), the research assistant then scheduled payment of the incentive.

The weekly messages—written and scheduled before the study launched—corresponded with all possible scenarios a participant could face each week. For example, in groups 1-3, participants who met the criteria for an incentive in a particular week received the following message: "Great job on [logging your food and/or losing weight] this week! $[XX] will be added to your debit card. Keep up the good work!," whereas participants who did not meet the criteria for an incentive in a particular week received the following message: "You did not [log your food or lose enough weight] this week. If you had, you would have earned $[XX]. Please log your food and lose weight for a chance to earn money." The message to participants who did not meet the criteria emphasized loss to enhance loss aversion [15]. Information in brackets was tailored by week and by the group to which participants belonged. These messages were 160 characters or less so that they appeared in a single message instead of breaking into 2 messages on mobile phones. They were loaded into Prompt and matched to the incentive requirement thresholds. We performed quality assurance testing to validate the message content and timing before the study began.

Participants also received a text message reminder to attend their in-person group session via Prompt. Study staff also sent a message to participants to let them know if there was a cancellation or class rescheduling because of weather. All participants in each condition received the same reminders.

Results

Technical Issues Encountered by the Research Team

Log2Lose successfully and consistently automated analysis of incoming data, applied algorithms to those data, and composed and sent text messages on a predetermined schedule. Study staff maintained weekly logs of the technical assistance they provided to participants. From enrollment to completion, technical assistance was needed in screening participants for the appropriate cell phone, aiding with the devices and apps, and running Prompt. Many of the technical issues involved not receiving data from the software apps or devices or being unable to connect to third-party software servers to collect data.

One challenge we encountered was that the cellular scale could only transmit data where a cellular signal existed. Thus, we could not know if data would transmit until after participants enrolled in the study and placed the scales in their homes. Fortunately, this only affected a handful of participants, who were able to move the scale to other locations in their homes for the scales to work.

Another challenge we encountered is that, by the time cohort 3 began, cellular service companies began shutting down 2G networks. A few of our scales were still transmitting on the 2G network, and we had to exchange them for new models. Nevertheless, similar issues could have occurred if we used Bluetooth- or Wi-Fi-enabled scales. As technology evolves, it is a challenge to predict when older technology will be phased out. This unanticipated cost delayed incentives for some participants by a couple of weeks. One lesson learned is that, if technologies are going to be used over a period of time, then they should be bought in batches rather than during study start-up. A couple of participants also stated that they did not like the design of the scale in that they had to turn it on by tapping it before stepping on. This was necessary to tare the scale.

There were also a handful of issues with message formatting. Several times, messages were sent out on different days and at different times than indicated in the protocol because of human error by a member of the research team or a participant. A full description of challenges and solutions is presented in Table 2.

Patient Acceptability of Technology Solution

Following the intervention, we asked participants to complete a survey regarding their experiences with the technologies. Table 3 describes the questions we asked regarding the diet app and scale. According to self-report, more than half of participants tracked their diet and weight loss over the 6-month study, although participation decreased over time. Participants also reported reading the incentive text messages, and, for the most part, immediately upon receipt.

When comparing follow-up survey responses with the corresponding app data, we found that self-reports generally were consistent. Among the 77 participants for whom follow-up survey responses were available, 18 (18/77, 23%) had no recorded diet app data (Table 3), 14 (14/77, 18%) overestimated the timeframe during which they tracked their diet on the app, and 15 (15/77, 19%) underestimated it. The 46 (46/77, 60%) participants who reported using the BodyTrace scale “every day” averaged 5.4 days per week (SD 1.6) according to the cellular data, the 28 (28/77, 36%) who reported “more than once a week, but not every day” averaged 3.5 days/week (SD 1.5), and the 2 (2/77, 3%) who reported “less than once a week” averaged 1 day a week (SD 1.3). These data suggest high but decreasing adherence to instructions to weigh regularly (at least twice per week) and to log food and drink in MyFitnessPal.

After the study, we interviewed a subset of participants (n=33) from each cohort to ask about their experiences. We purposefully selected individuals who had lost at least 5% of their body weight and those who did not. When asked about their experiences using the cellular scale, most participants stated satisfaction with the scale throughout their use or after a period of adjustment. A few participants said they did not like the scale and did not trust the weight readings because they varied too much day to day. Other individuals reported that the scale was difficult to use at first but that they got used to it. Another complaint was that the scale had to be kept on an even surface. If the scale was placed on an uneven surface, such as carpet, weights were inaccurate. We instructed participants at baseline to place the scale in a place with even surfaces, yet this was still a challenge at times. Furthermore, 1 participant had to exchange a scale 2 times because of unknown issues; these were returned to the manufacturer. Regarding the self-monitoring process, participants mentioned weighing less frequently because they did not enjoy daily weighing or that they became unmotivated when not seeing significant weight loss.

We also asked participants about their experiences using MyFitnessPal. Although 1 individual chose to use paper instead (which would have eliminated them from an incentive for recording food), most participants stated that they liked recording in the app. The most common feature that participants liked was the ability to use the bar scan for commercially available products. However, that was not always possible. If participants went to a restaurant or someone else cooked, they did not always know what ingredients were in the food, making it challenging to record in the app. Complaints from participants who did not use the app were as follows: did not need to use the app because they already knew the calorie content of food, did not like the design of the app, and too many food choices in the app made scrolling difficult.

Table 2. Log2Lose software troubleshooting.

Issue type	Example of issue	Potential problem	Solution or solutions implemented
No data	No diet data; no weight data; and incentive text messages do not populate in Prompt	Participant has a previous account and could not remember the password; participant is not tracking or weighing; data are not transmitting from the device or app to server; data are not transmitting from server to Prompt; user accounts are not connected correctly; and programming bug in Prompt	Retrain participant on correct tracking methods; check connection between device/app and third-party server; check API[a] and Prompt connections; reconnect user accounts; and check and resolve bugs in Prompt
Incorrect information sent to participants	Minor formatting errors in text messages; incorrect incentive amount listed in text message; messages sent out at different days/times than indicated in the protocol; and text messages do not populate in Prompt	Bugs in algorithms and bug in Prompt	Fix/edit original protocols and implement change in Prompt code; manually override information and manually edit and send text messages; and fix bug in Prompt and test fix
Prompt user interface unavailable	Staff unable to log in to user interface	Staff entering incorrect password and bugs in Prompt	Reset passwords and fix bug, test, and implement
Connection error (device/app to the third-party server, to the API, or to Prompt)	Participant tracks in MyFitnessPal, but data are not viewable in Fitbit server, and participant says that she/he weighs, but weight does not show up in server	Prompt reported an error retrieving data from cellular scale and/or Fitbit APIs; participant not tracking; errors appear on the scale; and bugs in Prompt	Check connection between device/app and server; check API-Prompt connections; and resolve bugs in Prompt and test fixes and implement fixes

[a]API: application programming interface.

Table 3. Postintervention survey questions.

Question (n=77)	Count, n (%)	Count, n (%)	Count, n (%)	Count, n (%)	Count, n (%)	Count, n (%)
For how long did you track what you ate and drank during this study?	1 month or less, 6 (7%)	Up to 2 months, 9 (11%)	Up to 3 months, 4 (5%)	Up to 4 months, 9 (11%)	Up to 5 months, 6 (7%)	All 6 months, 43 (55%)
I tracked it (diet) with the MyFitnessPal app on my cell phone	Never, 2 (2%)	Rarely, 5 (6%)	Sometimes, 9 (11%)	Often, 11 (14%)	Always, 50 (64%)	__[a]
I tracked it (diet) with a different app on my cell phone	Never, 70 (90%)	Rarely, 1 (1%)	Sometimes, 3 (3%)	Often, 2 (2%)	Always, 1 (1%)	—
How often did you use the BodyTrace scale to weigh yourself during the study?	Every day, 46 (59%)	More than once a week but not every day, 28 (36%)	Less than once a week, 3 (3%)	—	—	—

[a]Indicates there were no answers to select from.

Discussion

Principal Findings

Innovative methods are needed to assist individuals in losing weight and maintaining weight loss. As weight loss is a daily undertaking, interventions that reward individuals closer in time to when weight loss behaviors occur are more likely to be effective than those that delay rewards [16]. Incentivizing individuals based on their weight loss behaviors in their daily environments is possible via use of mobile phone technology. This study demonstrates that it is feasible to engage in regular incenting for diet and weight loss in individuals' daily environments. Although our team incentivized participants with money, many other types of incentives could be transmitted through software systems, including, but not limited to, coupons, discounts, or home delivery of products. In addition, other types of streaming data could be collected from wearable sensors, phone-tethered devices, and sensors placed in the environment.

Our system was sophisticated in its ability to analyze data from multiple sources, to collate and process data to put into appropriate text messages, and to inform study staff of the appropriate incentives. In this development phase of our research, staff spent time monitoring incentives weekly and approving payments. Through our experiences and lessons

learned, we anticipate that future versions of the software will be more automated. If the software can apply incentives securely, appropriately, and automatically, then it can reduce the potential for human error as well as costs.

We approached this feasibility study from a developmental standpoint, anticipating frequent corrections to the algorithms and technology platforms. Despite the need for validation checks and test runs, we were able to develop this system in only 4 months. We quickly recognized the importance of having processes in place to help participants resolve bugs and other issues. One of the most important lessons learned is that although we were not able to process incentives immediately because of institutional requirements for processing payments, our system enables this to be achieved in environments with less restrictions. Thus, we have created a methodology that will allow for widespread scaling and dissemination.

Limitations

First, because this intervention required in-person group meetings, participation was limited to individuals who live close to the medical center. Although research indicates interventions with in-person contact are more efficacious for weight loss [17], requiring continuous in-person contact may affect dissemination potential. Second, because this intervention required a mobile phone with a data plan, access was limited to individuals with a mobile phone and financial means. More than 77% of Americans own mobile phones, and ownership is greater among racial minorities [18]. Third, this study required participants to use diet apps that only work on certain brands of mobile phones, eliminating individuals with less common models of phones (ie, Jitterbug). Fourth, participants may not have recorded dietary intake accurately. As we are focusing on the process rather than the content of recording dietary intake, we believe that incenting this behavior will result in improved outcomes [4]. Finally, because the maximum value of incentives was the same in all 3 incentive conditions, participants in the combined incentive group received half the reward for each singular behavior than the other 2 groups that only rewarded one behavior.

Conclusions

The study design allowed us to evaluate the feasibility and acceptability of incentivizing dietary self-monitoring on a mobile phone app. These results provide the foundation for a comprehensive, randomized controlled trial to evaluate the impact of incentivizing dietary self-monitoring and weight loss. Due to the proliferation of mobile phones [18] and the increased availability of wireless scales at a more affordable price, this approach looks promising to aid individuals trying to lose weight or maintain weight loss.

Acknowledgments

This study was funded by a grant awarded to CV and RS by the NIH National Heart, Lung, and Blood Institute (NHLBI; 1R34HL125669). This manuscript was also made possible by a Research Career Scientist award (RCS 14-443) to CV from the Department of Veterans Affairs Health Services Research & Development service, a career development award from the NHLBI to MM (K23 HL127334), and the Center of Innovation for Health Services Research in Primary Care (CIN 13-410) award to WY at the Durham VA Health Care System.

Conflicts of Interest

None declared.

Multimedia Appendix 1

CONSORT - EHEALTH checklist (V 1.6.1).

[PDF File (Adobe PDF File), 2MB - mhealth_v7i2e11972_fig.pdf]

References

1. Ogden CL, Carroll MD, Kit BK, Flegal KM. Prevalence of childhood and adult obesity in the United States, 2011-2012. J Am Med Assoc 2014 Feb 26;311(8):806-814. [doi: 10.1001/jama.2014.732] [Medline: 24570244]
2. World Health Organization. 2009. Global Health Risks-Mortality and burden of disease attributable to selected major risks URL: https://www.who.int/healthinfo/global_burden_disease/GlobalHealthRisks_report_full.pdf [accessed 2019-01-22] [WebCite Cache ID 75cmoXMbC]
3. Knowler W, Barrett-Connor E, Fowler S, Hamman RF, Lachin JM, Walker EA, Diabetes Prevention Program Research Group. Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. N Engl J Med 2002 Feb 07;346(6):393-403 [FREE Full text] [doi: 10.1056/NEJMoa012512] [Medline: 11832527]
4. Burns RJ, Donovan AS, Ackermann RT, Finch EA, Rothman AJ, Jeffery RW. A theoretically grounded systematic review of material incentives for weight loss: implications for interventions. Ann Behav Med 2012 Dec;44(3):375-388. [doi: 10.1007/s12160-012-9403-4] [Medline: 22907712]
5. Leahey TM, Subak LL, Fava J, Schembri M, Thomas G, Xu X, et al. Benefits of adding small financial incentives or optional group meetings to a web-based statewide obesity initiative. Obesity (Silver Spring) 2015 Jan;23(1):70-76 [FREE Full text] [doi: 10.1002/oby.20937] [Medline: 25384463]

6. Volpp KG, John LK, Troxel AB, Norton L, Fassbender J, Loewenstein G. Financial incentive-based approaches for weight loss: a randomized trial. J Am Med Assoc 2008 Dec 10;300(22):2631-2637 [FREE Full text] [doi: 10.1001/jama.2008.804] [Medline: 19066383]
7. Kullgren JT, Troxel AB, Loewenstein G, Asch DA, Norton LA, Wesby L, et al. Individual- versus group-based financial incentives for weight loss: a randomized, controlled trial. Ann Intern Med 2013 Apr 2;158(7):505-514 [FREE Full text] [doi: 10.7326/0003-4819-158-7-201304020-00002] [Medline: 23546562]
8. Almeida FA, You W, Harden SM, Blackman KCA, Davy BM, Glasgow RE, et al. Effectiveness of a worksite-based weight loss randomized controlled trial: the worksite study. Obesity (Silver Spring) 2015 Apr;23(4):737-745 [FREE Full text] [doi: 10.1002/oby.20899] [Medline: 25678325]
9. Shaw PA, Yancy Jr JW, Wesby L, Ulrich V, Troxel A, Huffman D, et al. The design and conduct of Keep It Off: an online randomized trial of financial incentives for weight-loss maintenance. Clinical trials Feb 2017;14(1):29-36. [doi: 10.1177/1740774516669679] [Medline: 27646508]
10. Yawn B, Madison S, Bertram S. Automated patient and medication payment method for clinical trials. Open Access J Clin Trials 2013;5(1):23-31. [doi: 10.2147/OAJCT.S38489]
11. Voils CI, Levine E, Gierisch JM, Pendergast J, Hale SL, McVay MA, et al. Study protocol for Log2Lose: a feasibility randomized controlled trial to evaluate financial incentives for dietary self-monitoring and interim weight loss in adults with obesity. Contemp Clin Trials 2018 Feb;65:116-122. [doi: 10.1016/j.cct.2017.12.007] [Medline: 29289702]
12. Leventhal H, Diefenbach M, Leventhal E. Illness cognition: using common sense to understand treatment adherence and affect cognition interactions. Cognitive therapy and research 1992 Apr;16(2):143-163. [doi: 10.1007/BF01173486]
13. Steinberg DM, Christy J, Batch BC, Askew S, Moore RH, Parker P, et al. Preventing weight gain improves sleep quality among black women: results from a RCT. Ann Behav Med 2017 Aug;51(4):555-566 [FREE Full text] [doi: 10.1007/s12160-017-9879-z] [Medline: 28213632]
14. Novello S, Cash S, Roberts S, Masters W. Self-weighing increases weight loss in free-living adults: a double-blind randomized field trial among 200,000 health app users. FASEB J 2015;29(1_supplement):594-599 [FREE Full text]
15. Kahneman D. Prospect theory: an analysis of decisions under risk. Econometrica 1979;47(2):263-291. [doi: 10.1142/9789814417358_0006]
16. Colwell N, Egan S, Sheppard A, Walsh K. 43 Real time longitudinal monitoring of patient diet, exercise and weight loss has substantial positive effects on metabolic syndrome parameters. BMJ Heart 2015 Oct 07;101(Suppl 5):A23.3-A2324. [doi: 10.1136/heartjnl-2015-308621.43]
17. Jensen MD, Ryan DH, Apovian CM, Ard JD, Comuzzie AG, Donato KA, American College of Cardiology/American Heart Association Task Force on Practice Guidelines, Obesity Society. 2013 AHA/ACC/TOS guideline for the management of overweight and obesity in adults: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines and The Obesity Society. J Am Coll Cardiol 2014 Jul 1;63(25 Pt B):2985-3023 [FREE Full text] [doi: 10.1016/j.jacc.2013.11.004] [Medline: 24239920]
18. Pew Research Center. 2018. Mobile Fact Sheet URL: http://www.pewinternet.org/fact-sheet/mobile/ [accessed 2019-01-22] [WebCite Cache ID 75coN6U2y]

Abbreviations

API: application programming interface
UI: user interface

Edited by G Eysenbach; submitted 19.08.18; peer-reviewed by X Xu, J Lei, J Ceasar; comments to author 14.10.18; revised version received 06.12.18; accepted 09.12.18; published 13.02.19.

Please cite as:
Shaw R, Levine E, Streicher M, Strawbridge E, Gierisch J, Pendergast J, Hale S, Reed S, McVay M, Simmons D, Yancy W, Bennett G, Voils C
Log2Lose: Development and Lessons Learned From a Mobile Technology Weight Loss Intervention
JMIR Mhealth Uhealth 2019;7(2):e11972
URL: http://mhealth.jmir.org/2019/2/e11972/
doi:10.2196/11972
PMID:30758297

Original Paper

Evaluating Motivational Interviewing and Habit Formation to Enhance the Effect of Activity Trackers on Healthy Adults' Activity Levels: Randomized Intervention

Laura D Ellingson[1], PhD; Jeni E Lansing[1], MS; Kathryn J DeShaw[1], MS; Karissa L Peyer[2], PhD; Yang Bai[3], PhD; Maria Perez[1], MS; L Alison Phillips[4], PhD; Gregory J Welk[1], PhD

[1]Department of Kinesiology, Iowa State University, Ames, IA, United States

[2]Department of Health and Human Performance, University of Tennessee-Chatanooga, Chattanooga, TN, United States

[3]Deparment of Rehabilitation and Movement Science, University of Vermont, Burlington, VT, United States

[4]Department of Psychology, Iowa State University, Ames, IA, United States

Corresponding Author:
Laura D Ellingson, PhD
Department of Kinesiology
Iowa State University
534 Wallace Road
Ames, IA, 50011
United States
Phone: 1 5154990663
Fax: 1 5152948740
Email: ellingl@iastate.edu

Abstract

Background: While widely used and endorsed, there is limited evidence supporting the benefits of activity trackers for increasing physical activity; these devices may be more effective when combined with additional strategies that promote sustained behavior change like motivational interviewing (MI) and habit development.

Objective: This study aims to determine the utility of wearable activity trackers alone or in combination with these behavior change strategies for promoting improvements in active and sedentary behaviors.

Methods: A sample of 91 adults (48/91 female, 53%) was randomized to receive a Fitbit Charge alone or in combination with MI and habit education for 12 weeks. Active and sedentary behaviors were assessed pre and post using research-grade activity monitors (ActiGraph and activPAL), and the development of habits surrounding the use of the trackers was assessed postintervention with the Self-Reported Habit Index. During the intervention, Fitbit wear time and activity levels were monitored with the activity trackers. Linear regression analyses were used to determine the influence of the trial on outcomes of physical activity and sedentary time. The influence of habits was examined using correlation coefficients relating habits of tracker use (wearing the tracker and checking data on the tracker and associated app) to Fitbit wear time and activity levels during the intervention and at follow-up.

Results: Regression analyses revealed no significant differences by group in any of the primary outcomes (all $P>.05$). However, personal characteristics, including lower baseline activity levels (beta=–.49, $P=.01$) and lack of previous experience with pedometers (beta=–.23, $P=.03$) were predictive of greater improvements in moderate and vigorous physical activity. Furthermore, for individuals with higher activity levels at the baseline, MI and habit education were more effective for maintaining these activity levels when compared with receiving a Fitbit alone (eg, small increase of ~48 steps/day, $d=0.01$, vs large decrease of ~1830 steps/day, $d=0.95$). Finally, habit development was significantly related to steps/day during ($r=.30$, $P=.004$) and following the intervention ($r=.27$, $P=.03$).

Conclusions: This study suggests that activity trackers may have beneficial effects on physical activity in healthy adults, but benefits vary based on individual factors. Furthermore, this study highlights the importance of habit development surrounding the wear and use of activity trackers and the associated software to promote increases in physical activity.

Trial Registration: ClinicalTrials.gov NCT03837366; https://clinicaltrials.gov/ct2/show/NCT03837366

(JMIR Mhealth Uhealth 2019;7(2):e10988) doi:10.2196/10988

KEYWORDS

activity tracker; habit; mHealth; motivational interviewing; mobile phone; physical activity; wearable electronic devices

Introduction

Wearable technology remains popular [1], with industry experts projecting continued growth of the consumer sector [2,3]. However, despite widespread use, the utility of wearable activity trackers for improving physical activity (PA) is equivocal as highlighted in several reviews [4-6]. Thus, an important behavioral consideration is to determine how to optimally use the feedback from these monitors to most effectively facilitate behavior change.

Reportedly, activity trackers may be more effective when combined with additional behavior change strategies [7]; several studies have explored this possibility by testing the added benefits of education-based counseling and goal setting [8], as well as by integration with short message service text messaging and incentives through mobile health apps [9-11]. However, these strategies have not markedly enhanced the effectiveness of trackers. While a variety of additional strategies may be able to address this need, a recent review [12] concluded that monitors are more likely to be effective if they are accompanied by coaching or counseling from personnel with expertise in promoting behavior change.

This study aims to evaluate this recommendation by assessing the utility of a low-dose health coaching intervention to enhance outcomes associated with the use of wearable trackers. A low-dose coaching format was chosen as this would be more cost-effective and, thus, more suitable for broader translation through mobile health apps. The health coaching was based on the principles of motivational interviewing (MI). The use of MI offers advantages for clinically based health coaching, as it is designed to help build intrinsic motivation for behavior change [13]. In addition, MI has been widely used to positively influence behavior with notable apps for PA in several large intervention trials [14,15] and has been shown to improve adherence to and retention of health behaviors when combined with other behavior change strategies [16]. A novel aspect of our evaluation is in assessing whether MI-based coaching can promote more effective behavior change, incorporating the use of activity trackers. We hypothesized that those receiving health coaching along with their tracker would have greater improvements in PA and sedentary behavior than those receiving the tracker alone.

In concert with the MI-based approach, we also used contemporary theories of habit formation to promote and evaluate participants' adoption and use of the trackers and the relationship of these behaviors to PA and sedentary time. A habit has been defined as "a goal-directed sequence of actions that becomes automatic in response to learned, contextual cues" [17]. Though evidence has demonstrated that the development of habits can be involved in adoption and adherence to PA behaviors [18-21], studies, to date, have not specifically used habit theory to understand or promote PA along with activity tracker usage. The development of habits surrounding the tracker and associated app, such as regular wear of the monitor and checking data at a time where PA could be performed (eg, during the lunch hour), could lead to more consistent use and promote sustained improvements in activity levels.

Existing theory and empirical findings from observational studies showed that behavior change occurs in 2 distinct action phases—a motivational phase, which results in the formation of a behavioral goal or intention, and a volitional phase, which results in the enactment of behavior once a goal or intention is formed [22-25]. MI in combination with habit formation may enhance both initial motivation and the completion of the volitional phase of behavioral action (ie, carrying through with one's intentions, over time). Thus, as a secondary hypothesis, we anticipated that individuals receiving health coaching would establish stronger habits than those receiving an activity tracker alone and that greater habit development would be associated with the more consistent use of the tracker and more positive outcomes. The systematic evaluation of strategies to enhance the utilization of wearable trackers is an essential step in the more effective use of these devices in behavioral research.

Methods

Participants

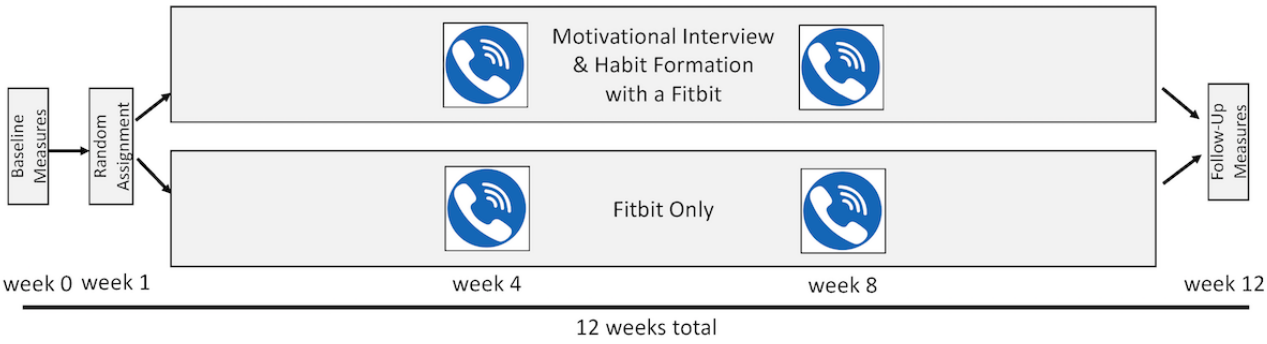
All procedures were approved by the relevant Institutional Review Board. Participants were recruited from the campus community using electronic mailing lists. The inclusion criteria were as follows: aged 24-65 years; regular access to a computer or smartphone; and a willingness to wear an activity tracker for the study duration. The exclusion criteria were we as follows: current use of an activity tracker; the presence of health conditions that prevented safe engagement in PA; current participation in a structured exercise program; or self-reported activity levels sufficient to meet the aerobic component of PA guidelines of 150 minutes of moderate or 75 minutes of vigorous activity per week.

Procedures

This study was designed as a feasibility trial [26] to refine methods prior to implementation in larger clinical trials. The protocol involved 3 laboratory visits and 2 phone calls over 3 months (Figure 1). Data were collected in 5 cohorts, ranging from 15 to 20 participants each. The first 3 began in summer, and the fourth and fifth began in the fall. No cohorts started or finished within 1 week of a major holiday. Prior to data collection, participants read and signed the informed consent document and completed the Physical Activity Readiness Questionnaire [27] to assess eligibility. Participants then completed a demographic questionnaire, and the International Physical Activity Questionnaire [28] was given in interview format to further determine eligibility. Participants who self-reported being sufficiently active to meet guidelines (n=2) were excluded from remaining study procedures.

To characterize the sample, blood pressure and heart rate were assessed with an automated blood pressure monitor (Omron HEM712C; Omron Healthcare, Inc, Hoffman Estates, IL, USA).

Figure 1. Timeline of study procedures.



While height was measured using a standard stadiometer, weight and body composition were estimated using bioelectrical impedance (InBody 720; InBody, Cerritos, CA, USA). Two research-grade activity monitors (described below) were then provided for participants to wear for 1 week to obtain baseline measures of PA and sedentary behaviors. Participants were verbally encouraged to behave as usual during this baseline assessment, and visits were rescheduled if participants indicated that the week would likely be unrepresentative of their typical behavior patterns owing to travel, illness, or other obligations.

Following the baseline monitoring period, participants were randomly assigned to receive a Fitbit alone (FB) or in combination with MI and habit education provided by a trained health coach (FB+). The FB group was intended to be comparable to a real-world setting wherein individuals purchase and utilize activity trackers on their own. Regardless of group assignment, all participants received a Fitbit Charge wrist-worn activity monitor and were instructed to use it at their discretion for the duration of the 3-month intervention. The study staff assisted participants in setting up a Web-based Fitbit account and demonstrated features of the Fitbit itself and the associated computer software and app (eg, logging food, dashboard displays, and alarms). In addition, participants assigned to the FB+ group discussed their self-determined goals regarding PA and principles of habit formation with a trained staff member. Habit education included a brief definition of habits and their relevance for sustained behavior change, followed by working with participants to determine salient cues to remember to wear their Fitbit and regularly check their data on the Fitbit itself and through the app at a time when PA was feasible. During the 3-month intervention, behaviors were monitored through Fitabase software (Small Steps, Inc; San Diego, CA, USA) as a measure of implementation, including metrics for Fitbit wear time and use and to track PA behaviors (eg, steps/day) throughout the intervention.

During the fourth and eighth weeks of the study, all participants were contacted by phone for a brief (~10-30 minute) conversation regarding their experience using the Fitbit; this included likes, dislikes, and any technical difficulties they were experiencing. In addition, individuals in the FB+ group revisited their self-selected PA goals and motivation for change. After 12 weeks, during which participants still had access to their Fitbits, the same research-grade PA monitors were distributed to assess behavior change. Participants then returned to the lab for their final visit, repeating all baseline assessments and completing a measure of habit strength surrounding the use of the Fitbit and associated software (described below).

Measures

Evaluation of Physical Activity and Sedentary Time

Pre- and postintervention, PA and sedentary behaviors were assessed objectively using ActiGraph GT3X+ (ActiGraph, LLC, Fort Walton Beach, FL, USA) and activPAL3 (Physical Activity Technologies, Glasgow, UK) activity monitors. Participants were instructed to place the ActiGraph on the hip using the elastic belt and the activPAL on the midline of the thigh on either leg. For both monitors, participants were instructed to wear them during all waking hours, except when bathing or swimming. To accompany the monitors, participants were given a log sheet and asked to record time on or off for each monitor and waking hours (eg, sleep and wake times).

The ActiGraph and activPAL data were initially processed using proprietary software (ActiLife 6.13.3 and Performance Analysis of Logs Analysis 7.2.32) with outputs at the 1-second epoch. These data were then integrated through their timestamps and further analyzed using the Sojourns Including Posture method

[29]. The Sojourns Including Posture method has been shown to provide a high-level validity for the full range of behaviors from sedentary to vigorous in comparison to direct observation and indirect calorimetry [29]. The output from this processing method was used to calculate the primary outcomes for the study, including average steps, minutes of moderate-to-vigorous intensity PA (MVPA; total and in 10-minute bouts) and sedentary time (total and in 30-minute bouts) per day. Change scores for these metrics were calculated by subtracting baseline scores from follow-up scores, and percent changes were then calculated using the following formula: [(follow-up–baseline)/baseline]×100.

Evaluation of Habit Formation

The habit development was assessed postintervention using the Automaticity Subscale from the Self-Reported Habit Index (SRHI) [30]. The SRHI was designed to be flexible such that it can be used with a variety of potentially habitual behaviors to fit study needs and has been shown to predict PA behaviors [31]. The Automaticity Subscale includes 4 of the 12 items from the full questionnaire and has been validated as a measure of habitual behavior [32]; each item is scored on a 5-point Likert scale with anchors ranging from strongly agree to strongly disagree. Behaviors in question are used as the stems of sentences that are followed by the 4 automaticity-related sentence endings. The behaviors assessed in this study were wearing the Fitbit, checking data on the Fitbit itself, and checking data on the Fitbit software or app. Thus, participants were asked to indicate the extent to which they agreed with statements like "Wearing my Fitbit is something I do without thinking" and "Using the Fitbit mobile app is something I start doing without realizing I'm doing it." Scores on this measure are summed across the 4 items to create an automaticity index for each behavior. In this study, each of the 3 behaviors noted above was examined separately and a habit automaticity total was also created by summing the totals from each of the separate behaviors.

Physical Activity Behaviors During the Intervention

Fitbit wear, as well as activity levels, were monitored during the study using the Fitabase software mentioned above. The data from Fitabase were downloaded in 1-minute epochs and summarized to determine the frequency for wearing the tracker, as well as accumulation of daily steps and minutes of MVPA averaged over each week of the intervention. A valid monitor wear day was defined as accumulating steps during, at least, 10 hours/day, indicating that the monitor was being worn. Days on which the Fitbit was worn for <10 hours were excluded from analyses. Thus, metrics of interest (eg, average steps/day) were calculated using only days with sufficient wear time. Participants' wear time, daily steps, and minutes of MVPA were tracked throughout the 12 weeks to enable indicators of habit to be directly related to the objective data.

Statistical Analyses

Descriptive statistics were used to characterize the sample with regards to demographic variables, as well as baseline PA and sedentary behaviors. Group differences at baseline were analyzed using chi-square tests for categorical variables and independent-samples *t* tests for continuous variables. In addition, data from Fitabase were examined descriptively to assess behavior and use of the monitors during each week of the study. Data were then compared between groups using independent-samples *t* tests for each week of the intervention.

To examine the influence of the MI-based health coaching on outcomes related to active and sedentary behaviors, we first descriptively compared within- and between-groups changes using effect size calculations (Cohen *d*). To examine the influence of baseline levels of activity on the effects of the FB or in combination with MI, the sample was also further subdivided into high and low active with 7500 steps/day at the baseline, serving as the cutoff point based on the established range for being considered "active" from Tudor-Locke et al [33]. Outcomes for steps, MVPA, and sedentary time were again compared within these subgroups using effect size calculations. Finally, a series of linear regression analyses were performed for various outcomes (eg, steps, minutes of MVPA, and sedentary time) with percent change as the dependent variable in each analysis. Predictors were group, study cohort, age, gender, BMI, previous pedometer use, and baseline value of the selected outcome variable. In addition, an interaction term between the group and baseline level of each outcome variable was included in each analysis; values for baseline activity levels were centered around the grand mean prior to calculation of this term. The alpha level was set at .05 for all analyses.

To evaluate the impact of habit formation on outcomes, we first compared group scores on the SRHI Automaticity Subscale for each of the 3 Fitbit-related habits using independent-sample *t* tests. For descriptive purposes, participants were also divided into those with high and low habit strength based on a median split of total Automaticity scores (across the 3 behaviors). Then, wear time (average valid wear days/week) was plotted over the course of the intervention. Finally, as individuals in both groups reported developing habits surrounding their Fitbits, correlation coefficients were calculated across all participants to examine associations among habit formation scores (eg, Automaticity scores for each of the 3 habits, including wearing the Fitbit, checking data on the Fitbit, checking data on the app, and the habit total score), FB usage averaged across the intervention period, for example, wear time and activity (eg, steps/day), and the active and sedentary behavior-related outcomes at follow-up.

Results

Participants were primarily white, college-educated, and overweight. As shown in Table 1, groups were similar with respect to basic demographic characteristics and resting heart rate and blood pressure. In addition, groups were similar in their active and sedentary behaviors at the baseline, as shown in Table 2.

Across the 12 weeks of the study, participants in both groups decreased the number of days/week the Fitbit was worn, with more notable declines in the final weeks (Figure 2). Similarly, even on days when the Fitbit was worn, steps/day during the intervention (measured via the Fitbit) decreased toward the end of the intervention period. There were no significant group differences at any time-point in any of these measures.

Table 1. Baseline participants' characteristics.

Characteristics	Full sample (n=91)	Fitbit plus motivational interviewing and habit education (n=45)	Fitbit alone (n=46)	Group differences, *P* value
Age in years, mean (SD)	41.7 (9.3)	41.1 (9.2)	42.2 (9.4)	.56
Male, n (%)	43 (47)	20 (44)	23 (50)	.68
Education with college degree, n (%)	87 (97)	43 (96)	44 (96)	.83
Married, n (%)	73 (80)	36 (80)	37 (80)	.77
Income>50,000/year, n (%)	79 (87)	38 (84)	40 (87)	.69
White individuals, n (%)	72 (79)	33 (73)	39 (85)	.44
Employed full-time, n (%)	88 (97)	43 (96)	46 (100)	.34
Previous pedometer use reported as yes, n (%)	36 (40)	17 (38)	19 (41)	.83
Heart rate (bpm), mean (SD)	70.5 (11.6)	70.9 (11.9)	70.1 (9.1)	.74
Blood pressure				
Systolic blood pressure (mm HG), mean (SD)	121.7 (16.5)	12.1 (17.5)	123.2 (15.6)	.37
Diastolic blood pressure (mm HG), mean (SD)	76.5 (12.5)	74.9 (13.6)	78.1 (11.2)	.24
Height (cm), mean (SD)	171.9 (8.5)	17.9 (9.13)	172.8 (7.7)	.40
Weight (kg), mean (SD)	87.8 (21.4)	82.7 (21.4)	9.5 (18.9)	.21
Body mass index, mean (SD)	29.6 (6.3)	28.7 (6.3)	3.4 (6.2)	.20
% Body fat, mean (SD)	34.8 (8.2)	34.1 (9.0)	35.5 (8.6)	.43

As shown in Table 2, results comparing outcomes from pre to post demonstrated that participants in FB+ had small improvements in steps/day and MVPA in bouts of, at least, 10 minutes, while participants in FB had small decreases in steps per day and MVPA. For sedentary time, FB+ had small increases, and FB had small decreases. These changes were all nonsignificant. Effect size calculations comparing changes between groups showed that group differences in change over the intervention were also small in magnitude (d_{range}=0.13-0.29). However, the effectiveness of the intervention was variable among participants with some individuals improving markedly and others decreasing activity levels over the 12-week period, as highlighted by the large SDs for change scores shown in Table 2.

This variability was partially explained when participants were further subdivided into high and low active using their baseline activity levels, as shown in Figure 3. Effect size calculations demonstrated that participants assigned to FB, who were low active at baseline (shown in the gray striped bars) had moderate improvements in both active and sedentary behaviors over time (d_{range}=0.36-0.66). Those who were high active at baseline (shown in the solid gray bars), however, became less active and more sedentary across the 3 months (d_{range}=0.18-0.95). A comparison of these changes within the FB group demonstrated that activity status at baseline had a moderate to large effect (d_{range}=0.46-1.23) on the benefits of using FB. Differences within the FB+ group (shown in the solid and striped black bars) based on activity level at baseline were small in magnitude (d_{range}=0.02-0.46) and generally favored the lower active group for variables related to PA and the higher active group for variables related to sedentary time.

Table 3 presents results from the linear regression analyses. Regressions demonstrated no significant difference in any of the primary outcomes (all *P*>.05) between the FB+ and FB groups. With respect to individual differences, lower baseline steps/day significantly predicted increases in daily steps over the intervention (*P*=.002). In addition, the interaction term (Group×Baseline steps; *P*<.001) was statistically significant. Specifically, as illustrated in Figure 3, participants with higher steps at baseline benefited more from being assigned to the FB+ group than the FB group. For changes in MVPA, significant predictors were previous experience with a pedometer (*P*=.03), and baseline minutes of MVPA (*P*=.01). Participants with no previous pedometer experience and lower levels of baseline MVPA showed greater improvements. For changes in sedentary time, the only significant predictor was baseline levels of sedentary time, such that participants with higher baseline levels of sedentary time had greater decreases in sedentary time postintervention (*P*=.003).

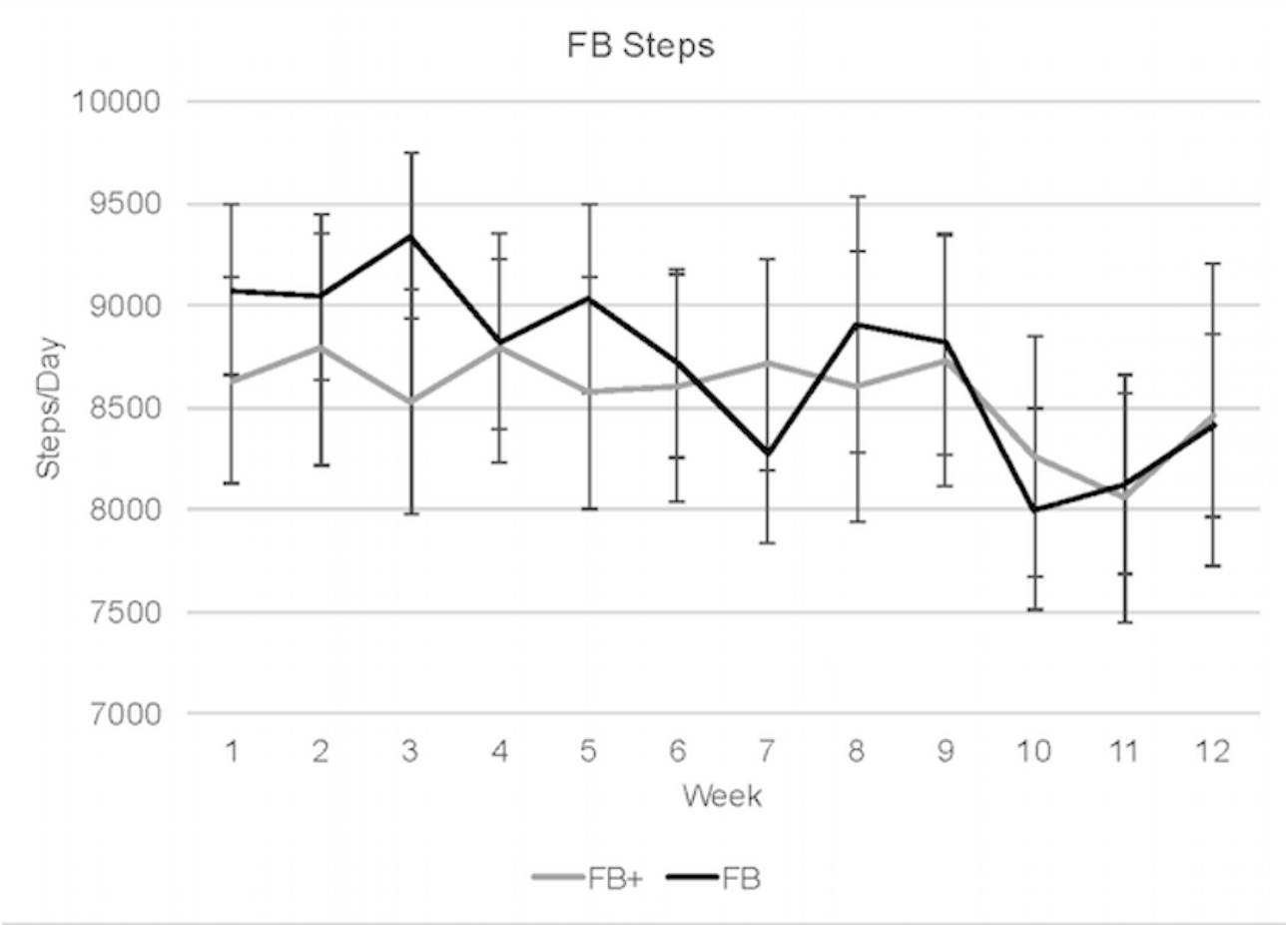
Table 2. Within- and between-groups comparisons of active and sedentary behaviors assessed using ActiGraph and activPAL monitors over the intervention.

Comparisons	Fitbit plus motivational interviewing and habit education				Fitbit alone				Between-groups effect size for change (Cohen *d*)
	Baseline, mean (SD)	12 weeks, mean (SD)	Change[a] (post-pre), mean (SD)	Within-group effect size (Cohen *d*)	Baseline, mean (SD)	12 weeks, mean (SD)	Change[a] (post-pre), mean (SD)	Within-group effect size (Cohen *d*)	
Average steps/day (n=91)	7496.88 (2895.94)	7574.38 (3499.38)	77.50 (2293.78)	.02	7519.60 (2259.13)	7221.80 (2106.18)	–297.80 (2658.84)	0.14	0.15
Average minutes of MVPA[b]/day (n=81)	68.57 (3.93)	68.91 (31.51)	.35 (3.96)	.13	75.06 (23.41)	69.93 (21.74)	–5.12 (24.40)	0.06	0.20
Average minutes of MVPA in 10+ min bouts/week (n=81)	77.99 (118.83)	93.16 (122.95)	15.17 (122.86)	.13	86.19 (78.99)	81.13 (77.55)	–5.07 (87.51)	0.17	0.19
Average minutes of sedentary time/day (n=91)	550.69 (102.74)	563.81 (106.53)	13.11 (99.68)	.14	547.96 (106.52)	529.94 (111.61)	–18.02 (112.82)	0.04	0.29
Average minutes of sedentary time in 30+ min bouts (n=91)	290.63 (128.34)	308.63 (127.37)	18.00 (88.41)	.01	279.35 (117.62)	284.22 (103.72)	4.87 (115.27)	0.23	0.13

[a]Positive values for change scores indicate an increase from pre- to postintervention.

[b]MVPA: moderate-to-vigorous intensity physical activity.

Figure 2. Wear data and steps collected from the Fitbit via Fitabase during the 12-week intervention. FB+: Fitbit plus motivational interviewing and habit education; FB: Fitbit alone.



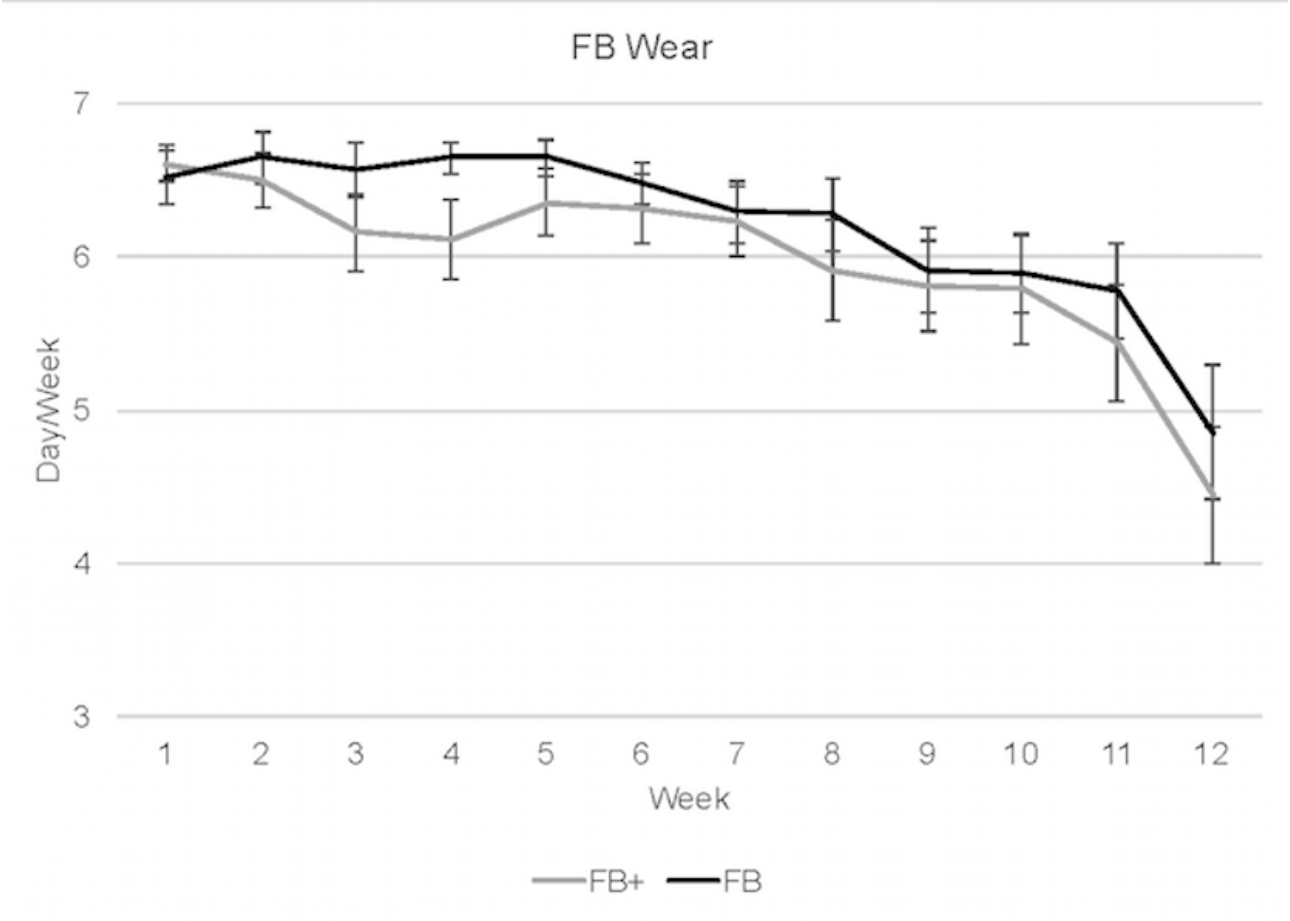


Figure 3. Within- and between-group comparisons of changes in active and sedentary behaviors measured via ActiGraph and activPAL pre-post based on original grouping (FB+, FB) and baseline activity level. Positive values for change scores indicate an increase from pre- to postintervention. FB+: Fitbit with motivational interviewing; FB: Fitbit alone.

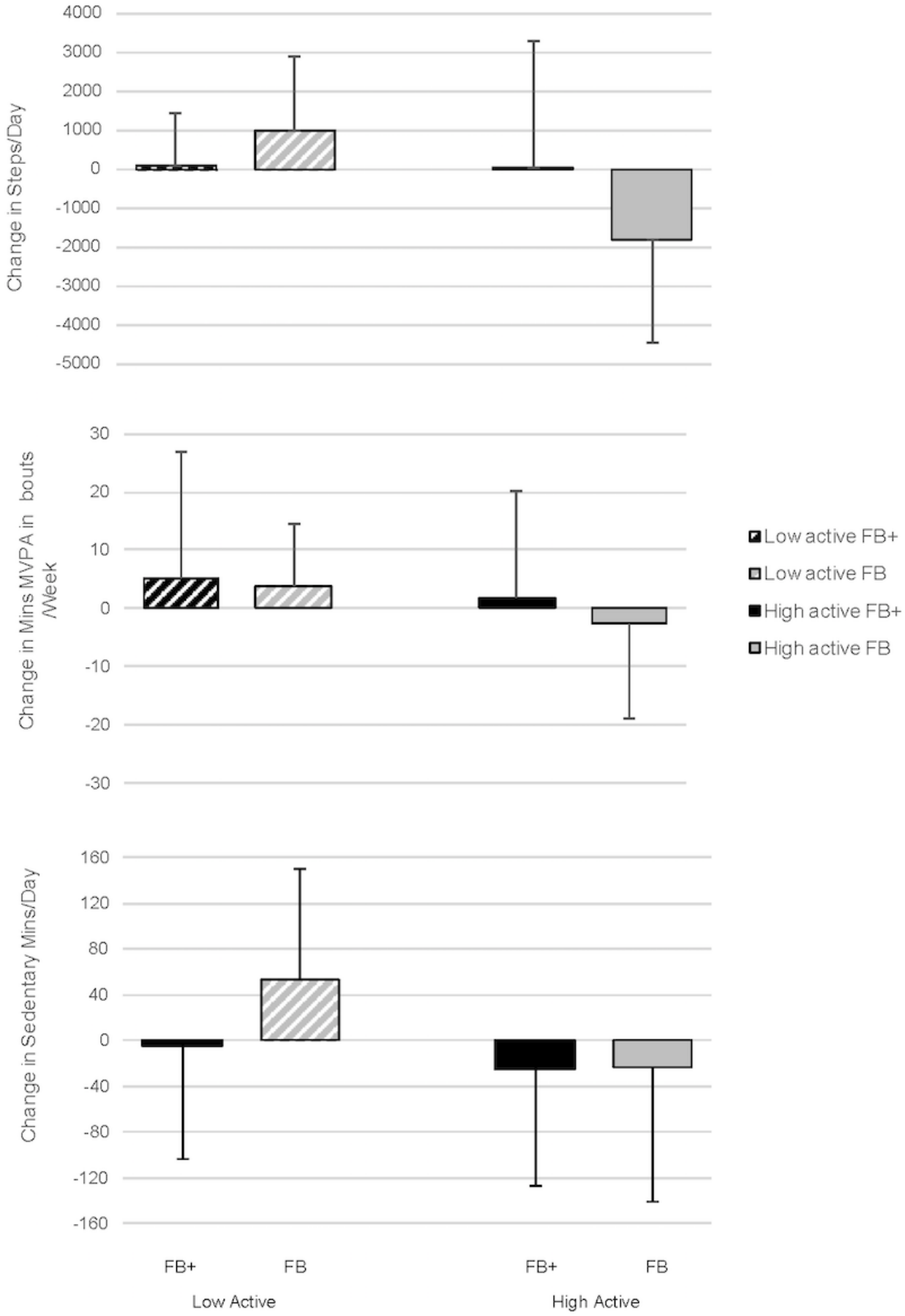


Table 3. Results from the regressions examining predictors of change in active and sedentary time measured using ActiGraph and activPAL monitors.

Independent variables	Dependent variables: percent change across the intervention					
	Δ Steps		Δ MVPA[a]		Δ Sedentary time	
	Standard beta	*P* value	Standard beta	*P* value	Standard beta	*P* value
Group	.01	.89	–.06	.57	–.15	.15
Cohort	–.15	.15	–.18	.10	.09	.34
Sex	–.01	.89	–.17	.11	–.14	.17
Age	.02	.83	.16	.15	–.03	.78
Body mass index	–.03	.76	–.04	.70	–.08	.41
Previous pedometer experience	–.13	.18	–.23	.03	–.02	.83
Baseline level of DV[b] (steps, MVPA, sedentary time)	–.79	.002	–.49	.006	–.42	.003
Group×baseline level of DV	.49	<.001	–.04	.79	–.07	.61
Overall model	R^2=0.27	<.001	R^2=0.27	.003	R^2=0.25	.003

[a]MVPA: moderate-to-vigorous intensity physical activity.

[b]DV: dependent variable.

Table 4. Descriptive statistics (mean [SD]) for the Automaticity Subscale of the Self-Reported Habit Index.

Self-reported habit index Automaticity Subscale	Full Sample (n=91)	Fitbit with motivational interviewing (n=45)	Fitbit alone (n=46)	Group differences (*P* value)
Wearing the Fitbit	16.3 (3.9)	16.7 (3.9)	15.8 (3.8)	.28
Checking data on the Fitbit	13.4 (4.8)	13.8 (5.2)	13.0 (4.4)	.44
Checking data using software or app	12.3 (4.9)	12.1 (5.1)	12.4 (4.8)	.78

Figure 4. Fitbit wear time in relation to habit automaticity over the intervention period by the group (FB+, FB). High versus low habit was defined using a median split of the data. FB+: Fitbit with motivational interviewing; FB: Fitbit alone.

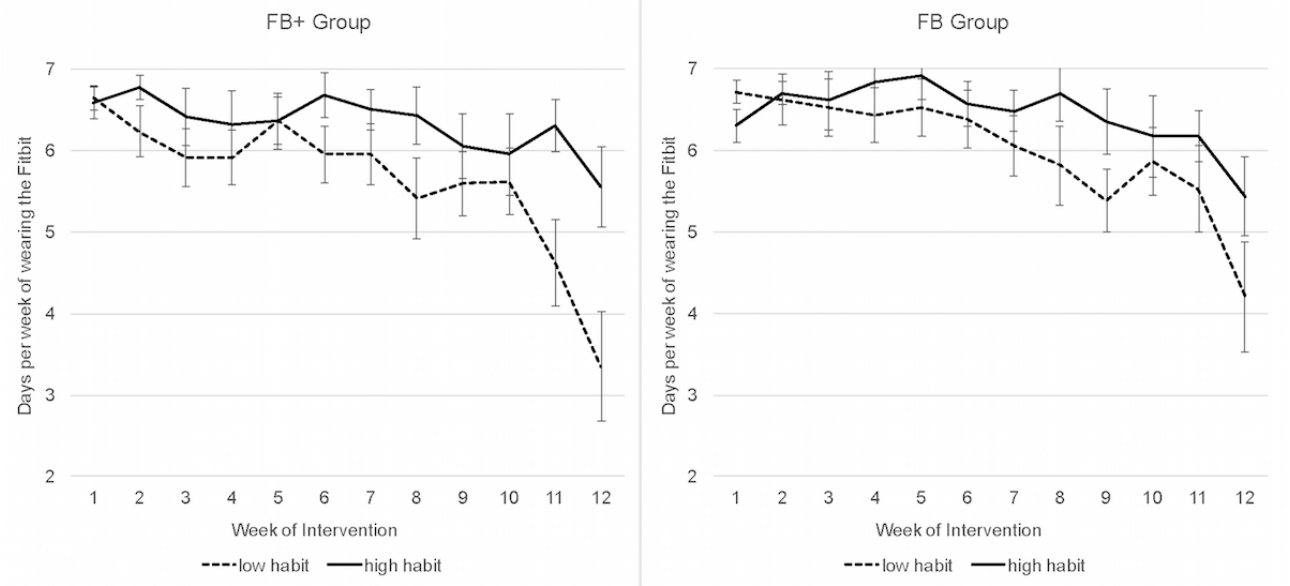


Table 5. The correlation matrix showing relationships among habits associated with wear and use of Fitbits, objectively measured Fitbit wear and steps during the intervention (data collected from Fitbits via Fitabase), and primary outcomes for active and sedentary behaviors (data collected via ActiGraph and activPAL monitors).

Habit and Outcome Variables	Habits				Fitbit (1-12 weeks), mean		MVPA[b]		Steps	Sedentary	
	Wearing FB[a]	Checking data on FB	Checking data on app	Total	Days of wear	Steps	Total	In bouts	Steps	Total	In bouts
Habits for wearing FB	1	N/A[c]	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Habits for checking data on FB	.51[d]	1	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Habits for checking data on app	.41[d]	.62[d]	1	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Habit total	.74[d]	.88[d]	.84[d]	1	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Days of FB wear (1-12 weeks), mean	.48[d]	.22[e]	.38[d]	.43[d]	1	N/A	N/A	N/A	N/A	N/A	N/A
FB steps (1-12 weeks), mean	.10	.27[e]	.35[d]	.30[d]	.21[e]	1	N/A	N/A	N/A	N/A	N/A
MVPA total	.02	.23[e]	.19	.18	.06	.78[d]	1	N/A	N/A	N/A	N/A
MVPA in bouts	.10	.21	.27[e]	.24[e]	.15	.67[d]	.86[d]	1	N/A	N/A	N/A
Steps	.02	.25[b]	.27[c]	.23[e]	.04	.69[d]	.85[d]	.71[e]	1	N/A	N/A
Sedentary total	.25[a]	−.06	−.11	.01	.10	−.40[d]	−.38[d]	−.24[e]	−.38[d]	1	N/A
Sedentary bouts	.11	−.18	−.17	−.12	.12	−.43[d]	−.47[d]	−.27[d]	−.47[d]	.83[d]	1

[a]FB: Fitbit.

[b]MVPA: moderate-to-vigorous intensity physical activity.

[c]N/A: not applicable.

[d]Significant at $P<.01$.

[e]Significant at $P<.05$.

Groups were not significantly different with respect to automaticity of habits surrounding the wearing and use of Fitbit, as shown in Table 4 (all $P>.05$). As shown in Figure 4, regardless of group, individuals reporting greater habit strength regarding Fitbit use were more likely to wear the Fitbit regularly during the second half of the intervention than those with lower habit strength. Results from the correlation analyses, shown in Table 5, demonstrated that habits surrounding the Fitbit were predictive of actual Fitbit wear and steps/day during the intervention (data from Fitabase). Postintervention, automaticity of habits surrounding checking the data on the Fitbit itself was most predictive of total minutes of MVPA and habit surrounding checking the data on the Fitbit App were most predictive of MVPA in bouts and average steps/day (data from ActiGraph and activPAL monitors). Correlations between Fitbit-related habits and sedentary time (total and in bouts) were largely small and nonsignificant, except habits regarding wearing the Fitbit being significantly associated with total minutes of sedentary time.

Discussion

The primary purpose of this study was to evaluate the utility of a low-dose, MI-based health coaching intervention to enhance outcomes associated with the use of wearable activity trackers. The results demonstrated that the effects of using FB on PA behaviors in healthy adults were modest and highly variable. On average, for participants who received a Fitbit without additional support, PA levels declined slightly from pre- to postintervention. However, results showed that individuals who were less active at the baseline had moderate improvements in PA (eg, ~19% increase in steps/day); this is in contrast to those who were higher active at baseline who showed decreases of ~17% in both steps and minutes of MVPA/day over the 12 weeks. Thus, there may be some benefits to simply providing activity trackers to generally healthy adults who are insufficiently active. However, these results also highlight an important consideration in that active adults may not benefit from these devices without additional support.

In addition, these results demonstrated that the provision of low-dose health coaching had little effect on the sample as a whole. On average, participants receiving coaching increased their steps and minutes of MVPA, but the effects were small in magnitude. However, somewhat surprisingly, health coaching did provide an added benefit for those who were more active at the baseline. While active individuals who received a Fitbit without additional support had declines in their activity levels, those who received health coaching were able to maintain their high levels of activity throughout the intervention. Previous research has examined the effects of wearable technology in combination with a variety of other behavior change strategies, including education and goal setting, short message service text messages, and incentives. These studies found that the additional

strategies were ineffective for improving outcomes associated with the provision of activity trackers [8-11].

In contrast to the additional behavior change strategies tested previously, the MI-based health coaching used in this study focused on the development of intrinsic motivation for behavior change, which should theoretically be more effective than strategies that focus on external motivators such as incentives [34]. Our results suggest that the health coaching did have benefits, but those were not equally distributed across participants. Future research examining different doses of coaching in concert with longer duration interventions will be critical for addressing these important questions.

This study adds to the growing body of research examining the effectiveness of wearable activity trackers and suggest that individual factors may influence the effectiveness of these devices for promoting increases in PA. Though there were several early studies showing that activity trackers had promise for increasing activity levels [35,36], there have also been larger trials that have failed to show that they significantly influenced behavior or health [10,37-39]. However, as described above, our data demonstrate that some individuals benefited a great deal from these devices, particularly those who were less active to start and, thus, potentially stood to gain the most from improved activity levels. Previous research in low-active samples is equivocal with respect to this finding with some evidence that activity trackers are beneficial [35] and other evidence that they do not substantively improve PA [39]. However, to the best of our knowledge, baseline activity levels have not been explicitly examined as a predictor of success with activity monitors. In addition, data from this study suggest that the lack of previous experience with a pedometer was also a potentially important predictor of improvements in PA. Thus, while these devices may not be effective for all individuals, they do appear to have substantial benefits for some. Additional research looking at individual differences in success with utilizing these devices is warranted to determine the profile of those who may benefit the most.

A possible explanation for the lack of impact in the studies noted above is the documented decline in interest and use of monitors over time after the novelty wears off. In line with this, investigations of wearable abandonment have shown that one-third of consumers discontinue the use of activity trackers after 6 months [40]; this is highlighted by a prominent and highly publicized study from Jakicic et al [37], which reported limited value from the inclusion of wearable activity trackers in standard weight loss programming. In this study, participants were provided with monitors for 18 months (~550 days), but the median number of days the monitor was worn was only 170, approximately 31% of available days, and the median wear time on the days the device was worn was 240 minutes out of a possible 1440 (~17% of the days). Similarly, in the Le et al study [39], monitors were only worn 19 days each month during the first 3 months of the intervention and usage declined to 15 days each month during the subsequent 3 months. Thus, the lack of impact in these and other studies may be attributed to reduced interest or lack of attention given to helping participants learn how to effectively use wearable monitors to promote sustained behavior change; our results reflected this as well. As highlighted in Figure 2, the number of days the Fitbits were worn each week declined from an average of 6.5-4.4 by the end of the intervention.

In addition, this study suggests that the habit development surrounding wear and regular interactions with these devices and their associated apps may be a key factor in promoting more sustained behavior change. Specifically, the habit development was predictive of wearing activity trackers more consistently during the final weeks of the intervention. Furthermore, habits, particularly those regarding the use of the Fitbit app, were predictive of greater levels of PA both during and following the intervention. Based on the design of the trial, it is difficult to determine whether habit developments played a causal role in behavior change. For example, it is possible that developing habits surrounding the use of the Fitbit led to improvements in PA. However, it is equally likely that those who were more successful at increasing PA were more likely to develop Fitbit-related habits. In addition, it should be noted that tying PA habits to an external device, like Fitbit, could be potentially problematic. These devices break, get lost, and people's preferences for wearing them likely change over time. Moreover, the devices and associated apps evolve with some features being added or removed, altering the user experience. Thus, when using activity trackers to develop PA habits, the focus should be on establishing cues that are separate from the device itself to promote sustained improvements in PA [41].

To the best of our knowledge, the effects of activity trackers on sedentary behavior have not been examined. Although this study was not specifically targeted towards sedentary time, our results showed that in a group of generally healthy adults, this intervention had little effect on this set of behaviors, even in participants who improved their levels of PA; this is not surprising as previous research has demonstrated that interventions focusing on PA have little effect on the lower end of the intensity continuum [42]. Furthermore, the device used (Fitbit Charge) did not provide specific feedback on sedentary behaviors to the wearer. Sedentary behavior is an important contributor to health outcomes, and newer models of Fitbit and other similar devices do include features like idle alerts, which are designed to alert the wearer when they have been inactive for a prolonged period (eg, 50 minutes). Future studies using wearables with the intention of deceasing sedentary time should explicitly target sedentary behavior and include the use of a tracker that provides cues related to sedentary behavior and some level of feedback specific to these behaviors.

This study has a number of limitations. The sample was relatively small, generally healthy, and the duration of the trial was only 3 months. As such, the generalizability is limited to healthy adults and examination of factors predicting long-term adherence was not possible. Furthermore, we used a single, low-dose of health coaching; this was done purposely with an eye towards future translatability. However, it is possible that a higher dose of health coaching would have been more effective. Finally, our results demonstrating that lower active individuals improved more while individuals who were more active became less so could be interpreted as classic regression to the mean. However, the fact that higher active individuals receiving health coaching maintained their levels of activity

across the intervention makes this a less likely explanation for these findings.

In summary, results from this study demonstrate that wearable activity trackers may have beneficial effects on PA in healthy adults. However, these benefits do not occur across the board and are more likely to be observed in individuals with lower levels of baseline activity. Furthermore, this study suggests that to engage more active individuals, additional behavior change strategies, like the provision of health coaching, may be needed. In addition, this study highlights the importance of habit development surrounding the wear and use of activity monitors and engagement with the associated software to promote increases in PA; this may be an important area to pursue, as making PA habitual is critical for realizing associated health benefits. As such, further research is warranted to determine whether explicit efforts to develop activity monitor-related habits can improve the effectiveness of these devices. Finally, additional research examining the effects of these devices alone or in combination with health coaching in subclinical and clinical populations is needed to determine whether they can be used in conjunction with more traditional strategies to prevent and treat chronic health conditions.

Conflicts of Interest

None declared.

Editorial notice: This randomized study was only retrospectively registered. The editor granted an exception from ICMJE rules mandating prospective registration of randomized trials because the risk of bias appears low. However, readers are advised to carefully assess the validity of any potential explicit or implicit claims related to primary outcomes or effectiveness, as retrospective registration does not prevent authors from changing their outcome measures retrospectively.

References

1. Bogaty S. 2015. The demographic divide: Fitness trackers and smartwatches attracting very different segments of the market, according to the NPD group Internet URL: https://www.npd.com/wps/portal/npd/us/news/press-releases/2015/the-demographic-divide-fitness-trackers-and-smartwatches-attracting-very-different-segments-of-the-market-according-to-the-npd-group [accessed 2015-11-30] [WebCite Cache ID 72tmWYtT7]
2. Llamas R. IDC Media Center. Worldwide Wearables Market to Nearly Double by 2021, According to IDC Internet URL: https://www.idc.com/getdoc.jsp?containerId=prUS42818517 [accessed 2018-11-22] [WebCite Cache ID 7470S2vKH]
3. Witkowski W. marketwatch. Smartwatch growth surges ahead of Apple, Fitbit offerings - MarketWatch Internet URL: https://www.marketwatch.com/story/smartwatch-growth-surges-ahead-of-apple-fitbit-offerings-2017-08-31 [accessed 2018-01-16] [WebCite Cache ID 72tpT2fLj]
4. Bort-Roig J, Gilson ND, Puig-Ribera A, Contreras RS, Trost SG. Measuring and influencing physical activity with smartphone technology: a systematic review. Sports Med 2014 May;44(5):671-686. [doi: 10.1007/s40279-014-0142-5] [Medline: 24497157]
5. Sullivan AN, Lachman ME. Behavior Change with Fitness Technology in Sedentary Adults: A Review of the Evidence for Increasing Physical Activity. Front Public Health 2016;4:289 [FREE Full text] [doi: 10.3389/fpubh.2016.00289] [Medline: 28123997]
6. Vandelanotte C, Müller AM, Short CE, Hingle M, Nathan N, Williams SL, et al. Past, Present, and Future of eHealth and mHealth Research to Improve Physical Activity and Dietary Behaviors. J Nutr Educ Behav 2016 Mar;48(3):219-228.e1. [doi: 10.1016/j.jneb.2015.12.006] [Medline: 26965100]
7. Patel MS, Asch DA, Volpp KG. Wearable devices as facilitators, not drivers, of health behavior change. JAMA 2015 Feb 3;313(5):459-460. [doi: 10.1001/jama.2014.14781] [Medline: 25569175]
8. Thompson WG, Kuhle CL, Koepp GA, McCrady-Spitzer SK, Levine JA. "Go4Life" exercise counseling, accelerometer feedback, and activity levels in older people. Arch Gerontol Geriatr 2014;58(3):314-319. [doi: 10.1016/j.archger.2014.01.004] [Medline: 24485546]
9. Choi J, Lee JH, Vittinghoff E, Fukuoka Y. mHealth Physical Activity Intervention: A Randomized Pilot Study in Physically Inactive Pregnant Women. Matern Child Health J 2016 May;20(5):1091-1101. [doi: 10.1007/s10995-015-1895-7] [Medline: 26649879]
10. Finkelstein EA, Haaland BA, Bilger M, Sahasranaman A, Sloan RA, Nang EEK, et al. Effectiveness of activity trackers with and without incentives to increase physical activity (TRIPPA): a randomised controlled trial. Lancet Diabetes Endocrinol 2016 Dec;4(12):983-995. [doi: 10.1016/S2213-8587(16)30284-4] [Medline: 27717766]
11. Wang JB, Cadmus-Bertram LA, Natarajan L, White MM, Madanat H, Nichols JF, et al. Wearable Sensor/Device (Fitbit One) and SMS Text-Messaging Prompts to Increase Physical Activity in Overweight and Obese Adults: A Randomized Controlled Trial. Telemed J E Health 2015 Oct;21(10):782-792. [doi: 10.1089/tmj.2014.0176] [Medline: 26431257]
12. Lobelo F, Kelli HM, Tejedor SC, Pratt M, McConnell MV, Martin SS, et al. The Wild Wild West: A Framework to Integrate mHealth Software Applications and Wearables to Support Physical Activity Assessment, Counseling and Interventions for Cardiovascular Disease Risk Reduction. Prog Cardiovasc Dis 2016;58(6):584-594. [doi: 10.1016/j.pcad.2016.02.007] [Medline: 26923067]

13. Miller WR, Rollnick S. Meeting in the middle: motivational interviewing and self-determination theory. Int J Behav Nutr Phys Act 2012;9:25 [FREE Full text] [doi: 10.1186/1479-5868-9-25] [Medline: 22385872]
14. Brodie DA, Inoue A. Motivational interviewing to promote physical activity for people with chronic heart failure. J Adv Nurs 2005 Jun;50(5):518-527. [doi: 10.1111/j.1365-2648.2005.03422.x] [Medline: 15882368]
15. Hardcastle SJ, Taylor AH, Bailey MP, Harley RA, Hagger MS. Effectiveness of a motivational interviewing intervention on weight loss, physical activity and cardiovascular disease risk factors: a randomised controlled trial with a 12-month post-intervention follow-up. Int J Behav Nutr Phys Act 2013 Mar 28;10:40 [FREE Full text] [doi: 10.1186/1479-5868-10-40] [Medline: 23537492]
16. Hettema J, Steele J, Miller WR. Motivational interviewing. Annu Rev Clin Psychol 2005;1:91-111. [doi: 10.1146/annurev.clinpsy.1.102803.143833] [Medline: 17716083]
17. Verplanken B, Aarts H, van Knippenberg A, Moonen A. Habit versus planned behaviour: a field experiment. Br J Soc Psychol 1998 Mar;37 (Pt 1):111-128. [Medline: 9554090]
18. Gardner B, Lally P, Wardle J. Making health habitual: the psychology of 'habit-formation' and general practice. Br J Gen Pract 2012 Dec;62(605):664-666 [FREE Full text] [doi: 10.3399/bjgp12X659466] [Medline: 23211256]
19. Kaushal N, Rhodes RE, Meldrum JT, Spence JC. The role of habit in different phases of exercise. Br J Health Psychol 2017 Dec;22(3):429-448. [doi: 10.1111/bjhp.12237] [Medline: 28374463]
20. Kaushal N, Rhodes RE, Spence JC, Meldrum JT. Increasing Physical Activity Through Principles of Habit Formation in New Gym Members: a Randomized Controlled Trial. Ann Behav Med 2017 Aug;51(4):578-586. [doi: 10.1007/s12160-017-9881-5] [Medline: 28188586]
21. Phillips LA, Gardner B. Habitual exercise instigation (vs. execution) predicts healthy adults' exercise frequency. Health Psychol 2016 Jan;35(1):69-77. [doi: 10.1037/hea0000249] [Medline: 26148187]
22. Gollwitzer PM. Goal Achievement: The Role of Intentions. European Review of Social Psychology 1993 Jan;4(1):141-185. [doi: 10.1080/14792779343000059]
23. Orbell S, Sheeran P. Motivational and Volitional Processes in Action Initiation: A Field Study of the Role of Implementation Intentions1. J Appl Social Pyschol 2000 Apr;30(4):780-797. [doi: 10.1111/j.1559-1816.2000.tb02823.x]
24. Scholz U, Keller R, Perren S. Predicting behavioral intentions and physical exercise: a test of the health action process approach at the intrapersonal level. Health Psychol 2009 Nov;28(6):702-708. [doi: 10.1037/a0016088] [Medline: 19916638]
25. Schwarzer R, Lippke S, Luszczynska A. Mechanisms of health behavior change in persons with chronic illness or disability: the Health Action Process Approach (HAPA). Rehabil Psychol 2011 Aug;56(3):161-170. [doi: 10.1037/a0024509] [Medline: 21767036]
26. Bowen DJ, Kreuter M, Spring B, Cofta-Woerpel L, Linnan L, Weiner D, et al. How we design feasibility studies. Am J Prev Med 2009 May;36(5):452-457 [FREE Full text] [doi: 10.1016/j.amepre.2009.02.002] [Medline: 19362699]
27. Adams R. Revised Physical Activity Readiness Questionnaire. Can Fam Physician 1999 Apr;45:992, 995, 1004-992, 995, 1005 [FREE Full text] [Medline: 10216799]
28. Craig CL, Marshall AL, Sjöström M, Bauman AE, Booth ML, Ainsworth BE, et al. International physical activity questionnaire: 12-country reliability and validity. Med Sci Sports Exerc 2003 Aug;35(8):1381-1395. [doi: 10.1249/01.MSS.0000078924.61453.FB] [Medline: 12900694]
29. Ellingson LD, Schwabacher IJ, Kim Y, Welk GJ, Cook DB. Validity of an Integrative Method for Processing Physical Activity Data. Med Sci Sports Exerc 2016 Dec;48(8):1629-1638. [doi: 10.1249/MSS.0000000000000915] [Medline: 27015380]
30. Verplanken B, Orbell S. Reflections on past behavior: A self-report index of habit strength. J Appl Soc Psychol 2003;33(6):1330.
31. Gardner B, de Bruijn G, Lally P. A systematic review and meta-analysis of applications of the Self-Report Habit Index to nutrition and physical activity behaviours. Ann Behav Med 2011 Oct;42(2):174-187. [doi: 10.1007/s12160-011-9282-0] [Medline: 21626256]
32. Gardner B, Abraham C, Lally P, de Bruijn G. Towards parsimony in habit measurement: testing the convergent and predictive validity of an automaticity subscale of the Self-Report Habit Index. Int J Behav Nutr Phys Act 2012 Aug 30;9:102 [FREE Full text] [doi: 10.1186/1479-5868-9-102] [Medline: 22935297]
33. Tudor-Locke C, Craig CL, Brown WJ, Clemes SA, De CK, Giles-Corti B, et al. How many steps/day are enough? For adults. Int J Behav Nutr Phys Act 2011;8:79 [FREE Full text] [doi: 10.1186/1479-5868-8-79] [Medline: 21798015]
34. Gardner B, Lally P. Does intrinsic motivation strengthen physical activity habit? Modeling relationships between self-determination, past behaviour, and habit strength. J Behav Med 2013 Oct;36(5):488-497. [doi: 10.1007/s10865-012-9442-0] [Medline: 22760451]
35. Cadmus-Bertram LA, Marcus BH, Patterson RE, Parker BA, Morey BL. Randomized Trial of a Fitbit-Based Physical Activity Intervention for Women. Am J Prev Med 2015 Sep;49(3):414-418. [doi: 10.1016/j.amepre.2015.01.020] [Medline: 26071863]
36. Rowe-Roberts D, Cercos R, Mueller F. Preliminary results from a study of the impact of digital activity trackers on health risk status. Stud Health Technol Inform 2014;204:143-148. [Medline: 25087541]

37. Jakicic JM, Davis KK, Rogers RJ, King WC, Marcus MD, Helsel D, et al. Effect of Wearable Technology Combined With a Lifestyle Intervention on Long-term Weight Loss: The IDEA Randomized Clinical Trial. JAMA 2016 Sep 20;316(11):1161-1171. [doi: 10.1001/jama.2016.12858] [Medline: 27654602]
38. Kim Y, Lumpkin A, Lochbaum M, Stegemeier S, Kitten K. Promoting physical activity using a wearable activity tracker in college students: A cluster randomized controlled trial. J Sports Sci 2018 Aug;36(16):1889-1896. [doi: 10.1080/02640414.2018.1423886] [Medline: 29318916]
39. Le A, Mitchell H, Zheng DJ, Rotatori J, Fahey JT, Ness KK, et al. A home-based physical activity intervention using activity trackers in survivors of childhood cancer: A pilot study. Pediatr Blood Cancer 2017 Dec;64(2):387-394. [doi: 10.1002/pbc.26235] [Medline: 27615711]
40. Maslakovic M. Gadgets Wearables. Wearable sales are booming, but are we wearing them? Internet URL: http:/ /gadgetsandwearables.com/2015/12/05/wearable-sales-booming-wearing/ [WebCite Cache ID 72tpXuWV4]
41. Rothman AJ, Gollwitzer PM, Grant AM, Neal DT, Sheeran P, Wood W. Hale and Hearty Policies: How Psychological Science Can Create and Maintain Healthy Habits. Perspect Psychol Sci 2015 Nov;10(6):701-705. [doi: 10.1177/1745691615598515] [Medline: 26581721]
42. Martin A, Fitzsimons C, Jepson R, Saunders DH, van der Ploeg HP, Teixeira PJ, EuroFIT consortium. Interventions with potential to reduce sedentary time in adults: systematic review and meta-analysis. Br J Sports Med 2015 Aug;49(16):1056-1063. [doi: 10.1136/bjsports-2014-094524] [Medline: 25907181]

Abbreviations

FB: Fitbit
FB+: Fitbit plus motivational interviewing and habit education
MI: motivational interviewing
MVPA: moderate-to-vigorous intensity physical activity
PA: physical activity
SRHI: Self-Reported Habit Index

Edited by G Eysenbach; submitted 07.05.18; peer-reviewed by J Maher, A Sullivan; comments to author 01.09.18; revised version received 04.10.18; accepted 20.10.18; published 14.02.19.

Please cite as:
Ellingson LD, Lansing JE, DeShaw KJ, Peyer KL, Bai Y, Perez M, Phillips LA, Welk GJ
Evaluating Motivational Interviewing and Habit Formation to Enhance the Effect of Activity Trackers on Healthy Adults' Activity Levels: Randomized Intervention
JMIR Mhealth Uhealth 2019;7(2):e10988
URL: https://mhealth.jmir.org/2019/2/e10988/
doi:10.2196/10988
PMID:30762582

Original Paper

Impact of a Novel Smartphone App (CureApp Smoking Cessation) on Nicotine Dependence: Prospective Single-Arm Interventional Pilot Study

Katsunori Masaki[1], MD, PhD; Hiroki Tateno[1,2], MD, PhD; Naofumi Kameyama[1], MD, PhD; Eriko Morino[3], MD, PhD; Riri Watanabe[4], MD, PhD; Kazuma Sekine[5,6], MD; Tomohiro Ono[7], MD, PhD; Kohta Satake[8,9], MD, MBA, MPH; Shin Suzuki[9], MD; Akihiro Nomura[8,10,11], MD, PhD; Tomoko Betsuyaku[1], MD, PhD; Koichi Fukunaga[1], MD, PhD

[1]Division of Pulmonary Medicine, Department of Medicine, Keio University School of Medicine, Tokyo, Japan
[2]Department of Internal Medicine, Saitama City Hospital, Saitama, Japan
[3]Department of Respiratory Medicine, Disease Control and Prevention Center, National Center for Global Health and Medicine, Tokyo, Japan
[4]Department of General Internal Medicine, National Center for Global Health and Medicine, Tokyo, Japan
[5]Akao Clinic, Yokohama, Japan
[6]Department of Respiratory Medicine, Nippon Koukan Hospital, Kawasaki, Japan
[7]Ono Clinic, Kakegawa, Japan
[8]CureApp Institute, Karuiwaza, Japan
[9]CureApp Inc, Tokyo, Japan
[10]Innovative Clinical Research Center, Kanazawa University, Kanazawa, Japan
[11]Department of Cardiovascular and Internal Medicine, Kanazawa University Graduate School of Medicine, Kanazawa, Japan

Corresponding Author:
Koichi Fukunaga, MD, PhD
Division of Pulmonary Medicine, Department of Medicine
Keio University School of Medicine
35 Shinanomachi, Shinjuku-ku
Tokyo, 160-8582
Japan
Phone: 81 3 3353 1211
Fax: 81 3 3353 2502
Email: km-fuku@cpnet.med.keio.ac.jp

Abstract

Background: Mobile apps have been considered to provide active and continuous support for smoking cessation. However, it is yet to be known whether a smoking cessation smartphone app improves long-term abstinence rates in nicotine-dependent patients.

Objective: This study aimed to evaluate the long-term abstinence effect of a novel smartphone app, CureApp Smoking Cessation (CASC), in patients with nicotine dependence.

Methods: In this prospective, interventional, multicenter, single-arm study, we provided the CASC app to all the participants, who used it daily for 24 weeks. The CASC app includes features to maximize the therapeutic effect of pharmacological therapies and counseling at outpatient clinics for smoking cessation. The primary endpoint was a continuous abstinence rate (CAR) from weeks 9 to 24, whereas secondary endpoints were CARs from weeks 9 to 12 and 9 to 52.

Results: Of the 56 adult smokers recruited, 1 did not download the app; therefore, 55 participants constituted the full analysis sample. The CAR from weeks 9 to 24 was 64% (35/55, 95% CI 51%-76%), whereas the CARs from weeks 9 to 12 and 9 to 52 were 76% (42/55, 95% CI 65%-88%) and 58% (32/55, 95% CI 46%-71%), respectively. These CARs were better than the results of the national survey on outpatient clinics with regard to smoking cessation under the National Health Insurance Program and that of the varenicline phase 3 trial in Japan and the United States. There was only 1 participant who dropped out during the 12 weeks of the treatment period. This treatment decreased the scores related to withdrawal and craving symptoms.

Conclusions: The addition of CASC to usual smoking cessation therapies resulted in high CARs, high patient retention rates, and improvement of cessation-related symptoms. The smartphone app CASC is a feasible and useful tool to help long-term continuous abstinence that can be combined with a standard smoking cessation treatment program.

(JMIR Mhealth Uhealth 2019;7(2):e12694) doi:10.2196/12694

KEYWORDS

digital therapeutics; nicotine dependence; smoking cessation; smartphone application; telemedicine

Introduction

Background: Smoking Cessation Therapy

Smoking is a risk factor for cancer, respiratory disease, heart disease, and cerebral vascular disease [1]. Quitting smoking lowers the risk of smoking-related illnesses and can add years to the lives of ex-smokers [1,2]. However, only 3% to 5% of self-quitters achieve prolonged abstinence for 6 to 12 months after an attempt to quit [3]. Although the smoking prevalence continues to decline in Japan, according to the data of the Ministry of Health, Labour and Welfare, the rate of smokers among the entire Japanese population was still as high as 17.7% (males: 29.4%, females: 7.2%) in 2017. In Japan, patients diagnosed with nicotine dependence and those who wish to quit smoking are eligible for treatment with varenicline or nicotine patches under the National Health Insurance Program (NHIP) [4]. Under the NHIP, physicians see their patients and provide pharmacotherapy and counseling 5 times during the 12-week treatment period according to the national guidelines (The Standard Procedure Manual for Smoking Cessation, 6th edition, 2014) [5]. Although pharmacotherapy, which helps patients stop smoking, has been covered by NHIP since 2006, continuous abstinence rate (CAR) after 12 weeks has still remained low mainly because of the high dropout rate. Only 30% of the patients completed all 5 visits to see their physicians, and the average number of visits per patient was 3.3 [6]. Dropouts and the lack of visits mean that physicians cannot provide their patients sufficient support for smoking cessation. In fact, there was a negative correlation between the number of visits and the patients' CARs [6]. It is well known that insufficient support leads to misunderstandings with regard to nicotine dependence and side effects of varenicline and nicotine patches, which eventually result in re-smoking [7-9]. Therefore, preventing dropouts and maintaining a high level of motivation and confidence to continue to remain engaged in the treatment are essential for smoking cessation therapy.

Mobile Health and Study Objectives

Recently, mobile phone, Web-based, or smartphone apps have been considered to provide effective and continuous support for smoking cessation [10-17]. For example, Tweet2Quit, a unique social networking service–based smoking cessation Twitter program and a 12-week short message service–based intervention for university students might be effective on smoking cessation [13,14]. In addition, a smartphone smoking cessation app that offers momentary ecological assessments could be useful in providing users with timely adaptive interventions by sending various treatment messages over a 3-week period [16]. One smartphone app for smoking cessation helped 30% of its users abstain from smoking for 8 weeks [17]. However, it is yet to be known whether a smoking cessation smartphone app that provides continuous evidence-based behavioral support and personalized counseling programs improves long-term abstinence rates among nicotine-dependent patients. No official or evidence-based smoking cessation app that provides advice and scientific information and offers various alternative approaches to smokers has yet been developed in Japan.

In this research, we performed a single-arm interventional study to evaluate whether a novel smartphone app for smoking cessation, CureApp Smoking Cessation (CASC), was useful and effective for long-term abstinence in patients with nicotine dependence.

Methods

Study Design

This was a prospective, interventional, multicenter, single-arm study. We provided the CASC smartphone app to all the participants, and they used it daily for 24 weeks. If they wished to, participants could use this app during the entire research period (52 weeks from their registration). We asked them to keep an electronic diary regarding smoking cessation, browse messages and tutorial videos, and report withdrawal symptoms and cravings through a chat system with an artificial intelligence (AI) nurse. Participants also visited their physicians' clinics at weeks 0, 2, 4, 8, and 12 according to the Smoking Cessation Standard Protocol in Japan [5]. Their CAR was the primary endpoint from weeks 9 to 24.

This study was approved by the Ethics Committee of Keio University School of Medicine and other facilities. Written informed consent was obtained from all the participants. This study was registered at the University Hospital Medical Information Network (UMIN) Clinical Trials Registry (UMIN000020123).

Recruitment

We recruited adult smokers who visited the institutions for nicotine dependence treatment and were eligible for nationally insured treatment under the NHIP. Participants had to be at least 20 years of age, be diagnosed with a nicotine dependence score of greater than or equal to 5 points on the Tobacco Dependence Screener (TDS) [4], smoke cigarettes with a Brinkman index at least 200 (only for those not younger than 35 years), wish to quit smoking, own their iPhone or Android smartphone, and provide their written consent. These criteria were the same as those used to diagnose nicotine dependence under NHIP in Japan. We excluded the participants who seemed to find it difficult to use their smartphones as instructed or if their

physicians ruled that they could not complete the study owing to severe mental illness.

The registration period was from March 2016 to March 2017. We followed up on the participants for 52 weeks from the time of registration. In total, 5 institutions were involved in the study, including a university hospital, 3 general hospitals, and a primary care clinic, all of which were located near the Tokyo metropolitan area.

"CureApp Smoking Cessation" App Software

The CASC was developed by CureApp Inc (Tokyo, Japan) in collaboration with the Division of Pulmonary Medicine, Department of Medicine, Keio University School of Medicine. Smoking cessation specialists at the institution supervised the development of the app's content. It was compatible with iOS and Android smartphones and met the software inspection criteria and security requirements of Apple's App Store and Google Play. In outpatient clinics, physicians provided participants the app prescription codes. The participants downloaded the app from the App Store or Google Play using these codes, and then input their information including their age, sex, years of smoking, number of cigarettes smoked per day, medication (ie, varenicline or nicotine patch), and motivation and self-confidence regarding smoking cessation onto their smartphones. This information was securely stored on the cloud system, and our AI system created appropriate and personalized counseling advice for each participant to support their smoking cessation in accordance with the national guidelines [5]. The text of the app was presented in English for convenience (Figure 1).

Physicians could refer to the participants' data on the website through the cloud system and offer advice to the participants during their clinic visits based on the data, as described below. The app system comprised 4 features to maximize the therapeutic effect of pharmacological therapies and counseling at outpatient clinics for smoking cessation as follows (Multimedia Appendix 1).

Diary of Smoking Cessation (Once a Day)

Instead of making entries in paper diaries, participants filled in the electronic diary within this app regarding cessation status, physical condition, medication use, and adverse events, if any (Figure 2).

Messages and Educational Videos to Help Users Quit Smoking

Participants received several messages every day and watched videos (1- to 3-min animations). A total of 20 videos were delivered during the pharmacotherapy period (12 weeks). The educational videos were delivered frequently in the first few weeks, after which the frequency gradually decreased. The timing of the delivery of messages and videos varied depending on the drug prescribed (varenicline or nicotine patch) and each individual participant's life cycle. Participants could view these videos even after week 12 or anytime they wished (Figure 3).

Counseling Chat Sessions Between the Users and the Artificial Intelligence Nurse

Whenever the participants experienced cravings or withdrawal symptoms, they could tap "Call" and send a message to an AI nurse. The AI nurse would immediately reply and provide personalized advice on how to deal with the symptoms such as a chatbot. This nurse also provided encouraging messages for smoking cessation to the participants at appropriate times (Figure 4).

Advice for Physicians

The CASC displayed recommendations to physicians on the website screen, which could be used by the physicians to offer appropriate advice and counseling support to the participants based on the national guidelines (Figure 5) [5].

Figure 1. Scheme of the CureApp Smoking Cessation support program system for participants with nicotine dependence.

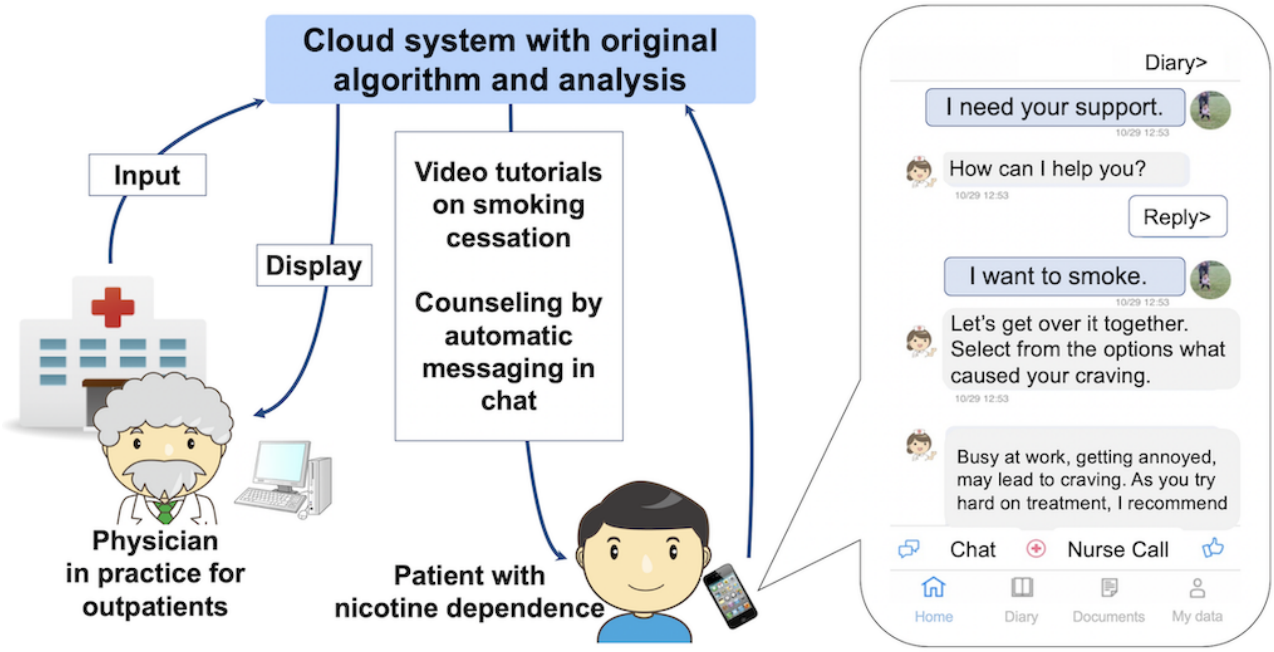


Figure 2. Example of the screen display for the smoking cessation diary feature.

Data Collection

We collected baseline profiles of each participant including their age, sex, height, weight, Brinkman index, TDS score, and Fagerström Test for Nicotine Dependence (FTND) score [18] on their first visit. Smoking status, medication use (varenicline or nicotine patch), exhaled carbon monoxide (CO) concentration, as well as their scores on the Mood and Physical Symptoms Scale (MPSS) [19], 12-item French version of the Tobacco Craving Questionnaire (FTCQ-12) [20], and Kano Test for Social Nicotine Dependence (KTSND) [21] were collected at every visit (weeks 2, 4, 8, and 12). The MPSS is a validated questionnaire that assesses the severity of participants' withdrawal symptoms and the strength and frequency of their urges to smoke, with scores ranging from 5 to 35. The FTCQ-12 is a valid 12-item self-report instrument that assesses tobacco cravings based on 4 factors: emotionality, expectancy, compulsivity, and purposefulness. The FTCQ-12 general craving score is derived by summing all items and dividing the total by 12, yielding a score ranging from 1 to 7. The KTSND is a method of evaluating the psychological aspects of smoking and

comprises 10 questions with a total score range of 0 to 30. In these 3 scoring systems, higher scores indicate more severe symptoms and features. We also gathered the app usage status such as the number of days the participants updated their diaries, the number of behaviors they actually modified as a result of counseling, and the number of educational videos they viewed completely. The primary endpoint was the CAR from weeks 9 to 24. Secondary endpoints were CARs from weeks 9 to 12 and 9 to 52 and changes in MPSS, FTCQ-12, and KTSND scores.

Confirmation of Continuous Smoking Cessation

Self-reported continuous abstinence was confirmed if participants recorded a breath CO concentration of less than 8 ppm during their clinic visits until week 12. Confirmation of abstinence at weeks 24 and 52 was determined by face-to-face or telephone interviews.

Statistical Analysis

We analyzed the primary and secondary endpoints based on the full analysis set (FAS) and compared our data with those of the national survey of Japan [22]. CI for the binomial proportion was computed using the Agresti-Coull method.

Figure 3. Example of the screen display for the lecture and educative videos feature.

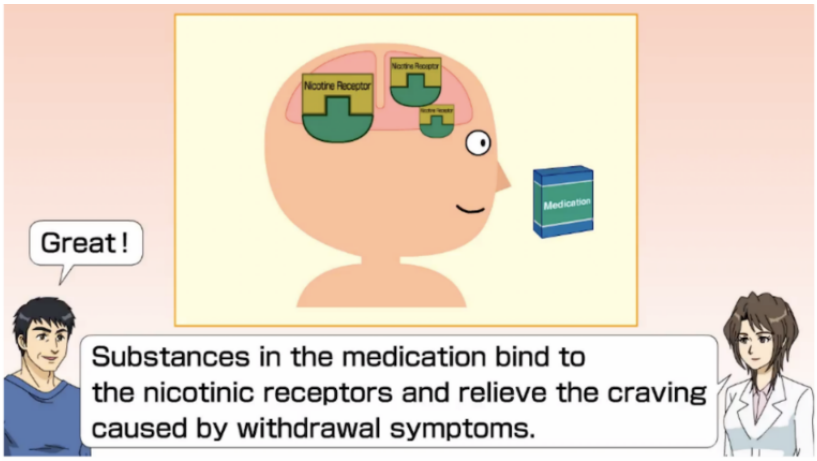


Figure 4. Example of the screen display for the counseling chat feature between the users and the artificial intelligence nurse.

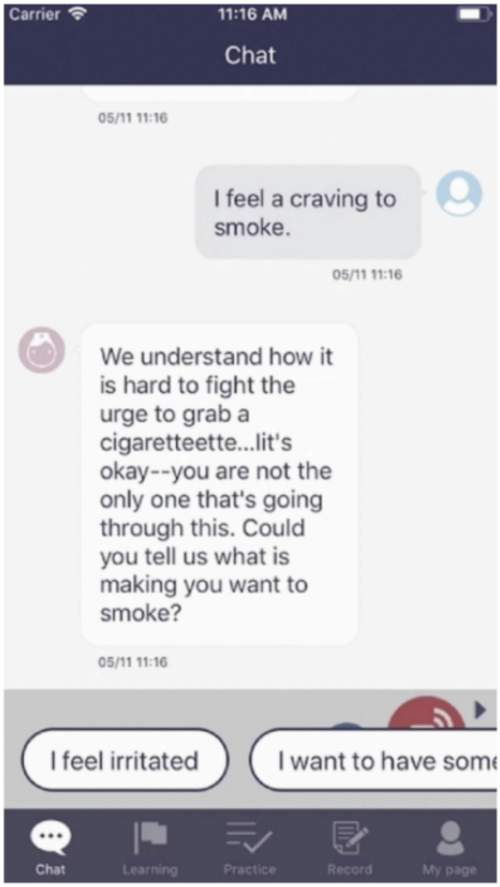
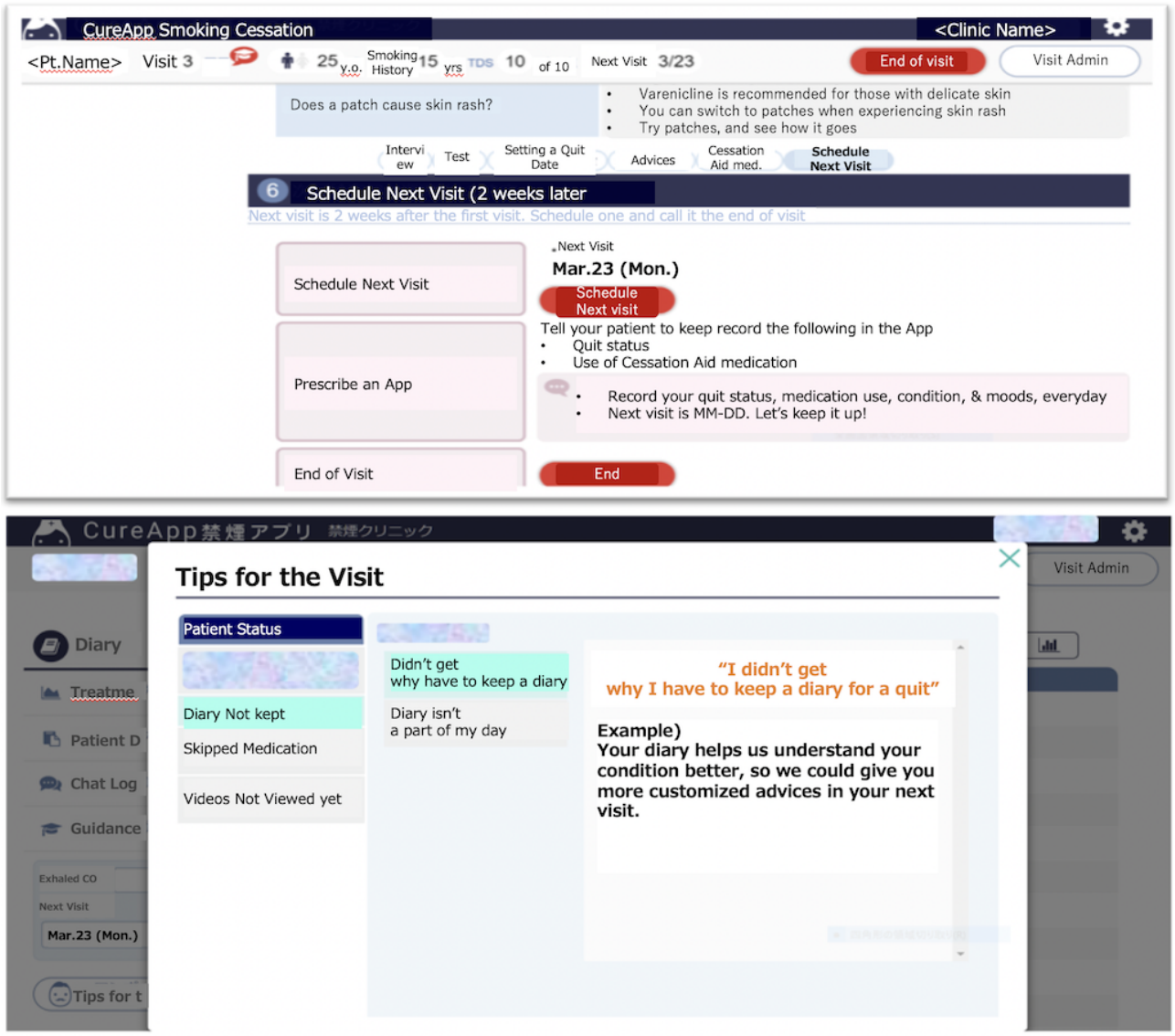


Figure 5. Example of the screen display feature offering advice for the physicians.



Results

Recruitment and Baseline Characteristics

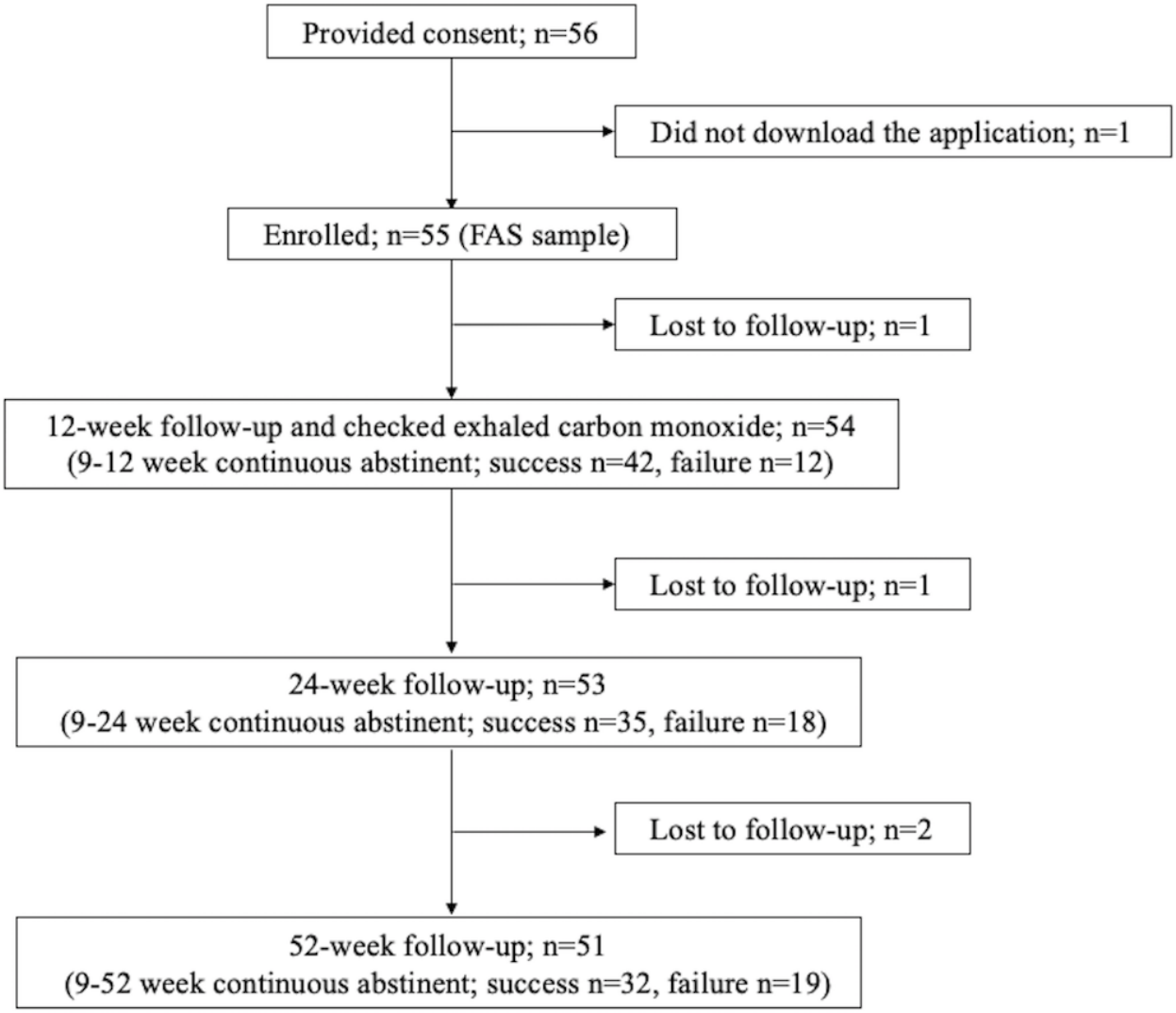
We recruited 56 participants with nicotine dependence who met the inclusion criteria. Of these participants, because 1 did not download the app, the FAS sample comprised 55 participants. From this FAS sample, 1 participant dropped out during weeks 0 to 12, 1 dropped out during weeks 12 to 24, and 2 dropped out during weeks 24 to 52. These dropouts were treated as failures. A flowchart of the study is depicted in Figure 6.

The baseline characteristics of the study participants are depicted in Table 1. The participants' mean age was 43.3 years and 71% of them were male. Some of the study participants had comorbidities, such as chronic obstructive pulmonary disease (13%), hypertension (7%), cancer (7%), diabetes mellitus (5%), and dyslipidemia (5%), and these comorbidities were all under control and in stable conditions. The mean of Brinkman index was 486, baseline exhaled CO concentration was 20.6 ppm, and TDS was 7.9. Varenicline was administered to 87% (48/55) of the participants, and 13% (7/55) received nicotine patches. Of the participants, 1 used both varenicline and a nicotine patch sequentially, and another participant used neither varenicline nor a nicotine patch. The mean FTND, the MPSS total score, the FTCQ-12 general craving score, and the KTSND score at the first visit were 5.5, 20.6, 3.2, and 16.6, respectively.

Evaluation Outcomes

Of the FAS sample, 35 participants succeeded in continuously abstaining from smoking until the 24th week. The CAR from weeks 9 to 24 was 64% (35/55, 95% CI 51%-76%); see Table 2). The CARs from weeks 9 to 12 and 9 to 52 were 76% (42/55, 95% CI 65%-88%) and 58% (32/55, 95% CI 46%-71%), respectively.

Figure 6. Flowchart depicting the selection of study participants.



The total score of the MPSS at week 12 was 13.7, which decreased by 6.4 points from the baseline. The mean MPSS scores related to “depressed,” “irritable,” “restless,” “hungry,” and “poor concentration” at week 12 were 1.3, 1.4, 1.3, 2.2, and 1.7, respectively. Similarly, the mean score of the FTCQ-12 (general craving score) at week 12 was 2.5, which had reduced by 0.6 points from the baseline. The mean FTCQ-12 scores related to “emotionality” and “compulsivity” at week 12 were 1.2 and 1.6, respectively. They had decreased from the baseline scores. On the other hand, the mean FTCQ-12 score related to “expectancy” barely changed from baseline, whereas the mean FTCQ-12 score related to “purposefulness” at week 12 was 4.7 on average and had increased from the baseline. The KTSND score at week 12 was 10.5 and had decreased by 6.7 points from the baseline.

App Usage

There were several indicators of the app’s usage status (Table 3). On average, the participants tapped “Like!” 26.5 times from weeks 0 to 12, which implied that they agreed with the advice provided by the AI nurse. They tapped “Call,” which summoned the AI nurse when participants had smoking impulses or side effects, an average of 1.7 times from weeks 0 to 12. They mean number of days on which participants made diary entries from weeks 0 to 12 was 56.1, and 45.5% (25/55) of the participants updated their diaries every day. The mean number of educational videos viewed from start to finish by the participants was 12.6.

Adverse Events

In total, 3 participants, all of whom took varenicline reported 3 adverse events. Skin eruptions appeared in 1 participant and mild nausea appeared in another. Both of these participants could continue to take varenicline. Depressive symptoms appeared in 1 participant in the 9th week, and the physician decided to stop the varenicline treatment for this participant. On the basis of their judgment, the physicians concluded that these adverse events were caused by the drugs rather than by the app.

Table 1. Baseline characteristics of study participants (N=55).

Participant characteristics	Statistics
Age (years), mean (SD)	43.3 (10.5)
Males, n (%)	39 (71)
Body weight (kg), mean (SD)	63.1 (10.7)
Body mass index (kg/cm^2), mean (SD)	23.3 (3.4)
Years of smoking, mean (SD)	23.6 (10.8)
Number of cigarettes smoked per day, mean (SD)	19.9 (7.3)
Brinkman index, mean (SD)	486 (329)
Number of lifelong cessation attempts, mean (SD)	1.1 (1.2)
Exhaled CO[a] concentration (ppm), mean (SD)	20.6 (15.9)
FTND[b], mean (SD)	5.5 (1.9)
TDS[c], mean (SD)	7.9 (1.4)
KTSND[d], mean (SD)	16.6 (4.3)
MPSS[e] total, mean (SD)	20.6 (4.7)
FTCQ-12[f] general craving score, mean (SD)	3.2 (0.7)
Varenicline, n (%)	48 (87)
Nicotine patch, n (%)	7 (13)
No pharmacotherapy, n (%)	1 (2)

[a]CO: carbon monoxide.
[b]FTND: Fagerström Test for Nicotine Dependence.
[c]TDS: Tobacco Dependence Screener.
[d]KTSND: Kano Test for Social Nicotine Dependence.
[e]MPSS: Mood and Physical Symptoms Scale.
[f]FTCQ-12: 12-item French version of the Tobacco Craving Questionnaire.

Table 2. Primary and secondary endpoints.

Endpoints	Statistics
CAR[a] from weeks 9 to 12, % (95% CI)	76 (65-88)
CAR from weeks 9 to 24 (primary endpoint), % (95% CI)	64 (51-76)
CAR from weeks 9 to 52, % (95% CI)	58 (46-71)
MPSS[b] total at week 12, mean (SD)	13.7 (3.5)
ΔMPSS total from baseline to week 12, mean (SD)	–6.4 (5.8)
FTCQ-12[c] general craving score at week 12, mean (SD)	2.5 (1.1)
ΔFTCQ-12 general craving score from baseline to week 12, mean (SD)	–0.6 (1.5)
KTSND[d] at week 12, mean (SD)	10.5 (5.9)
ΔKTSND from baseline to week 12, mean (SD)	–6.7 (5.2)

[a]CAR: continuous abstinence rate.
[b]MPSS: Mood and Physical Symptoms Scale.
[c]FTCQ-12: 12-item French version of the Tobacco Craving Questionnaire.
[d]KTSND: Kano Test for Social Nicotine Dependence.

Table 3. App usage from weeks 0 to 12.

Engagement with the app	Mean (SD)
Number of times “Like!” was tapped	26.5 (63.8)
Number of times “Call” was tapped	1.7 (2.4)
Number of days a diary entry was made	56.1 (31.3)
Number of educational videos viewed	12.6 (6.8)
Practices for environmental improvement	67.5 (135.6)
Behaviors actually modified	74.3 (158.5)
Practices for compensatory behaviors	97.4 (374.3)
Practices for self-assertiveness	8.2 (17.3)

Discussion

Principal Findings

This was the first prospective study to evaluate whether a novel smartphone app, CASC, could achieve long-term CARs from weeks 9 to 24 in nicotine-dependent patients. Our results showed that more than three-fifths of participants succeeded in quitting smoking at week 24. The combination of CASC and conventional pharmacotherapy resulted in more than three-fourths of all participants maintaining their CARs from weeks 9 to 12, and 76% (32/42) of quitters at week 12 continued to abstain from smoking at week 52. This combination treatment also decreased the participants’ MPSS, FTCQ-12 (general craving), and KTSND scores by week 12 of the study period. Except 1 participant who dropped out during the 12 weeks of the treatment period, all others remained, and many participants tapped “Like!” in response to the advice they received from the AI nurse and updated the diary daily, indicating that the usage of CASC was feasible.

The CAR from weeks 9 to 24 obtained in this study appeared to be substantially better than those of previous studies and of the national survey (Figure 7) [22-24]. According to the national survey data and the varenicline phase 3 trial in Japan, the CAR from weeks 9 to 24 were 40.81% (1039/2546) and 37.7% (49/130), respectively [22,23]. The CARs from weeks 9 to 12, 9 to 24, and 9 to 52 were 76% (42/55, 95% CI 65%-88%), 64% (35/55, 95% CI 51%-76%), and 58% (32/55, 95% CI 46%-71%), respectively.

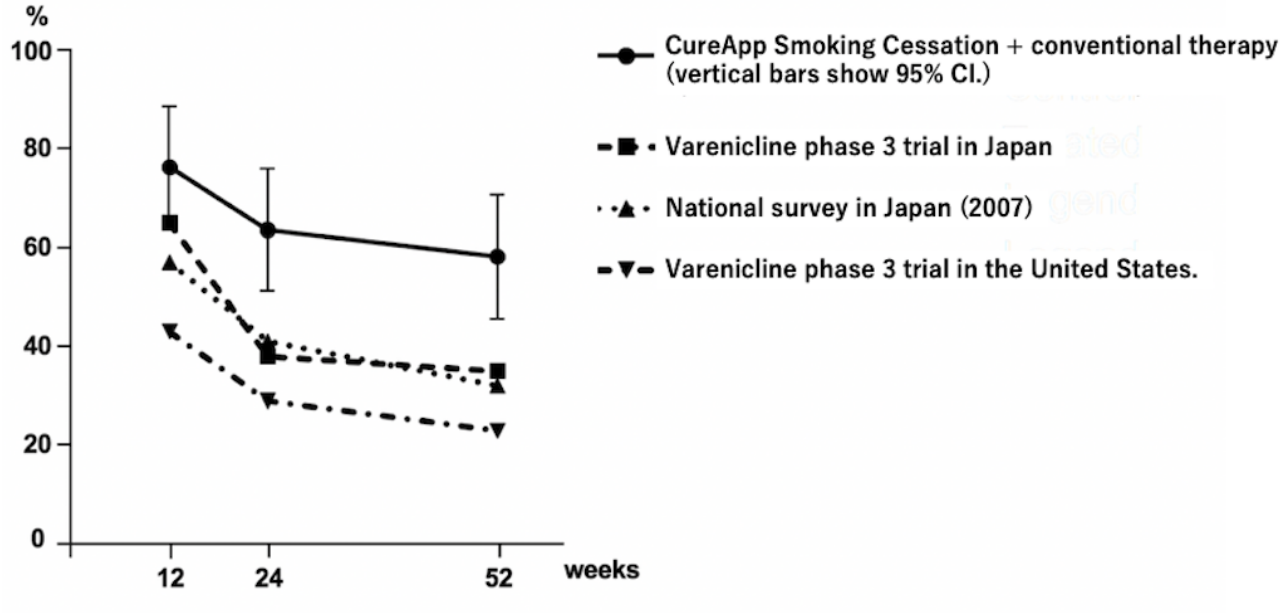
There are several reasons why our smartphone app could improve long-term CARs. First, it was designed to promote understanding of nicotine dependence through several evidence-based approaches. For example, educational videos included the following treatment programs: education about nicotine dependence, explanation of nicotine withdrawal symptoms, and behavioral therapy. Consequently, the quitters at week 12 did not relapse frequently because they might have understood that smoking behavior was caused by the dependence on nicotine and because they were aware of the characteristics of addictive behaviors. In fact, the KTSND score, which was considered as one of the indices reflecting the understanding of nicotine dependence [21], decreased more than 6 points from weeks 0 to 12 (Table 2).

Second, continuous encouragement and timely advice by an AI nurse might prevent dropouts during the treatment. All but 1 participant (92%) completed the outpatient smoking cessation treatment for 12 weeks in our study, whereas only 30% of the participants completed the treatment that was a part of the domestic survey data [6]. Notably, the existence of an AI nurse might keep participants from dropping out because many studies have shown that smoking cessation success rates improve with interventions not only by physicians but also by nurses and pharmacists [25-30]. In a single fight, patients easily abandon smoking cessation because of cravings and withdrawal symptoms. Besides, maintaining one’s motivation to make multiple attempts considering several aspects was reported to be important in smoking cessation [31]. These studies indicated the importance of encouragement and support during the pharmacotherapy period. Therefore, individualized support and encouragement for each patient is essential for an app to maximize the effect of the medication and counseling. The AI nurse advised the participants of this study about how to deal with cravings for smoking as if it were an expert partner who was accompanying them while they were trying to quit smoking. In this study, not only the physician but also the CASC may have functioned as a supporter of and advisor to participants.

Third, adequate advice through educational videos to prevent relapse might help the quitters continue to abstain from smoking easily. In fact, the MPSS total score and FTCQ-12 general craving score decreased from baseline to week 12 (Table 2). The decrease in the urge to smoke might contribute to the continuation of smoking cessation from weeks 12 to 24 and 52.

Finally, the app might prevent treatment failure because of the adverse effects of the varenicline or nicotine patch. It was decided that only 1 participant would stop varenicline treatment owing to its adverse effects. On our app, the alarm feature reminding the participants to take their medicines and an explanation of the side effects were different for participants who received varenicline and for those who received the nicotine patch. Participants could learn how to deal with the adverse effects by watching the educational videos and reading the messages sent by the AI nurse.

Figure 7. Continuous abstinence rates from weeks 9 to 12, 9 to 24, and 9 to 52.



Strengths and Limitations

The strengths of this study are as follows: (1) this study is the first to measure the long-term effectiveness and feasibility of smartphone app intervention for more than 24 weeks; (2) the features of our app were designed based on the national treatment guidelines [5], and it was different from other apps that aimed merely to record the participants' trajectories; and (3) our results also highlighted the additional effect of the app on pharmacotherapy. The limitations of this study are as follows: (1) because the sample size was small, the study participants might not have represented all Japanese smokers; (2) participating institutes were restricted only to the area around Tokyo in Japan; (3) the education level of the study participants was not surveyed, which might have influenced the results; and (4) we did not confirm abstinence with exhaled CO concentration at weeks 24 and 52. In previous intervention studies, confirmation of self-assessment through telephone surveys was a standard when the study design was minimally invasive [32]. In our study, all 35 successful quitters at week 24 had shown normal exhaled CO levels during their week 12 visits, indicating the reliability of the self-reports in our sample.

Conclusions

In conclusion, a novel smartphone app for smoking cessation, CACS, benefitted 63.6% of the users who maintained their CARs from weeks 9 to 24. Our 95% CI ranges of CARs were superior to those of the historical cohorts [22-24] (Figure 7). This app might be a useful tool for improving long-term CAR combined with a standard smoking cessation treatment program. Currently, a randomized placebo app–controlled trial is ongoing to evaluate the efficacy of the app for nicotine dependence (UMIN000031589). We will continue to improve this app further to make it an available and essential tool in smoking cessation therapy.

Acknowledgments

The authors thank all the participants of this project. The authors also express their gratitude to Tomoyasu Muto for his assistance and Dr Marc Moss for her outstanding support. They also thank Editage for English language editing.

Conflicts of Interest

KS is a founder and shareholder of CureApp Inc and a patent holder of the CASC smartphone app. SS is an employee and shareholder of CureApp Inc and a patent holder of the CASC smartphone app. AN received consulting fees from CureApp Inc. HT and KF have received honoraria from CureApp Inc. The other authors have no conflicts of interest to declare.

Multimedia Appendix 1

How the app appears on a patient's smartphone screen.

[MOV File, 36MB - mhealth_v7i2e12694_app1.mov]

References

1. US Department of Health & Human Services: Surgeon General. 2014. The Health Consequences of Smoking—50 Years of Progress URL: https://www.surgeongeneral.gov/library/reports/50-years-of-progress/full-report.pdf [accessed 2019-02-09] [WebCite Cache ID 764I74FQG]
2. Centers for Disease Control and Prevention. 2010. How tobacco smoke causes disease: what it means to you URL: https://www.cdc.gov/tobacco/data_statistics/sgr/2010/consumer_booklet/index.htm [accessed 2019-02-09] [WebCite Cache ID 764I7zKeX]
3. Hughes JR, Keely J, Naud S. Shape of the relapse curve and long-term abstinence among untreated smokers. Addiction 2004 Jan;99(1):29-38 [FREE Full text] [doi: 10.1111/j.1360-0443.2004.00540.x] [Medline: 14678060]
4. Kawakami N, Takatsuka N, Inaba S, Shimizu H. Development of a screening questionnaire for tobacco/nicotine dependence according to ICD-10, DSM-III-R, and DSM-IV. Addict Behav 1999;24(2):155-166. [doi: 10.1016/S0306-4603(98)00127-0] [Medline: 10336098]
5. The Japanese Circulation Society. The Standard Procedure Manual for Smoking Cessation, 6th edition, 2014 URL: http://www.j-circ.or.jp/kinen/anti_smoke_std/pdf/anti_smoke_std_rev6.pdf [accessed 2019-02-09] [WebCite Cache ID 764IGK4S6]
6. Ministry of Health, Labour and Welfare. 2017. Report on the efficacy of smoking cessation treatment based on nicotine dependency management fee URL: https://www.mhlw.go.jp/file/05-Shingikai-12404000-Hokenkyoku-Iryouka/0000192293.pdf [accessed 2019-02-09] [WebCite Cache ID 764IKHmJP]
7. Drovandi AD, Chen CC, Glass BD. Adverse effects cause varenicline discontinuation: a meta-analysis. Curr Drug Saf 2016;11(1):78-85. [Medline: 26412667]
8. Song F, Maskrey V, Blyth A, Brown TJ, Barton GR, Aveyard P, et al. Differences in longer-term smoking abstinence after treatment by specialist or nonspecialist advisors: secondary analysis of data from a relapse prevention trial. Nicotine Tob Res 2016 May;18(5):1061-1066 [FREE Full text] [doi: 10.1093/ntr/ntv148] [Medline: 26152558]
9. McDermott MS, Beard E, Brose LS, West R, McEwen A. Factors associated with differences in quit rates between. Nicotine Tob Res 2013 Jul;15(7):1239-1247. [doi: 10.1093/ntr/nts262] [Medline: 23239841]
10. Boccio M, Sanna RS, Adams SR, Goler NC, Brown SD, Neugebauer RS, et al. Telephone-based coaching. Am J Health Promot 2017 Dec;31(2):136-142 [FREE Full text] [doi: 10.4278/ajhp.140821-QUAN-424] [Medline: 26559720]
11. Puckett M, Neri A, Thompson T, Underwood JM, Momin B, Kahende J, Centers for Disease Control and Prevention. Tobacco cessation among users of telephone and web-based interventions--four states, 2011-2012. MMWR Morb Mortal Wkly Rep 2015 Jan 2;63(51):1217-1221 [FREE Full text] [Medline: 25551593]
12. Whittaker R, McRobbie H, Bullen C, Rodgers A, Gu Y. Mobile phone-based interventions for smoking cessation. Cochrane Database Syst Rev 2016 Apr 10;4:CD006611. [doi: 10.1002/14651858.CD006611.pub4] [Medline: 27060875]
13. Pechmann C, Delucchi K, Lakon CM, Prochaska JJ. Randomised controlled trial evaluation of Tweet2Quit: a social network quit-smoking intervention. Tob Control 2016 Feb 29 [FREE Full text] [doi: 10.1136/tobaccocontrol-2015-052768] [Medline: 26928205]
14. Müssener U, Bendtsen M, Karlsson N, White IR, McCambridge J, Bendtsen P. Effectiveness of short message service text-based smoking cessation intervention among university students: a randomized clinical trial. JAMA Intern Med 2016 Mar;176(3):321-328. [doi: 10.1001/jamainternmed.2015.8260] [Medline: 26903176]
15. Ubhi H, Michie S, Kotz D, Wong W, West R. A mobile app to aid smoking cessation: preliminary evaluation of SmokeFree28. J Med Internet Res 2015 Jan 16;17(1):e17 [FREE Full text] [doi: 10.2196/jmir.3479] [Medline: 25596170]
16. Businelle M, Ma P, Kendzor D, Frank S, Vidrine D, Wetter D. An Ecological Momentary Intervention for Smoking Cessation: Evaluation of Feasibility and Effectiveness. J Med Internet Res 2016 Dec 12;18(12):e321 [FREE Full text] [doi: 10.2196/jmir.6058] [Medline: 27956375]
17. Iacoviello B, Steinerman J, Klein D, Silver T, Berger A, Luo S, et al. Clickotine, A Personalized Smartphone App for Smoking Cessation: Initial Evaluation. JMIR Mhealth Uhealth 2017 Apr 25;5(4):e56 [FREE Full text] [doi: 10.2196/mhealth.7226] [Medline: 28442453]
18. Heatherton TF, Kozlowski LT, Frecker RC, Fagerström KO. The Fagerström test for nicotine dependence: a revision of the Fagerstrom Tolerance Questionnaire. Br J Addict 1991 Sep;86(9):1119-1127. [doi: 10.1111/j.1360-0443.1991.tb01879.x] [Medline: 1932883]
19. West R, Hajek P. Evaluation of the mood and physical symptoms scale (MPSS) to assess cigarette withdrawal. Psychopharmacology (Berl) 2004 Dec;177(1-2):195-199. [doi: 10.1007/s00213-004-1923-6] [Medline: 15179542]
20. Berlin I, Singleton EG, Heishman SJ. Validity of the 12-item French version of the Tobacco Craving Questionnaire in treatment-seeking smokers. Nicotine Tob Res 2010 May;12(5):500-507 [FREE Full text] [doi: 10.1093/ntr/ntq039] [Medline: 20335281]
21. Otani T, Yoshii C, Kano M, Kitada M, Inagaki K, Kurioka N, et al. Validity and reliability of Kano Test for Social Nicotine Dependence. Ann Epidemiol 2009 Nov;19(11):815-822. [doi: 10.1016/j.annepidem.2009.03.016] [Medline: 19457683]
22. Ministry of Health, Labour and Welfare. Survey report on the actual status of smoking cessation success rate for smoking cessation outpatient clinic URL: https://www.mhlw.go.jp/shingi/2007/04/dl/s0418-3d.pdf [accessed 2019-02-09] [WebCite Cache ID 764IYjxbX]
23. Nakamura M, Oshima A, Fujimoto Y, Maruyama N, Ishibashi T, Reeves KR. Efficacy and tolerability of varenicline, an alpha4beta2 nicotinic acetylcholine receptor partial agonist, in a 12-week, randomized, placebo-controlled, dose-response

study with 40-week follow-up for smoking cessation in Japanese smokers. Clin Ther 2007 Jun;29(6):1040-1056. [doi: 10.1016/j.clinthera.2007.06.012] [Medline: 17692720]
24. Gonzales D, Rennard SI, Nides M, Oncken C, Azoulay S, Billing CB, Varenicline Phase 3 Study Group. Varenicline, an alpha4beta2 nicotinic acetylcholine receptor partial agonist, vs sustained-release bupropion and placebo for smoking cessation: a randomized controlled trial. J Am Med Assoc 2006 Jul 5;296(1):47-55. [doi: 10.1001/jama.296.1.47] [Medline: 16820546]
25. Rice V, Heath L, Livingstone-Banks J, Hartmann-Boyce J. Cochrane Database of Systematic Reviews. 2013. Nursing interventions for smoking cessation URL: http://cochranelibrary-wiley.com/doi/10.1002/14651858.CD001188.pub5/epdf [accessed 2019-02-09] [WebCite Cache ID 764IaX2Qk]
26. Armero AM, Hernandez MA, Perez-Vicente S, Martinez-Martinez F. Pharmaceutical care in smoking cessation. Patient Prefer Adherence 2015;9:209-215 [FREE Full text] [doi: 10.2147/PPA.S67707] [Medline: 25678779]
27. Cantor SB, Deshmukh AA, Luca NS, Nogueras-González GM, Rajan T, Prokhorov AV. Cost-effectiveness analysis of smoking-cessation counseling training for physicians and pharmacists. Addict Behav 2015 Jun;45:79-86 [FREE Full text] [doi: 10.1016/j.addbeh.2015.01.004] [Medline: 25644592]
28. Saba M, Diep J, Bittoun R, Saini B. Provision of smoking cessation services in Australian community pharmacies: a simulated patient study. Int J Clin Pharm 2014 Jun;36(3):604-614. [doi: 10.1007/s11096-014-9944-7] [Medline: 24718946]
29. Tremblay M, Comtois D, O'Loughlin J. Pharmacists' smoking cessation counseling practices: a comparison between 2005 and 2010. Nicotine Tob Res 2013 Dec;15(12):2114-2119. [doi: 10.1093/ntr/ntt118] [Medline: 23943845]
30. Khan N, Anderson JR, Du J, Tinker D, Bachyrycz AM, Namdar R. Smoking cessation and its predictors: results from a community-based pharmacy tobacco cessation program in New Mexico. Ann Pharmacother 2012 Sep;46(9):1198-1204. [doi: 10.1345/aph.1P146] [Medline: 22911338]
31. Hughes JR. Motivating and helping smokers to stop smoking. J Gen Intern Med 2003 Dec;18(12):1053-1057 [FREE Full text] [Medline: 14687265]
32. Stead L, Hartmann-Boyce J, Perera R. Cochrane Database of Systematic Reviews. Telephone counseling for smoking cessation URL: http://cochranelibrary-wiley.com/doi/10.1002/14651858.CD002850.pub3/epdf [accessed 2019-02-09] [WebCite Cache ID 764Ie1L0s]

Abbreviations

AI: artificial intelligence
CAR: continuous abstinence rate
CASC: CureApp Smoking Cessation
CO: carbon monoxide
FAS: full analysis set
FTCQ-12: 12-item French version of the Tobacco Craving Questionnaire
FTND: Fagerström Test for Nicotine Dependence
KTSND: Kano Test for Social Nicotine Dependence
MPSS: Mood and Physical Symptoms Scale
NHIP: National Health Insurance Program
TDS: Tobacco Dependence Screener
UMIN: University Hospital Medical Information Network

Edited by G Eysenbach; submitted 05.11.18; peer-reviewed by J Xu, Y Zhang; comments to author 14.12.18; revised version received 30.12.18; accepted 07.01.19; published 19.02.19.

Please cite as:
Masaki K, Tateno H, Kameyama N, Morino E, Watanabe R, Sekine K, Ono T, Satake K, Suzuki S, Nomura A, Betsuyaku T, Fukunaga K
Impact of a Novel Smartphone App (CureApp Smoking Cessation) on Nicotine Dependence: Prospective Single-Arm Interventional Pilot Study
JMIR Mhealth Uhealth 2019;7(2):e12694
URL: http://mhealth.jmir.org/2019/2/e12694/
doi:10.2196/12694
PMID:30777848

Original Paper

Use of Weight-Management Mobile Phone Apps in Saudi Arabia: A Web-Based Survey

Ghadeer S Aljuraiban[1], PhD

Department of Community Health Sciences, College of Applied Medical Sciences, King Saud University, Riyadh, Saudi Arabia

Corresponding Author:
Ghadeer S Aljuraiban, PhD
Department of Community Health Sciences
College of Applied Medical Sciences
King Saud University
Turki Alawwal
Riyadh, 145111
Saudi Arabia
Phone: 966 14651122
Email: galjuraiban@ksu.edu.sa

Abstract

Background: In recent years, the use of mobile phone weight-management apps has increased significantly. Weight-management apps have been found effective in promoting health and managing weight. However, data on user perception and on barriers to app usage are scarce.

Objective: This study aimed to investigate the use of weight-management apps and barriers to use as well as reasons for discontinuing use in a sample of mobile phone users in Saudi Arabia.

Methods: Mobile phone users aged 18 years and above from the general public in Saudi Arabia completed a Web-based survey. The survey included questions on weight-management app usage patterns, user perceptions concerning weight management, efficacy of weight-management apps, and reasons for discontinuing use. Participants were classified into normal weight (body mass index [BMI]: 18.5 to 24.9 kg/m^2) and overweight or obese (BMI: ≥25.0 kg/m^2).

Results: The survey included 1191 participants; 513 of them used weight-management apps. More overweight or obese respondents used these apps compared with normal weight respondents (319/513, 62.2% vs 194/513, 37.8%, respectively). App features that overweight or obese users were most interested in were mainly the possibility to be monitored by a specialist and barcode identification of calorie content, whereas normal weight users mostly preferred availability of nutrition information of food items. Reasons for discontinuing use among overweight or obese respondents were mainly that monitoring by a specialist was not offered (80/236, 33.9%) and the app was not in the local language (48/236, 20.3%). Among normal weight users, the main reason for noncontinuance was the app language (45/144, 31.3%) and difficulty of use (30/144, 20.8%).

Conclusions: To better address the needs of both normal weight and overweight or obese adults, improved app designs that offer monitoring by a specialist are needed. Developers may consider ways of overcoming barriers to use, such as language, by developing local language apps, which can improve the efficacy of such apps and help spread their use.

(JMIR Mhealth Uhealth 2019;7(2):e12692) doi:10.2196/12692

KEYWORDS

lifestyle; mobile app; weight loss; mobile phone

Introduction

Background

Long-term analyses of trends in body mass index (BMI) levels show that obesity increased globally between 1980 and 2008 [1]. In 2013, a national survey reported that the prevalence of obesity in Saudi adults was 29% and was higher in women compared with men (33.5% vs 24.1%, respectively) [2]; by 2022, prevalence is expected to increase to 40% for men and 78% for women [3]. Several interventions including lifestyle and dietary modifications have been shown to be effective in the management of obesity; however, limitations such as financial cost and accessibility may restrict their use [4,5].

A recent, encouraging development is the marked increase in the number of mobile phone apps for weight loss or weight management [6-9]. According to the 2018 Global Mobile Market Report, Saudi Arabia has one of the highest mobile phone penetration rates in the world [10]. In 2017, Saudi Arabia had 21 million mobile phone users (84% of total adults); this is expected to increase to nearly 24 million in 2022 [11]. Mobile phone apps that focus on health have achieved worldwide popularity; over 325,000 health, fitness, and medical apps were available for download on all major app stores in 2017 [12].

With the projected epidemic in increasing obesity, frequent recording of both food intake and physical activity with weight-management apps may be a useful, potentially cost-effective and conveniently available method for modifying health-related behaviors [13]. A systematic review and meta-analysis of 12 randomized controlled trials and case-control studies showed that use of weight-management apps affected BMI by a net reduction of 0.4 kg/m^2 more than control groups (ie, traditional interventions or intensive counseling) [14].

Rational and Aim

Development and use of weight-management apps is still in an early phase, and sample populations in various countries have reported barriers such as limited functionality, lack of input options for tracking data, no evidence-based guidelines, and improper app classification (ie, only one-fifth of the apps classified as *health and fitness* or *medical* actually facilitated behavioral or physical activity; many simply offered health information) [15]. There is a paucity of evidence regarding usage and barriers of use in Saudi Arabia. One qualitative study restricted to Saudi women only reported that the available weight-management apps offered diet programs that lacked locally available foods and suggested physical activities that were not applicable to their environment [16].

Despite growing research on weight-management apps, evidence on user perspective—especially among users in Saudi Arabia, whose perspective varies vastly from that of the West—has been little explored. To the author's knowledge, this is the first survey aimed to identify the sociodemographic characteristics of weight-management app users, app usage patterns, user perceptions, efficacy, and reasons for continuing or discontinuing use of a weight-management phone app in Saudi Arabia.

Methods

Design and Sample

From May to July 2018, mobile phone users aged 18 years and older from the general public in the city of Riyadh filled out an open Web-based survey. The survey was designed using *Microsoft* Forms (Microsoft Corp). Authentication cookies prevented multiple entries from the same individual. The survey adhered to advanced Web-based survey methodology, such as internet protocol address verification. The multifaceted recruitment strategy included university portals, social media, newsletters, and posters, all citing a quick response code for declaring interest. Those outside of the internet-based survey were not contacted. Of the 1302 visitors to the survey link, 1191 agreed to participate; all of the participants resided in Riyadh. Of these 1191, a total of 513 reported having previously downloaded an app to track their weight and completed all survey items. No personal identifying information was collected. The King Saud University Institutional Review Board approved all study procedures (reference # KSU-HE-18-1006).

Survey Items

The survey comprised 28 items in the following domains: (1) sociodemographic characteristics (gender, age, employment, education, and income), (2) general health questions (weight, height, physical activity, smoking habits, and medical history), (3) usage patterns (number of apps used, frequency of use, reasons for use, and desired features), (4) user conceptions and efficacy of weight-management apps (effectiveness of weight-reduction apps, effectiveness of fitness apps, and effectiveness of long-term use), and (5) reasons for discontinuing use (number of apps that the respondent no longer uses and reasons for discontinuing use).

Krebs et al previously used the survey questions in their national survey throughout the United States [17]. As the survey items on app use had no known precedent, a team of experts with expertise in questionnaire design field-tested the items among researchers. The Web-based survey administration system was also pretested and then pilot-tested live for 1 day, attracting 86 respondents. These data, however, were not included in the final analysis of this study but were analyzed to adjust the Web-based survey for errors.

The items were presented in a logical order, and consideration was taken to reduce response set bias. Each question required a response before the respondent was allowed to advance to the next screen. The Web-based survey displayed 1 item per screen, with the first screen asking respondents for their informed consent to participate in the study. Following consent, participants were allowed to begin the survey. Participation was completely voluntary. A *back* button on each screen allowed participants to edit previous answers. The average time needed to complete the survey was about 8 min.

Statistical Analysis

Statistical analyses were conducted using JMP (version 12; SAS Institute). No statistical correction procedures or weightings were applied. Categorical variables are presented as percentages, whereas continuous variables are presented as means and SDs. Participants were stratified by gender and by BMI, where participants were considered of normal weight if their BMI was 18.5 to 24.9 kg/m^2 and overweight or obese if their BMI was ≥25 kg/m^2.

Sample size was calculated based on the current population of Riyadh—the capital of Saudi Arabia has 5.5 million adults aged above 18 years [18] and 84% are mobile phone users [11]—using a confidence level of 95% and a precision of 5%. With these parameters, the minimum sample required for the analyses to have a power of 95% would be 207 individuals.

Results

Sociodemographic Characteristics and Health Status

Of the respondents, 43.07% (513/1191) reported that they had previously used an app to manage their weight; only analyses for weight-management app users are presented (Table 1). More overweight or obese respondents reported that they had previously used an app to track their weight compared with normal-weight respondents (319/513, 62.2% vs 194/513, 37.8%, respectively; Table 2). Most users were Saudi nationals (487/513, 94.9%), women (305/513, 59.5%), had a mean age of 26.6 (SD 8.1) years, and a mean BMI of 28.6 (SD 6.7) kg/m^2. Most users were also educated and had a high monthly income. The majority of the female respondents were overweight or obese (178/305, 59.5%). Most users (399/513, 77.8%) considered their general health to be good, very good, or excellent. The analyses found gender differences in lifestyle: men were more likely to smoke than women, and more women reported that they never engaged in physical activity (Multimedia Appendix 1).

Usage Patterns

Most weight-management app users used 1 to 5 apps (476/513, 92.8%; Table 2). The frequency of use was less than once a month for 34.9% (179/513) or a few times a week for 22.0% (113/513). In overweight or obese users, the most common reasons for wanting to download a weight-management app were to lose weight (134/319, 42.0%) and monitor food intake (88/319, 27.6%; Table 2). Normal-weight participants, on the other hand, downloaded a weight-management app mainly to monitor food intake (82/194, 42.3%) and track physical activity (52/194, 26.8%). The most common reasons for downloading a particular weight-management app were its rank in the app store, recommendations from friends and family (154/513, 30.0%), and social media influencers (93/513, 18.1%; Table 2).

The app features that overweight or obese users were most interested in were (1) the possibility to be monitored by a specialist (126/319, 39.5%), (2) barcode identification of calorie content (77/319, 24.1%), (3) availability of nutrition information on numerous food items (56/319, 17.6%), (4) a weekly or monthly progress report (44/319, 13.8%), and (5) constant reminders to follow a chosen diet or exercise (16/319, 5.0%). For normal-weight users, the desired features were (1) availability of nutrition information on numerous food items (67/194, 34.5%), (2) barcode identification of calorie content (54/194, 27.8%), (3) the possibility to be monitored by a specialist (31/194, 16.0%), (4) a weekly or monthly progress report (26/194, 13.4%), and (5) constant reminders to follow a chosen diet or exercise (16/194, 8.3%; Table 2). No gender differences in usage patterns were observed (Multimedia Appendix 2).

User Conceptions and Efficacy of Weight-Management Apps

Most weight-management app users agreed or strongly agreed that apps were helpful in losing weight (Table 3). More normal-weight users agreed or strongly agreed that weight-management apps had helped them lose weight compared with overweight or obese users (98/194, 50.5% vs 148/319, 46.4%; Table 3). As for apps that suggested exercise plans, more normal-weight users agreed or strongly agreed that such apps helped them lose weight compared with overweight or obese respondents (116/194, 59.8% vs 177/319, 55.5%, respectively; Table 3).

Concerning efficacy, overweight or obese users were unsure whether weight-management apps would be effective in the long term (170/319, 53.3%), whereas normal-weight users agreed or strongly agreed that they were effective (102/194, 52.6%; Table 3). No gender differences were observed (Multimedia Appendix 3).

Reasons for Discontinuing Use

Most users (380/513, 74.1%) reported that they have downloaded weight-management apps they no longer use, and the majority of these were overweight or obese (236/380, 62.1%; Table 4). Among overweight or obese users, reasons for discontinuing use included (1) monitoring by a specialist was not offered (80/236, 33.9%), (2) the app language was not the local language (48/236, 20.3%), (3) some costs of app usage were hidden (37/236, 15.7%), (4) the app was confusing to use (34/236, 14.4%), and (5) the respondent had lost interest (30/236, 12.7%). Normal-weight users stopped using an app because (1) it was not in the local language (45/144, 31.3%), (2) it was confusing to use (30/144, 20.8%), (3) it did not offer monitoring by a specialist, and (4) some costs were hidden (21/144, 14.6%). Reasons for discontinuing use were similar between men and women (Multimedia Appendix 4).

Table 1. Sociodemographic characteristics and health status of participants stratified by body mass index.

Characteristics	Body mass index		Total
	18.5 to 24.9 kg/m^2	>25 kg/m^2	
Previously downloaded an *app* to track anything related to weight, n (%)			
No	283 (59.33)	395 (55.32)	678 (56.93)
Yes	194 (40.67)	319 (44.68)	513 (43.07)
Gender, n (%)			
Female	127 (65.46)	178 (55.80)	305 (59.45)
Male	67 (34.54)	141(44.20)	208 (40.55)
Nationality, n (%)			
Non-Saudi	8 (4.12)	18 (5.64)	26 (5.07)
Saudi	186 (95.88)	301 (94.36)	487 (94.93)
Age (years), mean (SD)	24.4 (6.9)	27.7 (8.3)	—
Employment, n (%)			
In school	122 (62.89)	161 (50.47)	283 (55.17)
Not working or retired	7 (3.61)	20 (6.27)	27 (5.26)
Working full-time	65 (33.51)	138 (43.26)	203 (39.57)
Education, n (%)			
High school degree	42 (21.65)	45 (14.11)	87 (16.96)
Bachelor's degree	122 (62.89)	189 (59.25)	311 (60.62)
Graduate degree (Masters, PhD, or MD)	30 (15.46)	85 (26.65)	115 (22.42)
Household income, n (%)			
Less than 5,000 SR/month	7 (3.61)	20 (6.27)	27 (5.26)
5,100 to 10,000 SR/month	55 (28.35)	93 (29.15)	148 (28.85)
10,100 to 20,000 SR/month	76 (39.18)	105 (32.92)	181 (35.28)
>20,000 SR/month	56 (28.87)	101 (31.66)	157 (30.60)
General health status, n (%)			
Poor	2 (1.03)	14 (4.39)	16 (3.12)
Average	32 (16.49)	66 (20.69)	98 (19.10)
Good	22 (11.34)	43 (13.48)	65 (12.67)
Very good	80 (41.24)	112 (35.11)	192 (37.43)
Excellent	58 (29.90)	84 (26.33)	142 (27.68)
Frequency of exercise or physical activity for at least 15 min in the past week, n (%)			
Never	55 (28.35)	119 (37.30)	174 (33.92)
1 day	28 (14.43)	23 (7.21)	51 (9.94)
2 days	39 (20.10)	49 (15.36)	88 (17.15)
3-4 days	39 (20.10)	66 (20.69)	105 (20.47)
5-7 days	33 (17.01)	62 (19.44)	95 (18.52)
Overall nutritional status of the diet, n (%)			
Poor	22 (11.34)	80 (25.08)	102 (19.88)
Fair	39 (20.10)	77 (24.14)	116 (22.61)
Good	62 (31.96)	111 (34.80)	173 (33.72)
Very good	59 (30.41)	45 (14.11)	104 (20.27)
Excellent	12 (6.19)	6 (1.88)	18 (3.51)

Characteristics	Body mass index		Total
	18.5 to 24.9 kg/m^2	>25 kg/m^2	
Smoking, n (%)			
Yes	23 (11.86)	53 (16.61)	76 (14.81)
No	171 (88.14)	266 (83.39)	437 (85.19)
Co-morbidities, n (%)			
None	135 (69.59)	212 (66.46)	347 (67.64)
Hypercholesterolemia	9 (4.64)	8 (2.51)	17 (3.31)
Hypertension	9 (4.64)	9 (2.82)	18 (3.51)
Depression	8 (4.12)	46 (14.42)	54 (10.53)
Diabetes	21 (10.82)	21 (6.58)	42 (8.19)
Other chronic disease	12 (6.19)	23 (7.21)	35 (6.82)

Table 2. Pattern of use of weight-management apps stratified by body mass index.

Pattern of use	Body mass index		Total
	18.5-24.9 kg/m^2	≥25 kg/m^2	
Number of weight-related phone apps used, n (%)			
1-5 apps	174 (89.6)	302 (94.6)	476 (92.79)
6-10 apps	15 (7.73)	12 (3.76)	27 (5.26)
>11 apps	5 (2.58)	5 (1.57)	10 (1.95)
Frequency of weight-management app use, n (%)			
2 or more times a day	29 (14.95)	40 (12.54)	69 (13.45)
About 1 time each day	17 (8.76)	32 (10.03)	49 (9.55)
A few times each week	43 (22.16)	70 (21.94)	113 (22.03)
A few times a month	43 (22.16)	60 (18.81)	103 (20.08)
Less than once a month	62 (31.96)	117 (36.6)	179 (34.89)
Reasons for downloading a weight-management app, n (%)			
Weight loss	26 (13.40)	134 (42.0)	160 (31.19)
Monitor food intake	82 (42.27)	88 (27.59)	170 (33.14)
Track how much activity or exercise I get	52 (26.80)	47 (14.73)	99 (19.30)
Show or teach me exercises	19 (9.79)	24 (7.52)	43 (8.38)
I want to kill time when bored	7 (3.61)	12 (3.76)	19 (3.70)
Other reasons	8 (4.12)	14 (4.39)	22 (4.29)
Reasons for downloading a particular weight-management app, n (%)			
Best ranked in the app store	53 (27.32)	101 (31.6)	154 (30.02)
Recommendations from friends or family	66 (34.02)	88 (27.59)	154 (30.02)
Social media influencers	32 (16.49)	61 (19.12)	93 (18.13)
Web searches (eg, Google)	15 (7.73)	45 (14.11)	60 (11.70)
Recommended by other apps	26 (13.40)	18 (5.64)	44 (8.58)
TV	2 (1.03)	6 (1.88)	8 (1.56)
Desired features of weight-management apps, n (%)			
Monitored by a specialist	31 (15.98)	126 (39.5)	157 (30.60)
Can identify calories using barcode	54 (27.84)	77 (24.14)	131 (25.54)
Nutrition information of many food items	67 (34.54)	56 (17.55)	123 (23.98)
Provides a weekly or monthly progress report	26 (13.40)	44 (13.79)	70 (13.65)
Reminders to follow a diet or exercise	16 (8.25)	16 (5.02)	32 (6.24)

Table 3. User conceptions and efficacy of weight-management apps stratified by body mass index.

User conceptions	Body mass index		Total
	18.5-24.9 kg/m^2	>25 kg/m^2	
Apps that provide a weight-reduction meal plan helped in managing weight, n (%)			
Strongly disagree	65 (33.51)	74 (23.20)	139 (27.10)
Disagree	11 (5.67)	19(5.96)	30 (5.85)
Unsure	20 (10.31)	78 (24.45)	98 (19.10)
Agree	82 (42.27)	125 (39.18)	207 (40.35)
Strongly agree	16 (8.25)	23 (7.21)	39 (7.60)
Apps that provide an exercise plan helped in managing weight, n (%)			
Strongly disagree	51 (26.29)	71 (22.26)	122 (23.78)
Disagree	9 (4.64)	15 (4.70)	24 (4.68)
Unsure	18 (9.28)	56 (17.55)	74 (14.42)
Agree	96 (49.48)	148 (46.39)	244 (47.56)
Strongly agree	20 (10.31)	29 (9.09)	49 (9.55)
Weight-management apps are effective for long-term use, n (%)			
Strongly disagree	50 (25.77)	52 (16.30)	102 (19.88)
Disagree	0 (0.00)	7 (2.19)	7 (1.36)
Unsure	42 (21.65)	170 (53.29)	212 (41.33)
Agree	80 (41.24)	60 (18.81)	140 (27.29)
Strongly agree	22 (11.34)	30 (9.40)	52 (10.14)

Table 4. Reasons for discontinuing use, stratified by body mass index.

Reasons for discontinuing use	Body mass index		Total
	18.5-24.9 kg/m^2	>25 kg/m^2	
Have you downloaded any weight management apps that you no longer use? n (%)			
No	50 (25.77)	83 (26.02)	133 (25.93)
Yes	144 (74.23)	236 (73.98)	380 (74.07)
Why do you no longer use them? n (%)			
Monitoring by a specialist was not offered	21 (14.58)	80 (33.90)	101 (26.58)
The app language was not in the local language	45 (31.25)	48 (20.34)	93 (24.47)
There were hidden costs	21 (14.58)	37 (15.68)	58 (15.26)
Loss of interest	17 (11.81)	30 (12.71)	47 (12.37)
The app was confusing to use	30 (20.83)	34 (14.41)	64 (16.84)
I no longer need it, or I met my goals	10 (6.94)	7 (2.97)	17 (4.47)

Discussion

Principal Findings

The survey found that only 43% of respondents used weight-management apps. Most users were young, educated women with a high income, and were overweight or obese. Overweight or obese users download a weight-management app mostly to lose weight, whereas normal-weight participants, on the contrary, download a weight-management app mainly to monitor food intake and track physical activity. The feature of the weight-management app that overweight or obese users desired most was the possibility to be monitored by a specialist, whereas the least desired feature was the constant reminders to follow a diet or exercise. On the contrary, normal-weight participants mainly desired apps that could provide nutrition information and calorie content of food items through barcode identification.

Among all users of a weight-management app, this study found that one of the main reasons for discontinuing use was that the language of the app was not the local language of the user.

Among overweight or obese users, another main reason was lack of monitoring by a specialist, and among normal-weight users, difficulty of use.

Comparison With Previous Work

In this study, 43% of the study sample currently used weight-management apps. A similar survey by Krebs and Duncan showed health app use among US adults to be about 60% [17]. Cultural differences and the newness of app use for health and weight management in the Saudi population may explain the difference [19].

Findings regarding the characteristics of weight-management app users in our survey are consistent with the results of a previous large-scale study in the United States [20]; we also found that most participants were young, had a higher income, had higher education, and reported very good or excellent health. Age and education were found to be important predictors for using mobile phones and apps as younger people are more exposed to mobile technology than the elderly, resulting in a higher utilization of mobile health-related apps [21]. We also found that the highest use of weight-management apps was among women. Women's high use of weight-management apps in our study is in line with previous investigations that showed women were more concerned about their health status and tended to follow a healthier dietary pattern than men [22]. We also found more overweight or obese participants used such apps (mostly to lose weight) compared with normal-weight participants, who mostly wanted to maintain their weight. This is consistent with a previous study showing that weight-management apps were mostly downloaded by overweight or obese users, which may be because these users are more concerned about their overall health than others [23].

We found that overweight or obese participants mostly preferred to be monitored by a specialist compared with other features of the app such as constant reminders to follow a diet or exercise. A previous intervention found that, for encouraging healthy behavior, individually tailored feedback and action plans were preferable to general health instructions [24]. A randomized controlled trial showed that individualized health apps were more effective in addressing patients' needs and provided the high-quality information and support required by patients [25]. In our survey, the most frequent reason for downloading a weight-management app, as reported by overweight or obese users, was to reduce caloric intake, whereas normal-weight users wanted to monitor their food intake. Among US health-related app users, Krebs et al found that the main reason for downloading apps was to track how much physical activity they were getting (52.8%) [17]. Their findings, however, did not differentiate between normal weight and overweight or obese users. The difference in preference may be related to cultural and social variables in Saudi Arabia rather than biological reasons, although activities such as walking for women is generally acceptable in major cities; this is not always the case in rural parts of the country [26].

As for efficacy of use, more normal-weight users found that weight-management apps helped them lose weight compared with overweight or obese users. A randomized controlled trial, however, found that apps for weight loss did not affect the weight of overweight primary care patients and were only useful for individuals who were willing to self-monitor their calories [27].

Participants in our survey also tended to discontinue use of the app for various reasons such as monitoring by a specialist was not offered and app was not in the local language. Similar barriers were reported in a previous qualitative study on the challenges of app use in supporting health behavior changes, showing that users generally found health apps to be time consuming, have unexpected costs, and lack the possibility of consulting an expert [28]. A previous qualitative study restricted to Saudi women only reported that only a few obese or overweight women were aware of the availability of weight-management apps and reported barriers of app use to be lack of motivational support and the language of the app [16].

Strengths and Limitations

The strength of this study is that it is the first of its kind to target app users in Saudi Arabia. Additionally, it provides information on the demographics of this particular region along with insight into the extent of weight-management app use and barriers of use, which varies from barriers in other regions. This study also highlighted that weight-management apps are not only used by overweight or obese adults but also by normal-weight adults who wish to maintain their weight. Data on the differences in perception of use between overweight and normal-weight adults are also provided.

Although this is the first study to characterize user perspective on weight-management apps in Riyadh, Saudi Arabia, there were a few limitations. The main limitation was study design; a causal relationship cannot be inferred, and usage patterns may change over time. Second, the study sample included the city of Riyadh only and was not representative of the general population. A third limitation was the self-recorded weight and height of participants, which is a possible source of error. Finally, as this was a convenience sample, it did not include individuals who are overweight or obese but do not own mobile phones, although a large percentage of the population use mobile phones.

Conclusions

Although the majority of users believe that weight-management apps are efficacious, considerable challenges remain. Given the increasing prevalence of overweight and obesity in Saudi Arabia and the high accessibility to mobile phones among adults, app developers should address findings of this study, especially regarding reasons for use and discontinuance. More specifically, app developers should consider the aforementioned barriers by developing apps in local languages and by developing apps which are linked to health professionals to provide individualized plans to manage the weight of both normal and overweight adults.

Acknowledgments

This research project was supported by a grant from the Research Center of the Female Scientific and Medical Colleges, Deanship of Scientific Research, King Saud University.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Sociodemographic characteristics and health status of participants stratified by gender.

[PDF File (Adobe PDF File), 114KB - mhealth_v7i2e12692_app1.pdf]

Multimedia Appendix 2

Pattern of use of weight-management apps stratified by gender.

[PDF File (Adobe PDF File), 96KB - mhealth_v7i2e12692_app2.pdf]

Multimedia Appendix 3

User conceptions and efficacy of weight-management apps stratified by gender.

[PDF File (Adobe PDF File), 89KB - mhealth_v7i2e12692_app3.pdf]

Multimedia Appendix 4

Reasons for discontinuing use stratified by gender.

[PDF File (Adobe PDF File), 85KB - mhealth_v7i2e12692_app4.pdf]

References

1. Finucane MM, Stevens GA, Cowan MJ, Danaei G, Lin JK, Paciorek CJ, Global Burden of Metabolic Risk Factors of Chronic Diseases Collaborating Group (Body Mass Index). National, regional, and global trends in body-mass index since 1980: systematic analysis of health examination surveys and epidemiological studies with 960 country-years and 9·1 million participants. Lancet 2011 Feb 12;377(9765):557-567 [FREE Full text] [doi: 10.1016/S0140-6736(10)62037-5] [Medline: 21295846]
2. Memish ZA, El Bcheraoui C, Tuffaha M, Robinson M, Daoud F, Jaber S, et al. Obesity and associated factors--Kingdom of Saudi Arabia, 2013. Prev Chronic Dis 2014 Oct 09;11:E174 [FREE Full text] [doi: 10.5888/pcd11.140236] [Medline: 25299980]
3. Al-Quwaidhi AJ, Pearce MS, Critchley JA, Sobngwi E, O'Flaherty M. Trends and future projections of the prevalence of adult obesity in Saudi Arabia, 1992-2022. East Mediterr Health J 2014 Oct;20(10):589-595. [Medline: 25356689]
4. Diabetes Prevention Program Research Group, Knowler WC, Fowler SE, Hamman RF, Christophi CA, Hoffman HJ, et al. 10-year follow-up of diabetes incidence and weight loss in the Diabetes Prevention Program Outcomes Study. Lancet 2009 Nov 14;374(9702):1677-1686 [FREE Full text] [doi: 10.1016/S0140-6736(09)61457-4] [Medline: 19878986]
5. Eddy DM, Schlessinger L, Kahn R. Clinical outcomes and cost-effectiveness of strategies for managing people at high risk for diabetes. Ann Intern Med 2005 Aug 16;143(4):251-264. [Medline: 16103469]
6. de Jongh T, Gurol-Urganci I, Vodopivec-Jamsek V, Car J, Atun R. Mobile phone messaging for facilitating self-management of long-term illnesses. Cochrane Database Syst Rev 2012 Dec 12;12:CD007459. [doi: 10.1002/14651858.CD007459.pub2] [Medline: 23235644]
7. Beratarrechea A, Lee AG, Willner JM, Jahangir E, Ciapponi A, Rubinstein A. The impact of mobile health interventions on chronic disease outcomes in developing countries: a systematic review. Telemed J E Health 2014 Jan;20(1):75-82 [FREE Full text] [doi: 10.1089/tmj.2012.0328] [Medline: 24205809]
8. Kim YJ, Rhee SY, Byun JK, Park SY, Hong SM, Chin SO, et al. A smartphone application significantly improved diabetes self-care activities with high user satisfaction. Diabetes Metab J 2015 Jun;39(3):207-217 [FREE Full text] [doi: 10.4093/dmj.2015.39.3.207] [Medline: 26124991]
9. Klasnja P, Pratt W. Healthcare in the pocket: mapping the space of mobile-phone health interventions. J Biomed Inform 2012 Feb;45(1):184-198 [FREE Full text] [doi: 10.1016/j.jbi.2011.08.017] [Medline: 21925288]
10. Newzoo. Top 50 Countries/Markets by Smartphone Users and Penetration URL: https://newzoo.com/insights/rankings/top-50-countries-by-smartphone-penetration-and-users/ [accessed 2019-02-13] [WebCite Cache ID 769WlDJ9W]

11. Statista. Number of smartphone users in Saudi Arabia from 2017 to 2023 (in millions)* URL: https://www.statista.com/statistics/494616/smartphone-users-in-saudi-arabia/ [accessed 2019-02-13] [WebCite Cache ID 769XQKbV0]
12. Research2Guidance. 2017. mHealth Economics 2017 – Current Status and Future Trends in Mobile Health URL: https://research2guidance.com/product/mhealth-economics-2017-current-status-and-future-trends-in-mobile-health/ [accessed 2019-02-13] [WebCite Cache ID 769XqTOFt]
13. Burke LE, Styn MA, Glanz K, Ewing LJ, Elci OU, Conroy MB, et al. SMART trial: a randomized clinical trial of self-monitoring in behavioral weight management-design and baseline findings. Contemp Clin Trials 2009 Nov;30(6):540-551 [FREE Full text] [doi: 10.1016/j.cct.2009.07.003] [Medline: 19665588]
14. Flores Mateo G, Granado-Font E, Ferré-Grau C, Montaña-Carreras X. Mobile phone apps to promote weight loss and increase physical activity: a systematic review and meta-analysis. J Med Internet Res 2015 Nov 10;17(11):e253 [FREE Full text] [doi: 10.2196/jmir.4836] [Medline: 26554314]
15. Murray A. Institute for Healthcare Informatics. 2013 Nov 1. Patient Apps for Improved Healthcare: From Novelty to Mainstream URL: http://moodle.univ-lille2.fr/pluginfile.php/215345/mod_resource/content/0/IIHI_Patient_Apps_Report.pdf [accessed 2019-02-17] [WebCite Cache ID 76G4msQuC]
16. Alnasser AA, Alkhalifa AS, Sathiaseelan A, Marais D. What overweight women want from a weight loss app: a qualitative study on arabic women. JMIR Mhealth Uhealth 2015 May 20;3(2):e41 [FREE Full text] [doi: 10.2196/mhealth.4409] [Medline: 25993907]
17. Krebs P, Duncan DT. Health app use among US mobile phone owners: a national survey. JMIR Mhealth Uhealth 2015 Nov 04;3(4):e101 [FREE Full text] [doi: 10.2196/mhealth.4924] [Medline: 26537656]
18. General Authority for Statistics. 2016. Demography Survey of the Kingdom of Saudi Arabia URL: https://www.stats.gov.sa/en
19. Alnasser AA, Amalraj RE, Sathiaseelan A, Al-Khalifa AS, Marais D. Do Arabic weight-loss apps adhere to evidence-informed practices? Transl Behav Med 2016 Dec;6(3):396-402 [FREE Full text] [doi: 10.1007/s13142-015-0347-7] [Medline: 27528528]
20. Carroll JK, Moorhead A, Bond R, LeBlanc WG, Petrella RJ, Fiscella K. Who uses mobile phone health apps and does use matter? A secondary data analytics approach. J Med Internet Res 2017 Dec 19;19(4):e125 [FREE Full text] [doi: 10.2196/jmir.5604] [Medline: 28428170]
21. Latif S, Rana R, Qadir J, Ali A, Imran MA, Younis MS. Mobile health in the developing world: review of literature and lessons from a case study. IEEE Access 2017;5:11540-11556. [doi: 10.1109/ACCESS.2017.2710800]
22. Beer-Borst S, Hercberg S, Morabia A, Bernstein M, Galan P, Galasso R, et al. Dietary patterns in six european populations: results from EURALIM, a collaborative European data harmonization and information campaign. Eur J Clin Nutr 2000 Mar;54(3):253-262. [Medline: 10713749]
23. Binks M, van Mierlo T. Utilization patterns and user characteristics of an ad libitum internet weight loss program. J Med Internet Res 2010;12(1):e9 [FREE Full text] [doi: 10.2196/jmir.1347] [Medline: 20350926]
24. de Vries H, Kremers SP, Smeets T, Brug J, Eijmael K. The effectiveness of tailored feedback and action plans in an intervention addressing multiple health behaviors. Am J Health Promot 2008 Jul;22(6):417-425. [doi: 10.4278/ajhp.22.6.417] [Medline: 18677882]
25. Weymann N, Härter M, Dirmaier J. A tailored, interactive health communication application for patients with type 2 diabetes: study protocol of a randomised controlled trial. BMC Med Inform Decis Mak 2013 Feb 13;13(1). [doi: 10.1186/1472-6947-13-24]
26. Al-Zalabani AH, Al-Hamdan NA, Saeed AA. The prevalence of physical activity and its socioeconomic correlates in Kingdom of Saudi Arabia: a cross-sectional population-based national survey. J Taibah Univ Med Sci 2015 Jun;10(2):208-215. [doi: 10.1016/j.jtumed.2014.11.001]
27. Laing BY, Mangione CM, Tseng C, Leng M, Vaisberg E, Mahida M, et al. Effectiveness of a smartphone application for weight loss compared with usual care in overweight primary care patients: a randomized, controlled trial. Ann Intern Med 2014 Nov 18;161(10 Suppl):S5-12 [FREE Full text] [doi: 10.7326/M13-3005] [Medline: 25402403]
28. Dennison L, Morrison L, Conway G, Yardley L. Opportunities and challenges for smartphone applications in supporting health behavior change: qualitative study. J Med Internet Res 2013 Apr 18;15(4):e86 [FREE Full text] [doi: 10.2196/jmir.2583] [Medline: 23598614]

Abbreviations

BMI: body mass index

Edited by G Eysenbach; submitted 05.11.18; peer-reviewed by S Boucher, J Alvarez Pitti, C Carrion, K Tamura, PH Lin; comments to author 05.01.19; revised version received 08.01.19; accepted 25.01.19; published 22.02.19.

Please cite as:
Aljuraiban GS
Use of Weight-Management Mobile Phone Apps in Saudi Arabia: A Web-Based Survey
JMIR Mhealth Uhealth 2019;7(2):e12692
URL: http://mhealth.jmir.org/2019/2/e12692/
doi:10.2196/12692
PMID:30794205

Original Paper

Comparing Self-Monitoring Strategies for Weight Loss in a Smartphone App: Randomized Controlled Trial

Michele L Patel[1,2,3], PhD; Christina M Hopkins[1,2], BS; Taylor L Brooks[1,2], BA; Gary G Bennett[1,2], PhD

[1]Department of Psychology and Neuroscience, Duke University, Durham, NC, United States

[2]Duke Digital Health Science Center, Duke Global Health Institute, Durham, NC, United States

[3]Stanford Prevention Research Center, Stanford University School of Medicine, Palo Alto, CA, United States

Corresponding Author:
Michele L Patel, PhD
Stanford Prevention Research Center
Stanford University School of Medicine
1070 Arastradero Road
Suite 100
Palo Alto, CA, 94304
United States
Phone: 1 650 549 7047
Fax: 1 650 725 6247
Email: michele.patel@stanford.edu

Abstract

Background: Self-monitoring of dietary intake is a valuable component of behavioral weight loss treatment; however, it declines quickly, thereby resulting in suboptimal treatment outcomes.

Objective: This study aimed to examine a novel behavioral weight loss intervention that aims to attenuate the decline in dietary self-monitoring engagement.

Methods: GoalTracker was an automated randomized controlled trial. Participants were adults with overweight or obesity (n=105; aged 21-65 years; body mass index, BMI, 25-45 kg/m^2) and were randomized to a 12-week stand-alone weight loss intervention using the MyFitnessPal smartphone app for daily self-monitoring of either (1) both weight and diet, with weekly lessons, action plans, and feedback (Simultaneous); (2) weight through week 4, then added diet, with the same behavioral components (Sequential); or (3) only diet (App-Only). All groups received a goal to lose 5% of initial weight by 12 weeks, a tailored calorie goal, and automated in-app reminders. Participants were recruited via online and offline methods. Weight was collected in-person at baseline, 1 month, and 3 months using calibrated scales and via self-report at 6 months. We retrieved objective self-monitoring engagement data from MyFitnessPal using an application programming interface. Engagement was defined as the number of days per week in which tracking occurred, with diet entries counted if ≥800 kcal per day. Other assessment data were collected in-person via online self-report questionnaires.

Results: At baseline, participants (84/100 female) had a mean age (SD) of 42.7 (11.7) years and a BMI of 31.9 (SD 4.5) kg/m^2. One-third (33/100) were from racial or ethnic minority groups. During the trial, 5 participants became ineligible. Of the remaining 100 participants, 84% (84/100) and 76% (76/100) completed the 1-month and 3-month visits, respectively. In intent-to-treat analyses, there was no difference in weight change at 3 months between the Sequential arm (mean −2.7 kg, 95% CI −3.9 to −1.5) and either the App-Only arm (−2.4 kg, −3.7 to −1.2; *P*=.78) or the Simultaneous arm (−2.8 kg, −4.0 to −1.5; *P*=.72). The median number of days of self-monitoring diet per week was 1.9 (interquartile range [IQR] 0.3-5.5) in Sequential (once began), 5.3 (IQR 1.8-6.7) in Simultaneous, and 2.9 (IQR 1.2-5.2) in App-Only. Weight was tracked 4.8 (IQR 1.9-6.3) days per week in Sequential and 5.1 (IQR 1.8-6.3) days per week in Simultaneous. Engagement in neither diet nor weight tracking differed between arms.

Conclusions: Regardless of the order in which diet is tracked, using tailored goals and a commercial mobile app can produce clinically significant weight loss. Stand-alone digital health treatments may be a viable option for those looking for a lower intensity approach.

Trial Registration: ClinicalTrials.gov NCT03254953; https://clinicaltrials.gov/ct2/show/NCT03254953 (Archived by WebCite at http://www.webcitation.org/72PyQrFjn).

(JMIR Mhealth Uhealth 2019;7(2):e12209) doi:10.2196/12209

KEYWORDS

weight loss; obesity, self-monitoring; technology; mobile app; mobile health; caloric restriction; treatment adherence and compliance; randomized controlled trial

Introduction

Background

Self-monitoring of dietary intake is a cornerstone of behavioral weight loss treatment [1], and past research has demonstrated that the frequency of self-monitoring is positively associated with weight loss [2]. Despite its utility, dietary self-monitoring typically declines over the course of treatment [2,3]. Novel strategies are needed to improve dietary self-monitoring engagement [4]. One strategy involves enriching self-monitoring with other theoretically and empirically supported behavior change techniques such as tailored goals and feedback, action plans, and skills training [5,6].

A second strategy includes building mastery, self-efficacy, and self-regulation—key constructs of behavior change in Carver's Control Theory [7] and Bandura's Social Cognitive Theory [8]— *before* asking participants to engage in dietary self-monitoring. Fostering self-regulatory skills may provide an opportunity for mastery and, in turn, strengthen self-efficacy [9], which has been linked to greater weight reduction [10]. Behavioral weight loss interventions typically involve self-monitoring multiple behaviors or outcomes simultaneously during treatment [2,11,12], which can serve as an efficient strategy for producing behavior change but may detrimentally impede performance on each item or result in greater treatment dropout [13,14]. Failing to develop mastery of a behavior can weaken self-efficacy [15], leading to worse treatment outcomes. In addition, when people begin to self-monitor a behavior less frequently, it is likely that they will also self-monitor a corresponding behavior less frequently, as demonstrated in an analysis of self-monitoring patterns [16].

We propose a novel solution that aims to attenuate the decline in engagement by employing a *sequential* [17] self-monitoring approach, wherein individuals track only body weight for a period of time and then begin to track diet. Tracking body weight was chosen as it requires minimal effort, provides an opportunity for habit formation (eg, track every morning upon waking), and is efficacious for weight loss [18]. We focused on only self-monitoring of body weight during the first month based on prior research demonstrating that enhanced engagement in the first month of treatment may have long-lasting repercussions [19]. Sequential approaches are based on the premise that mastery is more easily obtained if only 1 new behavior is targeted at a time [20,21]. Recent research finds that both sequential and simultaneous approaches are effective for behavior change [21] and that focusing on a single component instead of a multicomponent intervention produces comparable weight loss [22], suggesting that a simpler approach is sufficient to produce behavior change.

We tested this sequential strategy using a remotely delivered intervention that utilizes a popular commercially available mobile app—MyFitnessPal. Utilizing technology for self-monitoring dietary intake has been shown to produce greater adherence and less-pronounced declines in engagement than traditional paper-based tracking methods [23,24]. As demonstrated in recent reviews [25-27], smartphone apps can produce significant weight loss, although most existing trials that use commercial apps lack fully powered designs or are pilot studies. Interventions without counseling that utilize commercial technology for dietary self-monitoring have produced clinically meaningful weight losses between 2.5 kg and 5.5 kg at 3 months and beyond in recent studies [28-32].

Objectives

In the current randomized trial, we test a weight loss intervention among adults with overweight or obesity that targets the aforementioned strategies of including empirically supported behavior change techniques, promoting mastery and self-efficacy through self-monitoring of body weight before diet, and utilizing a free, commercially available app (MyFitnessPal). We hypothesize that a sequential approach will produce greater weight loss and self-monitoring engagement at 3 months compared with a traditional simultaneous approach and to an "off-the-shelf" app.

Methods

Study Design Overview

GoalTracker was a 3-arm randomized controlled trial comparing 3 stand-alone weight loss interventions: (1) a *Simultaneous* self-monitoring arm in which participants simultaneously tracked body weight and dietary intake each day and received additional empirically supported behavior change techniques (see Table 1) via email for the entirety of the intervention, (2) a *Sequential* arm, consisting of identical intervention components but allowing for mastery of 1 skill (ie, self-monitoring of body weight) before beginning self-monitoring of diet, and (3) an *App-Only* arm that tracked only diet with no additional behavior change components. Study evaluation visits were held at baseline, 1 month, and 3 months. Self-reported weight was collected at 6 months. All study procedures were approved by the Duke University Institutional Review Board (protocol #: D0822; date of approval: 10/14/16).

Table 1. Differences in intervention components between treatment arms.

Intervention component	Sim[a]	Seq[b]	App-Only arm	Theoretical construct	Behavior change technique [5]
Weight loss goal: 0.5-2.0 lb/week (tailored) and 5% by 12 weeks	✓[c]	✓	✓	Self-regulation	Goal setting (outcome)
Calorie goal: tailored based on individual factors and rate of weight loss; minimum 1200 kcal for women, 1500 kcal for men	✓	✓ (delayed)	✓	Self-regulation	Goal setting (behavior)
Self-monitoring of body weight: daily via the app	✓	✓	—[d]	Self-regulation	Prompt self-monitoring of behavioral outcome
Self-monitoring of dietary intake: daily via the app	✓	✓ (delayed)	✓	Self-regulation	Prompt self-monitoring of behavior
Facilitate mastery experience by first tracking weight then tracking diet	—	✓	—	Self-efficacy; self-regulation; mastery	Set graded tasks
In-app real-time feedback	✓	✓	✓	Self-regulation; self-efficacy	Provide feedback on performance
Out-of-app summary feedback via weekly email (tailored)	✓	✓	—	Self-regulation; self-efficacy	Provide feedback on performance; prompt review of outcome goals; prompt review of behavioral goals
Skills training via weekly email with structured behavioral lesson and tips on how to use features of the app	✓	✓	—	Outcome expectancies; self-efficacy	Provide information on consequences of behavior in general; prompt generalization of a target behavior; provide information on when and where to perform the behavior; provide instruction on how to perform the behavior; environmental restructuring; plan social support/social change; relapse prevention/coping planning
Action plans via weekly email	✓	✓	—	Self-regulation; self-efficacy	Action planning; motivational interviewing [33]; barrier identification/problem solving [34]; prompt practice; plan social support/social change
Reminder of goals	✓	✓	—	Self-regulation	Prompt review of outcome goals; prompt review of behavioral goals
In-app automated reminders to track diet and/or weight sent daily (App-Only received reminders to track diet, Simultaneous received both diet and weight tracking reminders, and Sequential received weight tracking for all 12 weeks and diet tracking reminders starting in week 5)	✓	✓	✓	Self-regulation	Teach to use prompts/cues

[a]Sim: Simultaneous self-monitoring intervention arm.

[b]Seq: Sequential self-monitoring intervention arm.

[c]The component is present.

[d]The component is not present.

Participants

Inclusion criteria comprised men and women aged 21 to 65 years with a body mass index (BMI) between 25.0 and 45.0 kg/m^2 who were interested in losing weight through dietary change. We required participants to have an iPhone or Android smartphone, email address, access to a bathroom scale, and written English fluency. Participants needed to be willing to download the mobile app on their phone and not track diet or body weight using any other modality (eg, other health or weight tracking apps, websites, and paper diaries) for the duration of the intervention. We excluded participants if they were enrolled in another weight loss intervention, had used MyFitnessPal to track diet in the past 6 months, had lost ≥10 lb or used a weight loss medication in the past 6 months, had previous or planned bariatric surgery, or if weight loss would be contraindicated (eg, pregnancy or <12 months postpartum or in need of medical or psychiatric intervention such as for cancer diagnosis, eating disorder, uncontrolled hypertension, diabetes mellitus, cardiovascular event, or congestive heart failure). A total of 2 criteria were amended during the trial recruitment to promote generalizability of findings: the BMI criteria were expanded to include participants in the 40.0 to 45.0 kg/m^2 range, and the weight change criteria were adjusted to no longer exclude individuals who gained more than 10 lb in the past 6 months. The institutional review board approved both amendments.

Recruitment

Recruitment occurred between April and September 2017 in central North Carolina via a university-affiliated research website and listservs, social media postings (Twitter and Facebook), ClinicalTrials.gov registry, and community advertisements (Craigslist, Nextdoor, and paper flyers). Advertisements provided a description of the study and eligibility criteria. Participants were enrolled on a rolling basis until we met our intended sample size.

Procedure

We directed interested individuals to a study website with descriptive information and a screening questionnaire that assessed all eligibility criteria, including participants' height and weight. Study personnel contacted eligible candidates within 3 business days to schedule an in-person baseline visit. During the baseline visit, trained study staff obtained written informed consent, confirmed eligibility, collected anthropometric measurements, and assisted participants in installing and navigating the MyFitnessPal mobile app; participants then completed an online survey.

Using simple random assignment, participants were then randomized by study staff to 1 of 3 treatment arms using Excel's random number generator to allocate participants equally (1:1:1) across conditions. Randomization was revealed to participants by study personnel; as such, study staff were not blinded to treatment allocation but were blinded to the allocation sequence. Participants then reviewed materials describing their treatment condition and goals (see Intervention Design section below) in writing and with study staff to reduce contamination.

In-person follow-up visits occurred at 1 month and 3 months. At 1 month, study staff provided participants with information on their goals for the remainder of the intervention (see Intervention Design section below for details). We compensated participants with Amazon electronic gift cards (US $12 at baseline, US $6 for each follow-up visit, and US $5 bonus for completing dietary measures). Questionnaires were administered in English via a desktop computer. There was no contact with participants from months 3 to 6, and participants were not asked to self-monitor in MyFitnessPal during this time (though they could still do so if desired). At 6 months, study staff contacted participants via email and text message to collect self-reported body weight. Data collection ended in March 2018.

Intervention Design

Participants were randomized to 1 of 3 conditions: (1) Simultaneous, (2) Sequential, or (3) App-Only, as outlined below and in Table 1. The CALO-RE taxonomy is used to describe behavior change techniques [5]. The intervention period lasted 12 weeks.

Common Components

All treatment arms self-monitored dietary intake using MyFitnessPal, a free commercial app that allows users to log food and beverages and provides nutritional information from a database with over 6 million foods [35]. This app has high acceptability [36]. In-app feedback in both graphical and text format provides users with real-time progress updates. When setting up participants' MyFitnessPal accounts, study staff entered an end goal weight that corresponded with losing 5% of their initial body weight by 12 weeks. On the basis of this goal and the participant's current weight, a weekly weight loss goal between 0.5 and 2.0 pounds was calculated. Along with the Mifflin-St. Jeor equation that factors in basal metabolic rate [37], this value was used to determine a tailored daily calorie goal, with a minimum caloric goal of 1200 kcal/day for women and 1500 kcal/day for men. During the baseline visit, in-app push-reminders were programmed to be sent each day if tracking had not occurred by a prespecified time in the evening. No structured dietary advice (eg, follow a low carbohydrate diet) was given to participants. Of note, study staff also created a Fitbit account for each person via the platform's website and linked this account with MyFitnessPal. Participants were not given a Fitbit device and they were never asked to use this Fitbit account; its sole purpose was for accessing MyFitnessPal's data using Fitbit's application programming interface (API). In the App-Only arm, MyFitnessPal served as an "off-the-shelf," self-guided approach that the general US population can already access for free in the commercial marketplace.

Both Simultaneous and Sequential Arms

In addition to the common intervention components, participants in the Simultaneous and Sequential arms were asked to self-weigh and enter their body weight in the app each day. Each week, study staff sent participants an email with tailored feedback that was automatically generated using Microsoft Word's Mail Merge feature. This feedback email described the participants' overall weight loss progress and their progress on each goal in the past week, including track weight daily, meet weekly weight loss goal, track diet daily, and meet daily calorie goal (the latter 2 goals not given to Sequential participants until week 5). Feedback on weight outcomes was provided as long as 1 weight was recorded in the past week. Individuals who did not track their weight in the past week received a message stating "Make sure to enter your weight in MyFitnessPal so that we can give you helpful insights!" Feedback pertaining to the calorie goal included only days with complete food diaries (≥800 kcal) [38]. This calorie feedback was not given to the Sequential arm until week 5.

Each week, participants were also sent skills training materials via email on a different day, including a researcher-designed tip on using different features of the app (eg, using the barcode scanner) accompanied by step-by-step screenshots of the app, a lesson on nutrition or behavior change (eg, reducing sugary foods and managing food intake on vacations; see Multimedia Appendix 1 for an example) adapted from gold-standard weight loss curriculum [39,40], and a brief online action plan to reinforce the weekly lesson. Accessed via a link to a Qualtrics survey, action plans incorporated motivational interviewing and problem-solving strategies [33,34] and included the following types of components: identifying current behaviors and beliefs; evaluating confidence and reasons for change; thinking about the *when*, *where*, and *what* of each action; brainstorming potential barriers that may arise and crafting solutions; identifying a support person; and reviewing past action plans (see Tables 2 and 3 for lesson and tip topics).

Table 2. Topics of structured lessons.

Week	Lesson topic
1	Overview of the program (losing 5% weight, self-monitoring); calorie balance[a]
2	Red zone foods; green zone foods
3	Reading food labels[a]
4	Reducing sugar
5	Portion control
6	Preparing meals at home; shopping tips
7	Eating out
8	Social support
9	Environmental cues; vacations and holidays
10	Emotional eating
11	Slippery slope; weight loss maintenance; relapse prevention

[a]In these 2 lessons for the Sequential treatment arm, there was no discussion of tracking diet or adhering to a calorie goal. Otherwise, all lesson content was identical between the arms.

Table 3. Tips for using the MyFitnessPal app.

Week	Tips for Simultaneous arm	Tips for Sequential arm
Sent after baseline visit	A: How to track body weight; B: How to view weight progress; C: How to track a food item; D: How to view your calorie goal and the foods you have tracked	A: How to track body weight; B: How to view weight progress
1	How to use the barcode scanner	How to delete a weight entry
2	How to use multiadd to speed up food tracking	How to add progress photos
3	How to view nutrition progress	How to change reminders to track (weight)
4	A: How to delete a weight entry; B: How to add progress photos	How to recruit a friend to use MyFitnessPal
Sent after 1MV[a]	—[b]	A: How to track a food item; B: How to view your calorie goal and the foods you have tracked
5	How to track food from a restaurant	How to use the barcode scanner
6	A: How to create a meal; B: How to log a meal	How to use multiadd to speed up food tracking
7	A: How to add a recipe; B: How to log a recipe	How to view nutrition progress
8	How to use the “Complete Diary” feature	How to track food from a restaurant
9	How to change reminders to track (weight and food)	A: How to create a meal; B: How to log a meal
10	How to customize meal names	A: How to add a recipe; B: How to log a recipe
11	How to recruit a friend to use MyFitnessPal	A: How to use the “Complete Diary” feature; B: How to customize meal names

[a]1MV: 1-month visit.

[b]Data are not applicable (no tips were provided to the Simultaneous arm directly after the 1-month visit).

Sequential Arm's Self-Monitoring and Feedback

Individuals in the Sequential arm received the same intervention components as the Simultaneous arm, but they did not begin self-monitoring dietary intake until week 5 of the intervention. They did not receive a calorie goal until their 1-month evaluation visit, nor were their in-app reminders for tracking diet set up before this time point. The Sequential arm's weekly feedback emails did not mention diet tracking or the calorie goal until they began tracking diet. In addition, their weekly app usage tips did not describe diet-tracking tips until after the first month (see Table 3). Like the Simultaneous arm, they were still encouraged to make healthy dietary changes during the first month as suggested in the weekly lessons and action plans, but these lessons did not mention tracking diet or adhering to a calorie goal.

Outcome Measures

Primary Outcome: Change in Weight

The primary outcome was weight change at 3 months. We measured body weight using a calibrated electronic scale (SECA

876) at baseline, 1 month, and 3 months in light clothing with shoes removed. Height was measured to the nearest 0.1 cm using a calibrated, wall-mounted stadiometer (SECA 222). Baseline height was used for calculation of BMI at all time points. We collected self-reported body weight at 6 months and asked participants to send a photo with their feet on the scale displaying the value in either kg or lb. We assessed the proportion of participants at 3 months who achieved weight loss of ≥3% and ≥5% from baseline.

Self-Monitoring Engagement Data

We used a software engine developed at Duke—Prompt—to collect participants' objective MyFitnessPal self-monitoring data; Prompt retrieved these data using the API of Fitbit, which was linked to each participant's MyFitnessPal account. Primary outcomes for self-monitoring engagement span from day 1 (the day after participants' baseline visit) to day 83 and were categorized into the first 4 weeks in the intervention (days 1-28), the final 2 months (days 29-83), and the entire 83-day intervention period. Exploratory analyses examined engagement data after the intervention ended up to 6 months (day 183) post randomization.

For all self-monitoring data, we only counted days with complete diet entries (ie, recording ≥800 kcal/day [38]). We examined the median number of days per week that participants self-monitored weight and diet, as well as the percentage of days that entries were recorded (ie, number of days with entries recorded divided by number of days instructed to record an entry, multiplied by 100).

Engagement in Action Plans

Percentage of action plans completed was examined through objective Qualtrics survey data: each action plan was coded as "completed" or "not completed." The completion status of each action plan was combined to generate a summary score with a possible range of 0% to 100% (indicative of 11 out of 11 action plans completed).

Engagement in Feedback Email

In the 3-month survey, we assessed participants' self-reported frequency of reading their weekly feedback email, with the question "How frequently did you read your weekly Progress Reports (sent via email), on average?" and 5 response options (*Several times per week*, *One time per week*, *Less than 1 time per week*, *Less than 1 time per month*, or *Never*).

Sociodemographic and Clinical Characteristics

At baseline, we collected data on participant demographics, socioeconomic status, and type of smartphone. To assess past MyFitnessPal use, we asked the Pew Research Center's question: "What kind of health apps do you currently have on your phone?" [41]; if the "Diet, Food, Calorie Counter" response option was selected, then the open-ended question "What are the names of the diet, food, or calorie-counting apps that you used on your phone?" was asked.

We also assessed whether participants had ever been told by a doctor or other health professional that they had prediabetes or hypertension. Self-monitoring of weight and self-monitoring of diet in the month before baseline were each measured with a 7-point scale ranging from *several times per day* to *never* [42].

Statistical Analysis

Sample size was calculated based on power to detect a 3.5 kg difference in weight change at 3 months between the Sequential arm and the App-Only arm (our primary comparison) using 3-month results from previous remotely delivered weight loss interventions for our Sequential arm [43] (results in kilograms were provided upon request by the author) and our App-Only arm [44]. Our power analysis (G*Power 3.1.9.2.) determined that 31 participants per group were needed to achieve 80% power for a 2-sided test with an alpha level of .05. To account for attrition of 10% and to obtain equal-size groups, we aimed to recruit 105 participants (35 per group). In exploratory analyses, we compared weight change between the Sequential arm and the Simultaneous arm, although we were not adequately powered to detect a significant effect.

For the baseline characteristics, we computed descriptive statistics stratified by treatment arm. To determine whether baseline characteristics differed by retention status, we used the Pearson chi-square test for categorical variables, analysis of variance for continuous variables, and Fisher exact tests with small cell counts. All analyses were 2-tailed. Participants who became ineligible during the study period up to 3 months were excluded from the analyses. Investigators remained blinded to outcomes until the completion of the 6-month trial.

We used intent-to-treat analyses to test our primary aim using linear mixed modeling with an unstructured covariance matrix and restricted maximum likelihood estimates to examine changes in weight over time by treatment arm. We did not control for any additional variables, and we assumed missing at random and used SAS 9.4 PROC MIXED (SAS Institute) for these analyses. For 6-month weight values sent via photo, we subtracted 0.172 kg (0.4 lb) to account for participants holding a device on the scale to take the photo. To account for the 6-month self-reported weight data without photos, we used a regression model to adjust for age, gender, and race/ethnicity [45]. Participants who sent a photo of their 6-month weight did not differ on any measured sociodemographic characteristics from those who did not send a photo (data not shown). We used chi-square tests to assess proportion of participants achieving ≥3% and ≥5% weight loss; we assumed noncompleters did not achieve this clinical threshold.

Given non-normally distributed intervention engagement data, we reported medians and interquartile ranges (IQR). To examine differences between treatment arms, we used Wilcoxon Mann-Whitney *U* tests (if 2 arms) and the Kruskal-Wallis tests (if 3 arms). We used Spearman rank correlation coefficients (r_s) to examine the relation between self-monitoring engagement and change in weight. We also assessed for contamination by exploring whether participants self-monitored when they were not expected to do so.

Results

Participant Enrollment and Retention

Of the 670 individuals who completed the online screen for eligibility, 58.3% (391) were ineligible, whereas 23.7% (159) were invited to attend the baseline visit. We enrolled 105 participants and randomized them equally to 3 treatment arms (n=35, for each; see Figure 1 for Consolidated Standards of Reporting Trials [CONSORT] diagram). During the trial, 5 participants became ineligible (3 because of pregnancy, 1 because of cancer diagnosis, 1 because of previously undisclosed eating disorder). Of the remaining 100 participants, 84.0% (84 of 100) completed the 1-month visit and 76.0% (76 of 100) completed the 3-month visit. Between 3 and 6 months, one additional participant became ineligible because of pregnancy. At 6 months, 78% (77 of 99) of participants self-reported their weight. Participant retention did not differ significantly between arms at any time point (1 month: *P*=.84; 3 months: *P*=.23; 6 months: *P*=.32). We had no missing self-monitoring engagement data.

Baseline Characteristics

Table 4 illustrates the baseline characteristics of GoalTracker participants. At baseline, participants had a mean age (SD) of 42.7 (11.7) years and BMI of 31.9 (4.5) kg/m^2 and were predominantly female (84/100) and employed (78/100). One-third (33/100) were racial or ethnic minorities, most were married or living with a partner (64/100), and the majority had at least a college education (83/100). The majority (56/100) did not track diet in the month before baseline, although most had experience tracking body weight (87.0%). Completers at 3 months differed from noncompleters in race or ethnic minority status (*P*=.03), with 16% (11/67) of non-Hispanic white participants and 39% (13/33) of racial or ethnic minority participants missing the visit.

Weight Loss

Figure 2 displays weight change over time by treatment arm. Weight change was significant over time for all arms (see Table 5). In our primary analysis, the Sequential arm did not significantly differ from the App-Only arm in weight change at 1 month (*P*=.06), 3 months (*P*=.78), or 6 months (*P*=.72). In exploratory analyses, the Sequential arm did not differ from the Simultaneous arm in weight change at 1 month (*P*=.36), 3 months (*P*=.92), or 6 months (*P*=.45).

The proportion of participants achieving at least 3% weight loss at 3 months was similar between arms (Sequential: 44%, 15/34 vs App-Only: 29%, 10/34, *P*=.21; exploratory analysis: Sequential vs Simultaneous 41%, 13/32, *P*=.77). Likewise, weight loss of at least 5% at 3 months occurred in 21% of Sequential participants (7/34) and 15% of App-Only participants (5/34), which was not significantly different (*P*=.52). In exploratory analyses, the proportion of participants with ≥5% weight loss did not significantly differ between the Sequential arm and the Simultaneous arm (31% of participants, 10/32, *P*=.32).

Intervention Engagement

As expected, the Sequential arm tracked *weight* significantly more days than the App-Only arm (who was not asked to track weight) over the 12-week intervention; median (IQR) 70%, (28%-90%) vs 1% (0%-8%), respectively (*P*<.001). In exploratory analyses, the frequency of days participants self-monitored *weight* did not differ between the Simultaneous arm (73%; 25%-90%) and the Sequential arm over 12 weeks (*P*=.92).

As expected, the frequency with which the Sequential arm and the App-Only arm tracked *diet* in weeks 1 to 4 was significantly different (see Table 6). Between weeks 5 and 12 (once the Sequential arm began tracking diet), there was no longer a significant difference between the Sequential and App-Only arms; 27% (4%-80%) vs 21% (0%-62%), respectively (*P*=.54). In exploratory analyses, there were no significant differences in the frequency of days of self-monitoring *diet* between the Simultaneous and App-Only arms over the intervention period; 77% (27%-96%) vs 42% (17%-75%), *P*=.10. Table 6 displays additional self-monitoring and action plan completion outcomes (see Figures 3 and 4 for weekly data).

Relation Between Self-Monitoring Frequency and Weight Change

The percentage of days weight was tracked was significantly associated with 3-month weight change in both the Simultaneous arm (r_s=−.48, *P*=.02) and the Sequential arm (r_s=−.47, *P*=.01). In the same time period, the association between weight change and the percentage of days with complete diet entries was significant in the App-Only arm (r_s=−.58, *P*=.003) but not for the Simultaneous arm (r_s=−.25, *P*=.24). The percentage of days diet was tracked starting in week 5 for the Sequential arm was significantly associated with weight change at 3 months (r_s=−.44, *P*=.02; see Table 7 for additional details).

Contamination

The median (IQR) frequency of days that App-Only participants tracked weight in the MyFitnessPal app during the 3-month intervention was 1% (0%-8%), and the frequency of days that Sequential participants tracked diet during month 1 was 0% (0%-0%; see Table 6 for absolute values).

Action Plan Completion

In the Simultaneous arm, the median (IQR) number of action plans completed was 7.7 of 11—70% (14%-91%)—compared with 3 of 11—27% (9%-82%)—in the Sequential arm; in exploratory analyses, this difference was not statistically significant (*P*=.21). Percent action plan completion was significantly related to weight change at 3 months in the Sequential arm (r_s=−.60, *P*<.001) but not in the Simultaneous group (r_s=−.07, *P*=.75).

Review of Feedback Email

Most participants (35/52; 67%) reported reading their weekly feedback email at least once per week, whereas 12% (6/52) of participants reported never reading them. In exploratory analyses, there were no significant differences between the Simultaneous and the Sequential arm (*P*=.90).

Figure 1. Consolidated Standards of Reporting Trials (CONSORT) flow diagram. BMI: body mass index.

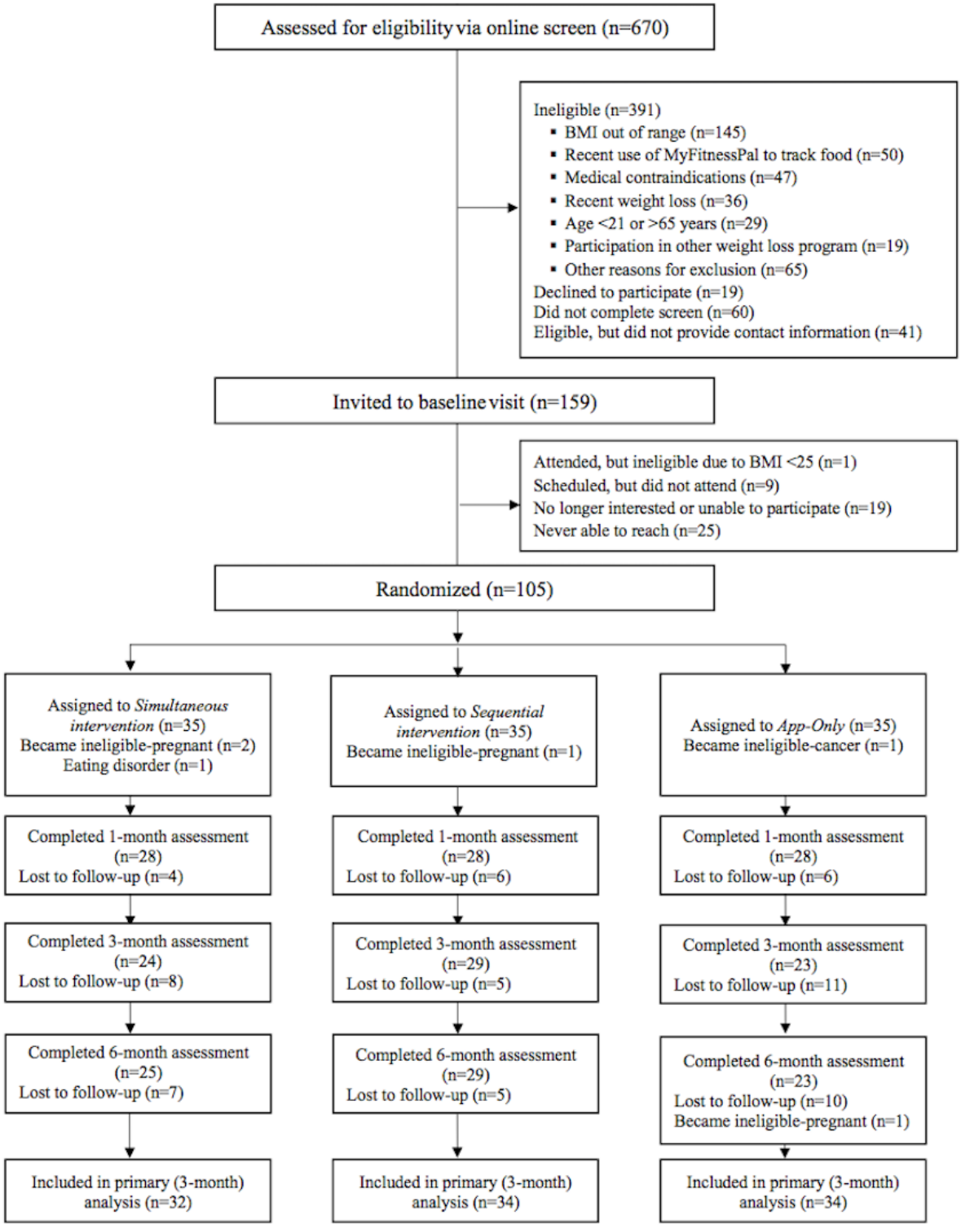


Table 4. Baseline characteristics by treatment arm.

Characteristic at baseline	Total (N=100[a])	Simultaneous (n=32)	Sequential (n=34)	App-Only (n=34)
Age (years), mean (SD)	42.7 (11.7)	43.8 (13)	42.1 (11)	42.3 (12)
Gender, n (%)				
Male	16 (16.0)	8 (25)	4 (12)	4 (12)
Female	84 (84.0)	24 (75)	30 (88)	30 (88)
Marital status, n (%)				
Married or living with partner	64 (64.0)	22 (69)	18 (53)	24 (71)
Not married or living with partner	36 (36.0)	10 (31)	16 (47)	10 (29)
Race/ethnicity, n (%)				
Non-Hispanic white	67 (67.0)	21 (66)	23 (68)	23 (68)
Non-Hispanic black	22 (22.0)	9 (28)	6 (18)	7 (21)
Hispanic (all races)	3 (3.0)	0 (0)	2 (6)	1 (3)
Non-Hispanic other	8 (8.0)	2 (6)	3 (9)	3 (9)
Education, n (%)				
Less than college graduate	17 (17.0)	7 (22)	6 (18)	4 (12)
College graduate or above	83 (83.0)	25 (78)	28 (82)	30 (88)
Employment status, n (%)				
Employed, full-time	67 (67.0)	20 (63)	27 (79)	20 (59)
Employed, part-time	11 (11.0)	3 (9)	1 (3)	7 (21)
Not employed	22 (22.0)	9 (28)	6 (18)	7 (21)
Annual household income, in US dollars, n (%)				
$0-$49,999	26 (26.0)	8 (25)	9 (27)	9 (27)
$50,000-$99,999	36 (36.0)	14 (44)	12 (35)	10 (29)
$100,000 or greater	34 (34.0)	9 (28)	11 (32)	14 (41)
Unknown/not reported	4 (4.0)	1 (3)	2 (6)	1 (3)
Weight, mean (SD), kg	89.6 (16.0)	89.3 (17)	90.8 (17)	88.6 (15)
Body mass index, mean (SD), kg/m^2	31.9 (4.5)	31.3 (4)	32.6 (5)	31.7 (4)
Body mass index category, n (%)				
Overweight, 25-29.9 kg/m^2	40 (40.0)	13 (41)	12 (35)	15 (44)
Class 1 obesity, 30-34.9 kg/m^2	38 (38.0)	14 (44)	13 (38)	11 (32)
Class 2 obesity, 35-39.9 kg/m^2	17 (17.0)	4 (13)	6 (18)	7 (21)
Class 3 obesity, 40+ kg/m^2	5 (5.0)	1 (3)	3 (99)	1 (3)
Self-monitoring of diet frequency, n (%)				
Daily	6 (6.0)	1 (3)	2 (6)	3 (9)
1 to 6 times per week	14 (14.0)	6 (19)	4 (12)	4 (12)
Less than 1 time per week	24 (24.0)	4 (13)	11 (32)	9 (27)
Never	56 (56.0)	21 (66)	17 (50)	18 (53)
Self-monitoring of weight frequency, n (%)				
Daily	11 (11.0)	6 (19)	2 (6)	3 (9)
1 to 6 times per week	35 (35.0)	6 (19)	14 (41)	15 (44)
Less than 1 time per week	41 (41.0)	13 (41)	16 (47)	12 (35)

Characteristic at baseline	Total (N=100[a])	Simultaneous (n=32)	Sequential (n=34)	App-Only (n=34)
Never	13 (13.0)	7 (22)	2 (6)	4 (12)
Type of smartphone, n (%)				
iPhone	54 (54.0)	17 (53)	16 (47)	21 (62)
Android	46 (46.0)	15 (47)	18 (53)	13 (38)
MyFitnessPal already on phone before study, n (%)	20 (20.0)	6 (19)	9 (27)	5 (15)

[a]Five participants omitted because they became ineligible during the intervention period.

Figure 2. Weight change over time by treatment arm. Data were included for 100 participants; mean (SD) values were estimated using an intention-to-treat analysis with a linear mixed-model.

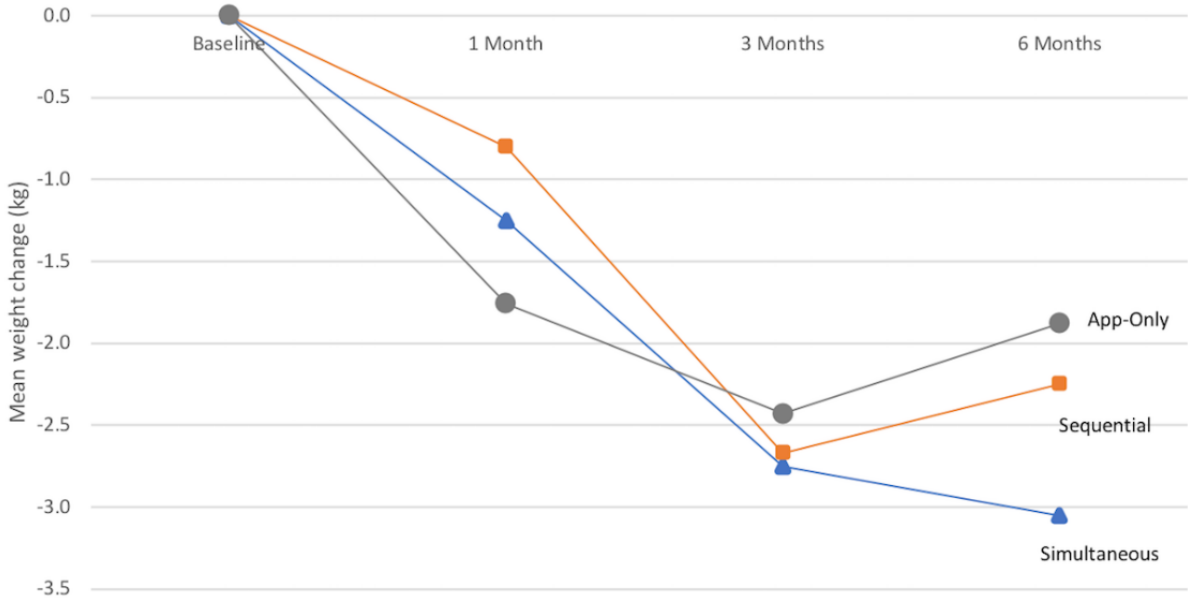


Table 5. Change in weight and body mass index (intent-to-treat).

Outcome by time point	Mean (95% CI)			
	Simultaneous	Sequential	App-Only	Between-group difference (Seq[a] vs App-Only)[b]
Weight change from baseline (kg)				
1 month	−1.25 (−1.97 to −0.53)	−0.80 (−1.49 to −0.10)	−1.76 (−2.48 to −1.05)	0.97 (−0.03 to 1.97)
3 months	−2.75 (−4.01 to −1.49)	−2.67 (−3.85 to −1.49)	−2.43 (−3.69 to −1.16)	−0.24 (−1.97 to 1.49)
6 months[c]	−3.05 (−4.57 to −1.52)	−2.25 (−3.66 to −0.85)	−1.88 (−3.41 to −0.34)	−0.38 (−2.46 to 1.71)
BMI[d] change from baseline (kg/m^2)				
1 month	−0.46 (−0.71 to −0.21)	−0.29 (−0.53 to −0.05)	−0.63 (−0.88 to −0.38)	0.35 (0 to 0.69)
3 months	−0.99 (−1.44 to −0.55)	−0.95 (−1.37 to −0.54)	−0.88 (−1.32 to −0.43)	−0.08 (−0.69 to 0.53)
6 months[c]	−1.06 (−1.60 to −0.52)	−0.81 (−1.31 to −0.32)	−0.67 (−1.21 to −0.13)	−0.15 (−0.88 to 0.59)

[a]Seq: Sequential self-monitoring intervention arm.

[b]This is the primary comparison; the App-Only arm is the reference group.

[c]One additional participant was omitted in analyses (App-Only arm) at 6-months due to becoming ineligible after the intervention period and before 6-months.

[d]BMI: body mass index.

Table 6. Self-monitoring engagement by treatment arm.

Self-monitoring engagement by time period	Median (interquartile range)			
	Simultaneous (n=32)	Sequential (n=34)	App-Only (n=34)	*P* value
Baseline to 4 weeks (out of 28 days)				
Number of days per week tracked weight	6.25 (2.63 to 6.75)	5.75 (3.50 to 6.50)	0 (0 to 0.75)	—[a]
Number of days per week tracked diet	6.50 (3.88 to 7.00)	0 (0)	5.38 (2.25 to 7.00)	—
Percentage of days tracked weight	89 (37 to 96)	82 (50 to 93)	0 (0 to 11)	.52[b]; <.001[c]
Percentage of days tracked diet	93 (55 to 100)	0 (0)	77 (32 to 100)	.37[d]; <.001[c]
5 to 12 weeks (out of 55 days)				
Number of days per week tracked weight	4.50 (0.69 to 6.13)	4.06 (0.75 to 6.63)	0 (0 to 0.25)	—
Number of days per week tracked diet	4.88 (0.44 to 6.56)	1.88 (0.25 to 5.50)	1.44 (0 to 4.25)	—
Percentage of days tracked weight	65 (10 to 89)	59 (11 to 95)	0 (0 to 4)	.95[b]; <.001[c]
Percentage of days tracked diet	70 (6 to 95)	27 (4 to 80)	21 (0 to 62)	.54[c]; .17[e]
Entire intervention (out of 83 days)				
Number of days per week tracked weight	5.08 (1.75 to 6.25)	4.83 (1.92 to 6.25)	0.08 (0 to 0.58)	—
Number of days per week tracked diet	5.33 (1.83 to 6.67)	—[f]	2.92 (1.17 to 5.17)	—
Percentage of days tracked weight	73 (25 to 90)	70 (28 to 90)	1 (0 to 8)	.92[b]; <.001[c]
Percentage of days tracked diet	77 (27 to 96)	—[f]	42 (17 to 75)	.10[d]
Percentage of action plans completed	70 (14 to 91)	27 (9 to 82)	—	.21[b]
13 weeks to 6 months (postintervention; out of 99 days)[g]				
Number of days per week tracked weight	0.32 (0 to 0.89)	0.43 (0 to 1.64)	0 (0 to 0.14)	—
Number of days per week tracked diet	0.14 (0 to 1.18)	0 (0 to 0.29)	0 (0 to 0.43)	—
Percentage of days tracked weight	5 (0 to 14)	7 (0 to 23)	0 (0 to 2)	.78[b]; .004[c]
Percentage of days tracked diet	3 (0 to 17)	0 (0 to 5)	0 (0 to 6)	.96[c]; .43[e]

[a]Not applicable.

[b]Simultaneous arm versus Sequential arm.

[c]Sequential arm versus App-Only arm.

[d]Simultaneous arm versus App-Only arm.

[e]All arms.

[f]As the Sequential arm did not track diet in the first 4 weeks of the intervention, their results for the "Entire intervention" section would be the same as the results in the "5 to 12 weeks" section above.

[g]One additional participant was omitted in analyses (App-Only arm) at 6 months because of becoming ineligible after the intervention period and before 6-months.

Figure 3. Self-monitoring of weight per intervention week by treatment arm.

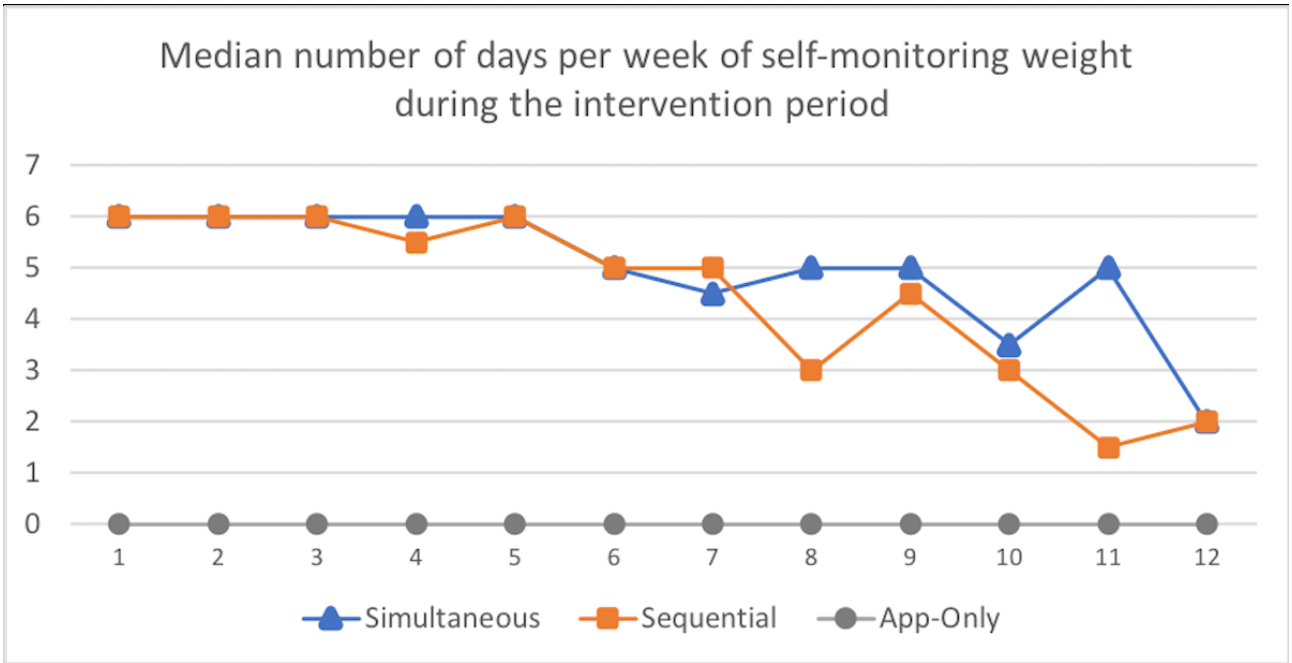


Figure 4. Self-monitoring of dietary intake per intervention week by treatment arm.

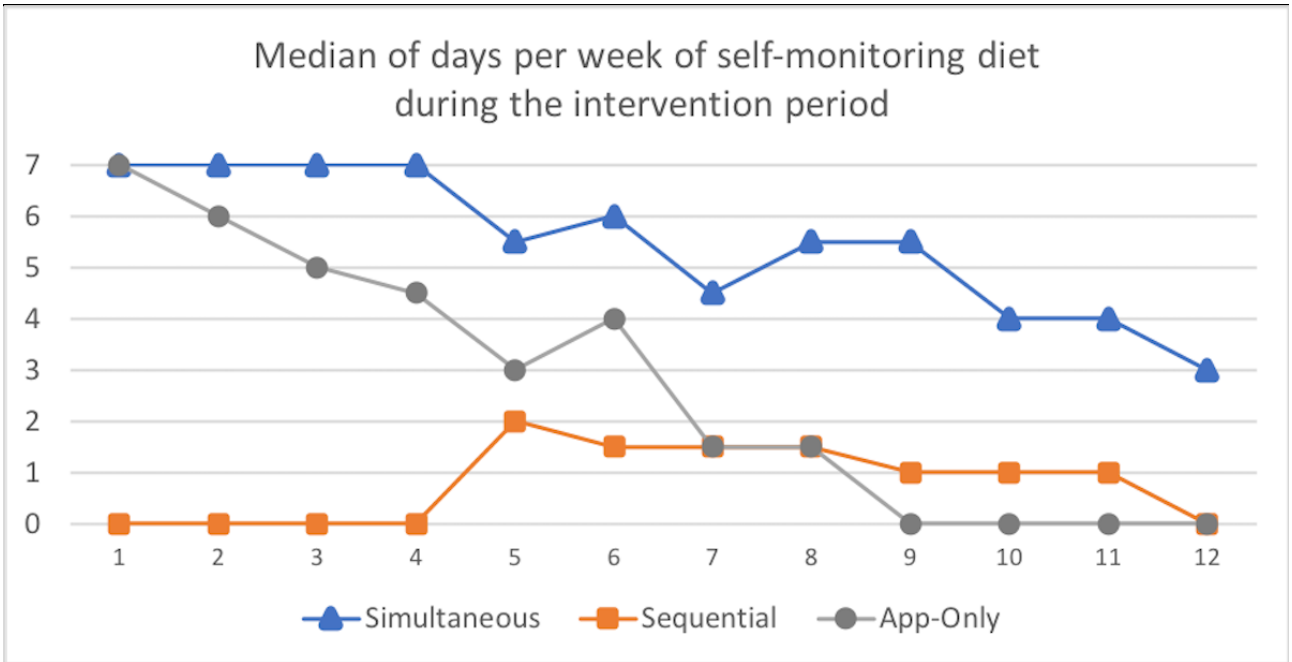


Table 7. Spearman rank correlation between engagement metrics and weight change. The table displays correlations and *P* values for the treatment arms that were asked to track during the given time period.

Engagement metric	Weight change by 1 month	Weight change by 3 months
Percentage of action plans completed	—[a]	Both: −.41[b]
	—	Sim[c]: −.07
	—	Seq[d]: −.60[b]
Baseline to 4 weeks		
Percentage of days tracked weight	Both: −.35[b]	Both: −.40[b]
	Sim: −.40[e]	Sim: −.51[e]
	Seq: −.29	Seq: −.34
Percentage of days tracked diet	Both: −.42[b]	Both: −.40[b]
	Sim: −.36	Sim: −.30
	App[f]: −.51[b]	App: −.48[b]
5 to 12 weeks		
Percentage of days tracked weight	Both: −.44[b]	Both: −.48[b]
	Sim: −.37	Sim: −.49[e]
	Seq: −.54[b]	Seq: −.46[e]
Percentage of days tracked diet	All[g]: −.37[b]	All: −.42[b]
	Sim: −.24	Sim: −.27
	Seq: −.50[b]	Seq: −.44[e]
	App: −.52[b]	App: −.55[b]
Entire intervention		
Percentage of days tracked weight	Both: −.44[b]	Both: −.47[b]
	Sim: −.40[e]	Sim: −.48[e]
	Seq: −.50[b]	Seq: −.47[e]
Percentage of days tracked diet[h]	Both: −.35[b]	Both: −.42[b]
	Sim: −.30	Sim: −.25
	App: −.52[b]	App: −.58[b]
13 weeks to 6 months (post intervention period)[i]		
Percentage of days tracked weight	Both: −.50[b]	Both: −.43[b]
	Sim: −.49[b]	Sim: −.43[e]
	Seq: −.59[b]	Seq: −.43[e]
Percentage of days tracked diet	Both: −.29[b]	All: −.35[b]
	Sim: −.27	Sim: −.17
	Seq: −.42[e]	Seq: −.47[e]
	App: −.20	App: −.39

[a]Not applicable.

[b]*P*<.01.

[c]Sim: Simultaneous self-monitoring arm.

[d]Seq: Sequential self-monitoring arm.

[e]*P*<.05.

[f]App: App-Only self-monitoring arm.

[g]All: All 3 treatment arms.

[h]As the Sequential arm did not track diet in the first 4 weeks of the intervention, their results for the "Entire intervention" section would be the same as the results in the "5 to 12 weeks" section above.

[i]One additional participant was omitted in analyses (App-Only arm) at 6 months because of becoming ineligible after the intervention period and before 6-months.

Discussion

Principal Findings

A low-intensity intervention utilizing a commercial app for self-monitoring resulted in comparable weight loss at 3 months, with no variability between the Sequential arm and the "off-the-shelf" App-Only arm. Nevertheless, loss of 3% to 5% of initial weight has been linked to improved health outcomes [46,47], suggesting that GoalTracker is an efficacious intervention for clinically meaningful weight loss.

The addition of evidence-based features such as weekly action plans, behavioral lessons, and tailored feedback did not substantially impact outcomes over and above the core intervention (ie, self-monitoring and in-app feedback) during our 12-week treatment, which parallels findings from several digital health weight loss trials [30,48] but not others [6]. Although most commercial weight loss apps do not include many evidence-based features [49], we suspect that weight loss might still occur with the inclusion of goals, daily self-monitoring, and a daily reminder to track. However, we suspect that had the trial been of a longer duration, the benefit of these enhanced features (eg, weekly lessons) would have become apparent in the findings; indeed, by 6 months, trends suggest continued weight loss at 6 months for the Simultaneous arm and relatively less weight regain in the Sequential arm compared with the App-Only arm.

Given that the GoalTracker trial compared 3 multicomponent interventions, we are unable to isolate the effect of self-monitoring diet and weight, along with each of the additional intervention components. Using a factorial design in consort with the multiphase optimization strategy (MOST) [50] would allow researchers to investigate the unique impact of each intervention component and then build and test an optimized intervention.

We found that self-monitoring engagement was high and that greater frequency of self-monitoring was related to greater weight loss. Contrary to our hypothesis, the Sequential arm did not demonstrate significantly greater engagement in self-monitoring dietary intake than the App-Only arm. In fact, once the Sequential arm was instructed to begin tracking diet after the first month, they tracked only 27% of days. It is possible that the Sequential arm participants' minimal weight loss in the first month may have negatively impacted their future engagement or that being asked to track diet after a 1-month period seemed like an additional burden that many were unwilling to begin (indeed, almost half or 47% of Sequential participants never or rarely [<20% of days] tracked diet in weeks 5 to 12 [data not shown]). Although the trial was not powered to detect differences between the Simultaneous and Sequential arms, it appears that the Simultaneous arm outperforms the Sequential arm in frequency of tracking diet. This finding suggests that concurrently tracking diet and weight may have reinforced use of the app and activated the overarching goal of losing weight [51], thus leading to high engagement in both entities.

Future studies could consider framing the initial period before self-monitoring diet as a time during which to build self-regulatory skills rather than focus on weight loss, as has been done previously [9]. We selected weight tracking as the precursor to diet tracking for the Sequential arm because we wanted a behavior that could target the theoretical constructs of mastery—considered the best way to strengthen self-efficacy—and self-regulation. Self-weighing is a well-accepted strategy [52] in which mastery can be achieved [16], and self-regulatory capacity can be strengthened [53]. It is possible that providing more rationale for using a sequential approach would have encouraged participants to engage in diet tracking once asked to do so.

Comparison With Prior Work

Notably, GoalTracker is the first weight loss trial to compare a sequential self-monitoring approach with a traditional approach that asks participants to track multiple components simultaneously. Previous work has compared a simultaneous approach with either a sequential or single component approach in other contexts, with mixed results [21,54]; no examination has focused on self-monitoring or digital approaches for weight loss.

The App-Only arm in GoalTracker performed better than expected. Intervention participants in Laing et al's trial used the same MyFitnessPal app for diet tracking and lost 0.27 kg at 3 months and 0.03 kg at 6 months and had poor intervention engagement [44]. Possible explanations for the difference in weight loss between Laing et al's trial and GoalTracker include GoalTracker's use of specific goals to track diet daily and to lose 5% of initial weight by a specified end date and usage of phone-based reminder notifications.

In comparison with other randomized trials of *commercial* [28,29,32,55-58] or *researcher*- designed [24,59,60] apps for self-monitoring of diet, GoalTracker's Simultaneous arm tended to have greater adherence to diet tracking, whereas the Sequential arm had lower adherence. Given that most weight loss trials of commercial apps are pilot studies and/or were not powered to detect an effect in weight change between treatment arms [26,28,29,38,56,61], more fully powered studies are needed that examine the efficacy of commercial apps for weight loss.

GoalTracker's Simultaneous arm had a comparable or higher proportion of participants achieving 5% weight loss compared with other weight loss interventions that used mobile apps for self-monitoring dietary intake (range: 26%-35%) [28,59,62,63]

but lower rates than some interventions including counseling (range: 42%-44%) [29,60,61].

Strengths

Strengths of this trial include the collection of objective self-monitoring data for all participants via an API, use of a popular commercially available smartphone app, and ability to isolate the effect of a sequential versus simultaneous self-monitoring approach. In addition, this trial mimicked real-world weight loss experience (ie, no run-in period, prebaseline visit, or orientation session); consequently, it is possible that removal of these treatment barriers allowed for inclusion of participants with lower motivation and readiness to change. This design may have greater external validity but may make it harder to detect an effect between arms. Another strength was that the trial had little contamination between arms, which has been a problem in past app-based trials where up to 50% of participants in no-treatment control arms were found to have used commercial apps during the study period [44,55,63].

Limitations

As this study was powered on superiority rather than equivalency, we cannot definitively assert that the treatment arms produce comparable weight loss. In addition, we collected self-reported weight at 6-months because of logistical reasons; however, we were encouraged to find that no additional attrition occurred between 3 and 6 months, despite no contact occurring during that period and no incentive given to provide a weight value. As is common in behavioral interventions, we provided minimal financial compensation to offset costs of attending study visits. Although we acknowledge that financial compensation can serve as an incentive for some to participate—and thus, may result in response bias on self-report measures—we expect this is unlikely, given that compensation was appropriately low. In addition, neither study staff nor participants were blinded to treatment arm, and we required participants to have access to a bathroom scale, although this mimics the real-world population who would track weight. Finally, this study did not include a pure control arm without an intervention, which may have led to an underestimation of treatment effects, as could the possibility of data not actually missing at random.

Conclusions

This study adds to the limited literature of randomized trials that assess the efficacy of commercially available mobile apps for weight loss [20,26]. In the GoalTracker trial, all 3 versions of the intervention produced weight loss and had high self-monitoring engagement, with no significant impact of additional features nor differential findings between a sequential versus simultaneous approach to self-monitoring. These results suggest that regardless of the order in which diet is tracked, using tailored weight and calorie goals and a commercial app can produce clinically significant weight loss in one-third of individuals. Stand-alone digital health treatments may be a viable option for those looking for a lower intensity approach who are willing and able to track.

Acknowledgments

The research described in this paper was supported by a grant to the first author from the American Psychological Association, the Duke Interdisciplinary Behavioral Research Center, and the Aleane Webb Dissertation Research Award provided by The Graduate School at Duke University. GGB serves on the scientific advisory board of Nutrisystem and Interactive Health; he holds equity in Coeus Health. These organizations had no role in study design, data collection, data analysis and interpretation of data, in the writing of the report, or in the decision to submit the article for publication. The remaining authors have no disclosures.

The authors would like to thank Tia Kelley, Azaria Anderson, and Kendra Pallin for their dedication as research assistants and are deeply grateful for many members of Duke Digital Health, including Martin Streicher, Sandy Askew, Dori Steinberg, Hira Ahmed, and Hallie Davis-Penders. Finally, the authors would like to thank all who participated in GoalTracker.

Authors' Contributions

MLP and GGB conceived of and designed the study. MLP, CMH, and TLB delivered the intervention and collected data. MLP analyzed the results and drafted the manuscript. All authors were involved in writing the paper and had final approval of the submitted and published versions.

MLP was affiliated with Duke University at the time of the trial and is currently affiliated with the Stanford University School of Medicine.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Example weekly lesson.

[PDF File (Adobe PDF File), 166KB - mhealth_v7i2e12209_app1.pdf]

Multimedia Appendix 2

CONSORT - EHEALTH checklist (V 1.6.1).

[PDF File (Adobe PDF File), 2MB - mhealth_v7i2e12209_app2.pdf]

References

1. Hartmann-Boyce J, Johns DJ, Jebb SA, Aveyard P, Behavioural Weight Management Review Group. Effect of behavioural techniques and delivery mode on effectiveness of weight management: systematic review, meta-analysis and meta-regression. Obes Rev 2014 Jul;15(7):598-609 [FREE Full text] [doi: 10.1111/obr.12165] [Medline: 24636238]
2. Burke LE, Wang J, Sevick MA. Self-monitoring in weight loss: a systematic review of the literature. J Am Diet Assoc 2011 Jan;111(1):92-102 [FREE Full text] [doi: 10.1016/j.jada.2010.10.008] [Medline: 21185970]
3. Yu Z, Sealey-Potts C, Rodriguez J. Dietary self-monitoring in weight management: current evidence on efficacy and adherence. J Acad Nutr Diet 2015 Dec;115(12):1931-1938. [doi: 10.1016/j.jand.2015.04.005] [Medline: 26028176]
4. Lentferink AJ, Oldenhuis HK, de Groot M, Polstra L, Velthuijsen H, van Gemert-Pijnen JE. 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 1;19(8):e277 [FREE Full text] [doi: 10.2196/jmir.7288] [Medline: 28765103]
5. Michie S, Ashford S, Sniehotta FF, Dombrowski SU, Bishop A, French DP. A refined taxonomy of behaviour change techniques to help people change their physical activity and healthy eating behaviours: the CALO-RE taxonomy. Psychol Health 2011 Nov;26(11):1479-1498. [doi: 10.1080/08870446.2010.540664] [Medline: 21678185]
6. Hutchesson MJ, Rollo ME, Krukowski R, Ells L, Harvey J, Morgan PJ, et al. eHealth interventions for the prevention and treatment of overweight and obesity in adults: a systematic review with meta-analysis. Obes Rev 2015 May;16(5):376-392. [doi: 10.1111/obr.12268] [Medline: 25753009]
7. Carver CS, Scheier MF. Control theory: a useful conceptual framework for personality-social, clinical, and health psychology. Psychol Bull 1982 Jul;92(1):111-135. [doi: 10.1037/0033-2909.92.1.111] [Medline: 7134324]
8. Bandura A. Health promotion from the perspective of social cognitive theory. Psychol Health 1998 Jul;13(4):623-649. [doi: 10.1080/08870449808407422]
9. Kiernan M, Brown SD, Schoffman DE, Lee K, King AC, Taylor CB, et al. Promoting healthy weight with "stability skills first": a randomized trial. J Consult Clin Psychol 2013 Apr;81(2):336-346 [FREE Full text] [doi: 10.1037/a0030544] [Medline: 23106759]
10. Teixeira PJ, Carraça EV, Marques MM, Rutter H, Oppert J, de Bourdeaudhuij I, et al. Successful behavior change in obesity interventions in adults: a systematic review of self-regulation mediators. BMC Med 2015 Apr 16;13:84 [FREE Full text] [doi: 10.1186/s12916-015-0323-6] [Medline: 25907778]
11. Shieh C, Knisely MR, Clark D, Carpenter JS. Self-weighing in weight management interventions: a systematic review of literature. Obes Res Clin Pract 2016;10(5):493-519. [doi: 10.1016/j.orcp.2016.01.004] [Medline: 26896865]
12. Lyzwinski LN. A systematic review and meta-analysis of mobile devices and weight loss with an intervention content analysis. J Pers Med 2014;4(3):311-385 [FREE Full text] [doi: 10.3390/jpm4030311] [Medline: 25563356]
13. King AC, Castro CM, Buman MP, Hekler EB, Urizar GG, Ahn DK. Behavioral impacts of sequentially versus simultaneously delivered dietary plus physical activity interventions: the CALM trial. Ann Behav Med 2013 Oct;46(2):157-168 [FREE Full text] [doi: 10.1007/s12160-013-9501-y] [Medline: 23609341]
14. Schulz DN, Kremers SP, Vandelanotte C, van Adrichem MJ, Schneider F, Candel MJ, et al. Effects of a web-based tailored multiple-lifestyle intervention for adults: a two-year randomized controlled trial comparing sequential and simultaneous delivery modes. J Med Internet Res 2014;16(1):e26 [FREE Full text] [doi: 10.2196/jmir.3094] [Medline: 24472854]
15. Muretta RJ. Exploring the four sources of self-efficacy. Dissertation Abstracts International, Section B: The Sciences and Engineering 2005;65(10-B):5447 [FREE Full text]
16. Zheng Y, Burke LE, Danford CA, Ewing LJ, Terry MA, Sereika SM. Patterns of self-weighing behavior and weight change in a weight loss trial. Int J Obes (Lond) 2016 Dec;40(9):1392-1396. [doi: 10.1038/ijo.2016.68] [Medline: 27113642]
17. Vandelanotte C. Are simultaneously delivered health behavior change interventions the way of the future?: a comment on King et al. Ann Behav Med 2013 Oct;46(2):133-134. [doi: 10.1007/s12160-013-9511-9] [Medline: 23637073]
18. Zheng Y, Klem ML, Sereika SM, Danford CA, Ewing LJ, Burke LE. Self-weighing in weight management: a systematic literature review. Obesity (Silver Spring) 2015 Feb;23(2):256-265. [doi: 10.1002/oby.20946] [Medline: 25521523]
19. Unick JL, Pellegrini CA, Demos KE, Dorfman L. Initial weight loss response as an indicator for providing early rescue efforts to improve long-term treatment outcomes. Curr Diab Rep 2017 Sep;17(9):69 [FREE Full text] [doi: 10.1007/s11892-017-0904-1] [Medline: 28726155]
20. Pellegrini CA, Pfammatter AF, Conroy DE, Spring B. Smartphone applications to support weight loss: current perspectives. Adv Health Care Technol 2015 Jul;1:13-22. [doi: 10.2147/AHCT.S57844] [Medline: 26236766]
21. James E, Freund M, Booth A, Duncan MJ, Johnson N, Short CE, et al. Comparative efficacy of simultaneous versus sequential multiple health behavior change interventions among adults: a systematic review of randomised trials. Prev Med 2016 Dec;89:211-223. [doi: 10.1016/j.ypmed.2016.06.012] [Medline: 27311332]

22. Ma Y, Olendzki BC, Wang J, Persuitte GM, Li W, Fang H, et al. Single-component versus multicomponent dietary goals for the metabolic syndrome: a randomized trial. Ann Intern Med 2015 Feb 17;162(4):248-257. [doi: 10.7326/M14-0611] [Medline: 25686165]
23. Spring B, Duncan JM, Janke EA, Kozak AT, McFadden HG, DeMott A, et al. Integrating technology into standard weight loss treatment: a randomized controlled trial. JAMA Intern Med 2013 Jan 28;173(2):105-111 [FREE Full text] [doi: 10.1001/jamainternmed.2013.1221] [Medline: 23229890]
24. Carter MC, Burley VJ, Nykjaer C, Cade JE. Adherence to a smartphone application for weight loss compared to website and paper diary: pilot randomized controlled trial. J Med Internet Res 2013;15(4):e32 [FREE Full text] [doi: 10.2196/jmir.2283] [Medline: 23587561]
25. Schippers M, Adam PC, Smolenski DJ, Wong HT, de Wit JB. A meta-analysis of overall effects of weight loss interventions delivered via mobile phones and effect size differences according to delivery mode, personal contact, and intervention intensity and duration. Obes Rev 2017 Feb 10. [doi: 10.1111/obr.12492] [Medline: 28187246]
26. Semper HM, Povey R, Clark-Carter D. A systematic review of the effectiveness of smartphone applications that encourage dietary self-regulatory strategies for weight loss in overweight and obese adults. Obes Rev 2016 Sep;17(9):895-906. [doi: 10.1111/obr.12428] [Medline: 27192162]
27. McCarroll R, Eyles H, Ni MC. Effectiveness of mobile health (mHealth) interventions for promoting healthy eating in adults: a systematic review. Prev Med 2017 Dec;105:156-168. [doi: 10.1016/j.ypmed.2017.08.022] [Medline: 28882743]
28. Burke LE, Zheng Y, Ma Q, Mancino J, Loar I, Music E, et al. The SMARTER pilot study: testing feasibility of real-time feedback for dietary self-monitoring. Prev Med Rep 2017 Jun;6:278-285 [FREE Full text] [doi: 10.1016/j.pmedr.2017.03.017] [Medline: 28409090]
29. Ross KM, Wing RR. Impact of newer self-monitoring technology and brief phone-based intervention on weight loss: a randomized pilot study. Obesity (Silver Spring) 2016 Aug;24(8):1653-1659. [doi: 10.1002/oby.21536] [Medline: 27367614]
30. Thomas JG, Raynor HA, Bond DS, Luke AK, Cardoso CC, Wojtanowski AC, et al. Weight loss and frequency of body-weight self-monitoring in an online commercial weight management program with and without a cellular-connected 'smart' scale: a randomized pilot study. Obes Sci Pract 2017 Dec;3(4):365-372 [FREE Full text] [doi: 10.1002/osp4.132] [Medline: 29259794]
31. Morgan PJ, Callister R, Collins CE, Plotnikoff RC, Young MD, Berry N, et al. The SHED-IT community trial: a randomized controlled trial of internet- and paper-based weight loss programs tailored for overweight and obese men. Ann Behav Med 2013 Apr;45(2):139-152. [doi: 10.1007/s12160-012-9424-z] [Medline: 23129021]
32. Turner-McGrievy GM, Wilcox S, Boutté A, Hutto BE, Singletary C, Muth ER, et al. The Dietary Intervention to Enhance Tracking with Mobile Devices (DIET Mobile) study: a 6-month randomized weight loss trial. Obesity (Silver Spring) 2017 Dec;25(8):1336-1342 [FREE Full text] [doi: 10.1002/oby.21889] [Medline: 28600833]
33. Armstrong MJ, Mottershead TA, Ronksley PE, Sigal RJ, Campbell TS, Hemmelgarn BR. Motivational interviewing to improve weight loss in overweight and/or obese patients: a systematic review and meta-analysis of randomized controlled trials. Obes Rev 2011 Sep;12(9):709-723. [doi: 10.1111/j.1467-789X.2011.00892.x] [Medline: 21692966]
34. Perri MG, Nezu AM, McKelvey WF, Shermer RL, Renjilian DA, Viegener BJ. Relapse prevention training and problem-solving therapy in the long-term management of obesity. J Consult Clin Psychol 2001 Aug;69(4):722-726. [Medline: 11550740]
35. MyFitnessPal. URL: https://www.myfitnesspal.com/ [accessed 2019-02-05] [WebCite Cache ID 75xFgNLZ2]
36. Bardus M, van Beurden SB, Smith JR, Abraham C. A review and content analysis of engagement, functionality, aesthetics, information quality, and change techniques in the most popular commercial apps for weight management. Int J Behav Nutr Phys Act 2016;13(1):35 [FREE Full text] [doi: 10.1186/s12966-016-0359-9] [Medline: 26964880]
37. Mifflin MD, St Jeor ST, Hill LA, Scott BJ, Daugherty SA, Koh YO. A new predictive equation for resting energy expenditure in healthy individuals. Am J Clin Nutr 1990 Feb;51(2):241-247. [doi: 10.1093/ajcn/51.2.241] [Medline: 2305711]
38. Wharton CM, Johnston CS, Cunningham BK, Sterner D. Dietary self-monitoring, but not dietary quality, improves with use of smartphone app technology in an 8-week weight loss trial. J Nutr Educ Behav 2014 Oct;46(5):440-444. [doi: 10.1016/j.jneb.2014.04.291] [Medline: 25220777]
39. Look AHEAD Research Group, Pi-Sunyer X, Blackburn G, Brancati FL, Bray GA, Bright R, et al. Reduction in weight and cardiovascular disease risk factors in individuals with type 2 diabetes: one-year results of the look AHEAD trial. Diabetes Care 2007 Jun;30(6):1374-1383 [FREE Full text] [doi: 10.2337/dc07-0048] [Medline: 17363746]
40. Centers for Disease Control and Prevention. 2018 Jun 19. National Diabetes Prevention Program: PreventT2 Curriculum and Handouts URL: https://www.cdc.gov/diabetes/prevention/lifestyle-program/curriculum.html [accessed 2018-07-26] [WebCite Cache ID 71CV9Ujz1]
41. Pew Research Center. 2012. Pew Internet & American Life Project: Mobile Health 2012 URL: http://www.pewinternet.org/wp-content/uploads/sites/9/media/Files/Reports/2012/PIP_MobileHealth2012_FINAL.pdf [accessed 2018-12-19] [WebCite Cache ID 74m3NPsBF]
42. Foley P, Steinberg D, Levine E, Askew S, Batch BC, Puleo EM, et al. Track: a randomized controlled trial of a digital health obesity treatment intervention for medically vulnerable primary care patients. Contemp Clin Trials 2016 Dec;48:12-20 [FREE Full text] [doi: 10.1016/j.cct.2016.03.006] [Medline: 26995281]

43. Steinberg DM, Tate DF, Bennett GG, Ennett S, Samuel-Hodge C, Ward DS. The efficacy of a daily self-weighing weight loss intervention using smart scales and e-mail. Obesity (Silver Spring) 2013 Sep;21(9):1789-1797 [FREE Full text] [doi: 10.1002/oby.20396] [Medline: 23512320]
44. Laing BY, Mangione CM, Tseng C, Leng M, Vaisberg E, Mahida M, et al. Effectiveness of a smartphone application for weight loss compared with usual care in overweight primary care patients: a randomized, controlled trial. Ann Intern Med 2014 Nov 18;161(10 Suppl):S5-12. [doi: 10.7326/M13-3005] [Medline: 25402403]
45. Jain RB. Regression models to predict corrected weight, height and obesity prevalence from self-reported data: data from BRFSS 1999-2007. Int J Obes (Lond) 2010 Nov;34(11):1655-1664. [doi: 10.1038/ijo.2010.80] [Medline: 20386548]
46. Jensen MD, Ryan DH, Apovian CM, Ard JD, Comuzzie AG, Donato KA, American College of Cardiology/American Heart Association Task Force on Practice Guidelines, Obesity Society. 2013 AHA/ACC/TOS guideline for the management of overweight and obesity in adults: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines and The Obesity Society. J Am Coll Cardiol 2014 Jul 1;63(25 Pt B):2985-3023 [FREE Full text] [doi: 10.1016/j.jacc.2013.11.004] [Medline: 24239920]
47. Wing RR, Lang W, Wadden TA, Safford M, Knowler WC, Bertoni AG, et al. Benefits of modest weight loss in improving cardiovascular risk factors in overweight and obese individuals with type 2 diabetes. Diabetes Care 2011 Jul;34(7):1481-1486 [FREE Full text] [doi: 10.2337/dc10-2415] [Medline: 21593294]
48. Collins CE, Morgan PJ, Hutchesson MJ, Callister R. Efficacy of standard versus enhanced features in a Web-based commercial weight-loss program for obese adults, part 2: randomized controlled trial. J Med Internet Res 2013;15(7):e140 [FREE Full text] [doi: 10.2196/jmir.2626] [Medline: 23876832]
49. 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]
50. Collins LM, Murphy SA, Strecher V. The multiphase optimization strategy (MOST) and the sequential multiple assignment randomized trial (SMART): new methods for more potent eHealth interventions. Am J Prev Med 2007 May;32(5 Suppl):S112-S118 [FREE Full text] [doi: 10.1016/j.amepre.2007.01.022] [Medline: 17466815]
51. Kruglanski AW, Shah JY, Fishbach A, Friedman R, Chun WY, Sleeth-Keppler D. A theory of goal systems. Adv Exp Soc Psychol 2002;34:331-378 [FREE Full text] [doi: 10.1016/S0065-2601(02)80008-9]
52. Dalcin AT, Jerome GJ, Fitzpatrick SL, Louis TA, Wang N, Bennett WL, et al. Perceived helpfulness of the individual components of a behavioural weight loss program: results from the Hopkins POWER Trial. Obes Sci Pract 2015 Dec;1(1):23-32. [doi: 10.1002/osp4.6] [Medline: 27668085]
53. Wing RR, Tate D, LaRose JG, Gorin AA, Erickson K, Robichaud EF, et al. Frequent self-weighing as part of a constellation of healthy weight control practices in young adults. Obesity (Silver Spring) 2015 May;23(5):943-949. [doi: 10.1002/oby.21064] [Medline: 25865175]
54. Johns DJ, Hartmann-Boyce J, Jebb SA, Aveyard P, Behavioural Weight Management Review Group. Diet or exercise interventions vs combined behavioral weight management programs: a systematic review and meta-analysis of direct comparisons. J Acad Nutr Diet 2014 Oct;114(10):1557-1568 [FREE Full text] [doi: 10.1016/j.jand.2014.07.005] [Medline: 25257365]
55. Jospe MR, Roy M, Brown RC, Williams SM, Osborne HR, Meredith-Jones KA, et al. The effect of different types of monitoring strategies on weight loss: a randomized controlled trial. Obesity (Silver Spring) 2017 Sep;25(9):1490-1498. [doi: 10.1002/oby.21898] [Medline: 28703448]
56. Allen JK, Stephens J, Dennison HC, Stewart KJ, Hauck S. Randomized controlled pilot study testing use of smartphone technology for obesity treatment. J Obes 2013;2013:151597 [FREE Full text] [doi: 10.1155/2013/151597] [Medline: 24392223]
57. Turner-McGrievy G, Tate D. Tweets, apps, and pods: results of the 6-month Mobile Pounds Off Digitally (Mobile POD) randomized weight-loss intervention among adults. J Med Internet Res 2011;13(4):e120 [FREE Full text] [doi: 10.2196/jmir.1841] [Medline: 22186428]
58. Rogers R, Lang W, Barone Gibbs B, Davis K, Burke L, Kovacs S, et al. Applying a technology-based system for weight loss in adults with obesity. Obes Sci Pract 2016;2(1):3-12. [doi: 10.1002/osp4.18] [Medline: 27812375]
59. Spring B, Pellegrini CA, Pfammatter A, Duncan JM, Pictor A, McFadden HG, et al. Effects of an abbreviated obesity intervention supported by mobile technology: the ENGAGED randomized clinical trial. Obesity (Silver Spring) 2017 Jul;25(7):1191-1198 [FREE Full text] [doi: 10.1002/oby.21842] [Medline: 28494136]
60. Fukuoka Y, Gay CL, Joiner KL, Vittinghoff E. A novel diabetes prevention intervention using a mobile app: a randomized controlled trial with overweight adults at risk. Am J Prev Med 2015 Aug;49(2):223-237 [FREE Full text] [doi: 10.1016/j.amepre.2015.01.003] [Medline: 26033349]
61. Hartman SJ, Nelson SH, Cadmus-Bertram LA, Patterson RE, Parker BA, Pierce JP. Technology-and phone-based weight loss intervention: pilot RCT in women at elevated breast cancer risk. Am J Prev Med 2016 Nov;51(5):714-721. [doi: 10.1016/j.amepre.2016.06.024] [Medline: 27593420]
62. Block G, Azar KM, Romanelli RJ, Block TJ, Hopkins D, Carpenter HA, et al. Diabetes prevention and weight loss with a fully automated behavioral intervention by email, web, and mobile phone: a randomized controlled trial among persons

with prediabetes. J Med Internet Res 2015 Oct 23;17(10):e240 [FREE Full text] [doi: 10.2196/jmir.4897] [Medline: 26499966]
63. Svetkey LP, Batch BC, Lin PH, Intille SS, Corsino L, Tyson CC, et al. Cell phone intervention for you (CITY): a randomized, controlled trial of behavioral weight loss intervention for young adults using mobile technology. Obesity (Silver Spring) 2015 Nov;23(11):2133-2141. [doi: 10.1002/oby.21226] [Medline: 26530929]

Abbreviations

API: application programming interface
BMI: body mass index
CONSORT: Consolidated Standards of Reporting Trials
IQR: interquartile range

Edited by G Eysenbach; submitted 14.09.18; peer-reviewed by B Turner-McGrievy, B Nezami, S Toukhy, M Bardus; comments to author 29.11.18; revised version received 19.12.18; accepted 06.01.19; published 28.02.19.

Please cite as:
Patel ML, Hopkins CM, Brooks TL, Bennett GG
Comparing Self-Monitoring Strategies for Weight Loss in a Smartphone App: Randomized Controlled Trial
JMIR Mhealth Uhealth 2019;7(2):e12209
URL: https://mhealth.jmir.org/2019/2/e12209/
doi:10.2196/12209
PMID:30816851

Original Paper

Repeated Automated Mobile Text Messaging Reminders for Follow-Up of Positive Fecal Occult Blood Tests: Randomized Controlled Trial

Revital Azulay[1,2*], MSc, PhD; Liora Valinsky[3*], RN, MPH; Fabienne Hershkowitz[3*], PhD; Racheli Magnezi[1*], MHA, PhD

[1]Master of Health Administration Program, Department of Management, Bar Ilan University, Ramat Gan, Israel

[2]Central Laboratory, Meuhedet Health Care, Lod, Israel

[3]Quality Department, Meuhedet Health Care, Tel Aviv, Israel

[*]all authors contributed equally

Corresponding Author:
Revital Azulay, MSc, PhD
Central Laboratory
Meuhedet Health Care
Pesach Lev 5
Lod,
Israel
Phone: 972 97289130421
Fax: 972 97289130422
Email: eazulay39@gmail.com

Abstract

Background: Fecal occult blood tests (FOBTs) are recommended by the US Preventive Services Task Force as a screening method for colorectal cancer (CRC), but they are only effective if positive results are followed by colonoscopy. Surprisingly, a large proportion of patients with a positive result do not follow this recommendation.

Objective: The objective of this study was to examine the effectiveness of text messaging (short message service, SMS) in increasing adherence to colonoscopy follow-up after a positive FOBT result.

Methods: This randomized controlled trial was conducted with patients who had positive CRC screening results. Randomization was stratified by residential district and socioeconomic status (SES). Subjects in the control group (n=238) received routine care that included an alert to the physician regarding the positive FOBT result. The intervention group (n=232) received routine care and 3 text messaging SMS reminders to visit their primary care physician. Adherence to colonoscopy was measured 120 days from the positive result. All patient information, including test results and colonoscopy completion, were obtained from their electronic medical records. Physicians of study patients completed an attitude survey regarding FOBT as a screening test for CRC. Intervention and control group variables (dependent and independent) were compared using chi-square test. Logistic regression was used to calculate odds ratios (ORs) and 95% CIs for performing colonoscopy within 120 days for the intervention group compared with the control group while adjusting for potential confounders including age, gender, SES, district, ethnicity, and physicians' attitude.

Results: Overall, 163 of the 232 patients in the intervention group and 112 of the 238 patients in the control group underwent colonoscopy within 120 days of the positive FOBT results (70.3% vs 47.1%; OR 2.17, 95% CI 1.49-3.17; P<.001); this association remained significant after adjusting for potential confounders (P=.001).

Conclusions: A text message (SMS) reminder is an effective, simple, and inexpensive method for improving adherence among patients with positive colorectal screening results. This type of intervention could also be evaluated for other types of screening tests.

Trial Registration: ClinicalTrials.gov NCT03642652; https://clinicaltrials.gov/ct2/show/NCT03642652 (Archived by WebCite at http://www.webcitation.org/74TlICijl)

(JMIR Mhealth Uhealth 2019;7(2):e11114) doi:10.2196/11114

KEYWORDS

adherence; cancer screening; colonoscopy; fetal occult blood test; patient-physician relationship; positive colorectal cancer screening; SMS; text reminder

Introduction

Colorectal cancer (CRC) is a major cause of morbidity and mortality throughout the world. It is the second most common malignant disease [1,2], with 90% survival when it is identified early and immediate surgical intervention is performed. In Israel, the screening policy for average-risk individuals aged 50-75 years is an annual fecal occult blood test (FOBT) [3]. A patient with positive FOBT result requires immediate follow-up with colonoscopy, and surgery should be performed when CRC is detected. A delay in follow-up markedly undermines the benefits of CRC screening, including incidence, mortality, life-years saved, and net costs of screening [4-7]. Recommendations regarding the time between a positive result and colonoscopy vary across countries, ranging between 30 and 180 days [8-13]. In Israel, the Ministry of Health guidelines define the standard period between a positive FOBT result and a follow-up colonoscopy as 90 days [3,14]. From the literature, we learned that 40%-60% of individuals who undergo FOBT screening do not continue with follow-up after a positive result [7,15,16]. In the Israeli population, follow-up rates after a positive FOBT result are 71%, and the proportion of patients who complete follow-up varies across health care organizations, with the proportion in Meuhedet being about 50%. The median time to follow-up, nationally, is 112 days, which is significantly longer than the recommended 90 days [17].

Barriers to follow-up after a positive FOBT result have been identified in the literature [18-23] and are divided into 4 general groups based on whether they are related to patients, physicians, providers, or information technology [24-26]. Intervention programs have been developed to target each of these groups [27]. Patient-targeted interventions include educational strategies, peer counselors or navigators, reminders, and coupons. The interventions were effective for short-term, but not long-term, follow-up. Several studies have described interventions aimed at increasing follow-up rates and have showed mixed results [13,20,28-41]. When the patient-physician relation is good and based on clear communication and trust, the patient is more likely to adhere to and fulfill the physician's instructions [42,43].

Studies addressing patient interventions have used letters, emails, telephone calls, and nurse navigators. Mailed invitations were found to be as effective as telephone reminders and increased follow-up rates by approximately 30% in an Italian study [40]. In a Scottish study, a reminder phone call that included making an appointment for colonoscopy increased adherence by 4.7% [13]. In a British study aimed at minority populations, the study nurse called patients repeatedly, invited them to the clinic, and scheduled a colonoscopy appointment for each one; this intervention increased follow-up by 8.4% [39]. All these interventions were found to be effective, but some are extremely labor-intensive and time-consuming.

To date, few studies have examined the effectiveness of short message service (SMS) text messaging as an intervention tool within community medical settings. SMS text message reminders have been shown to be effective in improving preparation for colonoscopy in Korea [44]; they increased *Streptococcus pneumonia* immunization in a primary care setting in Lebanon [45] by 7.2% (less than phone calls, but more than emails). A literature review examining the effect of SMS text messages and emails to improve diabetes management showed that simple phone calls, letters, or SMS text message reminders can have a positive impact on clinical and behavioral outcomes [46]. A systematic review [47] indicated that SMS text messaging interventions improved patients' medication adherence rate. In a review of the factors associated with nonadherence to oral antiplatelet therapy in acute coronary syndrome and interventions that modify these factors, only reminder-based interventions, including SMS text messages, had consistently beneficial impacts on adherence outcomes at 3 and 12 months [48]. In a randomized controlled trial designed to assess the effects of SMS text message reminders on adherence to a healthy diet, medication, and smoking cessation among adult patients with cardiovascular disease, researchers found that SMS text messaging was effective in improving adherence to a healthy diet and medication but not smoking cessation [49]. A different study reported increased success for smoking cessation among patients attending control visits as a result of scheduled clinic appointments following SMS text message reminders. The smoking cessation rate was 24% in patients who did not respond to SMS text message reminders at all and 28.6% (n=28) in patients answering any SMS text message at least once (P=.001) [50]. In contrast, in an Australian study, SMS text message reminders were used to improve hepatitis B vaccination among high-risk sexual health center attendees, and it was found that this intervention was not effective [51].

To the best of our knowledge, no study has used SMS text messaging technology to increase adherence to recommendations for CRC screening follow-up after a positive FOBT result. Therefore, we aimed to examine the effectiveness of sending SMS text messages to patients as an automated tool to increase adherence to colonoscopy follow-up and to examine the influence of physicians' attitude toward FOBT as a screening test for CRC and patient adherence to follow-up.

Methods

Methodology

This study was conducted between January 2016 and March 2017 in Meuhedet, 1 of 4 health care organizations in Israel. It insures and provides care for 1.2 million members. The current rate of colorectal screening among 180,000 members aged 50-75 years is 60%, similar to the Israeli population. As per our data, in 2016, the rate of follow-up colonoscopy after a positive FOBT result in Meuhedet was 41%. The FOBT test used in Meuhedet

is OC-SENSOR Immunochemical Kit (Eiken Chemical Co, Ltd, Tokyo, Japan). The study was approved by the Meuhedet Institutional Review Board on March 23, 2016 (trial reference number: 01-023-03-016). The study was exempted from informed consent, and only agreement to receive SMS text messages was required. Figure 1 describes the study flow (Multimedia Appendix 1 presents the CONSORT checklist [52]).

Study Population

In 2016, 3397 patients aged 50-75 years had a positive FOBT result. Of them, 609 were randomly selected and, then, randomly allocated to the intervention or control group. The inclusion criteria included age 50-75 years and providing consent to receive SMS text messages from Meuhedet. By Israeli law, health messages via SMS text messages may only be sent to members who actively agree to receive them. The exclusion criteria were personal or family history of CRC, colonoscopy 10 years before the positive FOBT result, or diagnosis of any type of cancer during the study period.

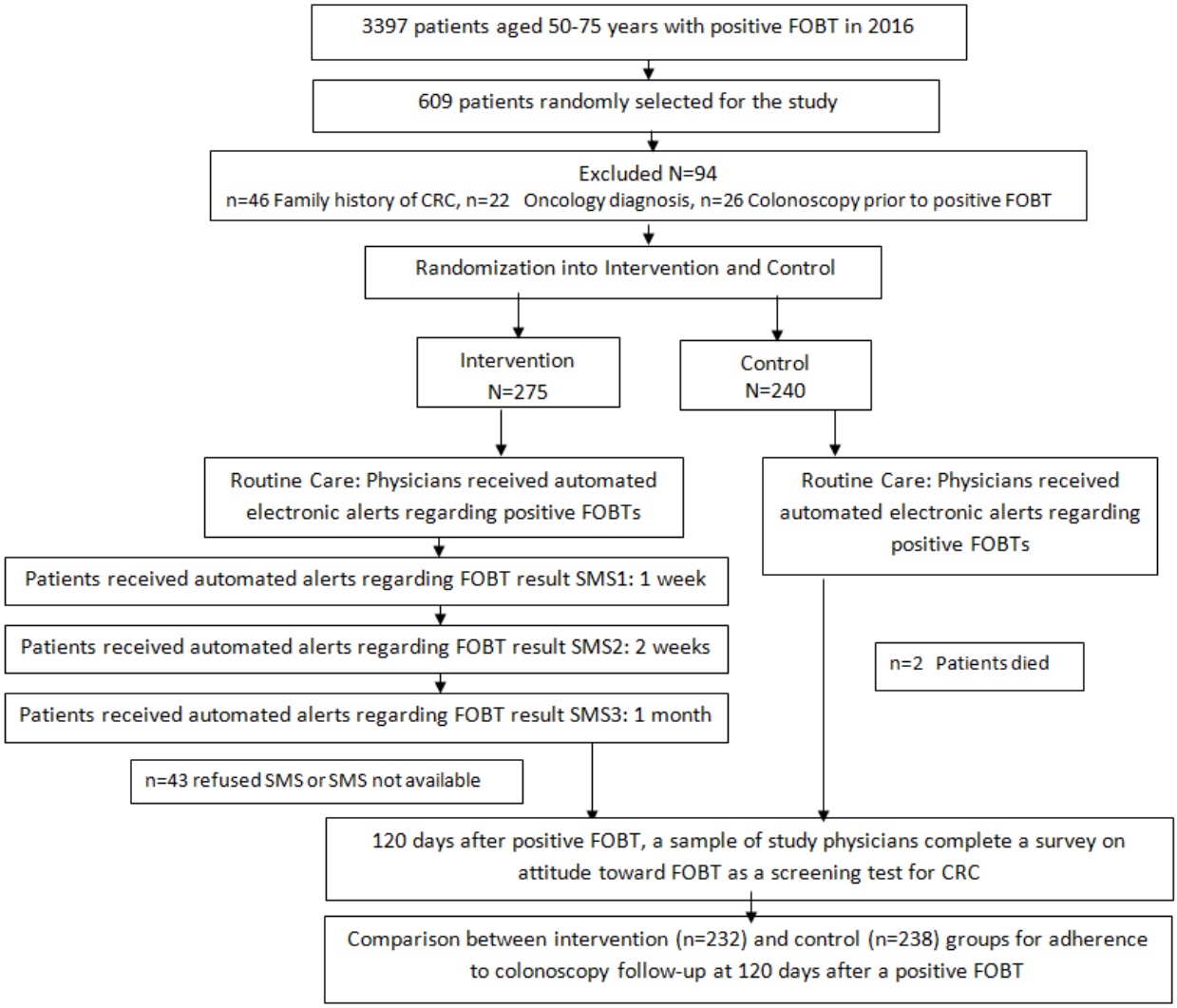
Physicians

All primary care physicians whose patients (control and intervention groups) had completed an FOBT during 2016 and were employed by Meuhedet were included in this study. Physicians no longer working at Meuhedet were excluded from the study. All primary care physicians were notified about the study.

Study Variables

The independent variables were the intervention or control group and physician attitude to FOBT. In addition, we examined potential confounders such as gender, age, socioeconomic status (SES), ethnicity, and residential district. The dependent variable was adherence to colonoscopy within 120 days of a positive FOBT result. This information was obtained from the patients' electronic medical records (EMRs). Positive FOBT results were obtained from the Meuhedet Central Laboratory. Patients' demographics and clinical characteristics including gender, age, district, SES, and ethnicity were obtained from the EMRs. SMS text messages were sent to patients via InforUMobile software (Shamir Systems, 1974, Rishon LeZion) with a track whether the SMS text message was received or rejected.

Figure 1. Study selection and design. FOBT: fecal occult blood test; CRC: colorectal cancer; SMS: short message service.



Study Protocol

Routine care for the control and intervention groups included an automated computer alert in a patient's EMR regarding the positive FOBT result, to his or her physician, with no indication of whether the patient had already visited the physician after the positive result. The automated alert was sent to physicians the moment the lab released the positive result. In addition, each patient in the intervention group received an automated SMS text message up to a week after the positive FOBT result. The text read: "Hello. There is a lab test result ready for you. Contact your physician for an explanation of the findings." Furthermore, 2 additional automated SMS text message reminders were sent to patients after 2 weeks and 1 month, reading, "Hello, This is a reminder. It is essential that you contact your physician if you have not already done so." As the SMS text messages were sent automatically to a randomly selected population, patients in the control group were not aware of the intervention. Physicians were blinded regarding which of their patients were in either group.

With no indication of whether a patient underwent colonoscopy, 120 days after the positive FOBT result, referring primary physicians (for both the control and intervention groups) answered a telephone survey regarding attitude toward FOBT as a screening test for CRC. We waited 120 days so as not to influence physicians' interactions with their patients. We focused on whether FOBT is a reliable screening test for CRC and whether physicians recommend a repeat FOBT after obtaining a positive FOBT result, instead of follow-up, such as colonoscopy (Figure 1). Physicians who stated that FOBT is not a reliable screening test for CRC or reported that they would advise the patient to repeat FOBT instead of sending him or her for colonoscopy were considered as having a negative attitude toward FOBT.

Statistical Analysis

Based on the statistical analysis of the Meuhedet patient database, we found that approximately 41% of patients with a positive FOBT result undergo colonoscopy within 120 days. To have an 80% chance of detecting a 20% increase in follow-up in the intervention group as significant (1-sided, 5% level), a sample of 77 was required in each group. The study sample of 232 in each group provided a 95% ability to detect a 20% increase in adherence to colonoscopy within 120 days.

Randomization was stratified by district and SES derived from members' home address and based on the Israeli Census Bureau locality definitions [53]. SES levels ranged from 1 (low) to 20 (high). We compared demographic variables between the intervention and control groups using chi-square test for discrete variables. In addition, logistic regression was used to calculate odds ratios (ORs) and 95% CIs for the rate of colonoscopy within 120 days of receiving a positive FOBT result, adjusting for the main potential and variables found to be significantly related to adherence in univariate analysis. Data were analyzed using IBM SPSS Statistics for Windows (IBM Corp, 2016, Version 24.0). $P<.05$ was considered significant for all analyses.

Results

Study Population

Of 609 patients randomly selected from the patient database, 94 (15.4%) were excluded from the study sample: 46 (7.6%) because of family history of CRC, 22 (3.6%) because of an oncology diagnosis, and 26 (4.3%) because they had undergone colonoscopy prior to the positive FOBT result. A total of 470 eligible patients were randomized: 232 (49.4%) into the intervention group and 238 (50.6%) into the control group. Of them, 2 patients in the control group died during the study period and 43 (9.1%) in the intervention group refused to receive an SMS text message after their initial approval or were unable to receive SMS text message although they had agreed to receive it. Table 1 summarizes patient characteristics.

Just over half of the final participants were male (52.3%, 246/470), and the mean age of participants was 62.0 (SD 6.6) years. Most participants were from the South district, and the fewest were from the North district. In both groups, most participants were in SES levels 9-13, but this value was missing for 11.9% (56/470) patients. Gender rates were similar between the intervention and control groups, and the groups were ethnically similar. Furthermore, geographic dispersion was similar.

Overall, 163 of 232 patients in the intervention group and 112 of 238 patients in the control group underwent colonoscopy within 120 days of the positive FOBT result (70.3% vs 47.1%; $P<.001$). The unadjusted OR for completion of colonoscopy for the intervention versus control group was 2.17 (95% CI 1.49-3.17; $P<.001$).

Table 2 presents the bivariate (unadjusted) association between patient characteristics and adherence to colonoscopy within 120 days of a positive FOBT result for the entire cohort (N=470). Adherence rates were similar across genders. Adherence rates were higher in the Central and North districts than in the South and Jerusalem districts. Adherence rates were higher among those aged 50-59 and 70-75 years and among those with higher SES levels. Adherence rates to colonoscopy within 120 days of a positive FOBT result were higher among patients who had physicians with a positive attitude toward FOBT than among those who had physicians with a negative attitude toward FOBT (241/399, 60.4% vs 24/53, 45.3%; $P=.04$).

Logistic regression analysis was performed to determine whether colonoscopy rates differed between the groups after adjusting for potential confounders. The adjusted OR for adherence by the intervention group versus control group was 2.9 (95% CI 1.92-4.48, $P=.001$; Table 3). We performed the analysis including the 43 patients who did not receive all 3 SMS text messages ("intention to treat") and found that adherence rates remained significantly higher both in the bivariate analysis (63.5% vs 47.1% in the control group; OR 1.96, 95% CI 1.37-2.79; $P<.001$) and in the multivariable model (OR 2.04, 95% CI 1.387-2.993; $P<.001$).

Table 1. Patient characteristics.

Characteristic	Control group (n=238), n (%)	Intervention group (n=232), n (%)	Total (N)
Gender			
Male	125 (52.5)	121 (52.2)	246
Female	113 (47.5)	111 (47.8)	224
Age (years)			
50-55	41 (17.2)	45 (19.5)	86
55-59	40 (16.8)	39 (16.8)	79
60-64	63 (26.5)	53 (22.8)	116
65-69	66 (27.7)	66 (28.4)	132
70-75	28 (11.8)	29 (12.5)	57
Ethnicity			
Jewish	237 (99.6)	227 (97.8)	464
Other	1 (0.4)	5 (2.2)	6
National district			
South	83 (34.9)	99 (42.7)	182
Jerusalem	60 (25.2)	49 (21.1)	109
Center	50 (21.0)	46 (19.8)	96
North	45 (18.9)	38 (16.4)	83
Socioeconomic status level[a]			
1-8	13 (6.3)	20 (9.6)	33
9-13	116 (56.6)	122 (58.4)	238
14-20	76 (37.1)	67 (32.0)	143

[a]n=414; data missing for 56 patients.

Table 2. Adherence within 120 days after a positive fecal occult blood test result for individual variables.

Variable	Did not adhere to colonoscopy[a] (n=195), n (%)	Adhered to colonoscopy[b] (n=275), n (%)	*P* value
Intervention versus control			<.001
Intervention	69 (29.7)	163 (70.3)	
Control	126 (52.9)	112 (47.1)	
Gender			.26
Male	96 (39)	150 (61)	
Female	99 (44.2)	125 (55.8)	
Age (years)	195 (41.5)	275 (58.5)	.91
50-54	35 (40.7)	51 (59.3)	
55-59	32 (40.5)	47 (59.5)	
60-64	49 (42.2)	67 (57.8)	
65-69	56 (42.4)	76 (57.6)	
70-75	23 (40.4)	34 (59.6)	
District	195 (41.5)	275 (58.5)	.36
South	79 (43.4)	103 (56.6)	
Jerusalem	51 (46.8)	58 (53.2)	
Center	33 (34.4)	63 (65.6)	
North	32 (38.6)	51 (61.4)	
Socioeconomic status level[c]			.10
1-8	18 (54.5)	15 (45.5)	
9-13	100 (42.0)	138 (58.0)	
14-20	58 (40.6)	85 (59.4)	
Physician attitude[d]			.04
Positive attitude	158 (39.6)	241 (60.4)	
Negative attitude	29 (54.7)	24 (45.3)	

[a]Of all, 41.5% patients did not adhere to colonoscopy.

[b]Of all, 58.5% patients adhered to colonoscopy.

[c]Data missing for 56 patients.

[d]Data missing for 18 patients.

Physicians

Among 282 primary physicians who referred participants to FOBT, 267 (94.7%) were interviewed, 4 (1.4%) refused, 9 (3.2%) could not be contacted, and 2 (0.9%) had left the organization. Most (83.5%, 223/267) physicians had a positive attitude toward FOBT as a tool for early detection of CRC. The adjusted OR for patients who had a physician with a positive attitude toward FOBT versus a negative attitude was 2.7 (95% CI 1.38-5.33; *P*=.004; Table 3). As some patients were referred for FOBT by the same primary care physician, we performed a binomial mixed model analysis to assess the impact of physician-level clustering. No difference was found, and the OR of the intervention group remained the same compared with that of the control group (2.9).

We conducted the intention-to-treat analysis including the 43 patients who did not receive the SMS text message and found that adherence rates remained significantly higher both in the bivariate (63.5% vs 47.1% in the control group; OR 1.96, 95% CI 1.37-2.79; *P*<.001) and the multivariable (OR 2.04, 95% CI 1.387-2.993; *P*<.001) analyses. As we only contacted physicians of patients who completed the study, we were unable to run the full multivariable model (including physician attitude).

Table 3. Multivariable analysis of colonoscopy rates.

Variable	Odds ratio	95% CI	*P* value
Intervention vs Control	2.93	1.92-4.48	<.001
Age (continuous)	1.01	0.98-1.04	.48
Female versus Male	0.81	0.53-1.22	.31
Socioeconomic status (continuous, 1-20)	0.99	0.92-1.07	.91
South district			
Jerusalem district	1.26	0.72-2.20	.41
Center district	1.92	1.04-3.55	.04
North district	1.59	0.87-2.91	.13
FOBT[a] attitude: Positive versus Negative	2.72	1.38-5.33	.004
Constant	0.09	—[b]	.15

[a]FOBT: fecal occult blood test.

[b]Not applicable.

Discussion

The effectiveness of mass cancer screening programs can be compromised by lack of follow-up of abnormal findings. Incomplete FOBT follow-up with colonoscopy is a significant problem that has been studied extensively. This study shows that sending SMS text message reminders to patients following a positive FOBT result is an effective way to increase adherence rates to follow-up colonoscopy. In this study, there was a relative increase of 49.2% in adherence in the intervention group using a simple, inexpensive means of communication. This surprisingly large increase in adherence could be a result of a combination of simplicity, repetition, and timeliness. Reminders were simply worded, with a clear message, sent immediately after results were obtained, and then repeated twice over the next month. Another potential advantage is that unlike telephone calls, which require active responses, the messages are "pushed." Probably, patients who responded to the message were those who were more likely to complete the follow-up but required a "nudge." Another possibility is that the message created cognitive dissonance in some patients who chose to ignore the results until they received the reminder. Future research should focus on how different contact methods are effective for different types of patients.

In this study, age, gender, SES, and geographic location were not significantly associated with adherence to follow-up of positive FOBT results either in the bivariate or multivariable analysis. Increasing age was previously associated with lack of follow-up in some studies [39], but not in others [21]. Similar to our findings, several studies did not find strong associations between gender and complete diagnostic follow-up [21,29,54]. However, others suggested that women are less likely than men to undergo follow-up testing [16,25,55]. Although the association between SES and follow-up was not significant, there appears to be a trend toward increasing follow-up rates at higher SES. This needs to be investigated further.

The physician plays an important role in a patient's decision regarding follow-up tests. Trust in the physician and good communication between the patient and physician will positively influence the adherence rates of patients after a positive FOBT result. In this study, 95.1% (254/267) of physicians declared that FOBT is a reliable screening test for CRC and that they would not advise their patients to undergo FOBT again to ensure that the positive result is reliable. The physician's attitude toward FOBT as a screening test significantly influences the patient's adherence to colonoscopy.

A new finding of this study is that physicians' attitude toward FOBT has a major influence on the rates of colonoscopy after a positive FOBT result. This finding may explain why intervention studies that provide physicians with knowledge and feedback have been effective. For example, Myers et al [56] showed that one-on-one physician training, audit, and feedback (physicians received lists of their patients with incomplete diagnostic evaluations) resulted in improved completion of diagnostic testing. Singh et al [57] assessed a clinic-based quality improvement activity that included provider education, a positive FOBT registry, and feedback; they found that it significantly decreased the time to colonoscopy referral and completion and increased colonoscopy completion by 18.7%. In spite of the fact that FOBT has been a standard practice for early detection of CRC, >10% of physicians in our study expressed mistrust in this method; this is possibly an underestimate, as some physicians may hesitate to express this view openly, which is contrary to organizational guidelines. It is, therefore, essential to educate physicians regarding the reliability and effectiveness of FOBT as a screening tool for CRC.

Owing to recent developments that have made digital communication within the health care sector readily available and inexpensive, SMS text message reminders could potentially be implemented in other areas. The finding of a large increase in follow-up after a positive FOBT result illustrates the need to further investigate different aspects of SMS text message usage, such as patients' age, gender, the type of action they aim to increase, and the wording of the message. Changing technology and patient preferences with regard to contact communication

should be considered when determining future interventions to improve usage and effectiveness.

A limitation of this study is that only patients who agreed to receive SMS text messages were included, which created a biased population in terms of age, cultural beliefs, and SES. An additional limitation is that SES was measured using zip code rather than a direct measure, and the SES of 56 participants was missing, although this is the standard method for measuring SES in Israel.

This study was conducted in 1 of the 4 health care provider organizations in Israel. Selection of providers was voluntary, and the member distribution in terms of age, gender, and SES was similar to that of the Israeli population. In addition, the Israeli population is very ethnically diverse and includes immigrants from many countries. Our findings are potentially generalizable to other populations.

In conclusion, this study is the first to directly evaluate SMS text message reminders for improving colonoscopy follow-up among Israeli CRC screening program participants following a positive FOBT result. We have shown that a simple, inexpensive intervention for patients improves colonoscopy follow-up after a positive FOBT result. It is important to maximize the potential of these findings by increasing the acceptance of SMS text messages within the population and to examine their use in other screening programs. In addition, it is important to examine, in future studies, the reasons because of which patients refuse to receive SMS text messages. The physician's attitude toward FOBT as a screening test significantly influences patient adherence to colonoscopy. Therefore, further work needs to be done among physicians to increase adherence.

Acknowledgments

We thank Professor Haim Reuveni for his help throughout the research process. This study was written as a PhD thesis at Bar Ilan University, Public Health and Health Systems Management Program, Department of Management, under the direction of Racheli Magnezi, PhD, MBA, MHA, and Professor Haim Reuveni. The study was approved on February 23, 2016.

Conflicts of Interest

None declared.

Multimedia Appendix 1

CONSORT - EHEALTH checklist (V 1.6.1).

[PDF File (Adobe PDF File), 134KB - mhealth_v7i2e11114_app1.pdf]

References

1. Ministry of Health, State of Israel. 2018. Colorectal Cancer in Israel: Updated March 2016 URL: https://www.health.gov.il/PublicationsFiles/ICR_04032016.pdf [accessed 2018-05-18] [WebCite Cache ID 74W4CDJtR]
2. Global Cancer Observatory. 2018. Colorectal cancer URL: http://gco.iarc.fr/today/data/factsheets/cancers/10_8_9-Colorectum-fact-sheet.pdf [accessed 2018-05-18] [WebCite Cache ID 74W4xP0Xe]
3. Ministry of Health, State of Israel. 2018. Prevention and early detection of malignant diseases URL: https://www.health.gov.il/hozer/mk08_2011.pdf [accessed 2018-05-18] [WebCite Cache ID 74W57kxxe]
4. Meester RGS, Doubeni CA, Lansdorp-Vogelaar I, Goede SL, Levin TR, Quinn VP, et al. Colorectal cancer deaths attributable to nonuse of screening in the United States. Ann Epidemiol 2015 Mar;25(3):208-213.e1 [FREE Full text] [doi: 10.1016/j.annepidem.2014.11.011] [Medline: 25721748]
5. Chubak J, Garcia MP, Burnett-Hartman AN, Zheng Y, Corley DA, Halm EA, PROSPR consortium. Time to Colonoscopy after Positive Fecal Blood Test in Four U.S. Health Care Systems. Cancer Epidemiol Biomarkers Prev 2016 Feb;25(2):344-350 [FREE Full text] [doi: 10.1158/1055-9965.EPI-15-0470] [Medline: 26843520]
6. Etzioni DA, Yano EM, Rubenstein LV, Lee ML, Ko CY, Brook RH, et al. Measuring the quality of colorectal cancer screening: the importance of follow-up. Dis Colon Rectum 2006 Jul;49(7):1002-1010. [doi: 10.1007/s10350-006-0533-2] [Medline: 16673056]
7. Meester RGS, Zauber AG, Doubeni CA, Jensen CD, Quinn VP, Helfand M, et al. Consequences of Increasing Time to Colonoscopy Examination After Positive Result From Fecal Colorectal Cancer Screening Test. Clin Gastroenterol Hepatol 2016 Dec;14(10):1445-1451.e8 [FREE Full text] [doi: 10.1016/j.cgh.2016.05.017] [Medline: 27211498]
8. Liss DT, Brown T, Lee JY, Altergott M, Buchanan DR, Newland A, et al. Diagnostic colonoscopy following a positive fecal occult blood test in community health center patients. Cancer Causes Control 2016 Dec;27(7):881-887. [doi: 10.1007/s10552-016-0763-0] [Medline: 27228991]
9. Oluloro A, Petrik AF, Turner A, Kapka T, Rivelli J, Carney PA, et al. Timeliness of Colonoscopy After Abnormal Fecal Test Results in a Safety Net Practice. J Community Health 2016 Dec;41(4):864-870 [FREE Full text] [doi: 10.1007/s10900-016-0165-y] [Medline: 26874943]

10. Paterson WG, Depew WT, Paré P, Petrunia D, Switzer C, Veldhuyzen van Zanten SJ, Canadian Association of Gastroenterology Wait Time Consensus Group. Canadian consensus on medically acceptable wait times for digestive health care. Can J Gastroenterol 2006 Jun;20(6):411-423 [FREE Full text] [Medline: 16779459]
11. Viiala CH, Tang KW, Lawrance IC, Murray K, Olynyk JK. Waiting times for colonoscopy and colorectal cancer diagnosis. Med J Aust 2007 Mar 19;186(6):282-285. [Medline: 17371207]
12. Lee HY, Choi KS, Jun JK, Hahm MI, Park EC. Time to follow up after an abnormal finding in organized gastric cancer screening in Korea. BMC Cancer 2012 Sep 10;12:400 [FREE Full text] [doi: 10.1186/1471-2407-12-400] [Medline: 22963347]
13. Steele RJC, McClements PL, Libby G, Black R, Morton C, Birrell J, et al. Results from the first three rounds of the Scottish demonstration pilot of FOBT screening for colorectal cancer. Gut 2009 Apr;58(4):530-535. [doi: 10.1136/gut.2008.162883] [Medline: 19036949]
14. Ministry of Health, State of Israel. Occult Blood Test-CRC Screening Program in Israel URL: https://www.health.gov.il/hozer/mr12_2014.pdf [accessed 2018-05-18] [WebCite Cache ID 74W5Doh5o]
15. Thamarasseril S, Bhuket T, Chan C, Liu B, Wong RJ. The Need for an Integrated Patient Navigation Pathway to Improve Access to Colonoscopy After Positive Fecal Immunochemical Testing: A Safety-Net Hospital Experience. J Community Health 2017 Jun;42(3):551-557. [doi: 10.1007/s10900-016-0287-2] [Medline: 27796633]
16. Singh H, Kadiyala H, Bhagwath G, Shethia A, El-Serag H, Walder A, et al. Using a multifaceted approach to improve the follow-up of positive fecal occult blood test results. Am J Gastroenterol 2009 Apr;104(4):942-952 [FREE Full text] [doi: 10.1038/ajg.2009.55] [Medline: 19293786]
17. The Israel National Institute for Health Policy Research. 2015. National Program for Quality Indicators in Community Health care 2013-2015 URL: http://www.israelhpr.org.il/1043/470.htm [accessed 2018-12-11] [WebCite Cache ID 74aPhvCi4]
18. Fiscella K, Humiston S, Hendren S, Winters P, Jean-Pierre P, Idris A, et al. Eliminating disparities in cancer screening and follow-up of abnormal results: what will it take? J Health Care Poor Underserved 2011 Feb;22(1):83-100 [FREE Full text] [doi: 10.1353/hpu.2011.0023] [Medline: 21317508]
19. Emery J, Trevena L, Mazza D, Fallon-Ferguson J, Shaw K, Williams B, et al. Sax Institute. 2012 Dec. The role of primary and communitybased healthcare professionals in early detection and follow-up in cancer care: a rapid review of best practice models URL: https://www.saxinstitute.org.au/wp-content/uploads/REPORT_Role-of-PHC-cancer-early-detection-2smallpdf.com_.pdf [accessed 2018-05-18] [WebCite Cache ID 6zVCoTkp2]
20. Martin J, Halm EA, Tiro JA, Merchant Z, Balasubramanian BA, McCallister K, et al. Reasons for Lack of Diagnostic Colonoscopy After Positive Result on Fecal Immunochemical Test in a Safety-Net Health System. Am J Med 2017 Jan;130(1):93.e1-93.e7 [FREE Full text] [doi: 10.1016/j.amjmed.2016.07.028] [Medline: 27591183]
21. Ferrat E, Le Breton J, Veerabudun K, Bercier S, Brixi Z, Khoshnood B, et al. Colorectal cancer screening: factors associated with colonoscopy after a positive faecal occult blood test. Br J Cancer 2013 Sep 17;109(6):1437-1444 [FREE Full text] [doi: 10.1038/bjc.2013.476] [Medline: 23989948]
22. Plumb AA, Ghanouni A, Rainbow S, Djedovic N, Marshall S, Stein J, et al. Patient factors associated with non-attendance at colonoscopy after a positive screening faecal occult blood test. J Med Screen 2017 Dec;24(1):12-19. [doi: 10.1177/0969141316645629] [Medline: 27216771]
23. Laing SS, Bogart A, Chubak J, Fuller S, Green BB. Psychological distress after a positive fecal occult blood test result among members of an integrated healthcare delivery system. Cancer Epidemiol Biomarkers Prev 2014 Jan;23(1):154-159 [FREE Full text] [doi: 10.1158/1055-9965.EPI-13-0722] [Medline: 24220914]
24. Humphrey LL, Shannon J, Partin MR, O'Malley J, Chen Z, Helfand M. Improving the follow-up of positive hemoccult screening tests: an electronic intervention. J Gen Intern Med 2011 Jul;26(7):691-697 [FREE Full text] [doi: 10.1007/s11606-011-1639-3] [Medline: 21327529]
25. Miglioretti DL, Rutter CM, Bradford SC, Zauber AG, Kessler LG, Feuer EJ, et al. Improvement in the diagnostic evaluation of a positive fecal occult blood test in an integrated health care organization. Med Care 2008 Sep;46(9 Suppl 1):S91-S96 [FREE Full text] [doi: 10.1097/MLR.0b013e31817946c8] [Medline: 18725839]
26. Partin MR, Burgess DJ, Burgess JF, Gravely A, Haggstrom D, Lillie SE, et al. Organizational predictors of colonoscopy follow-up for positive fecal occult blood test results: an observational study. Cancer Epidemiol Biomarkers Prev 2015 Feb;24(2):422-434 [FREE Full text] [doi: 10.1158/1055-9965.EPI-14-1170] [Medline: 25471345]
27. Yabroff KR, Washington KS, Leader A, Neilson E, Mandelblatt J. Is the promise of cancer-screening programs being compromised? Quality of follow-up care after abnormal screening results. Med Care Res Rev 2003 Sep;60(3):294-331. [doi: 10.1177/1077558703254698] [Medline: 12971231]
28. Callen JL, Westbrook JI, Georgiou A, Li J. Failure to follow-up test results for ambulatory patients: a systematic review. J Gen Intern Med 2012 Oct;27(10):1334-1348 [FREE Full text] [doi: 10.1007/s11606-011-1949-5] [Medline: 22183961]
29. Shields HM, Weiner MS, Henry DR, Lloyd JA, Ransil BJ, Lamphier DA, et al. Factors that influence the decision to do an adequate evaluation of a patient with a positive stool for occult blood. Am J Gastroenterol 2001 Jan;96(1):196-203. [doi: 10.1111/j.1572-0241.2001.03475.x] [Medline: 11197252]

30. Taplin SH, Yabroff KR, Zapka J. A multilevel research perspective on cancer care delivery: the example of follow-up to an abnormal mammogram. Cancer Epidemiol Biomarkers Prev 2012 Oct;21(10):1709-1715 [FREE Full text] [doi: 10.1158/1055-9965.EPI-12-0265] [Medline: 22911332]
31. Stock D, Rabeneck L, Baxter NN, Paszat LF, Sutradhar R, Yun L, et al. Mailed participant reminders are associated with improved colonoscopy uptake after a positive FOBT result in Ontario's ColonCancerCheck program. Implement Sci 2015 Mar 13;10:35 [FREE Full text] [doi: 10.1186/s13012-015-0226-0] [Medline: 25885531]
32. Chen ZJ, Kammer D, Bond JH, Ho SB. Evaluating follow-up of positive fecal occult blood test results: lessons learned. J Healthc Qual 2007;29(5):16-20, 34. [Medline: 17892078]
33. Zapka J, Taplin SH, Price RA, Cranos C, Yabroff R. Factors in quality care--the case of follow-up to abnormal cancer screening tests--problems in the steps and interfaces of care. J Natl Cancer Inst Monogr 2010;2010(40):58-71 [FREE Full text] [doi: 10.1093/jncimonographs/lgq009] [Medline: 20386054]
34. Bastani R, Yabroff KR, Myers RE, Glenn B. Interventions to improve follow-up of abnormal findings in cancer screening. Cancer 2004 Sep 01;101(5 Suppl):1188-1200 [FREE Full text] [doi: 10.1002/cncr.20506] [Medline: 15316914]
35. Green BB, Wang CY, Horner K, Catz S, Meenan RT, Vernon SW, et al. Systems of support to increase colorectal cancer screening and follow-up rates (SOS): design, challenges, and baseline characteristics of trial participants. Contemp Clin Trials 2010 Nov;31(6):589-603 [FREE Full text] [doi: 10.1016/j.cct.2010.07.012] [Medline: 20674774]
36. Lee J, Fulp W, Wells KJ, Meade CD, Calcano E, Roetzheim R. Effect of patient navigation on time to diagnostic resolution among patients with colorectal cancer-related abnormalities. J Cancer Educ 2014 Mar;29(1):144-150 [FREE Full text] [doi: 10.1007/s13187-013-0561-2] [Medline: 24113902]
37. Green BB, Anderson ML, Wang C, Vernon SW, Chubak J, Meenan RT, et al. Results of nurse navigator follow-up after positive colorectal cancer screening test: a randomized trial. J Am Board Fam Med 2014;27(6):789-795 [FREE Full text] [doi: 10.3122/jabfm.2014.06.140125] [Medline: 25381076]
38. Zheng Y, Saito T, Takahashi M, Ishibashi T, Kai I. Factors associated with intentions to adhere to colorectal cancer screening follow-up exams. BMC Public Health 2006 Nov 06;6:272 [FREE Full text] [doi: 10.1186/1471-2458-6-272] [Medline: 17083746]
39. Morris S, Baio G, Kendall E, von Wagner C, Wardle J, Atkin W, et al. Socioeconomic variation in uptake of colonoscopy following a positive faecal occult blood test result: a retrospective analysis of the NHS Bowel Cancer Screening Programme. Br J Cancer 2012 Aug 21;107(5):765-771 [FREE Full text] [doi: 10.1038/bjc.2012.303] [Medline: 22864455]
40. Zorzi M, Giorgi Rossi P, Cogo C, Falcini F, Giorgi D, Grazzini G, PARC Working Group. A comparison of different strategies used to invite subjects with a positive faecal occult blood test to a colonoscopy assessment. A randomised controlled trial in population-based screening programmes. Prev Med 2014 Aug;65:70-76. [doi: 10.1016/j.ypmed.2014.04.022] [Medline: 24811759]
41. Chiu H, Hung H, Lin H, Chen S. Effects of a health education and telephone counseling program on patients with a positive fecal occult blood test result for colorectal cancer screening: A randomized controlled trial. Psychooncology 2017 Oct;26(10):1498-1504. [doi: 10.1002/pon.4319] [Medline: 27859897]
42. Gupta S, Brenner AT, Ratanawongsa N, Inadomi JM. Patient trust in physician influences colorectal cancer screening in low-income patients. Am J Prev Med 2014 Oct;47(4):417-423 [FREE Full text] [doi: 10.1016/j.amepre.2014.04.020] [Medline: 25084682]
43. Hudson SV, Ferrante JM, Ohman-Strickland P, Hahn KA, Shaw EK, Hemler J, et al. Physician recommendation and patient adherence for colorectal cancer screening. J Am Board Fam Med 2012;25(6):782-791 [FREE Full text] [doi: 10.3122/jabfm.2012.06.110254] [Medline: 23136316]
44. Park J, Kim T, Lee N, Kim H, Seo EH, Heo N, et al. The Effectiveness of Short Message Service to Assure the Preparation-to-Colonoscopy Interval before Bowel Preparation for Colonoscopy. Gastroenterol Res Pract 2015;2015:628049 [FREE Full text] [doi: 10.1155/2015/628049] [Medline: 25792978]
45. Ghadieh AS, Hamadeh GN, Mahmassani DM, Lakkis NA. The effect of various types of patients' reminders on the uptake of pneumococcal vaccine in adults: A randomized controlled trial. Vaccine 2015 Oct 26;33(43):5868-5872. [doi: 10.1016/j.vaccine.2015.07.050] [Medline: 26232345]
46. Nuti L, Turkcan A, Lawley MA, Zhang L, Sands L, McComb S. The impact of interventions on appointment and clinical outcomes for individuals with diabetes: a systematic review. BMC Health Serv Res 2015 Sep 02;15:355 [FREE Full text] [doi: 10.1186/s12913-015-0938-5] [Medline: 26330299]
47. Ershad Sarabi R, Sadoughi F, Jamshidi Orak R, Bahaadinbeigy K. The Effectiveness of Mobile Phone Text Messaging in Improving Medication Adherence for Patients with Chronic Diseases: A Systematic Review. Iran Red Crescent Med J 2016 May;18(5):e25183 [FREE Full text] [doi: 10.5812/ircmj.25183] [Medline: 27437126]
48. Johnston N, Weinman J, Ashworth L, Smethurst P, El Khoury J, Moloney C. Systematic reviews: causes of non-adherence to P2Y12 inhibitors in acute coronary syndromes and response to intervention. Open Heart 2016;3(2):e000479 [FREE Full text] [doi: 10.1136/openhrt-2016-000479] [Medline: 27843565]
49. Akhu-Zaheya LM, Shiyab WY. The effect of short message system (SMS) reminder on adherence to a healthy diet, medication, and cessation of smoking among adult patients with cardiovascular diseases. Int J Med Inform 2017 Dec;98:65-75. [doi: 10.1016/j.ijmedinf.2016.12.003] [Medline: 28034414]

50. Önür ST, Uysal MA, İliaz S, Yurt S, Bahadır A, Hattatoğlu DG, et al. Does Short Message Service Increase Adherence to Smoking Cessation Clinic Appointments and Quitting Smoking? Balkan Med J 2016 Sep;33(5):525-531 [FREE Full text] [doi: 10.5152/balkanmedj.2016.151610] [Medline: 27761280]
51. McIver R, Dyda A, McNulty AM, Knight V, Wand HC, Guy RJ. Text message reminders do not improve hepatitis B vaccination rates in an Australian sexual health setting. J Am Med Inform Assoc 2016 Apr;23(e1):e88-e92 [FREE Full text] [doi: 10.1093/jamia/ocv145] [Medline: 26499103]
52. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. J Med Internet Res 2011 Dec 31;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]
53. State of Israel: Central Bureau of Statistics. 2018. URL: http://www.cbs.gov.il/reader/ [accessed 2018-05-18] [WebCite Cache ID 74W5JQkF8]
54. Gellad ZF, Almirall D, Provenzale D, Fisher DA. Time from positive screening fecal occult blood test to colonoscopy and risk of neoplasia. Dig Dis Sci 2009 Nov;54(11):2497-2502 [FREE Full text] [doi: 10.1007/s10620-008-0653-8] [Medline: 19093199]
55. Turner B, Myers RE, Hyslop T, Hauck WW, Weinberg D, Brigham T, et al. Physician and patient factors associated with ordering a colon evaluation after a positive fecal occult blood test. J Gen Intern Med 2003 May;18(5):357-363 [FREE Full text] [Medline: 12795734]
56. Myers RE, Turner B, Weinberg D, Hyslop T, Hauck WW, Brigham T, et al. Impact of a physician-oriented intervention on follow-up in colorectal cancer screening. Prev Med 2004 Apr;38(4):375-381. [doi: 10.1016/j.ypmed.2003.11.010] [Medline: 15020170]
57. Singh H, Petersen LA, Daci K, Collins C, Khan M, El-Serag HB. Reducing referral delays in colorectal cancer diagnosis: is it about how you ask? Qual Saf Health Care 2010 Oct;19(5):e27 [FREE Full text] [doi: 10.1136/qshc.2009.033712] [Medline: 20584706]

Abbreviations

CRC: colorectal cancer
EMR: electronic medical record
FOBT: fecal occult blood test
OR: odds ratio
SES: socioeconomic status
SMS: short message service

Edited by G Eysenbach; submitted 23.05.18; peer-reviewed by C Rutter, D Liss, J Chubak, M Anderson; comments to author 07.08.18; revised version received 25.08.18; accepted 26.10.18; published 05.02.19.

Please cite as:
Azulay R, Valinsky L, Hershkowitz F, Magnezi R
Repeated Automated Mobile Text Messaging Reminders for Follow-Up of Positive Fecal Occult Blood Tests: Randomized Controlled Trial
JMIR Mhealth Uhealth 2019;7(2):e11114
URL: http://mhealth.jmir.org/2019/2/e11114/
doi:10.2196/11114
PMID:30720439

Original Paper

Popular Nutrition-Related Mobile Apps: An Agreement Assessment Against a UK Reference Method

Rosalind Fallaize[1*], PhD; Rodrigo Zenun Franco[2*], PhD; Jennifer Pasang[1], BSc; Faustina Hwang[2], PhD; Julie A Lovegrove[1], PhD

[1]Hugh Sinclair Unit of Human Nutrition and Institute for Cardiovascular and Metabolic Research, Department of Food and Nutritional Sciences, University of Reading, Reading, United Kingdom

[2]Department of Biomedical Engineering, School of Biological Sciences, University of Reading, Reading, United Kingdom

[*]these authors contributed equally

Corresponding Author:
Julie A Lovegrove, PhD
Hugh Sinclair Unit of Human Nutrition and Institute for Cardiovascular and Metabolic Research
Department of Food and Nutritional Sciences
University of Reading
Whiteknights Campus
Reading,
United Kingdom
Phone: 44 1183786418
Email: j.a.lovegrove@reading.ac.uk

Abstract

Background: Nutrition-related apps are commonly used to provide information about the user's dietary intake, but limited research has been performed to assess how well their outputs agree with those from standard methods.

Objective: The objective of our study was to evaluate the level of agreement of popular nutrition-related apps for the assessment of energy and available macronutrients and micronutrients against a UK reference method.

Methods: We compared dietary analysis of 24-hour weighed food records (n=20) between 5 nutrition-related apps (Samsung Health, MyFitnessPal, FatSecret, Noom Coach, and Lose It!) and Dietplan6 (reference method), using app versions available in the United Kingdom. We compared estimates of energy, macronutrients (carbohydrate, protein, fat, saturated fat, and fiber), and micronutrients (sodium, calcium, iron, vitamin A, and vitamin C) using paired *t* tests and Wilcoxon signed-rank tests, correlation coefficients, and Bland-Altman plots. We obtained 24-hour weighed food records from 20 participants (15 female, 5 male participants; mean age 36.3 years; mean body mass index 22.9 kg/m^2) from previous controlled studies conducted at the Hugh Sinclair Unit of Human Nutrition, University of Reading, Reading, UK. Participants had recorded their food consumption over a 24-hour period using standard protocols.

Results: The difference in estimation of energy and saturated fat intake between Dietplan6 and the diet apps was not significant. Estimates of protein and sodium intake were significantly lower using Lose It! and FatSecret than using Dietplan6. Lose It! also gave significantly lower estimates for other reported outputs (carbohydrate, fat, fiber, and sodium) than did Dietplan6. Samsung Health and MyFitnessPal significantly underestimated calcium, iron, and vitamin C compared with Dietplan6, although there was no significant difference for vitamin A. We observed no other significant differences between Dietplan6 and the apps. Correlation coefficients ranged from *r*=–.12 for iron (Samsung Health vs Dietplan6) to *r*=.91 for protein (FatSecret vs Dietplan6). Noom Coach was limited to energy output, but it had a high correlation with Dietplan6 (*r*=.91). Samsung Health had the greatest variation of correlation, with energy at *r*=.79. Bland-Altman analysis revealed potential proportional bias for vitamin A.

Conclusions: The findings suggest that the apps provide estimates of energy and saturated fat intake comparable with estimates by Dietplan6. With the exception of Lose It!, the apps also provided comparable estimates of carbohydrate, total fat, and fiber. FatSecret and Lose It! tended to underestimate protein and sodium. Estimates of micronutrient intake (calcium, iron, vitamin A, and vitamin C) by 2 apps (Samsung Health and MyFitnessPal) were inconsistent and less reliable. Lose It! was the app least comparable with Dietplan6. As the use and availability of apps grows, this study helps clinicians and researchers to make better-informed decisions about using these apps in research and practice.

(JMIR Mhealth Uhealth 2019;7(2):e9838) doi:10.2196/mhealth.9838

KEYWORDS

weighed food records; smartphone application; dietary assessment; diet apps; nutrition apps; diet records; mobile applications

Introduction

Background

The advancement of technology has led to the development of novel electronic dietary assessment methods, for example, in the form of nutrition-related apps, which are commonly used on mobile phones. These apps provide information about the user's overall energy intake and expenditure. Benefits that arise from this method are the ease of use, convenience, and logging of food in real time [1,2], and the wide availability of these apps via downloading from app stores. The apps may also enable greater self-monitoring by individuals with chronic diseases such as obesity, cardiovascular disease, and type 2 diabetes [3-5], contributing to nutrition care. However, this form of dietary assessment has yet to be validated and wholly accepted by health care providers to confidently recommend its use. Known limitations of these apps include limited nutrient data, particularly for micronutrients, and inaccurate nutrient compositions [1,6]. Despite this, reported nutrition app use in dietetic practice in the United Kingdom, New Zealand, and Australia (5% response rate from the practitioners contacted) is high (62%) [7].

To date, limited research has been performed on nutrition-related apps, whether it be commercial or researcher developed, and hence the lack of standard methods for assessing validity and accuracy [8]. Furthermore, with the growing number of commercial nutrition-related apps, the difficulty of creating criteria to assess them increases with varying features and databases. This is often exacerbated by a lack of information on the source of nutritional data used by apps. Previous studies by Carter et al [6] and Raatz et al [9] comparing commercial nutrition-related apps with reference methods have used 24-hour recalls and weighed food records (WFRs), respectively. Both studies found no significant differences in mean energy and macronutrients between the nutrition-related app and the reference method. Carter et al used an app for mobile phones called My Meal Mate [6], whereas Raatz, et al used Tap & Track [9]. However, My Meal Mate is not considered a "popular" mobile phone app (with fewer than 500,000 installs), as identified by Franco et al [10], and Tap & Track is available only in Apple Inc's App Store, limiting its use.

Objective

Exploring the accuracy of popular apps would ensure that the findings from studies that used these apps are relevant for the commercial market. Thus, the aim of this study was to evaluate the extent to which popular nutrition-related apps give energy, macronutrient, and micronutrient data comparable with the research standard method using Dietplan6 version 6.0 (Forestfield Software Ltd) analysis program, using 24-hour WFR as an input.

Methods

Dietary Assessment

We obtained 20 handwritten 24-hour WFRs from a previous controlled diet study conducted at the Hugh Sinclair Unit of Human Nutrition, University of Reading, Reading, UK. Participants (female, n=15; male, n=5; mean age 36.3 years; mean body mass index 22.9 kg/m^2) had recorded their food consumption over a 24-hour period using standard protocols [11]. Briefly, the WFR consisted of recording the time, brand name, description of food or drink, cooking method, weight (grams), and leftovers (grams) prospectively. We entered the preexisting 24-hour WFRs into 5 nutrition-related mobile apps using a Samsung Galaxy Tab 4 tablet (Samsung Electronics Co, Ltd, Suwon, South Korea) or a Moto G smartphone (Motorola Mobility LLC, Chicago, IL, USA), using the app versions available in the UK Google Play Store (Google LLC, Mountain View, CA, USA) in January 2016. We compared these apps against a reference research method, Dietplan6. We selected the 5 apps based on popularity (a minimum of 500,000 installs from the Google Play Store), availability as a free download, and having a feature to provide energy (kcal) calculations. As defined and reported by Franco et al, the most popular apps meeting these criteria were Samsung Health (S Health; Samsung), MyFitnessPal (MyFitnessPal, Inc), FatSecret (Secret Industries Pty Ltd), Noom Coach (Noom Inc), and Lose It! (FitNow, Inc), which we therefore used in this study. Further details of the features available in these apps have been previously published [10].

Samsung Health

S Health features tracking of energy, macronutrients, micronutrients (including sodium, calcium, iron, vitamin A, and vitamin C), water (cups per day), and exercise. S Health compares intake data with the recommended intake and color codes the intake as low, average, and high.

MyFitnessPal

MyFitnessPal features tracking of energy, macronutrients, and micronutrients (including sodium, calcium, iron, vitamin A, and vitamin C). Together with S Health, it has the most extensive nutrient output of all 5 apps. MyFitnessPal compares intake data with the US recommended daily allowance. Tracking for water (cups per day) and exercise is also available. The user can create foods or recipes if they are not available in MyFitnessPal's database. Foods added by other users can also be selected for data entry [12].

FatSecret

FatSecret reports energy, macronutrients, and micronutrients (only sodium). It allows the user to add exercise and to save their own meals and add foods not available in the FatSecret database. Recipes added by other users can also be selected.

The app compares intake with the recommended intake for macronutrients only.

Noom Coach

Noom Coach reports only energy intake and compares it with the daily energy recommendation. In addition, exercise can be logged. Recipes can be saved onto the app for future use.

Lose It!

Lose It! reports energy, macronutrients, and sodium. Similar to the other 4 apps, this app allows exercise to be logged. The app compares intake with the recommended intake. Foods or recipes can be created and shared if they are not available in Lose It!'s own database.

Dietplan6

Dietplan6 is a nutrition analysis software package for professional dietitians and nutritionists that we used as the reference method in this study. Dietplan6 reports energy, and all macronutrients and micronutrients. The nutrient composition of foods selected was the *McCance and Widdowson's The Composition of Foods, 6th edition* [13].

Data Input

For each food in the handwritten WFRs, we selected the same food or a suitable (ie, similar) substitute from the databases of all the apps to ensure consistency in diet entry. Supplements were not included in the WFRs. All data input was completed by a single trained researcher (JP) and verified by a second researcher (RZF).

Statistical Analysis

We checked the energy and available macronutrient (carbohydrate, protein, total fat, saturated fat, and fiber) and micronutrient (sodium, calcium, iron, vitamin A, and vitamin C) data for normality using the D'Agostino-Pearson test. To explore differences in outputs between the diet apps and the reference method, Dietplan6, we used paired 2-sample *t* tests (2-tailed). We analyzed nonnormally distributed data using the Wilcoxon signed-rank test. We used Pearson correlations for normal distributions and Spearman correlation for nonnormally distributed data. We also analyzed data using Bland-Altman plots and correlation coefficients (*r*). Bland-Altman analysis examined the agreement between 2 samples using the standard deviation and mean to assess the linear relationship of the variables. We analyzed data using the SciPy 1.1.0 Stats package for Python [14], and we considered a *P* value smaller than .01 (Bonferroni correction applied based on the initial *P* value of .05) to be significant for hypothesis tests and correlation significances.

Results

App Outputs

All of the popular diet apps provided outputs for energy (kcal). With the exception of Noom Coach, outputs were also provided for carbohydrate (g), protein (g), fat (g), fiber (g), and sodium (mg). Additionally, S Health, MyFitnessPal, and Lose It! gave outputs for saturated fat (g). Of the 5 apps tested, only 2 apps had micronutrient outputs other than sodium (calcium, iron, vitamin A, and vitamin C): S Health and MyFitnessPal.

Comparison of Energy and Nutrient Intake Between Dietplan6 and the Apps

Multimedia Appendix 1 shows individual comparisons of energy and nutrient analysis between Dietplan6 and popular diet apps. We observed no significant difference in estimation of energy and saturated fat intake between Dietplan6 and the diet apps. Estimates of protein and sodium intake were significantly lower using Lose It! ($P<.001$) and FatSecret ($P=.004$ and $P=.007$, respectively) than using Dietplan6. Lose It! also gave significantly lower estimates for other reported outputs (carbohydrate, fat, fiber, and sodium) than did Dietplan6 ($P<.001$, $P=.003$, $P=.007$, and $P<.001$, respectively). S Health ($P<.001$) and MyFitnessPal ($P=.005$, $P=.002$, and $P=.008$, respectively) significantly underestimated calcium, iron, and vitamin C compared with Dietplan6, although there was no significant difference for vitamin A. We observed no other significant differences between Dietplan6 and the apps. Table 1 presents the correlation between estimates of energy and nutrients for Dietplan6 and the apps. Correlation coefficients ranged from $r=-.12$ for iron (S Health vs Dietplan6) to $r=.91$ for protein (FatSecret vs Dietplan6). Noom Coach was limited to energy output, but it had a high correlation with Dietplan6 ($r=.91$). S Health had the greatest variation of correlation, with energy at $r=.79$. Correlations were weakest for iron: S Health ($r=-.12$) and MyFitnessPal ($r=.13$), which were not significant. Correlations between Dietplan6 and both S Health and MyFitnessPal also were not significant for sodium, calcium, iron, vitamin A, and vitamin C.

Table 1. Correlation coefficients (*r*) for estimates of energy and nutrient intake between popular diet apps and Dietplan6 using 24-hour weighed food records (n=20)[a].

Nutrients	Samsung Health	MyFitnessPal	FatSecret	Noom Coach	Lose It!
Energy (kcal)	.79	.85	.86	.91	.87
Carbohydrates (g)	.90[b]	.85	.90	N/A[c]	.66
Protein (g)	.91	.82	.91[b]	N/A	.43[d]
Fat (g)	.84	.91[b]	.83	N/A	.79
Saturated fat (g)	.83[b]	.73b	N/A	N/A	.49[b,d]
Fiber (g)	.70[b]	.70[b]	.66[b]	N/A	.23[b,d]
Sodium (mg)	.44[d]	.44[d]	.47[b,d]	N/A	.51[d]
Calcium (mg)	.37[b,d]	.47[b,d]	N/A	N/A	N/A
Iron (mg)	–.12[d]	.13[b,d]	N/A	N/A	N/A
Vitamin A (μg)	.20[b,d]	.21[b,d]	N/A	N/A	N/A
Vitamin C (mg)	.49[b,d]	.54[b,d]	N/A	N/A	N/A

[a]Correlation assessed using Pearson rank correlation, significant at $P<.01$ (Bonferroni correction applied) (unless otherwise specified).

[b]Correlation assessed using Spearman correlation (r_s), significant at $P<.01$ (unless otherwise specified).

[c]N/A: not applicable.

[d]$P>.01$.

Multimedia Appendix 2 shows the difference, mean values, and limits of agreement between estimates of energy and nutrient intake using the apps compared with the reference (Dietplan6).

Figure 1 shows Bland-Altman plots for estimates of energy between Dietplan6 and the apps. Overall, less than 5% of cases fell outside of the limits of agreement for estimates of energy using S Health, FatSecret, and Noom Coach compared with Dietplan6, indicating good agreement between the methods. A total of 10% (2/20) of cases fell outside of the limits of agreement for MyFitnessPal and Lose It!, with Lose It! showing the greatest mean difference (bias) in energy compared with Dietplan6 at –146 kcal (Figure 1). The smallest difference in estimated energy was between Noom Coach and Dietplan6 (14.7 kcal). MyFitnessPal, FatSecret, and Lose It! generally reported less energy intake compared with Dietplan6, whereas S Health and FatSecret reported greater intake. The apps with the smallest bias for estimates of nutrients compared with Dietplan6 were as follows: carbohydrate: S Health (5.3 g); protein: S Health (–2.9 g); fat: S Health (–4.6 g); saturated fat: S Health (–3.8 g); fiber: Lose It! (–0.55 g); sodium: S Health (–197 mg); calcium: S Health (–360.9 mg); iron: S Health (–6.1 mg); vitamin A: S Health (180.1 μg); and vitamin C: S Health (–40.3 mg).

For carbohydrate (Figure 2), the Bland-Altman plots did not indicate clear proportional bias, although the greatest differences occurred in the range of measurements between 200 and 300 g. The Bland-Altman plots for protein (Figure 3) did not show clear clusters of agreement or disagreement, and the differences seemed to be random fluctuations around the mean. For total fat (Figure 4) and saturated fat (Figure 5), none of the apps had more than 5% of the estimates outside of the limits of agreement and no proportional bias was observed. Lose It! had the broadest limit of agreement for fiber (Figure 6) and all the estimates were within the limits of agreement for this app. We detected no proportional bias for sodium (Figure 7), calcium (Figure 8). or iron (Figure 9). On the other hand, for vitamin A (Figure 10) the data clustered together at low averages, but as the average intakes increased, the difference increased. This suggests a proportional bias in the vitamin A data, as both S Health and MyFitnessPal compared with Dietplan6 showed similar results. For vitamin C, the data clustered together at low averages (Figure 11) and no proportional bias was observed.

Figure 1. Bland-Altman plots of energy (kcal) difference and average between Samsung Health (S Health), MyFitnessPal (MFP), Fat Secret, Noom Coach, and Lose It! and Dietplan6. The limits of agreement are displayed as 2 SD.

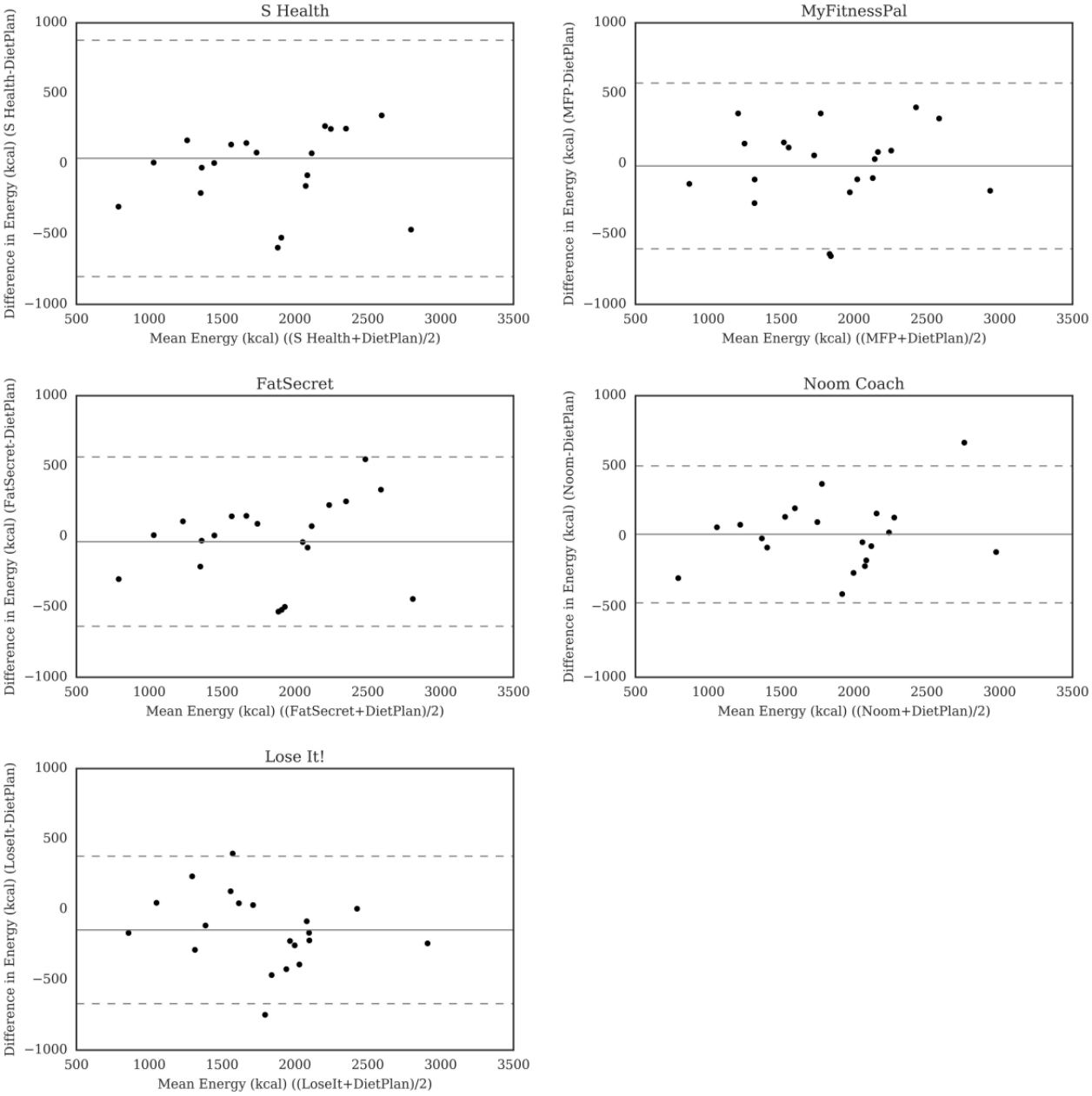


Figure 2. Bland-Altman plots of carbohydrate (g) difference and average between Samsung Health (S Health), MyFitnessPal (MFP), FatSecret, and Lose It! and Dietplan6. The limits of agreement are displayed as 2 SD.

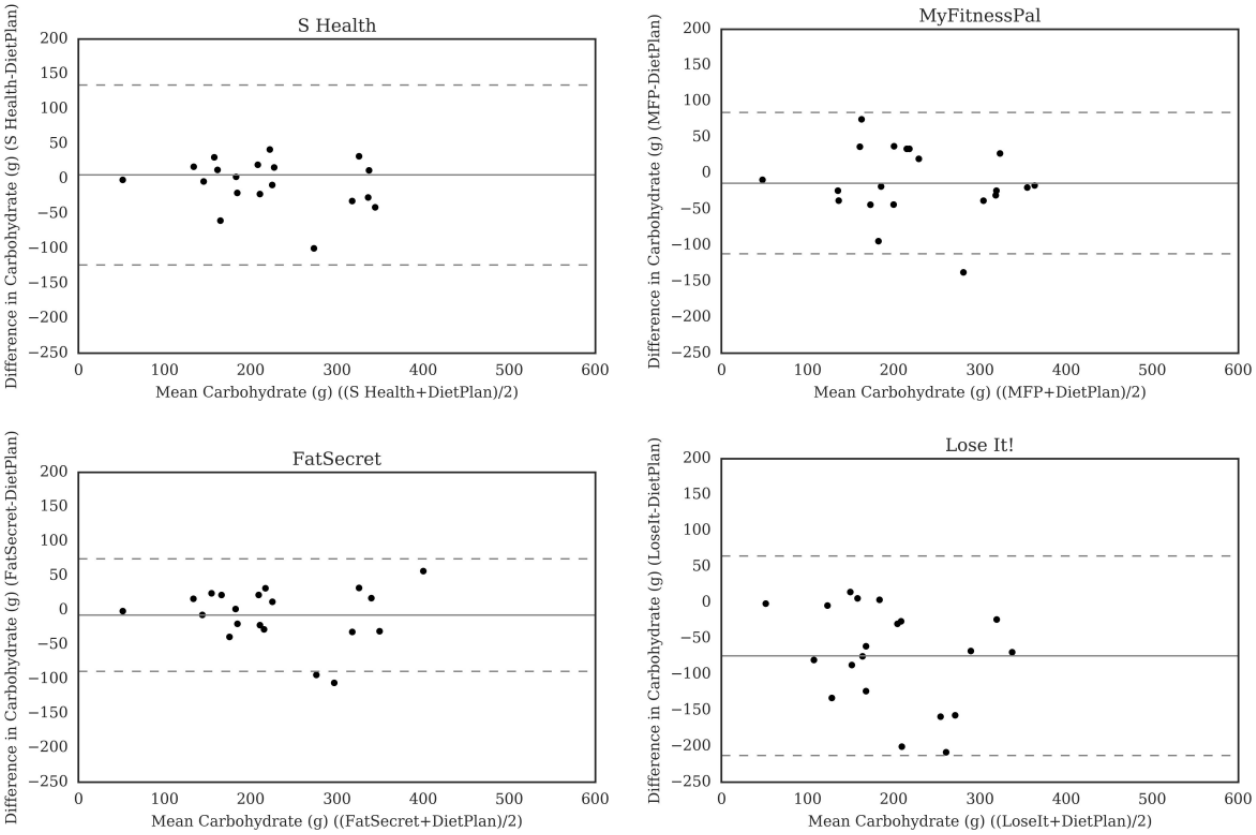


Figure 3. Bland-Altman plots of protein (g) difference and average between Samsung Health (S Health), MyFitnessPal (MFP), FatSecret, and Lose It! and Dietplan6. The limits of agreement are displayed as 2 SD.

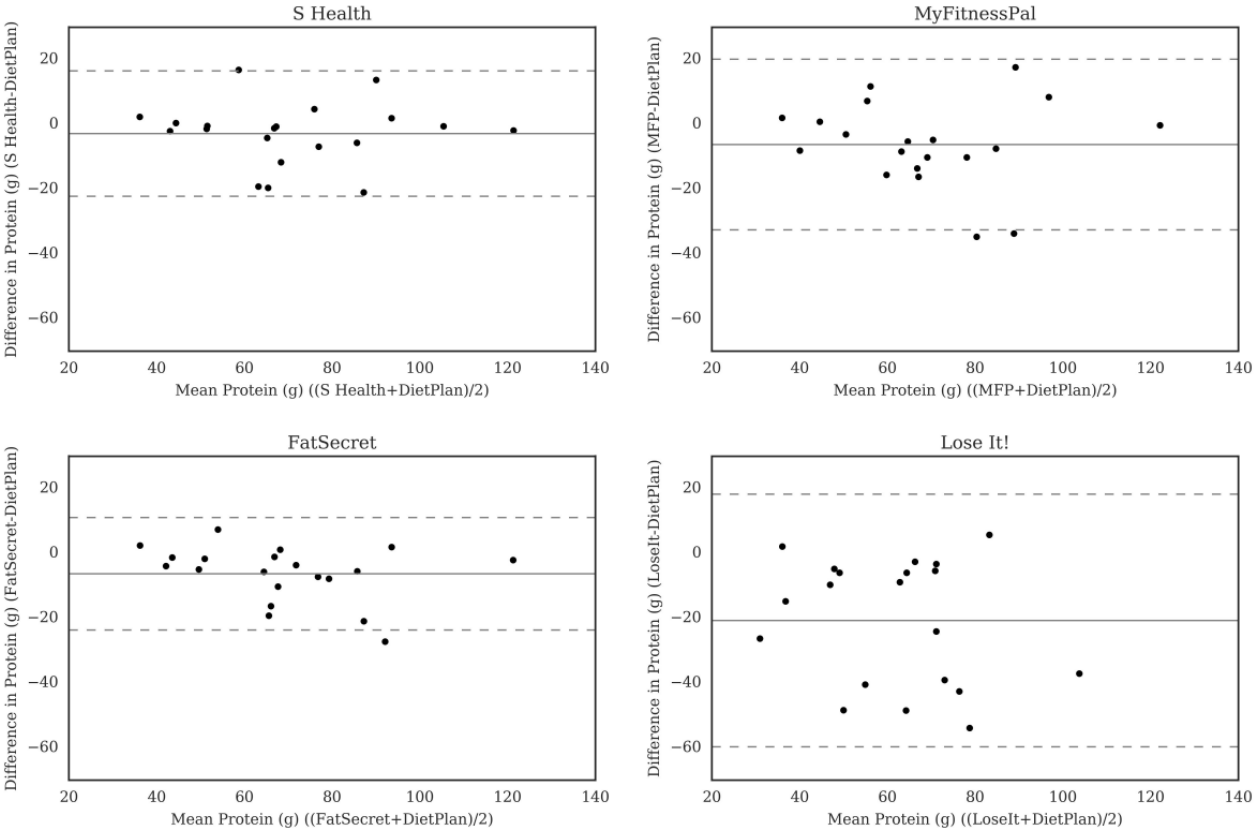


Figure 4. Bland-Altman plots of fat (g) difference and average between Samsung Health (S Health), MyFitnessPal (MFP), FatSecret, and Lose It! and Dietplan6. The limits of agreement are displayed as 2 SD.

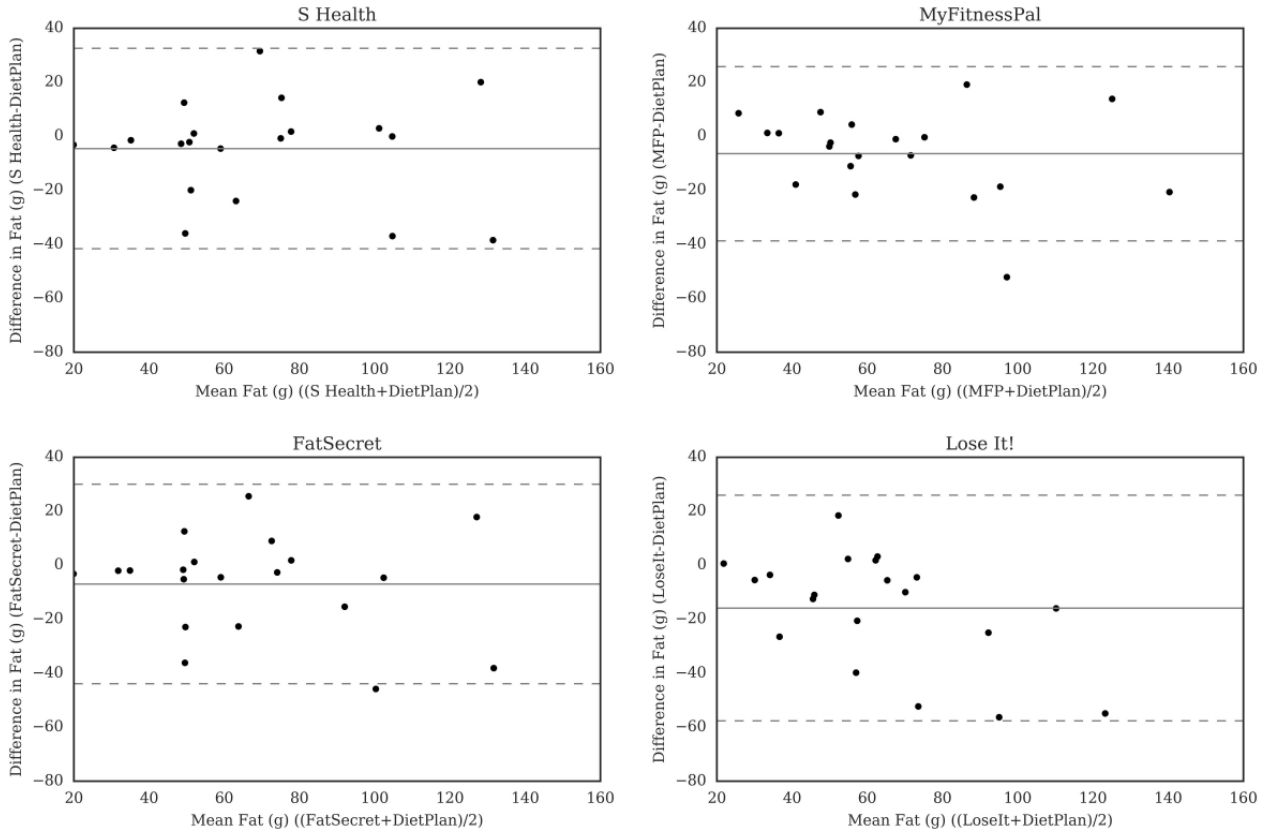


Figure 5. Bland-Altman plots of saturated fat (g) difference and average between Samsung Health (S Health), MyFitnessPal (MFP), and Lose It! and Dietplan6. The limits of agreement are displayed as 2 SD.

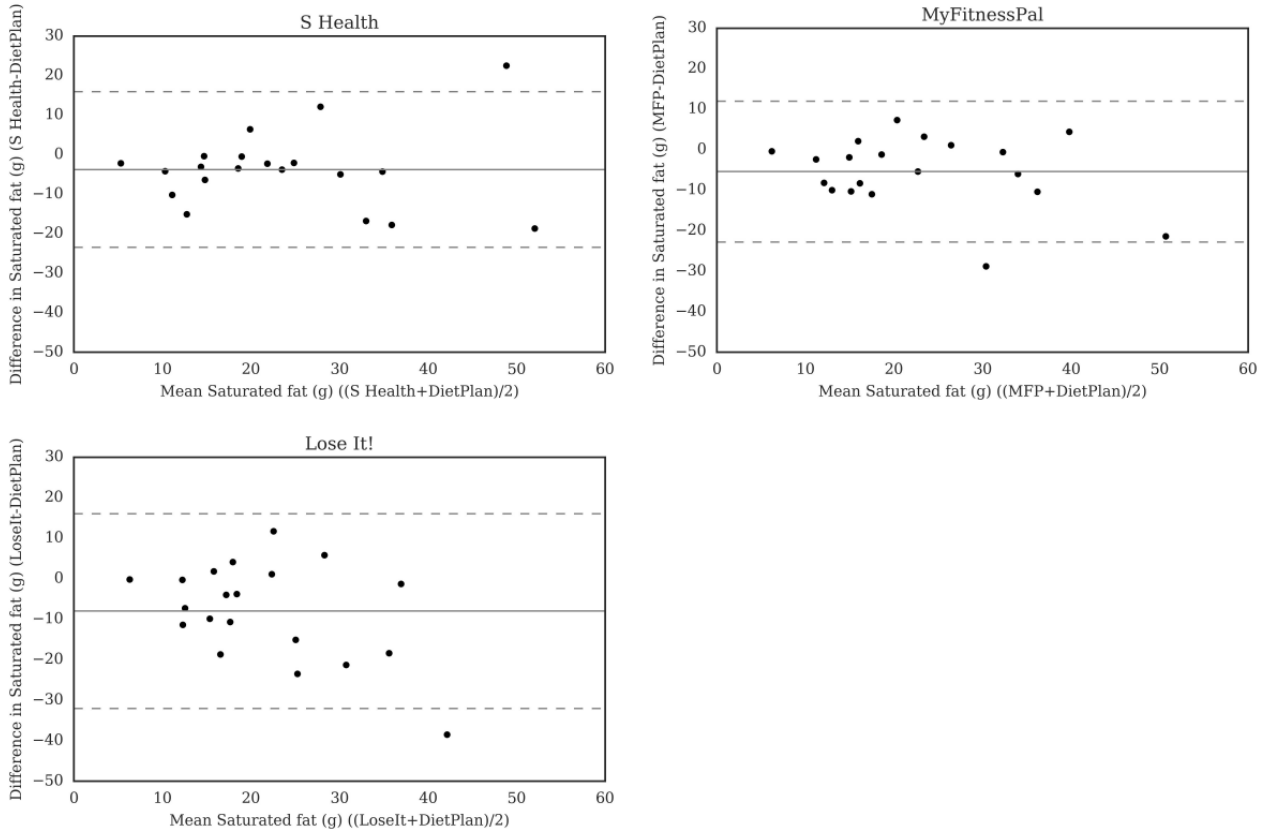


Figure 6. Bland-Altman plots of fiber (g) difference and average between Samsung Health (S Health), MyFitnessPal (MFP), FatSecret, and Lose It! and Dietplan6. The limits of agreement are displayed as 2 SD.

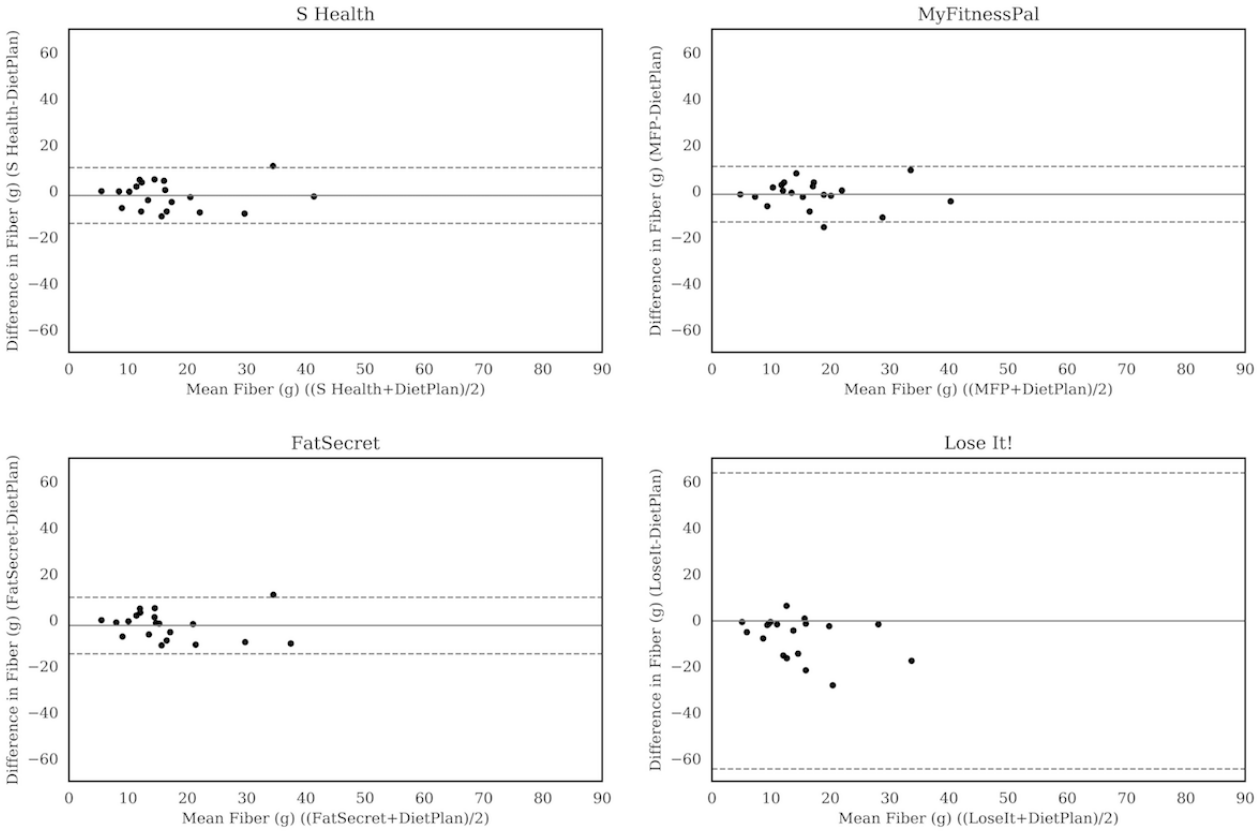


Figure 7. Bland-Altman plots of sodium (mg) difference and average between Samsung Health (S Health), MyFitnessPal (MFP), FatSecret, and Lose It! and Dietplan6. The limits of agreement are displayed as 2 SD.

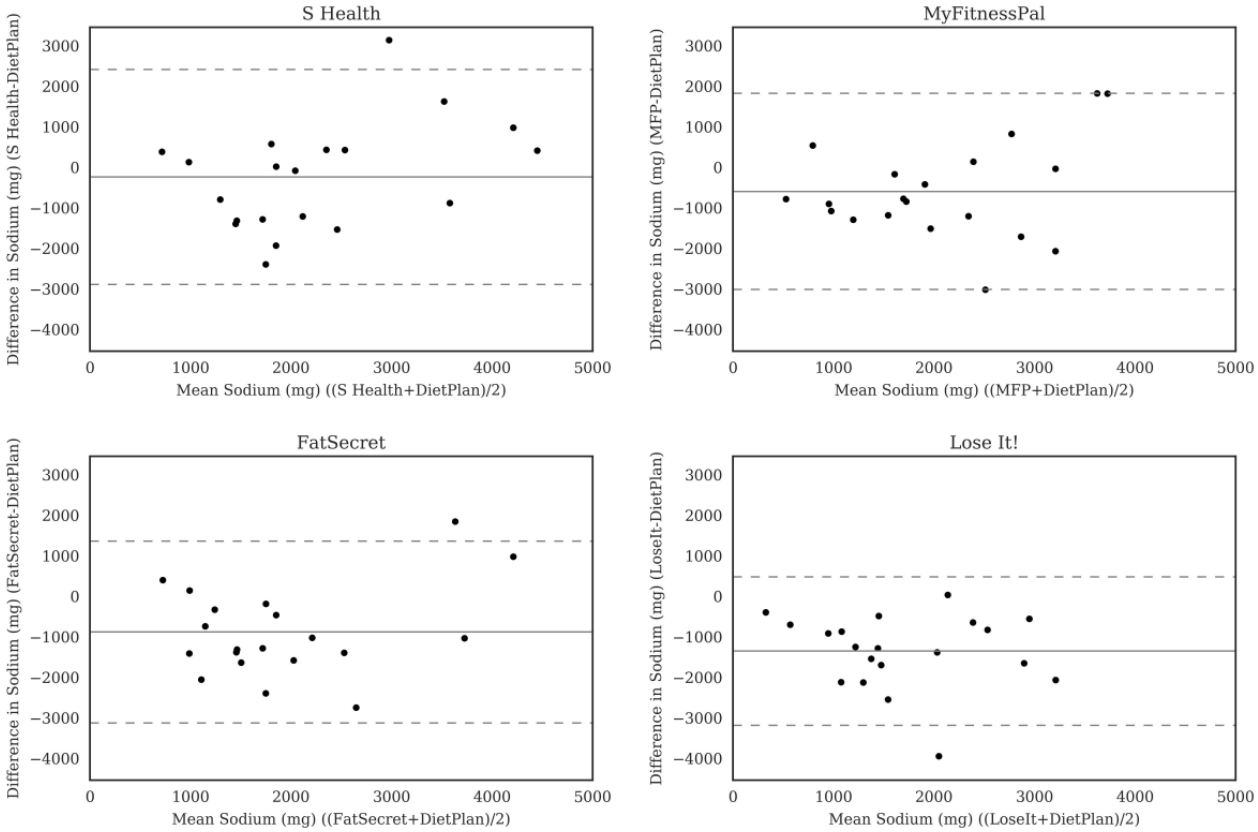


Figure 8. Bland-Altman plots of calcium (mg) difference and average between Samsung Health (S Health) and MyFitnessPal (MFP) and Dietplan6. The limits of agreement are displayed as 2 SD.

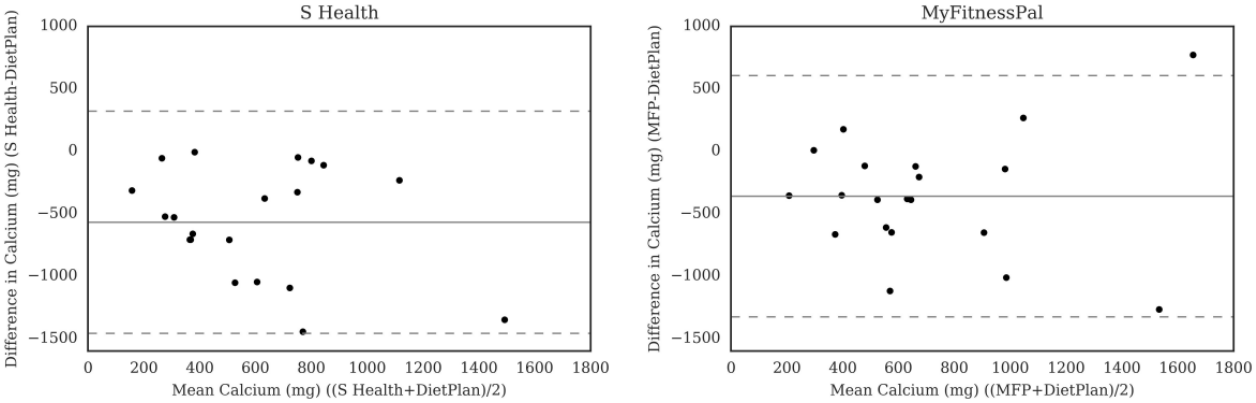


Figure 9. Bland-Altman plots of iron (mg) difference and average between Samsung Health (S Health) and MyFitnessPal (MFP) and Dietplan6. The limits of agreement are displayed as 2 SD.

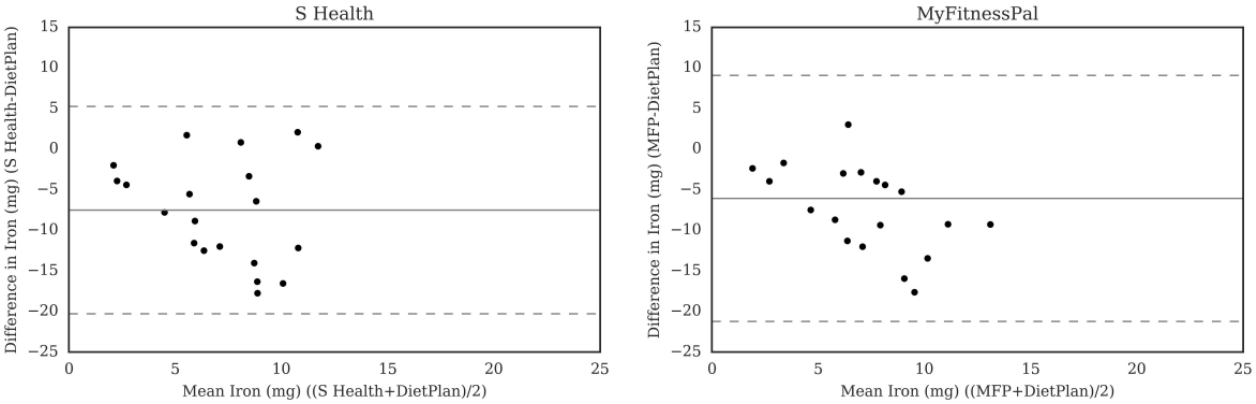


Figure 10. Bland-Altman plots of vitamin A (µg) difference and average between Samsung Health (S Health) and MyFitnessPal (MFP) and Dietplan6. The limits of agreement are displayed as 2 SD.

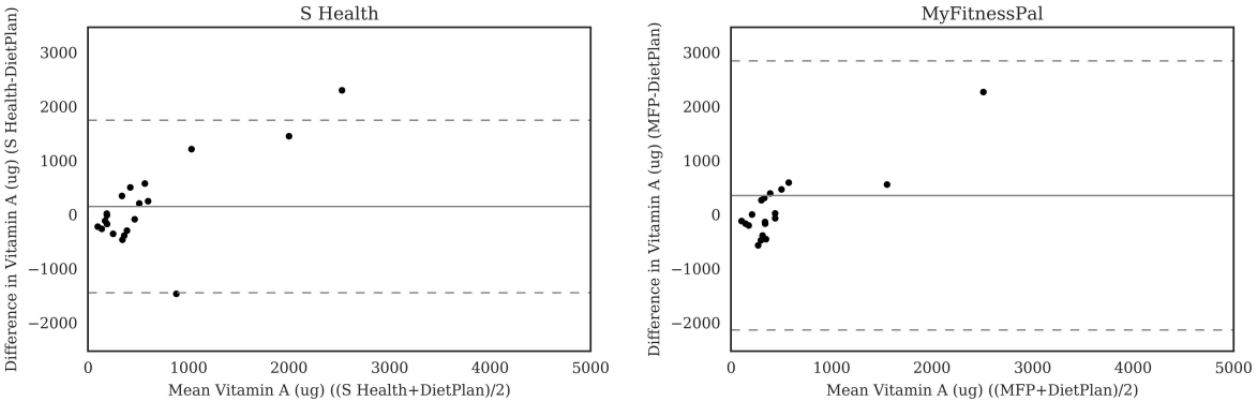
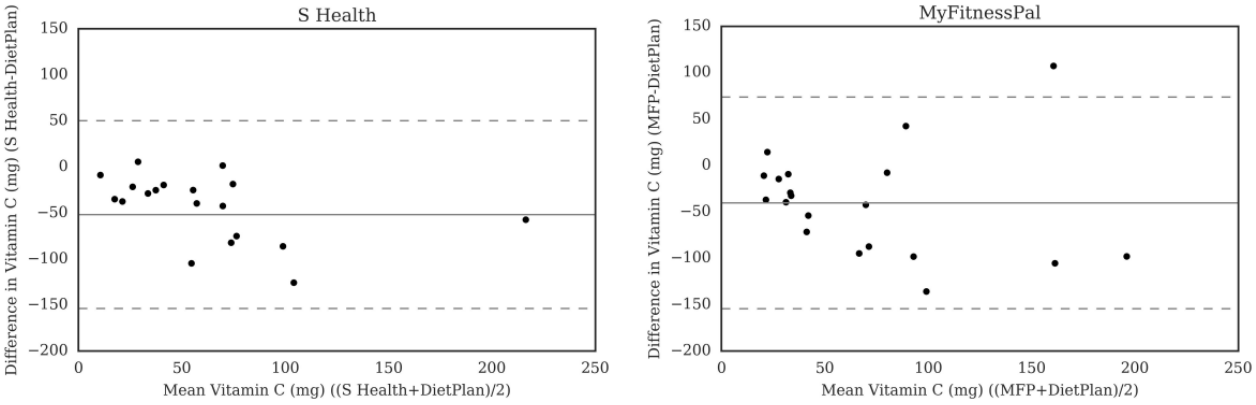


Figure 11. Bland-Altman plots of vitamin C (mg) difference and average between Samsung Health (S Health) and MyFitnessPal (MFP) and Dietplan6. The limits of agreement are displayed as 2 SD.



Discussion

Principal Findings

The primary finding of this study was that popular diet apps are generally comparable in their assessment of energy and key macronutrients (carbohydrate, total fat, and fiber) with a research standard for dietary analysis of WFRs (Dietplan6). However, we observed significant differences for certain apps and nutrients, and Lose It! significantly underestimated 5 of 7 outputs. We found apps to be less able to reliably estimate micronutrients when compared with a widely used and accepted method for dietary analysis.

For estimation of energy intake, S Health, MyFitnessPal, FatSecret, and Noom Coach demonstrated good agreement with Dietplan6, evidenced by small mean differences (14.7 to –36.9 kcal) and strong correlation coefficients (r=.79 to r=.91). Estimates of carbohydrate and total fat intake were also highly comparable between Dietplan6 and S Health, MyFitnessPal, and FatSecret, with no significant differences between the outputs and similarly small mean differences (5.3 to –13.8 g for carbohydrate and –4.6 to –15.9 g for total fat). Raatz et al also found no significant differences in estimates of energy, carbohydrate, and total fat when comparing nutrient analysis of a 3-day WFR using the Tap & Track app and US Department of Agriculture (USDA) nutrient analysis program (mean differences 85 kcal, 15.4 g, and 2.6 g, respectively) [9]. However, in contrast to our study, in the study of Raatz and colleagues, users (n=19) entered the WFR into Tap & Track and a registered dietitian entered data into the USDA nutrient analysis program [9]. The authors suggested that the observed variability in Bland-Altman plots between methods may have been associated with the ability of the users to pick appropriate food items and ensure that all data were recorded [9,15], which would not have been an issue in our study, as a single researcher entered data into all apps and the reference program.

Protein intake was significantly underestimated by FatSecret and Lose It! compared with Dietplan6, and correlations between the methods varied from r=.43 to r=.91. This indicates that these apps are less reliable at estimating protein intake than at estimating energy, carbohydrate, and total fat intake. Tap & Track's estimate of protein intake did not differ from the USDA nutrient analysis program output [9] and, in general, apps comparing traditional paper-based (eg, 24-hour recall, WFR) versus electronic-based methods of estimating energy intake and macronutrients have found comparable and acceptable results [6,16,17].

Correlations were weaker between S Health, MyFitnessPal, and FatSecret and Dietplan6 for sodium (r=.44 to r=.51), with significantly lower estimates given by FatSecret. S Health and MyFitnessPal were the only apps to assess iron, calcium, vitamin A, and vitamin C, which we observed were significantly underestimated by the apps (r=–.12 to r=.54). While there was no significant difference in estimated vitamin A between these apps and Dietplan6, results of the Bland-Altman analysis revealed a proportional bias, whereby higher intakes of vitamin A were less well estimated by the apps. These data suggest that the apps did not reliably estimate micronutrient intake in comparison with the reference method.

In this study, a single researcher was responsible for entering the WFRs into each of the methods (apps and Dietplan6) and, therefore, discrepancies in outputs were most likely due to variations in the apps' nutrient databases, as opposed to variations in data entry. For example, apps may use different sources to compile nutrient databases, including open source and manufacturer databases, and some apps involve users in the expansion of their database, which may result in inaccuracies and incomplete data [12]. The observation that micronutrients were less reliably estimated by the apps, compared with macronutrients, may be attributed to the fact that micronutrients are declared only voluntarily on nutritional labels of foods, and therefore apps using commercial databases are more likely to have missing composition data [18]. In addition, 4 of the 5 apps studied (ie, excluding S Health) have a barcode feature [10], which relies on nutrition data provided by manufacturers, which may have incomplete micronutrient data. This method is less reliable than the chemical analysis done in most of the food databases used by the reference method (McCance and Widdowson's nutrient database [13]).

For those apps unable to reliably assess intake of particular nutrients (eg, protein, sodium, vitamin A), the display of the data to the user is potentially misleading. For example, underestimation of sodium intake for a patient with high blood pressure (who would typically be advised to reduce their salt

intake) may result in the user unknowingly exceeding their recommended intake. Contrary to medical devices, apps available in app stores are not regulated by the US Food and Drug Administration. The level of validity of the nutrition-related apps available for commercial download is often uncertain, creating ambivalence for health organizations about their use. One reason for this is the absence of a standard method of assessing the validity of health-related apps [8]. Given the lack of transparency in the source of nutritional data used by app developers, it will be important to repeat these assessments in the future. With the increase in awareness of the possibilities of app-based dietary assessment and use of the most popular nutrition-related apps, our findings will help clinicians and researchers to be better informed about using these apps to facilitate nutrition care and research, and about whether their use can replace traditional research software. Reported barriers to app use in dietetic practice in the 3-country survey by Chen et al included lack of awareness about the best app to recommend [7] and, while MyFitnessPal was the most popular app recommended by dietitians in this survey, none of the other apps assessed in our study were recommended [7].

Strengths and Limitations

Strengths of this study include the use of a WFR to assess dietary intake and the entry of WFR data (reference method and apps) by a single researcher. Whereas previous studies have compared user versus professional entry of the same record [9], or user versus professional entry of records taken on separate occasions [19], this study eliminated the biases associated with user entry (eg, selection of an appropriate food item, estimation of portion size [15]). The strengths of using a WFR, as opposed to a 24-hour recall, are that dietary intake is recorded prospectively (ie, at the time of intake) and users are prompted to provide the level of detail required for analysis (eg, describe the brand, name, weight, and leftovers of each item). Prospective WFR recording potentially provides a greater level of detail than does a 24-hour recall, which is a retrospective method that relies on the participant's memory, as well as the skill of the researcher-interviewer to obtain sufficient detail on each food item [20].

Limitations include a relatively small sample size (n=20) and the use of 24-hour WFRs collected prior to the study. For example, it was not possible to check ambiguous entries (eg, unspecified brand) with the participant, although where we made assumptions, these were verified by an independent researcher and applied across all methods. Exploration of the impact of, for example, training on the accuracy of dietary data input by users would provide further evidence for the use of popular nutrition apps in the research and clinical setting.

Conclusions

This study compared the accuracy of the outputs of 5 nutrition-related apps (S Health, MyFitnessPal, FatSecret, Noom Coach, and Lose It!) against a research standard, using 24-hour WFRs as input. The findings suggest that the apps provide estimates of energy and saturated fat intake that are comparable with Dietplan6. With the exception of Lose It!, the apps also provided comparable estimates of carbohydrate, total fat, and fiber. Two apps displayed a tendency to underestimate protein and sodium (FatSecret and Lose It!). The estimates of micronutrient intake (calcium, iron, vitamin A, and vitamin C) by 2 apps (S Health and MyFitnessPal) were inconsistent and less reliable. Overall, the nutritional outputs provided by S Health and Noom Coach, which provided only energy, were the most reliable for this output. As the use and availability of nutrition-related apps grow, there are increasing opportunities for apps to support research in nutrition, care, and self-management of noncommunicable diseases. This study highlighted in which aspects the outputs from these apps are comparable with, and where they differ significantly from, widely accepted and professionally used nutrition analysis software, thereby making a valuable contribution toward helping clinicians and researchers to make better-informed decisions about using these apps in research and practice.

Acknowledgments

RZF was sponsored by the National Council of Technological and Scientific Development (CNPq) from the Brazilian government, via the Science without Borders program.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Mean estimates of energy and nutrient intake.

[DOCX File, 17KB - mhealth_v7i2e9838_app1.docx]

Multimedia Appendix 2

Difference and Bland-Altman limits of agreement.

[DOCX File, 14KB - mhealth_v7i2e9838_app2.docx]

References

1. Jospe MR, Fairbairn KA, Green P, Perry TL. Diet app use by sports dietitians: a survey in five countries. JMIR Mhealth Uhealth 2015 Jan 22;3(1):e7 [FREE Full text] [doi: 10.2196/mhealth.3345] [Medline: 25616274]
2. Turner-McGrievy GM, Beets MW, Moore JB, Kaczynski AT, Barr-Anderson DJ, Tate DF. Comparison of traditional versus mobile app self-monitoring of physical activity and dietary intake among overweight adults participating in an mHealth weight loss program. J Am Med Inform Assoc 2013 May 1;20(3):513-518 [FREE Full text] [doi: 10.1136/amiajnl-2012-001510] [Medline: 23429637]
3. El-Gayar O, Timsina P, Nawar N, Eid W. Mobile applications for diabetes self-management: status and potential. J Diabetes Sci Technol 2013;7(1):247-262 [FREE Full text] [Medline: 23439183]
4. Flores Mateo G, Granado-Font E, Ferré-Grau C, Montaña-Carreras X. Mobile phone apps to promote weight loss and increase physical activity: a systematic review and meta-analysis. J Med Internet Res 2015;17(11):e253 [FREE Full text] [doi: 10.2196/jmir.4836] [Medline: 26554314]
5. Pellegrini CA, Pfammatter AF, Conroy DE, Spring B. Smartphone applications to support weight loss: current perspectives. Adv Health Care Technol 2015 Jul;1:13-22. [doi: 10.2147/AHCT.S57844] [Medline: 26236766]
6. Carter MC, Burley VJ, Nykjaer C, Cade JE. 'My Meal Mate' (MMM): validation of the diet measures captured on a smartphone application to facilitate weight loss. Br J Nutr 2013 Feb 14;109(3):539-546. [doi: 10.1017/S0007114512001353] [Medline: 22717334]
7. Chen J, Lieffers J, Bauman A, Hanning R, Allman-Farinelli M. The use of smartphone health apps and other mobile health (mHealth) technologies in dietetic practice: a three country study. J Hum Nutr Diet 2017 Aug;30(4):439-452. [doi: 10.1111/jhn.12446] [Medline: 28116773]
8. Boudreaux ED, Waring ME, Hayes RB, Sadasivam RS, Mullen S, Pagoto S. Evaluating and selecting mobile health apps: strategies for healthcare providers and healthcare organizations. Transl Behav Med 2014 Dec;4(4):363-371 [FREE Full text] [doi: 10.1007/s13142-014-0293-9] [Medline: 25584085]
9. Raatz SK, Scheett AJ, Johnson LK, Jahns L. Validity of electronic diet recording nutrient estimates compared to dietitian analysis of diet records: randomized controlled trial. J Med Internet Res 2015 Jan 20;17(1):e21 [FREE Full text] [doi: 10.2196/jmir.3744] [Medline: 25604640]
10. Franco RZ, Fallaize R, Lovegrove JA, Hwang F. Popular nutrition-related mobile apps: a feature assessment. JMIR Mhealth Uhealth 2016 Aug 01;4(3):e85 [FREE Full text] [doi: 10.2196/mhealth.5846] [Medline: 27480144]
11. Markey O, Le Jeune J, Lovegrove JA. Energy compensation following consumption of sugar-reduced products: a randomized controlled trial. Eur J Nutr 2016 Sep;55(6):2137-2149 [FREE Full text] [doi: 10.1007/s00394-015-1028-5] [Medline: 26349919]
12. MyFitnessPal. Some food information in the database is inaccurate. Can I edit it?.: Under Armour, Inc; 2019 Jan 02. URL: https://myfitnesspal.desk.com/customer/en/portal/articles/13980-some-food-information-in-the-database-is-inaccurate-can-i-edit-it- [accessed 2019-01-22] [WebCite Cache ID 75cgxiBtS]
13. Food Standards Agency. McCance and Widdowson's The Composition of Foods. Sixth summary edition. Cambridge, UK: The Royal Society of Chemistry; 2002.
14. Jones E, Oliphant E, Peterson P. SciPy: open source scientific tools for Python.: SciPy developers; 2019. URL: https://scipy.org/about.html [accessed 2019-01-31] [WebCite Cache ID 75q50BXMD]
15. Hongu N, Hingle MD, Merchant NC, Orr BJ, Going SB, Mosqueda MI, et al. Dietary assessment tools using mobile technology. Top Clin Nutr 2011;26(4):300-311. [doi: 10.1097/TIN.0b013e3182379525]
16. Rangan AM, O'Connor S, Giannelli V, Yap ML, Tang LM, Roy R, et al. Electronic Dietary Intake Assessment (e-DIA): comparison of a mobile phone digital entry application for dietary data collection with 24-hour dietary recalls. JMIR Mhealth Uhealth 2015 Oct 27;3(4):e98 [FREE Full text] [doi: 10.2196/mhealth.4613] [Medline: 26508282]
17. Fallaize R, Forster H, Macready AL, Walsh MC, Mathers JC, Brennan L, et al. Online dietary intake estimation: reproducibility and validity of the Food4Me food frequency questionnaire against a 4-day weighed food record. J Med Internet Res 2014;16(8):e190 [FREE Full text] [doi: 10.2196/jmir.3355] [Medline: 25113936]
18. Population Health Division/10800. Technical guidance on nutrition labelling. London, UK: Department of Health, Health Behaviours; 2016 Sep. URL: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/595961/Nutrition_Technical_Guidance.pdf [accessed 2019-01-22] [WebCite Cache ID 75ch4sdnl]
19. Timon CM, Astell AJ, Hwang F, Adlam TD, Smith T, Maclean L, et al. The validation of a computer-based food record for older adults: the Novel Assessment of Nutrition and Ageing (NANA) method. Br J Nutr 2015 Feb 28;113(4):654-664. [doi: 10.1017/S0007114514003808] [Medline: 25630436]
20. Freudenheim JL. A review of study designs and methods of dietary assessment in nutritional epidemiology of chronic disease. J Nutr 1993 Dec;123(2 Suppl):401-405. [doi: 10.1093/jn/123.suppl_2.401] [Medline: 8429394]

Abbreviations

S Health: Samsung Health
USDA: US Department of Agriculture

WFR: weighed food record

Edited by G Eysenbach; submitted 14.01.18; peer-reviewed by V Marcil, K Davison, A Paglialonga, C Mamillapalli; comments to author 12.04.18; revised version received 10.07.18; accepted 24.11.18; published 20.02.19.

Please cite as:
Fallaize R, Zenun Franco R, Pasang J, Hwang F, Lovegrove JA
Popular Nutrition-Related Mobile Apps: An Agreement Assessment Against a UK Reference Method
JMIR Mhealth Uhealth 2019;7(2):e9838
URL: http://mhealth.jmir.org/2019/2/e9838/
doi:10.2196/mhealth.9838
PMID:30785409

Original Paper

Mobile Phone App–Based Pulmonary Rehabilitation for Chemotherapy-Treated Patients With Advanced Lung Cancer: Pilot Study

Sojung Park[1], MD, PhD; Ji Youn Kim[2], BSc; Jae Cheol Lee[3], MD, PhD; Hyeong Ryul Kim[4], MD, PhD; Seungjae Song[5], PhD; Hee Kwon[5], BSc; Wonjun Ji[1], MD; Chang Min Choi[1,3], MD, PhD

[1]Department of Pulmonary and Critical Care Medicine, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Republic of Korea
[2]Department of Outpatient Nursing, Asan Medical Center, Seoul, Republic of Korea
[3]Department of Oncology, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Republic of Korea
[4]Department of Thoracic and Cardiovascular Surgery, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Republic of Korea
[5]Life Semantics Corp, Seoul, Republic of Korea

Corresponding Author:
Chang Min Choi, MD, PhD
Department of Pulmonary and Critical Care Medicine, Asan Medical Center
University of Ulsan College of Medicine
Songpa–gu
88 Olympic-ro 43-gil
Seoul, 05505
Republic of Korea
Phone: 82 2 3010 5902
Fax: 82 2 3010 6968
Email: ccm9607@gmail.com

Abstract

Background: Advanced lung cancer patients often have chronic lung disease with reduced exercise capacities and various symptoms leading to altered quality of life (QoL). No studies have assessed pulmonary rehabilitation (PR) employing a mobile app and an Internet of Things device in advanced lung cancer patients undergoing chemotherapy.

Objective: This study aimed to determine the feasibility and efficacy of smartphone app–based PR on exercise capacity, symptom management, and QoL in patients with advanced lung cancer undergoing chemotherapy.

Methods: A total of 100 patients were recruited in a prospective, single-arm intervention study using a smartphone app–based PR program for 12 weeks. Exercise capacity (6-min walking distance, 6MWD), QoL, symptom scale scores, and distress indexes were investigated.

Results: A total of 90 patients completed the PR program. The most common cause of drop out was hospitalization because of cancer progression. After PR, there was significant improvement in the 6MWD; 380.1 m (SD 74.1) at baseline, 429.1 m (SD 58.6) at 6 weeks (P<.001), and 448.1 m (SD 50.0) at 12 weeks (P<.001). However, the dyspnea scale score showed no significant improvement in the patients overall, but there was a trend for improvement in those with a stable tumor response (P=.07). Role (P=.02), emotional (P<.001), and social functioning (P=.002) scale scores showed significant improvement after PR. Symptom scale scores for fatigue (P<.001), anorexia (P=.047), and diarrhea (P=.01) also showed significant improvement. There was significant improvement in depression (P=.048) and anxiety (P=.01), whereas there was no significant change in QoL (P=.06) and severity of pain (P=.24).

Conclusions: Smartphone app–based PR represents an effective and feasible program to improve exercise capacity and to manage symptoms and distress in patients with advanced lung cancer who are undergoing chemotherapy.

(JMIR Mhealth Uhealth 2019;7(2):e11094) doi:10.2196/11094

KEYWORDS

chemotherapy; physical fitness; lung cancer; rehabilitation; quality of life

Introduction

Background

Lung cancer is the leading cause of death worldwide. A large number of patients are diagnosed with lung cancer in advanced stage [1]. Although the median overall survival of patients with metastatic lung cancer remains poor, advances in cancer-targeting agents and immune checkpoint inhibitors have extended life expectancy [2-6]. Patients with lung cancer often have comorbidities, such as chronic obstructive pulmonary disease (COPD) and interstitial lung disease (ILD), associated with decreased lung function and increased respiratory symptoms [7]. In addition, the adverse effects of chemotherapy and various cancer symptoms alter quality of life (QoL) and reduce the physical activity and exercise capacity of patients undergoing chemotherapy [8].

Exercise capacity is associated with lung cancer prognosis [9]. Pulmonary rehabilitation (PR) has emerged as a cost-effective intervention for managing chronic lung disease. PR improves the 6-min walking distance (6MWD), QoL, and respiratory symptoms in patients with COPD and ILD [10-13]. However, there are very limited data to demonstrate the feasibility and efficacy of PR in advanced lung cancer patients [14-17]. Currently, a large number of patients with lung cancer receive chemotherapy as outpatients. Moreover, many live in areas where an attending physician is not readily accessible. Therefore, home-based PR and symptom management is essential for patients with advanced lung cancer. The use of mobile health care may assist in overcoming these barriers, improving educational levels, facilitate frequent assessments, and allow more objective data collection [18,19].

Objectives

We investigated the feasibility and efficacy of a smartphone app–based PR and educational program on exercise capacity, symptom management, and QoL in patients with advanced non–small cell lung cancer (NSCLC) undergoing chemotherapy.

Methods

Participant Identification

Consecutive patients with histologically diagnosed advanced NSCLC were identified from September to November 2016 in Asan Medical Center, Seoul, South Korea. All patients were eligible to receive palliative chemotherapy or adjuvant chemotherapy. The inclusion criteria were as follows: patients aged 20 to 75 years with (1) NSCLC stage Ⅱ-Ⅳ, (2) Eastern Cooperative Oncology Group-Performance Status (ECOG-PS) 0-2, and (3) having an android smartphone and the ability to handle the app. Patients were excluded if they (1) had concurrent malignancies other than lung cancer, (2) had received prior PR or education on lifestyle modification, (3) had symptomatic heart disease, (4) had severe cognitive impairment, and (5) were not willing to make regular visits.

Written informed consent was obtained from all patients in accordance with the Declaration of Helsinki. This study was approved by the institutional review board of Asan Medical Center (proposal identification number 2016-0445).

App for Mobile Health Care

A comprehensive mobile health care app, the Smart Aftercare app, was newly developed for this study. The English version of the brochure introducing the function of the app is found in Multimedia Appendix 1. Patients were provided with the Smart Aftercare app, an Internet of Things (IoT) wearable device (URBAN S, Partron Co, Seoul, Korea), which interlocked with the app, a portable pulse oximeter, thermometer, scale, and resistance bands for physical therapy.

The app engineers (Life Semantics Co, Seoul, Korea) developed the app with support from medical personnel. The app had a to-do list, individual health information, and an in-app chat service. The to-do list provided an alarm notification for daily tasks related to taking medication, performing rehabilitation exercise, and visiting the clinic on schedule. The app provided patients' laboratory results, 1 to 3 key computed tomography images, and information on the efficacy and adverse events of the chemotherapy regimen patients were receiving, as well as general lung cancer information. Through the in-app chat service, a clinical nurse specialist also provided a counseling service.

Pulmonary Rehabilitation Program

This study consisted of a 12-week rehabilitation program. The Smart Aftercare app provided an animation video on 10-min stretching exercises, 30-min aerobic exercises, 30-min muscle strengthening exercises, and 5-min finishing (stretching) exercises. Patients were instructed to run the IoT device and app during their exercise. The IoT device, which facilitates accelerometer-based activity monitoring, recorded patients' activity, including the number of steps taken and walking distance, as well as heart rate (HR). This information was transmitted to the IoT platform, developed for this study, allowing data sharing with the attending physician to check patients' condition in real time, objectively. To promote regular exercise, push notifications were sent to patients when the app had not been used for a period.

All patients performed the 6-min walk test (6MWT) every time they visited the clinic. The attending physician prescribed individualized exercise duration and intensity, which were adjusted after every clinic visit according to the results of the 6MWT. Walking, bicycle ergometer, and treadmill use were recommended. Exercise was prescribed as follows: once a day for 30 to 60 min at least three days a week, a walking distance target of 60% to 80% of the 6MWT, with an HR target of 70% of HR reserve plus resting HR (target HR=70% x [HRmax-HRrest] + HRrest). HRmax stands for maximum HR and HRrest stands for resting HR. Patients were instructed to quit exercising if their oxygen saturation fell below 88% or they could not talk with others because of dyspnea.

The muscle strengthening exercise program comprised strengthening of all major muscle groups in the limbs and trunk. Patients were instructed to perform strengthening exercises once a day for 30 min. Patients' muscle strength was assessed by arm abduction test, 10 times. Various resistance bands with different

intensities were provided according to patients' muscle strength. The app and IoT device checked whether the patient completed daily exercise tasks and gave feedback on the amount of activity and calories expended in a day.

Symptom Management

The Smart Aftercare app provided an animation video on pain control, nutritional support, and symptom management. Furthermore, algorithms for the management of pain and adverse skin reactions, a frequently asked questions service, were also provided. Patients were instructed to record daily body weight and temperature on the app.

Data Collection

Demographic information was investigated at baseline, and the relevant medical data were collected from medical records. Patients visited the clinic every 4 or 6 weeks, depending on their schedule of chemotherapy (Table 1). At visits, they completed the questionnaires for QoL, symptom (the European Organization for Research and Treatment of Cancer Quality of Life Questionnaire-C30, EORTC QLQ-C30), pain (numeric rating scale, NRS), and distress, including anxiety (Generalized Anxiety Disorder-7, GAD-7) and depression (Patient Health Questionnaire-9, PHQ-9). We also evaluated service satisfaction using a self-developed questionnaire (Multimedia Appendix 2). For patients who visited the clinic every 4 weeks, the mean value of each scale at 4 and 8 weeks was used for investigation at midterm assessment. Permission to use the EORTC QLQ-C30 was obtained, and the authorized Korean version of the questionnaire was freely downloadable from the internet. No permission was required for NRS, PHQ-9, or GAD-7.

Unexpected Visits to the Emergency Department

As we performed a single-arm study using a mobile app, this study was limited in that it had no control group. Therefore, we retrospectively evaluated the number of unexpected visits to the emergency department (ED) comparing patients who participated in this study with those who did not. The number of visits to the ED is widely regarded as an indicator of the development of unexpected, uncontrolled symptoms. We expected that the mobile app could provide information regarding how to manage symptoms through an algorithm and in-app chat service, thereby reducing medical expenses and time. From September to November 2016, a total of 1091 patients underwent chemotherapy in Asan Medical Center. Of these, 100 patients participated in this study, whereas 991 did not. Their electrical medical records were retrospectively reviewed.

Table 1. Assessment schedule.

Schedule	Screening	4, 8, or 6 weeks (±1 week)[a]	12 weeks (±1 week)[a]
Informed consent	✓[b]	—[c]	—
Demographic characteristics	✓	—	—
Inclusion or exclusion criteria	✓	—	—
Vital signs and oxygen saturation	✓	✓	✓
Physical examination	✓	✓	✓
History taking	✓	—	—
Disease status evaluation[d]	✓	✓	✓
Satisfaction questionnaire for the Aftercare app	—	—	✓
6-min walk test and exercise prescription	✓	✓	✓
EORTC QLQ-C30[e]	✓	—	✓
Numeric pain rating scale	✓	✓	✓
GAD-7[f]	✓	✓	✓
PHQ-9[g]	✓	✓	✓
Assessment for adverse reaction of chemotherapy	—	✓	✓

[a]1 week before and after the visiting day was allowed.

[b]Data were obtained at the time marked with a check.

[c]Data not obtained at the time marked with an dash.

[d]Treatment response was assessed at least once in all patients during the study period depending on their schedule of chemotherapy.

[e]EORTC QLQ-C30: European Organization for Research and Treatment of Cancer Quality of Life Questionnaire-C30.

[f]GAD-7: Generalized Anxiety Disorder-7.

[g]PHQ-9: Patient Health Questionnaire-9.

Statistical Analysis

Categorical variables were analyzed using Pearson chi-square test or Fisher exact test. Continuous variables were analyzed using a Student *t* test. A paired *t* test and Bonferroni correction were used for comparison of pre- and post-PR assessment. All tests of significance were two sided, and differences between groups were considered to be significant when the *P* value was <.05. All statistical analyses were performed with SPSS software version 22.0 (IBM Corp, Armonk, NY).

Results

Patients

A total of 100 patients were enrolled, and 90 patients completed the 12-week rehabilitation program. The baseline characteristics of the patients are presented in Table 2. The mean age of the patients was 55.1 years (SD 8.7); 46.0% (46/100) were males. The most common cause of drop out was hospitalization because of cancer progression (6/100, 6.0%), followed by transfer to other hospitals (2/100, 2.0%) and difficulty in handling the app (2/100, 2.0%). Patients who had completed the PR program had a significant higher baseline body mass index and better performance status than patients who had dropped out.

Exercise Capacity

The mean exercise number per week was 3.8 (SD 1.2) at 1 week, 4.2 (SD 1.1) at 6 weeks, and 4.1 (SD 1.2) at 12 weeks, satisfying the exercise prescription. A total of 85 patients completed all 6MWTs according to the schedule. In addition to 10 patients who dropped out, 5 patients refused to perform 6MWT because of general weakness, paraplegia, knee pain, and dizziness. There was significant difference in the baseline 6MWD according to baseline ECOG-PS. The mean distance was 416.8 m (SD 55.4) in patients with ECOG-PS 0, 369.8 m (SD 80.3) in those with ECOG-PS 1, and 305.7 m (SD 89.1) in those with ECOG-PS 2 (*P*=.04). After PR, the 6MWD had improved significantly: 380.1 m (SD 74.1) at baseline, 429.1 m (SD 58.6, *P*<.001) at 6 weeks, and 448.1 m (SD 50.0, *P*<.001) at 12 weeks (Figure 1). We investigated the exercise capacity depending on their treatment response. Consequently, patients with stable disease showed significantly improved 6MWD: 384.2 m (SD 74.6) at baseline, 426.1 m (SD 6.5, *P*<.001) at 6 weeks, and 447.4 m (SD 50.4, *P*<.001 at 12 weeks (Figure 1). However, exercise capacity remained unimproved in patients with progressive disease. The dyspnea scale, evaluated using the EORTC QLQ-C30, did not show any significant improvement in the patients overall, but patients with stable disease tended to improve.

Quality of Life, Functional, and Symptom Scales

Of 90 patients who completed the study, 86 patients completed all questions in the questionnaire EORTC QLQ-C30. The EORTC QLQ-C30 consists of a global health status and QoL scale, functional scales, and symptom scales. Global health status and QoL tended to improve in patients overall, although not statistically significant (Table 3), whereas all functional scales, except the cognitive scale, improved significantly after PR. Symptom scales, in which a high score represents more severe symptoms, for fatigue, appetite loss, and diarrhea showed significant improvement in patients overall. In patients with stable disease, global health status and QoL scale did not improve significantly. Nevertheless, similar improvement in functional and symptom scales were observed in these patients as in the patients overall.

Pain Control

Overall, the pain severity significantly decreased at 6 weeks: 1.7 (SD 2.2) at baseline and 1.2 (SD 1.8, *P*=.02) at 6 weeks, but not at 12 weeks (mean 1.4, SD 1.9; *P*=.20; Figure 2). In patients with stable disease, the NRS score tended to improve at 6 weeks: 1.7 (SD 2.2) at baseline and 1.2 (SD 2.0, *P*=.06) at 6 weeks, but not at 12 weeks (mean 1.5, SD 1.9; *P*=.79).

Distress Index

Overall, baseline distress indexes showed mild anxiety and depression (Figure 3). Low indexes represent less distress. Anxiety significantly improved at 12 weeks: 3.9 (SD 4.1; baseline), 3.4 (SD 3.7; 6 weeks *P*=.11), 2.4 (SD 3.8; 12 weeks, *P*<.001; Figure 3). Depression worsen at 6 weeks: 4.7 (SD 4.9; baseline), 5.0 (SD 5.2; 6 weeks, *P*=.44), but significantly improved at 12 weeks (mean 3.5, SD 4.5; *P*=.02; Figure 3). Only the depression index was associated with the treatment response (*P*=.04).

Service Satisfaction

Of 90 patients who completed the PR program, 69 (69/90, 77%) patients reported that they were satisfied with the service and 79 (79/90, 88%) reported that they would recommend it to others. Neither age nor home region affected on service satisfaction. A total of 86 (86/90, 96%) patients reported that they were paying more attention to their health or disease status since using the app. In addition, all patients reported that the management algorithms for adverse events were helpful for controlling symptoms and determining when to visit the hospital. Patients who reported dissatisfaction with the service mostly cited difficulty in handling the app and frequent system error.

Unexpected Visits to the Emergency Department

Of 991 patients who did not participate in this study, 302 (302/991, 30.5%) visited the ED. However, 15 (15/100, 15.0%) of 100 patients included in this study visited the ED during the same period; this indicated a significant reduction in frequency (*P*=.001). The baseline characteristics, such as sex and tumor stage, were not significantly different between 2 groups; baseline age was more in patients who did not participate in the study than in patients who did (62.9 years, SD 10.3 vs 55.2 years, SD 8.7; *P*=.03). However, there was no significant difference in age between patients who visited the ED and those who did not (62.6 years, SD 10.6 vs 63.0 years, SD 0.2; *P*=.71).

Table 2. Baseline characteristics of patients who participated in the mobile comprehensive rehabilitation program.

Baseline characteristics	Total (N=100)	Complete (n=90)	Interrupted (n=10)	*P* value
Age in years, mean (SD)	55.1 (8.7)	54.9 (8.8)	56.8 (7.7)	.53
Sex, n (%)				
Male	46 (46.0)	41 (46)	5 (50)	>.99
BMI[a] in kg/m^2, mean (SD)	24.2 (3.5)	24.4 (3.5)	22.0 (1.8)	.04
History of smoking, n (%)				.59
Current smoker	11 (11.0)	9 (10)	2 (20)	
Ex-smoker	28 (28.0)	26 (29)	2 (20)	
Never smoker	61 (61.0)	55 (61)	6 (60)	
ECOG[b], n (%)				.02
Zero	13 (13.0)	11 (12)	2 (20)	
One	83 (83.0)	77 (86)	6 (60)	
Two	4 (4.0)	2 (2)	2 (20)	
Time of diagnosis of lung cancer, n (%)				.82
Within 1 year	56 (56.0)	49 (54)	7 (70)	
Within 1-2 years	16 (16.0)	15 (17)	1 (10)	
Within 2-3 years	16 (16.0)	15 (17)	1 (10)	
>2 years	12 (12.0)	11 (12)	1 (10)	
Stage, n (%)				>.99
II	5 (5.0)	5 (6)	0 (0)	
III	0 (0.0)	0 (0)	0 (0)	
IV	95 (95.0)	10 (100)	85 (94)	
Lung function, mean (SD)				
FVC[c] (%)	84.2 (17.7)	84.4 (17.3)	82.3 (23.6)	.79
FEV_1[d] (%)	80.7 (19.2)	80.9 (19.3)	79.2 (19.5)	.84
Diffusing lung capacity (%)	82.2 (15.1)	82.8 (15.0)	73.3 (15.2)	.22
Line of chemotherapy, n (%)				.76
First	73 (73.0)	66 (73)	7 (70)	
Second	12 (12.0)	10 (11)	2 (20)	
Third and more	10 (10.0)	9 (10)	1 (10)	
Adjuvant	5 (5.0)	5 (6)	0	
Histology, n (%)				.70
Adenocarcinoma	94 (94.0)	84 (93)	10 (100)	
Squamous cell carcinoma	4 (4.0)	4 (4)	0	
Others	2 (2.0)	2 (2)	0	
Regimen of chemotherapy, n (%)				.78
Tyrosine kinase inhibitor	40 (40.0)	36 (40)	4 (40)	
Platinum-based chemotherapy	38 (38.0)	35 (39)	3 (30)	
Others	22 (22.0)	19 (21)	3 (30)	
Response rate[e], n (%)				.001
Complete response	0 (0.0)	0 (0)	0 (0)	
Partial response	15 (15.0)	14 (16)	1 (10)	

Baseline characteristics	Total (N=100)	Complete (n=90)	Interrupted (n=10)	*P* value
Stable disease	69 (69.0)	66 (73)	3 (30)	
Progressive disease	16 (16.0)	10 (11)	6 (60)	

[a]BMI: body mass index.

[b]ECOG: Eastern Cooperative Oncology Group.

[c]FVC: forced vital capacity.

[d]FEV_1: forced expiratory volume in 1 second.

[e]Disease response was evaluated using response evaluation criteria in solid tumors criteria in patients with stage IV lung cancer. We considered patients with stage II lung cancer to have stable disease status.

Figure 1. Exercise capacity. Six-minute walking distance improved significantly after pulmonary rehabilitation in the patients overall (A) and patients with stable tumor response (B).

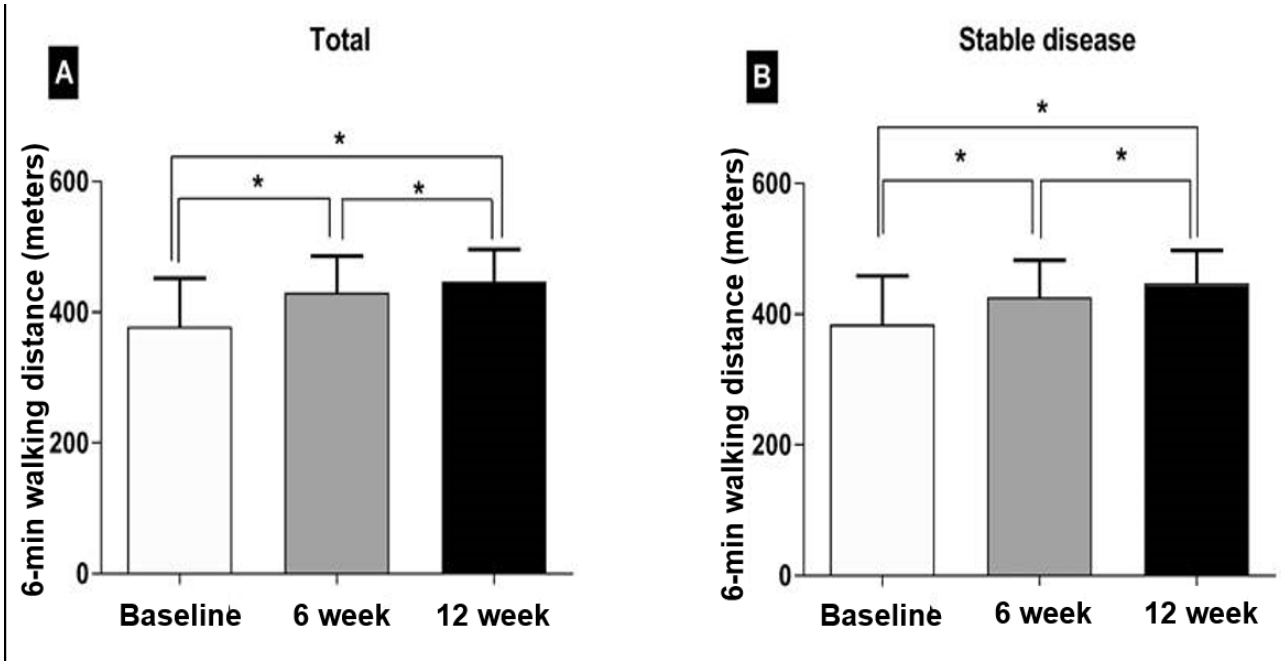


Table 3. Quality of life, functional, and symptom scale.

EORTC QLQ-C30[a]	Total (n=86)			Stable disease (n=65)			Progressive disease (n=7)		
	Baseline, mean (SD)	12 week, mean (SD)	*P* value	Baseline, mean (SD)	12 week, mean (SD)	*P* value	Baseline, mean (SD)	12 week, mean (SD)	*P* value
Global health status or QoL[b] scale	64.1 (24.7)	69.3 (21.2)	.06	66.3 (22.0)	70.5 (19.7)	.15	53.6 (35.0)	67.9 (26.5)	.24
Functional scales									
Physical functioning	78.2 (14.3)	81.1 (15.7)	.06	77.6 (14.5)	81.2 (15.5)	.06	79.0 (18.2)	77.1 (20.7)	.63
Role functioning	75.0 (22.8)	81.4 (23.3)	.02	75.6 (22.5)	82.1 (23.4)	.048	73.8 (30.2)	73.8 (33.1)	>.99
Emotional functioning	73.7 (19.6)	83.7 (18.7)	<.001	74.9 (18.8)	82.9 (19.2)	.002	76.2 (21.2)	86.9 (19.8)	.33
Cognitive functioning	81.4 (15.8)	83.9 (18.7)	.25	81.3 (15.5)	82.8 (20.2)	.54	88.1 (18.5)	85.7 (17.8)	.82
Social functioning	74.6 (25.1)	82.8 (20.2)	.002	73.6 (25.5)	83.3 (20.2)	.001	73.8 (23.3)	83.3 (21.5)	.23
Symptom scales									
Fatigue	35.7 (21.2)	27.1 (22.3)	<.001	35.0 (22.1)	26.8 (22.2)	.001	36.5 (24.6)	34.9 (29.0)	.86
Nausea or vomiting	8.9 (15.9)	10.5 (14.7)	.45	8.2 (15.6)	10.8 (15.4)	.27	9.5 (13.1)	7.1 (8.9)	.36
Pain	20.2 (20.9)	22.9 (23.6)	.33	19.0 (21.4)	25.4 (24.7)	.06	23.8 (16.3)	14.3 (15.0)	.23
Dyspnea	26.7 (23.3)	25.2 (25.0)	.56	27.2 (22.7)	25.6 (23.4)	.07	19.0 (26.2)	38.1 (44.8)	.10
Insomnia	26.0 (28.2)	21.3 (28.0)	.12	25.1 (29.5)	22.1 (27.2)	.39	19.0 (32.5)	28.6 (48.8)	.17
Appetite loss	21.7 (25.4)	16.3 (21.5)	.047	21.5 (24.6)	14.9 (21.3)	.03	33.3 (38.5)	19.0 (26.2)	.20
Constipation	15.5 (23.8)	17.1 (23.3)	.65	16.4 (25.8)	15.9 (22.9)	.90	14.3 (17.8)	23.8 (25.2)	.36
Diarrhea	19.4 (26.8)	11.6 (19.6)	.01	21.0 (28.6)	10.8 (17.8)	.01	28.6 (23.0)	14.3 (26.2)	.08
Financial difficulties	24.8 (26.2)	21.3 (25.0)	.14	26.7 (27.8)	21.5 (25.3)	.049	28.6 (23.0)	19.0 (26.2)	.36

[a]EORTC QLQ-C30: European Organization for Research and Treatment of Cancer Quality of Life Questionnaire-C30.

[b]QoL: quality of life.

Figure 2. Pain scale. The pain severity, assessed by numeric rating scale, significantly decreased at 6 weeks but increased at 12 weeks (A). There was no significant improvement in patients with stable tumor response (B).

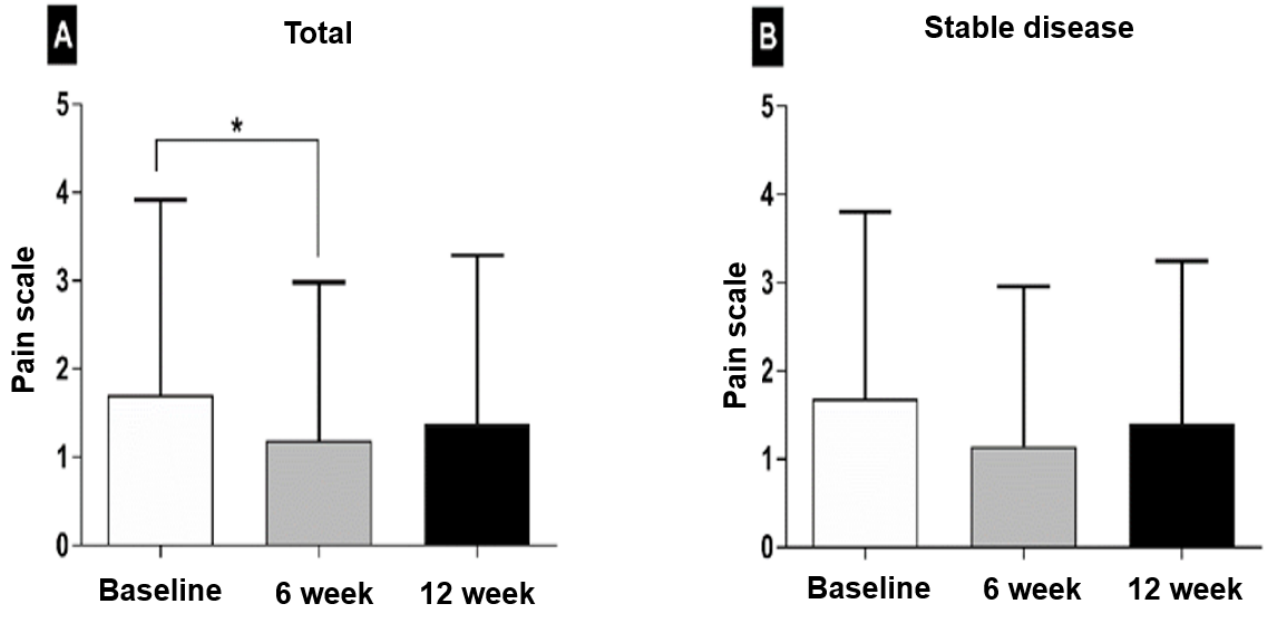
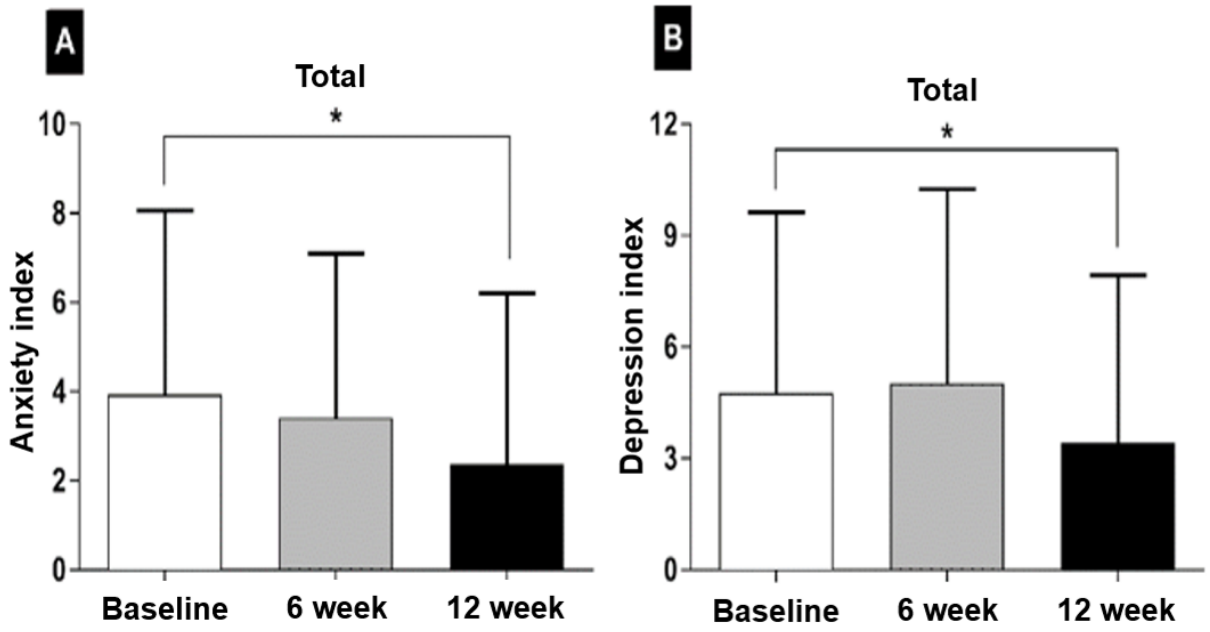


Figure 3. Distress indexes. Anxiety (A) and depression (B) significantly improved after pulmonary rehabilitation.



Discussion

Principal Findings

The 12-week Smart Aftercare PR program provided comprehensive management of common challenges of patients with advanced lung cancer. To the best of our knowledge, this is the first study that demonstrated the feasibility and efficacy of smartphone app–based PR for improving exercise capacity and symptoms in patients with advanced NSCLC during chemotherapy.

Because many lung cancer patients are elderly, there is some concern about difficulty in handling a smartphone app. However, only 2 patients ceased participation for this reason. The most important cause of drop out was cancer progression. A total of 16 patients suffered from cancer progression during the 12 weeks' PR program, and 2 patients died. Patients with advanced cancer will unavoidably suffer symptomatic deterioration with disease progression. However, exercise capacity, functional status, various symptoms, and distress indexes significantly improved in patients with stable disease. The objective response rate to estimated glomerular filtration rate tyrosine kinase inhibitors exceeds 70%, and the median duration of response to an immune checkpoint inhibitor is 10 to 25 months [2,3,5,6]. Therefore, the role of PR in advanced NSCLC seems increasingly important.

Exercise capacity is associated with advanced lung cancer prognosis [9]. Multiple small trials have demonstrated the benefit of hospital-based and home-based PR in patients with advanced lung cancer [14-17]. However, no studies have assessed PR employing a mobile app. The 6MWD significantly improved after PR in patients overall as well as in patients with stable disease. Although the dyspnea scale was not significantly improved, it tended to improve in patients with stable disease. Baseline performance status was significantly associated with baseline 6MWD but not with changes in absolute distance. Rather, patients with ECOG-PS 1 showed a mean increase in walking distance of 17 m more as compared with patients with ECOG-PS 0. One patient with ECOG-PS 2 showed an increase in walking distance of 132 m after the 12-week PR. Therefore, PR should be recommended to patients regardless of their disease status or PS, if feasible.

Smartphone-based PR provided significant improvement in functional scales, symptomatic scales, and distress indexes. Of 90 patients who completed the PR program, 86 (86/90, 96%) patients reported that smartphone-based PR promoted behavioral changes and facilitated self-monitoring of symptoms. This might have affected functional scale improvement, especially emotional scales and distress indexes. In addition, Shallwani et al reported that 6MWT distance was the predictor of change in the mental component of QoL. Therefore, improvement of exercise capacity affected emotional scales and distress indexes [20]. Moreover, 73 (73/90, 81%) patients reported that they felt that they were in contact with their health care team. Indeed, 78 (78/100, 78.0%) patients enrolled in this study lived outside of Seoul. This approach might be useful for patients living in areas where an attending physician is not readily accessible. Patients who participated in this study had significantly fewer ED visits than patients who did not participate in the study. In addition, there were no adverse events related to the smartphone-based PR, such as condition deterioration because of immoderate exercise or medicine overdose.

Anxiety and depression were significantly improved at 12 weeks. Depression was associated with the treatment response. In subgroup analysis, patients who received first-line chemotherapy showed a subsequent reduction in depression index, with a significant reduction in score at 12 weeks (5.6 [SD 5.8] at baseline, 5.5 [SD 5.8] at 6 weeks, and 3.5 [SD 4.8]

at 12 weeks; *P*=.04). Furthermore, patients who received second-line chemotherapy and beyond showed higher depression index at 6 weeks (4.1 [SD 3.0] at baseline vs 5.3 [SD 5.6] at 6 weeks; *P*=.24) but lower index at 12 weeks (3.4 [SD 3.9], *P*=.33). Therefore, disease status and duration of illness were important factors in the distress index.

QoL and severity of pain were not significantly improved through PR. Pain is a key factor affecting QoL in patients with lung cancer [21]. We provided various kinds of services to help reduce pain, but they were ineffective. Therefore, self-management of pain using a smartphone app is insufficient and frequent assessment and physical examination at a clinic is needed. In contrast, treatment response, an important factor affecting health-related QoL, showed no significant association with QoL in this study [22].

Limitations

This study has several limitations. First, there was no control group. Uhm et al reported that mobile health management did not show significant superiority over a conventional program in terms of physical function in patients with breast cancer [23]. However, the lungs are strongly associated with dyspnea and exercise capacity, and the usefulness of home-based PR for chronic lung diseases and lung cancer has previously been reported [17,24,25]. Further randomized controlled studies are needed to prove the superiority of the smartphone app–based PR over conventional education. Second, a small number of patients with heterogeneous disease status were included. In addition, the influence of the different chemotherapy regimens, which are associated with different adverse reactions that influence symptoms and QoL, was not considered. Third, monitoring of the amount and intensity of exercise relied on the IoT device only. If the IoT device had a systemic error, all data gathered on the IoT platform are unreliable. However, several studies have demonstrated the viability of a smartphone for step counting or gait analysis, and thus, they can be used to automate the 6MWT [26-29].

Conclusions

In conclusion, 12 weeks of comprehensive smartphone app–based individualized PR seems to be an effective and feasible approach for improving exercise capacity, symptom management, and distress in patients with advanced NSCLC undergoing systemic chemotherapy.

Acknowledgments

This work was supported by the National Information Society Agency funded by the Ministry of Science and Information and Communication Technology (Grant number: 2016-0-00586).

Authors' Contributions

CMC and SP had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. CMC contributed to the design of the study and final approval of the paper. SP contributed to data analysis and drafting the paper. JYK provided counseling service through the app. JCL, HRK, and WJ contributed to subject recruitment and offered counsel for development of the app. SS and HK contributed to develop the app and IoT platform.

Conflicts of Interest

SS and HK are employed by LifeSemantics. LifeSemantics developed the Aftercare and owns full right to use it for commercial purposes as fit by the company.

Multimedia Appendix 1

The English version of the brochure introducing the function of the Aftercare app.

[PNG File, 383KB - mhealth_v7i2e11094_app1.png]

Multimedia Appendix 2

Questionnaire for service quality and satisfaction.

[PDF File (Adobe PDF File), 153KB - mhealth_v7i2e11094_app2.pdf]

References

1. Park JY, Jang SH. Epidemiology of lung cancer in Korea: recent trends. Tuberc Respir Dis (Seoul) 2016 Apr;79(2):58-69 [FREE Full text] [doi: 10.4046/trd.2016.79.2.58] [Medline: 27064578]
2. Mok TS, Wu Y, Thongprasert S, Yang CH, Chu DT, Saijo N, et al. Gefitinib or carboplatin-paclitaxel in pulmonary adenocarcinoma. N Engl J Med 2009 Sep 03;361(10):947-957. [doi: 10.1056/NEJMoa0810699] [Medline: 19692680]
3. Maemondo M, Inoue A, Kobayashi K, Sugawara S, Oizumi S, Isobe H, North-East Japan Study Group. Gefitinib or chemotherapy for non-small-cell lung cancer with mutated EGFR. N Engl J Med 2010 Jun 24;362(25):2380-2388. [doi: 10.1056/NEJMoa0909530] [Medline: 20573926]

4. Chang YS, Choi C, Lee JC. Mechanisms of epidermal growth factor receptor tyrosine kinase inhibitor resistance and strategies to overcome resistance in lung adenocarcinoma. Tuberc Respir Dis (Seoul) 2016 Oct;79(4):248-256 [FREE Full text] [doi: 10.4046/trd.2016.79.4.248] [Medline: 27790276]
5. Herbst RS, Baas P, Kim DW, Felip E, Pérez-Gracia JL, Han JY, et al. Pembrolizumab versus docetaxel for previously treated, PD-L1-positive, advanced non-small-cell lung cancer (KEYNOTE-010): a randomised controlled trial. Lancet 2016 Apr 09;387(10027):1540-1550. [doi: 10.1016/S0140-6736(15)01281-7] [Medline: 26712084]
6. Horn L, Spigel DR, Vokes EE, Holgado E, Ready N, Steins M, et al. Nivolumab versus docetaxel in previously treated patients with advanced non-small-cell lung cancer: two-year outcomes from two randomized, open-label, phase III trials (CheckMate 017 and CheckMate 057). J Clin Oncol 2017 Dec 10;35(35):3924-3933. [doi: 10.1200/JCO.2017.74.3062] [Medline: 29023213]
7. de Torres JP, Marín JM, Casanova C, Cote C, Carrizo S, Cordoba-Lanus E, et al. Lung cancer in patients with chronic obstructive pulmonary disease--incidence and predicting factors. Am J Respir Crit Care Med 2011 Oct 15;184(8):913-919. [doi: 10.1164/rccm.201103-0430OC] [Medline: 21799072]
8. Granger CL, McDonald CF, Irving L, Clark RA, Gough K, Murnane A, et al. Low physical activity levels and functional decline in individuals with lung cancer. Lung Cancer 2014 Feb;83(2):292-299. [doi: 10.1016/j.lungcan.2013.11.014] [Medline: 24360323]
9. Kasymjanova G, Correa JA, Kreisman H, Dajczman E, Pepe C, Dobson S, et al. Prognostic value of the six-minute walk in advanced non-small cell lung cancer. J Thorac Oncol 2009 May;4(5):602-607 [FREE Full text] [doi: 10.1097/JTO.0b013e31819e77e8] [Medline: 19276833]
10. Liu XL, Tan JY, Wang T, Zhang Q, Zhang M, Yao L, et al. Effectiveness of home-based pulmonary rehabilitation for patients with chronic obstructive pulmonary disease: a meta-analysis of randomized controlled trials. Rehabil Nurs 2014;39(1):36-59. [doi: 10.1002/rnj.112] [Medline: 23780865]
11. Grosbois JM, Gicquello A, Langlois C, Le Rouzic O, Bart F, Wallaert B, et al. Long-term evaluation of home-based pulmonary rehabilitation in patients with COPD. Int J Chron Obstruct Pulmon Dis 2015;10(1):2037-2044 [FREE Full text] [doi: 10.2147/COPD.S90534] [Medline: 26445534]
12. Ryerson CJ, Cayou C, Topp F, Hilling L, Camp PG, Wilcox PG, et al. Pulmonary rehabilitation improves long-term outcomes in interstitial lung disease: a prospective cohort study. Respir Med 2014 Jan;108(1):203-210 [FREE Full text] [doi: 10.1016/j.rmed.2013.11.016] [Medline: 24332409]
13. Dowman L, Hill CJ, Holland AE. Pulmonary rehabilitation for interstitial lung disease. Cochrane Database Syst Rev 2014 Oct 06(10):CD006322. [doi: 10.1002/14651858.CD006322.pub3] [Medline: 25284270]
14. Temel JS, Greer JA, Goldberg S, Vogel PD, Sullivan M, Pirl WF, et al. A structured exercise program for patients with advanced non-small cell lung cancer. J Thorac Oncol 2009 May;4(5):595-601 [FREE Full text] [doi: 10.1097/JTO.0b013e31819d18e5] [Medline: 19276834]
15. Quist M, Rørth M, Langer S, Jones LW, Laursen JH, Pappot H, et al. Safety and feasibility of a combined exercise intervention for inoperable lung cancer patients undergoing chemotherapy: a pilot study. Lung Cancer 2012 Feb;75(2):203-208. [doi: 10.1016/j.lungcan.2011.07.006] [Medline: 21816503]
16. Olivier C, Grosbois JM, Cortot AB, Peres S, Heron C, Delourme J, et al. Real-life feasibility of home-based pulmonary rehabilitation in chemotherapy-treated patients with thoracic cancers: a pilot study. BMC Cancer 2018 Dec 13;18(1):178 [FREE Full text] [doi: 10.1186/s12885-018-4102-6] [Medline: 29433474]
17. Jastrzębski D, Maksymiak M, Kostorz S, Bezubka B, Osmanska I, Młynczak T, et al. Pulmonary rehabilitation in advanced lung cancer patients during chemotherapy. Adv Exp Med Biol 2015;861:57-64. [doi: 10.1007/5584_2015_134] [Medline: 26017725]
18. Mosa AS, Yoo I, Sheets L. A systematic review of healthcare applications for smartphones. BMC Med Inform Decis Mak 2012;12:67 [FREE Full text] [doi: 10.1186/1472-6947-12-67] [Medline: 22781312]
19. Dorsey ER, Yvonne Chan YF, McConnell MV, Shaw SY, Trister AD, Friend SH. The use of smartphones for health research. Acad Med 2017 Feb;92(2):157-160. [doi: 10.1097/ACM.0000000000001205] [Medline: 27119325]
20. Shallwani SM, Simmonds MJ, Kasymjanova G, Spahija J. Quality of life, symptom status and physical performance in patients with advanced non-small cell lung cancer undergoing chemotherapy: an exploratory analysis of secondary data. Lung Cancer 2016 Dec;99:69-75. [doi: 10.1016/j.lungcan.2016.06.018] [Medline: 27565917]
21. Wong WS, Fielding R. Quality of life and pain in Chinese lung cancer patients: is optimism a moderator or mediator? Qual Life Res 2007 Feb;16(1):53-63. [doi: 10.1007/s11136-006-9106-z] [Medline: 17091368]
22. Wu Y, Fukuoka M, Mok TS, Saijo N, Thongprasert S, Yang JC, et al. Tumor response and health-related quality of life in clinically selected patients from Asia with advanced non-small-cell lung cancer treated with first-line gefitinib: post hoc analyses from the IPASS study. Lung Cancer 2013 Aug;81(2):280-287. [doi: 10.1016/j.lungcan.2013.03.004] [Medline: 23540718]
23. Uhm KE, Yoo JS, Chung SH, Lee JD, Lee I, Kim JI, et al. Effects of exercise intervention in breast cancer patients: is mobile health (mHealth) with pedometer more effective than conventional program using brochure? Breast Cancer Res Treat 2017 Feb;161(3):443-452. [doi: 10.1007/s10549-016-4065-8] [Medline: 27933450]

24. Pradella CO, Belmonte GM, Maia MN, Delgado CS, Luise APT, Nascimento OA, et al. Home-based pulmonary rehabilitation for subjects with COPD: a randomized study. Respir Care 2015 Apr;60(4):526-532. [doi: 10.4187/respcare.02994] [Medline: 25269680]
25. Maltais F, Bourbeau J, Shapiro S, Lacasse Y, Perrault H, Baltzan M, Chronic Obstructive Pulmonary Disease Axis of Respiratory Health Network, Fonds de recherche en santé du Québec. Effects of home-based pulmonary rehabilitation in patients with chronic obstructive pulmonary disease: a randomized trial. Ann Intern Med 2008 Dec 16;149(12):869-878. [doi: 10.7326/0003-4819-149-12-200812160-00006}] [Medline: 19075206]
26. Huang Y, Zheng H, Nugent C, McCullagh P, Black N, Burns W, et al. An orientation free adaptive step detection algorithm using a smart phone in physical activity monitoring. Health Technol 2012 Oct 26;2(4):249-258 [FREE Full text] [doi: 10.1007/s12553-012-0035-2]
27. Naqvi NZ, Kumar A, Chauhan A, Sahni K. International Journal on Computer Science and Engineering. 2012 May. Step counting using smartphone-based accelerometer URL: http://www.enggjournals.com/ijcse/doc/IJCSE12-04-05-266.pdf [accessed 2018-11-30] [WebCite Cache ID 74JNgJTGh]
28. Nishiguchi S, Yamada M, Nagai K, Mori S, Kajiwara Y, Sonoda T, et al. Reliability and validity of gait analysis by android-based smartphone. Telemed J E Health 2012 May;18(4):292-296. [doi: 10.1089/tmj.2011.0132] [Medline: 22400972]
29. Carlson RH, Huebner DR, Hoarty CA, Whittington J, Haynatzki G, Balas MC, et al. Treadmill gait speeds correlate with physical activity counts measured by cell phone accelerometers. Gait Posture 2012 Jun;36(2):241-248 [FREE Full text] [doi: 10.1016/j.gaitpost.2012.02.025] [Medline: 22475727]

Abbreviations

6MWD: 6-min walking distance
6MWT: 6-min walk test
COPD: chronic obstructive lung disease
ECOG-PS: Eastern Cooperative Oncology Group-Performance Status
ED: emergency department
EORTC QLQ-C30: European Organization for Research and Treatment of Cancer Quality of Life Questionnaire-C30
FEV_1: forced expiratory volume in 1 second
FVC: forced vital capacity
GAD-7: Generalized Anxiety Disorder-7
HR: heart rate
ILD: interstitial lung disease
IoT: Internet of Things
NRS: numeric rating scale
NSCLC: non–small cell lung cancer
PHQ-9: Patient Health Questionnaire-9
PR: pulmonary rehabilitation
QoL: quality of life

Edited by G Eysenbach; submitted 19.05.18; peer-reviewed by YR Park, SY Shin, T Wang; comments to author 03.09.18; revised version received 23.10.18; accepted 10.11.18; published 04.02.19.

Please cite as:
Park S, Kim JY, Lee JC, Kim HR, Song S, Kwon H, Ji W, Choi CM
Mobile Phone App–Based Pulmonary Rehabilitation for Chemotherapy-Treated Patients With Advanced Lung Cancer: Pilot Study
JMIR Mhealth Uhealth 2019;7(2):e11094
URL: http://mhealth.jmir.org/2019/2/e11094/
doi:10.2196/11094
PMID:30714943

Original Paper

The Effectiveness and Safety of Utilizing Mobile Phone–Based Programs for Rehabilitation After Lumbar Spinal Surgery: Multicenter, Prospective Randomized Controlled Trial

Jingyi Hou[1*], MD; Rui Yang[1*], MD; Yaping Yang[2], PhD; Yiyong Tang[1], MS; Haiquan Deng[3], MS; Zhong Chen[1], MS; Yanfeng Wu[4*], MD; Huiyong Shen[5*], MD

[1]Department of Orthopedics, Sun Yat-sen Memorial Hospital, Sun Yat-sen University, Guangzhou, China

[2]Guangdong Provincial Key Laboratory of Malignant Tumor Epigenetics and Gene Regulation, Breast Tumor Center, Sun Yat-sen Memorial Hospital, Sun Yat-sen University, Guangzhou, China

[3]Department of Orthopedics, Guangxi Region People's Hospital, Nanning, China

[4]Department of Biotherapy Center, Sun Yat-sen Memorial Hospital, Sun Yat-sen University, Guangzhou, China

[5]Department of Orthopedics, 8th Affiliated Hospital of Sun Yat-sen University, Sun Yat-sen University, Shenzhen, China

[*]these authors contributed equally

Corresponding Author:
Jingyi Hou, MD
Department of Orthopedics
Sun Yat-sen Memorial Hospital
Sun Yat-sen University
No. 107, Yanjiangxi Road, Yuexiu District
Guangzhou,
China
Phone: 86 020 81332496
Email: billhou246@gmail.com

Abstract

Background: Rehabilitation is crucial for postoperative patients with low back pain (LBP). However, the implementation of traditional clinic-based programs is limited in developing countries, such as China, because of the maldistribution of medical resources. Mobile phone–based programs may be a potential substitute for those who have no access to traditional rehabilitation.

Objective: The aim of this study was to examine the efficacy of mobile phone–based rehabilitation systems in patients who underwent lumbar spinal surgery.

Methods: Patients who accepted spinal surgeries were recruited and randomized into 2 groups of rehabilitation treatments: (1) a mobile phone–based eHealth (electronic health) program (EH) or (2) usual care treatment (UC). The primary outcomes were (1) function and pain status assessed by the Oswestry Disability Index (ODI) and (2) the visual analog scale (VAS). Secondary outcomes were (1) general mental health and (2) quality of life (Likert scales, EuroQol-5 Dimension health questionnaire, and 36-item Short-Form Health Survey). All the patients were assessed preoperatively and then at 3, 6, 12, and 24 months postoperatively.

Results: A total of 168 of the 863 eligible patients were included and randomized in this study. Our analysis showed that the improvement of primary outcomes in the EH group was superior to the UC group at 24 months postoperatively (ODI mean 7.02, SD 3.10, P<.05; VAS mean 7.59, SD 3.42, P<.05). No significant difference of primary outcomes was found at other time points. A subgroup analysis showed that the improvements of the primary outcomes were more significant in those who completed 6 or more training sessions each week throughout the trial (the highest compliance group) compared with the UC group at 6 months (ODI mean 17.94, SD 5.24, P<.05; VAS mean 19.56, SD 5.27, P<.05), 12 months (ODI mean 13.39, SD 5.32, P<.05; VAS mean 14.35, SD 5.23, P<.05), and 24 months (ODI mean 18.80, SD 5.22, P<.05; VAS mean 21.56, SD 5.28, P<.05).

Conclusions: This research demonstrated that a mobile phone–based telerehabilitation system is effective in self-managed rehabilitation for postoperative patients with LBP. The effectiveness of eHealth was more evident in participants with higher compliance. Future research should focus on improving patients' compliance.

Trial Registration: Chinese Clinical Trial Registry ChiCTR-TRC-13003314; http://www.chictr.org.cn/showproj.aspx?proj=6245 (Archived by WebCite at http://www.webcitation.org/766RAIDNc)

(JMIR Mhealth Uhealth 2019;7(2):e10201) doi:10.2196/10201

KEYWORDS

mobile phone; low back pain; rehabilitation

Introduction

Background

Low back pain (LBP) is a common health problem with a point prevalence of 15% and a lifetime incidence as high as 85%. LBP is related to disability, work loss, and also accounts for high economic costs in society. For example, total annual back pain–related costs in the United States exceed US $200 billion and are still increasing [1,2].

Surgical intervention is an important treatment for LBP [3]. Due to the growing geriatric population, the number of lumbar spinal surgeries has increased rapidly. However, functional improvement and patient satisfaction after the surgery are varied; 25.0% (49/196) of the patients who underwent interbody fusion still suffer from back pain or thigh pain [4,5]. Of those who underwent discectomies, 40.0% (86/215) still remain unsatisfied with the operation and the recurrent rate can reach as high as 12.0% (24/200) [6].

Recent evidence suggests that patients with low back pain usually have inherent muscle dysfunction, which may be exacerbated by surgeries. Thus, postoperative rehabilitation becomes very critical [7,8]. Abbott et al suggested that if the patients carried out postoperative rehabilitation 3 months after surgery, their improvements in pain and function would be superior to those without any rehabilitation program [9]. A systemic review also indicated that postoperative rehabilitation would contribute to rapid functional recovery and pain alleviation of patients after lumbar spinal surgery [10]. It is also well accepted that, in order for the rehabilitation program to be fully effective, patients have to remain adherent and achieve total completion of the program [11,12]. However, there are a few efficient and effective strategies to help patients in maintaining and completing their rehabilitation therapy. On top of that, most of the recommended rehabilitation programs are clinic-based, in which patients have to visit the clinics for a total of 8 to 12 times in the span of 6 months. Personal visits to clinics are a major setback in these programs as patients in developing countries such as China have to travel a long distance to receive their treatments because of the maldistribution of medical resources [13]. Considering the cost and time, many patients eventually opt out of the rehabilitation program.

The rapid development of mobile phone–based programs provides a new option to promote health and prevent diseases [14-16]. Quinn et al used a mobile phone–based software to provide behavioral therapy for type 2 diabetes mellitus [14]. Other studies have also proved that the internet could be a useful tool to promote weight loss [17], increase physical activity [18], and improve self-management behaviors [19].

Objectives

This trial was conducted to investigate whether a mobile phone–based program (electronic health; eHealth), designed to provide telerehabilitation for patients with LBP, would reduce pain-related disability and improve prognosis among postoperative patients who have no access to traditional clinic-based rehabilitation.

Methods

Trial Design

This study was a multicenter, prospective, randomized controlled trial, approved by the Ethics Committees of the Sun Yat-sen Memorial Hospital. All the 3 hospitals that participated in this study were affiliated to the Sun Yat-sen University, where the surgery could be carried out safely and skillfully. All the patients were assessed for postoperative functional ability, pain, and general mental and health status at baseline and 3, 6, 12, and 24 months.

Inclusion and Exclusion Criteria

The researchers were required to explain the purposes, procedures, and possible risks of the trial in detail to the patients before inclusion. Written informed consents were obtained from all patients. The inclusion and exclusion criteria are shown in Textbox 1.

Sample Size

On the basis of previous studies, we anticipated that to have a 90% chance of detecting a between-group difference of 8 points on the Oswestry Disability Index (ODI) and declaring it statistically significant using a two-sided alpha=.05, an enrollment of 168 patients was required. This calculation allowed for a loss to follow-up of 23% [11].

Randomization

After completing the baseline survey, each participant was randomly allocated in a 1:1 ratio to the mobile phone–based eHealth program (EH) group or usual care treatment (UC) group according to a computer-generated randomization list. The allocation was stratified by a surgeon, operative procedures, and preoperative diagnosis. An email was sent to participants to inform them about their group assignment. The allocation sequence was concealed from the researchers enrolling and assessing patients.

Textbox 1. Inclusion and exclusion criteria for the study.

Inclusion criteria

- Aged between 18 and 64 years
- Agreed to receive lumbar spinal surgery and that the surgical intervention involved no more than 3 columns
- Diagnosed as lumbar disc herniation, spinal stenosis, or lumbar spondylolisthesis with imaging support
- Living at least 100 km or a 2 hours' drive away from the hospitals
- Signed the informed consent

Exclusion criteria

- Diagnosed as tuberculosis and tumor patients
- Those who accepted lumbar surgeries before this trial
- Patients with rheumatoid arthritis or ankylosing spondylitis
- Pregnancy
- Those who could not sign the informed consent or complete the rehabilitation exercise because of mental retardation or other reasons

Intervention

Usual Care

No specific rehabilitation program was provided to patients randomized to the UC control group. The relevant surgeons' usual practice was still provided, including advices to keep physically active and simple instructions to train the back muscles. Analgesia and other symptomatic treatments were also provided when necessary. All the postoperative regimes were documented.

eHealth

Besides the relevant surgeons'usual practice, patients randomized to the intervention group received telerehabilitation provided by eHealth, a mobile phone–based system developed by our group.

Moreover, eHealth was designed based on the user-centered theory, aimed to provide a platform for the delivery of self-management interventions [20,21]. It contained 2 interfaces: a mobile phone–based interface for patients and a Web-based interface for doctors. Through the mobile phone–based interface, patients were able to view the rehabilitation plans made by their physicians and conduct their rehabilitation following the video instructions. In addition, patients could receive daily reports about their exercise and alerts to prompt them to return to this system. They could also communicate with their doctors through this system. Through the Web-based interface, the doctors could adjust rehabilitation plans for patients and view reports about the patients' daily exercise. All data were synchronized and stored in a remote server. The eHealth system diagram is presented in Figure 1.

The exercises included in this software were designed based on core stability exercise principles, which were all aimed to restore normal muscle strength and mobility, to activate the deep core musculature and to promote balance and coordination of the patients' daily movements. The detailed plan of rehabilitation is shown in Multimedia Appendix 1.

The validation study was conducted with 10 healthy adults. Then, the information for the usability of the system was collected through paper-based questionnaires, 1 day after the tryout. The validation study confirmed that our system was well designed and easy to use, and the rehabilitation guidance was easy for users to understand. Combined with the results of previous studies and user preferences [11,22], our study set the rehabilitation for 20 min each time, twice a day (see Multimedia Appendix 2).

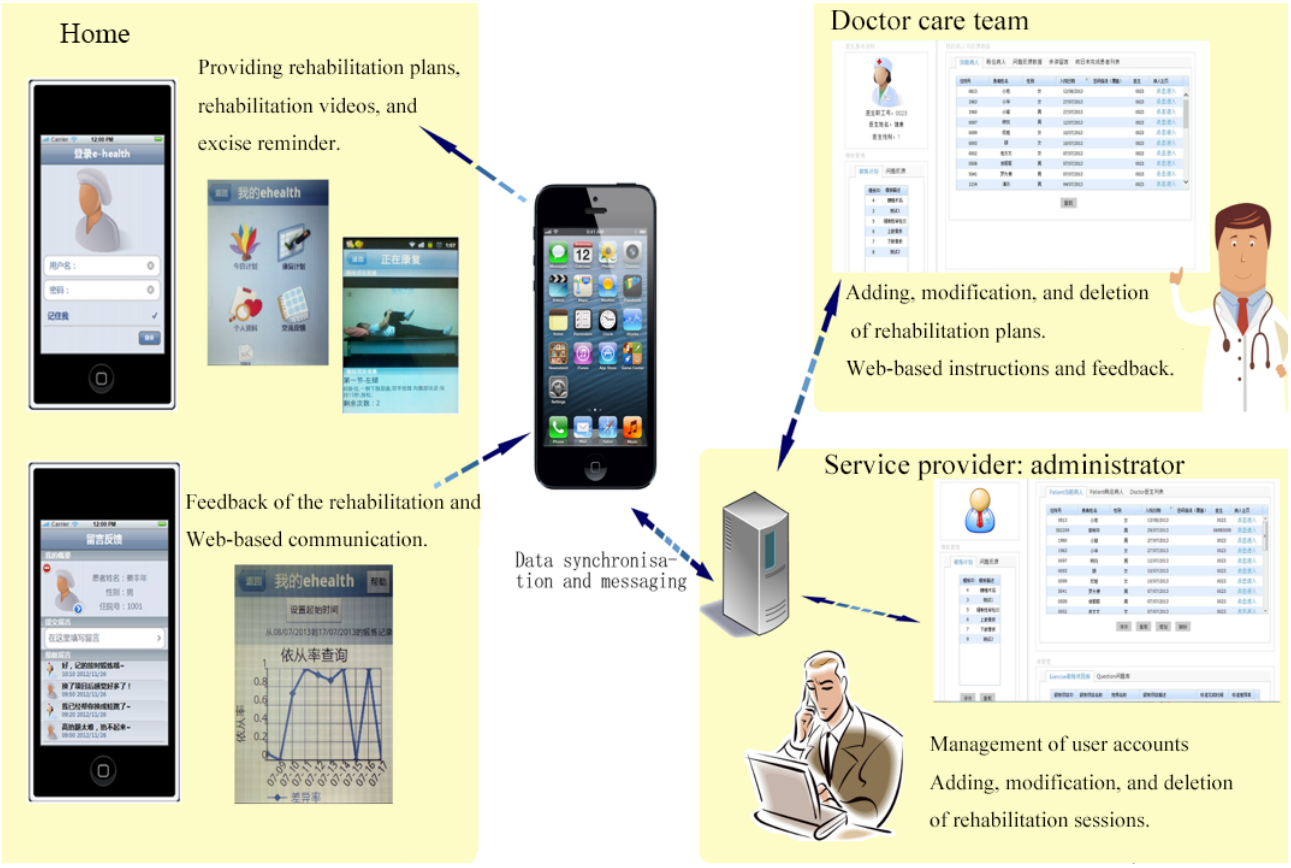
The software was installed into the patients' phones 3 months after the surgery. Two meetings were held to show the patients how to use this software and how to conduct the exercises. The patients were also evaluated to make sure they can conduct the rehabilitation exercise correctly. They were required to complete at least 2 months of training. After 2 months, the patients could still log on to the system, and those who completed 5 or more training sessions each week were considered as high adherence, 3 to 5 training sessions as medium adherence, and 2 training sessions and less as low adherence.

Outcome Measures

The primary outcome measures were the ODI, a disease-specific questionnaire documenting the function of known validity and reliability, and the visual analog scale (VAS) to record back pain [23,24]. The study was complemented by a series of secondary outcome measures of mental health and life status, which included the EuroQol 5-Dimension health questionnaire and 36-item Short-Form Health Survey (SF-36)—the Medical Outcomes Study SF-36 [25,26]. After baseline data collection, paper-based surveys of primary and secondary outcomes were conducted at 3, 6, 12, and 24 months. Since all outcome measurements were patient assessments, it was not possible to evaluate ODI, VAS, EQ-5D, SF-36 or Likert score blind to the randomized intervention.

At 12 months postoperative, an open survey was also conducted to detect the factors that affect patient compliance. All patients with medium and low compliance were asked to list 3 of the most important factors that they thought affected their compliance to the system.

Figure 1. The eHealth system contained 2 interfaces: mobile phone–based interface and Web-based interface. Through mobile phone–based interface, patients were able to view the rehabilitation plans and conduct their rehabilitation following the video instructions. Daily reports and an alert were sent to prompt them to return to this system. They could also communicate with doctors through this system. Through Web-based interface, the doctors could adjust rehabilitation plans for patients and view reports about their daily exercise. All data were synchronized and stored in a remote server.



Statistical Methods

The analyst assessing trial outcomes was blinded to the assignments. All analyses were conducted using an intent-to-treat approach with participants analyzed according to original group assignment. The baseline data for those lost to follow-up were included. Baseline characteristics were compared between the groups using chi-square tests for categorical data and a 2-sample *t* test for continuous data. Numeric data were represented by mean (SD).

For analyses of primary and secondary outcomes, a paired *t* test was applied to examine the changes within groups. A 2-sample *t* test was applied to compare changes between the groups. Missing data were not imputed. Only available data were analyzed. Compliance rates and lost to follow-up rates were compared with groups using chi-square tests.

All the analyses were conducted using Stata version 23.0 (StataCorp LLC) and a $P<.05$ was declared as significant.

Results

Study Population and Follow-Up

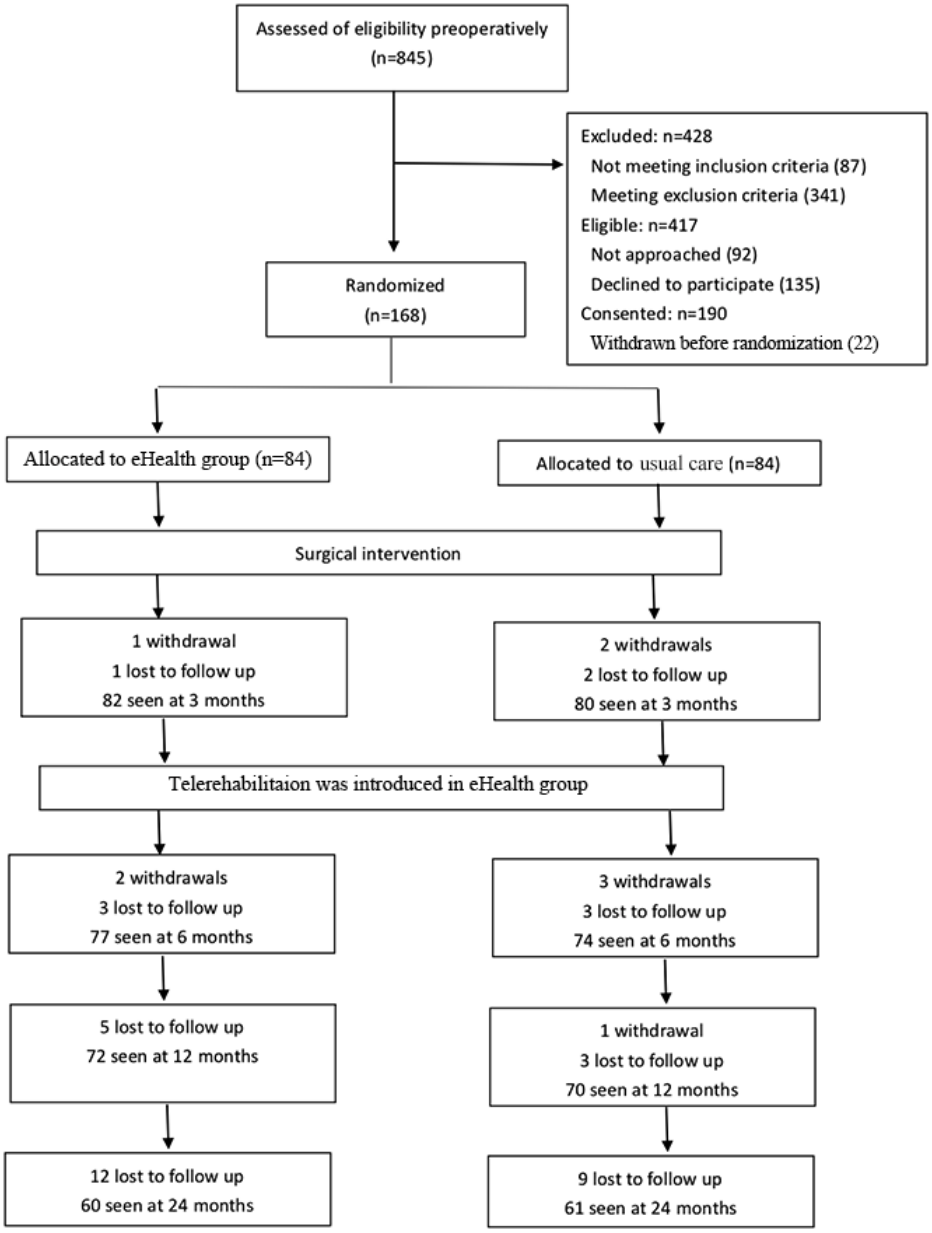
Recruitment occurred between August 2013 and November 2014 at 3 hospitals, and 845 patients were assessed for eligibility. Of those, 428 patients were excluded for not meeting the inclusion criteria or meeting the exclusion criteria. Of the 417 eligible patients, 92 were not approached, 135 declined to participate, and 22 consented patients withdrew before randomization. The final 168 consenting patients were then randomized in this study.

All the randomized patients received operation treatments and completed the required baseline assessments. However, during the study, 2 patients in the EH group and 4 patients in the UC group dropped out at 3 months. From the EH group, 82 patients entered the treatment phase, of which 77 finished the treatment and follow-up was done at 6 months. In the UC group, 80 and 74 patients were met for follow-up at 3 months and 6 months, respectively. The follow-up rate in the EH group was 97.62% (82/84) at 3 months, 91.67% (77/84) at 6 months, 85.71% (72/84) at 12 months, and 71.43% (60/84) at 24 months. In UC group, the follow-up rate was 95.24% (80/84) at 3 months, 88.10% (74/84) at 6 months, 83.33% (70/84) at 12 months, and 72.62% (61/84) at 24 months (see Figure 2).

Baseline Characteristics

Both the clinical and demographic characteristics of the patients were similar in the 2 groups ($P<.05$, see Table 1). Most of the study participants were married and had only finished high school or lower. On an average, participants reported moderate to severe pain and functional impairment based on ODI and VAS scores.

Figure 2. Flowchart.



Adherence to Interventions

Median eHealth attendance was 5 times per week (interquartile range, IQR, 4-6), 5 times per week (IQR 3-6), and 5 times per week (IQR 4-6) for 6, 12, and 24 months postoperatively. A total of 50, 37, and 38 patients were considered as high compliance at 6, 12, and 24 months, respectively, postoperatively. Although the high compliance rate was higher at 24 months (63.33%) compared with that at 12 months postoperatively (51.39%), it was not statistically significant (P>.05, see Table 2). Of these participants, 24 completed the whole trial with 6 or more training sessions each week and, therefore, were considered as the highest compliance (HC) group.

To determine the reason for low compliance, we carried out a brief survey focusing on the medium and low compliance group, asking them to list out the top 3 factors that affected their adherence at 12 months. A total of 33 out of 35 patients considered lack of communication with their doctors as the important factor. Through our record, we found that mean communication frequency was 2.54 (SD 0.89) for patients with medium or low compliance and 4.46 (SD 1.35) for patients with high compliance. The frequency of responses from doctors was significantly higher in patients with high compliance (P<.05). The other common reasons listed by patients were the concern about the accuracy of the action, limited symptom improvement, and lack of motivation (see Multimedia Appendix 3).

Primary Outcomes

The ODI and VAS for the EH and UC group were similar at baseline and 3 months postoperatively. At 6 and 12 months, the mean for change of the ODI from baseline was –7.27 (SD 5.31) and –18.43 (SD 23.92),respectively, for the EH,while it was –7.90 (SD 4.53) and –14.39 (SD 4.64), respectively, for the UC group (see Table 2). No significant difference was found between the EH and UC at 6 months and 12 months (P>.05).

Table 1. Demographics and baseline characteristics of all participants.

Characteristics	UC[a] (n=84)	EH[b] (n=84)	*P* value
Female, n (%)	42 (50)	48 (57)	.35
Age (years), mean (SD)	49.36 (9.52)	51.11 (9.54)	.24
Education status, n (%)			
High school or lower	63 (75)	60 (71)	.60
College degree or higher	21 (25)	24 (29)	.60
Currently employed	62 (74)	58 (69)	.50
Marriage status, n (%)			
Married	76 (90)	73 (87)	.72
Divorced	5 (6)	6 (7)	.72
Single	3 (4)	5 (6)	.72
Intervertebral discs involved in surgery, n (%)			
1 disc	37 (44)	36 (43)	.96
2 discs	41 (49)	41 (49)	.96
3 discs	6 (7)	7 (8)	.96
Mean ODI[c] score (SD)[d]	55.40 (14.78)	54.14 (15.18)	.59
Mean VAS[e] score (SD)[f]	60.11 (15.99)	57.71 (14.91)	.32
Mean Likert score (SD)[g]	59.14 (14.86)	59.71 (16.49)	.40
Mean EQ-5D[h] score (SD)[i]	35.75 (15.37)	34.26 (14.84)	.32
Mean SF-36 GH[j] score (SD)[k]	13.55 (5.58)	13.60 (6.02)	.96
Mean SF-36 PF[l] score (SD)[m]	21.11 (8.36)	21.52 (8.72)	.75

[a]UC: usual care.

[b]EH: eHealth program.

[c]ODI: Oswestry Disability Index.

[d]Rated to assess the patient's level of disability because of low back pain. Ranged from 0 to 100, with higher scores indicating more disability.

[e]VAS: visual analog scale.

[f]Rated between 0 and 100 with 100 representing worst pain possible.

[g]Measured using an 11-point numerical rating scale for average difficulty for movement in the previous week, where 0 indicated no difficulty and 10 indicated most difficulty.

[h]EQ-5D: EuroQol 5-Dimension health questionnaire.

[i]Rated between 0 and 100 with 100 representing a perfect health-related quality of life.

[j]SF-36 GH: General health for 36-item Short-Form Health Survey.

[k]Ranging from 0 to 100, with higher scores indicating better health-related quality of life.

[l]SF-36 PF: Physical functioning for 36-item Short-Form Health Survey.

[m]Ranging from 0 to 100, with higher scores indicating better health-related quality of life.

Table 2. Compliance status in different follow-ups.

Follow-up (month)	Low compliance, n (%)	Medium compliance, n (%)	High compliance, n (%)	*P* value for chi-square test
6	11 (14.47)	16 (19.74)	50 (65.79)	.23
12	8 (11.11)	27 (37.50)	37 (51.39)	.23
24	6 (11.48)	16 (26.23)	38 (62.29)	.23

Table 3. Primary outcomes change from baseline and between-group difference.

Measurements and follow-up (month)	UC[a] change from baseline		EH[b] change from baseline		Difference UC versus EH, mean (SD)	*P* value
	Participants	Mean (SD)	Participants	Mean (SD)		
ODI[c]						
3	80	–7.90 (4.53)	82	–7.27 (5.31)	–0.63 (0.78)	.42
6	74	–14.39 (4.64)	77	–18.43 (23.92)	4.0 (2.83)	.16
12	70	–22.07 (5.56)	72	–21.58 (24.64)	–0.49 (2.98)	.87
24	61	–23.41 (6.65)	60	–30.43 (23.75)	7.02 (3.18)	.03
VAS[d]						
3	80	–7.61 (5.15)	82	–7.02 (4.45)	–0.59 (0.76)	.44
6	74	–14.19 (5.11)	77	–17.49 (25.48)	3.30 (2.96)	.27
12	70	–21.94 (5.8)	72	–20.55 (25.92)	–1.39 (3.13)	.66
24	61	–22.36 (6.90)	60	–29.95 (25.60)	7.59 (3.42)	.03

[a]UC: usual care.

[b]EH: eHealth program.

[c]ODI: Oswestry Disability Index.

[d]VAS: visual analog scale.

However, at 24 months, improvement in the ODI was more significant in the EH group compared with the UC group (*P*<.05). For the VAS, the change was also not significantly different between the EH and UC at 6 months and 12 months (*P*>.05). At 24 months, the mean for change of the VAS from baseline was –22.36 (SD 6.90) in the EH and –29.95 (SD 25.60) in the UC. The improvement of the VAS was more significant in the EH than the UC (*P*<.05, see Table 3).

Secondary Outcomes

No difference in the Likert scale for movement was found at 3, 6, and 12 months postoperatively in the EH and UC. At 24 months, patients in the EH displayed superior results of Likert scale (mean of change: EH –32.51, SD 25.94; UC –22.54, SD 5.81; *P*<.05, see Table 4).

As for the EuroQol-5 Dimension (EQ-5D), the change was similar for EH and UC at 3 months. At 6 months, the improvement for EQ-5D was 0.23 (SD 0.03) for the EH and 0.13 (SD 0.08) for the UC. The patients in the EH got a significantly superior result over that in the UC. This advantage was sustained at subsequent time points (*P*<.05, see Table 4).

For the SF-36, the improvement was more significant in the EH compared with the UC at 3, 6, and 24 months (*P*<.05, see Table 4).

Subgroup Analysis

In the EH group, 24 patients completed all the follow-ups, with average eHealth attendance no less than 6 times per week, considered as the highest compliance. Thus, we conducted a subgroup analysis between the HC group with UC group.

Both the clinical and demographic characteristics were consistent between the 2 groups at baseline (see Multimedia Appendix 3). There were no significant differences in the changes of the ODI and VAS at 3 months. However, the HC group was superior to the UC group in the posttreatment ODI and VAS at 6 months (*P*<.05). This advantage was sustained at 12 and 24 months (see Table 5).

Adverse Events

Adverse events, mostly mild, self-limited joint and back pain, were reported in 9 EH and 6 UC participants. It did not differ significantly in frequency or severity of adverse events in these 2 groups.

Table 4. Secondary outcomes change from baseline and between-group difference outcomes.

Measurements and follow-up (month)	UC[a] change from baseline		EH[b] change from baseline		Difference UC versus EH, mean (SD)	*P* value
	Participants	Mean (SD)	Participants	Mean (SD)		
Likert score						
3	80	–7.79 (4.96)	82	–7.20 (4.74)	0.20 (0.78)	.80
6	74	–13.51 (5.39)	77	–19.66 (26.47)	6.15 (3.08)	.05
12	70	–21.09 (5.68)	72	–23.00 (27.12)	1.91 (3.31)	.56
24	61	–22.54 (5.81)	60	–32.51 (25.94)	9.98 (3.43)	.01
EQ-5D[c]						
3	80	0.09 (0.02)	82	0.09 (0.02)	0.00 (0.00)	.62
6	74	0.13 (0.08)	77	0.23 (0.03)	–0.10 (0.01)	<.001
12	70	0.17 (0.03)	72	0.24 (0.04)	–0.05 (0.01)	.003
24	61	0.22 (0.04)	60	0.35 (0.03)	–0.12 (0.01)	.001
SF-36 GH[d]						
3	80	38.16 (2.43)	82	40.01 (3.37)	–1.85 (0.46)	.004
6	74	45.85 (3.43)	77	54.75 (4.59)	–8.90 (0.66)	.002
12	70	55.53 (3.86)	72	56.25 (5.31)	–0.72 (0.78)	.36
24	61	57.98 (5.26)	60	62.80 (6.61)	–4.82 (1.09)	.002
SF-36 PF[e]						
3	80	30.58 (2.29)	82	40.76 (3.05)	–2.18 (0.42)	.004
6	74	46.35 (3.62)	77	56.12 (4.48)	–9.77 (0.66)	.003
12	70	56.13 (4.79)	72	56.74 (5.83)	–0.61 (0.89)	.49
24	61	59.07 (5.89)	60	62.45 (5.78)	–3.38 (1.06)	.02

[a]UC: usual care.

[b]EH: eHealth program.

[c]EQ-5D: EuroQol 5-Dimension health questionnaire

[d]SF-36 GH: General health for 36-item Short-Form Health Survey

[e]SF-36 PF: Physical functioning for 36-item Short-Form Health Survey

Table 5. Subgroup analysis of primary outcomes change from baseline and between-group difference outcomes.

Measurements and follow-up (month)	UC[a] change from baseline		HC[b] change from baseline		Difference UC versus HC, mean (SD)	*P* value
	Participants	Mean (SD)	Participants	Mean (SD)		
ODI[c]						
3	80	–7.90 (4.53)	24	–7.71 (5.29)	–0.19 (1.10)	.86
6	74	–14.39 (4.64)	24	–32.33 (25.56)	17.94 (5.24)	<.001
12	70	–22.07 (5.56)	24	–35.46 (25.88)	13.39 (5.32)	.02
24	61	–23.41 (6.65)	24	–42.21 (25.26)	18.80 (5.22)	.01
VAS[d]						
3	80	–7.61 (5.15)	24	–6.42 (4. 91)	–1.20 (1.19)	.32
6	74	–14.19 (5.11)	24	–33.75 (25.67)	19.56 (5.27)	.01
12	70	–21.94 (5.8)	24	–36.29 (25.38)	14.35 (5.23)	.01
24	61	–22.36 (6.90)	24	–43.92 (25.50)	21.56 (5.28)	.001

[a]UC: usual care.

[b]HC: highest compliance.

[c]ODI: Oswestry Disability Index.

[d]VAS: visual analog scale.

Discussion

Principal Findings

Much research had been done to explore the effect of rehabilitation on postoperative patients [9,27-29]. Compared with previous experiments, we aimed to explore the application of mobile phone–based rehabilitation in postoperative patients with LBP through a well-designed clinical trial. All the patients in our study lived far away from the hospital and were unable to accept the traditional clinic-based rehabilitation. This design is critical for developing countries. Due to the extremely uneven distribution of health resources, traditional clinic-based postoperative rehabilitation cannot be implemented in patients living faraway, whereas eHealth could be a solution to bridge the gap [30].

In this randomized controlled trial, we compared postoperative patients (EH group) with low back pain treated by eHealth, a mobile phone–based telerehabilitation system, with those that received nonspecific rehabilitation (UC group). We found that primary outcomes (ODI and VAS) in the EH group were superior to the UC group at 24 months postoperatively. However, no significant difference was found at all the other time points during follow-up. Furthermore, we compared 24 patients having an average eHealth attendance of no less than 6 times per week (HC group) with the UC group. Subgroup analysis showed that the improvements of the primary outcomes were more significant in the HC group compared with the UC group at 6, 12, and 24 months. These results suggest that patients with a higher compliance with our telerehabilitation system tend to have a better prognosis.

Adherence to postoperative rehabilitation in clinical practice is a serious problem [31]. Previous studies focused on clinic-based rehabilitation reported very low levels of compliance, with almost 30% failing to attend any classes, and of those attending, only 60% attended more than half of the classes [32,33]. Moreover, nonadherence of home-based rehabilitation could be as high as 50% [34]. Compared with previous studies, this study adopted video-based coaching and had a higher compliance: high compliance rates were 65.79%, 51.39%, and 62.29% at 6, 12, and 24 months, respectively. This paper supports the notion that well-designed connected health technologies, including digital, mobile health, and telehealth, could better support patients in their rehabilitation and provide an opportunity to increase adherence to exercise and rehabilitation [35].

This study also found that the 24 patients who completed 6 or more training sessions each week throughout the trial (HC group) had a better prognosis. This may be because the rehabilitation exercise needs to reach a certain length of time to achieve a more significant effect [36]. The rehabilitation protocols in previous studies were 40 to 60 min each day [11,22]. This study had set the rehabilitation for 20 min each time, twice a day, given the results of our previous validation study of user preferences (shown in Multimedia Appendix 2). However, in practice, many patients did not strictly conduct the exercise twice a day; therefore, the duration of the exercise for each day did not meet the requirements for a more efficient training session. Future research should pay more attention to the training duration of 1 single section and 1 day. Especially in the study of patients' self-rehabilitation, the training duration should be set for longer than what we expected. Meanwhile, the compliance to rehabilitation should be improved as much as possible. However, it must be noted that noncompliance is a complicated issue and the reasons for noncompliance are multifactorial. Our survey of patients with medium and low compliance at 12 months tried to find the main reasons for noncompliance, although most patients cited insufficient communication with doctors as the main cause for noncompliance. Further analysis implied that the patients' need

for communication with doctors was because of the patients' doubts during rehabilitation, such as whether the actions were standard, the intended goal of rehabilitation, or if more motivation was needed. All these reasons listed by the patients actually reflect that the doubt-free experience is very important for maintaining a high compliance to rehabilitation, especially for home-based rehabilitation. A well-designed system should possess the following features to create a doubt-free experience: (1) a clear goal, (2) comprehensive and detailed instructions, and (3) timely communication and motivation.

The abovementioned features could be achieved by optimizing the designs of a mobile phone–based system. Previous studies have revealed that the mobile phone is an effective tool in improving health behaviors of patients [37-39]. In a study conducted by Liu et al, patients with chronic obstructive pulmonary disease were encouraged to perform daily endurance walking while following the tempo of the music from a program installed on a mobile phone [40]. Lambert et al found that people with musculoskeletal conditions had better adherence to their home exercise programs provided on an app with remote support compared with paper handouts [41]. Meanwhile, the mobile phone can provide a lot of self-detecting methods, such as the Global Positioning System's positioning and gravity-sensing technology to track elderly patients and provide timely feedback to the medical staff [42]. The electrocardiogram, the blood glucose meter, and the blood pressure meter connected through Bluetooth could automatically upload health-related data to the medical system [43-45]. With the advancement of technology, motion-capture devices may also be applied in telerehabilitation [46]. Until then, the compliance could be better improved with remote monitoring of the patients' rehabilitation.

Limitations

The main limitation of this study was the high loss to follow-up rate. Compared with previous studies, the loss to follow-up rates were similar to the loss to follow-up rates in this study [47-49]. Further analysis showed that the baseline characteristics were similar among patients who were met or lost to follow-up at 24 months (P>.05, see Multimedia Appendix 3). Also, no significant difference was found in the improvement of primary outcomes at 12 months between patients who were met and were lost to follow-up between 12 and 24 months both in the UC and EH group (P>.05, see Multimedia Appendix 3). It implied that the reason for the loss to follow-up was the patients' own reason, which was randomized, instead of poor prognosis. The final result might not be seriously affected by those who were lost to follow-up.

The reason for the high loss to follow-up rate was probably because of the long distance from the patients' home to the hospital, which was 1 of the inclusion criteria. These inclusion criteria are consistent with our aim, which focused on remote self-rehabilitation. However, it also adds to the difficulty for following up. As our study was conducted with paper-based questionnaires, patients included in our study had to travel a long distance back to hospitals or mail back the questionnaires, causing more to be lost to follow-up. In order to overcome this problem, electronic surveys have great potential to improve data collection [50]. A major advantage of applying electronic surveys is that they could increase the amount of data collected at a lower cost [51]. Many researchers have explored the application of electronic versions of questionnaires. Terri et al found that electronic versions of the Faces Pain Scale-Revised and the Color Analog Scale on a mobile phone demonstrated good agreement with the original paper and plastic versions of these scales [52]. Pawar et al found that the software version of the Roland-Morris Disability Questionnaire was comparable with the paper version in patients with LBP [53]. Other researchers also found that Web- or mobile device–based systems could facilitate consecutive patient data collection in randomized controlled trials and could be used to increase response rates and enhance quality of research [54,55]. Therefore, future studies should consider applying electronic tools to simplify the follow-up process and reduce loss to follow-up.

Conclusions

In conclusion, eHealth, a mobile phone–based telerehabilitation system, may be an effective rehabilitation tool for postoperative patients with LBP, especially for those who have no access to traditional clinic-based rehabilitation. The effectiveness of eHealth was more evident in patients with higher adherence. However, more studies are still needed to find optimal methods to improve compliance.

Acknowledgments

This study was financially supported by the Guangdong Medical Research Fund Project (C2015048) and Guangdong Natural Science Foundation (2015A030310321). Phei Er Saw proofread and improved the language of this paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1

The detailed plan of rehabilitation.

[PDF File (Adobe PDF File), 1MB - mhealth_v7i2e10201_app1.pdf]

Multimedia Appendix 2

Results of the validation study.

[PDF File (Adobe PDF File), 53KB - mhealth_v7i2e10201_app2.pdf]

Multimedia Appendix 3

Supplementary tables.

[DOCX File, 27KB - mhealth_v7i2e10201_app3.docx]

Multimedia Appendix 4

CONSORT‐EHEALTH checklist (V 1.6.1).

[PDF File (Adobe PDF File), 505KB - mhealth_v7i2e10201_app4.pdf]

References

1. Juniper M, Le TK, Mladsi D. The epidemiology, economic burden, and pharmacological treatment of chronic low back pain in France, Germany, Italy, Spain and the UK: a literature-based review. Expert Opin Pharmacother 2009 Nov;10(16):2581-2592. [doi: 10.1517/14656560903304063] [Medline: 19874246]
2. Dagenais S, Caro J, Haldeman S. A systematic review of low back pain cost of illness studies in the United States and internationally. Spine J 2008;8(1):8-20. [doi: 10.1016/j.spinee.2007.10.005] [Medline: 18164449]
3. Issack PS, Cunningham ME, Pumberger M, Hughes AP, Cammisa Jr FP. Degenerative lumbar spinal stenosis: evaluation and management. J Am Acad Orthop Surg 2012 Aug;20(8):527-535. [doi: 10.5435/JAAOS-20-08-527] [Medline: 22855855]
4. Han X, Zhu Y, Cui C, Wu Y. A meta-analysis of circumferential fusion versus instrumented posterolateral fusion in the lumbar spine. Spine (Phila Pa 1976) 2009 Aug 1;34(17):E618-E625. [doi: 10.1097/BRS.0b013e3181a9beab] [Medline: 19644321]
5. Wiedenhöfer B, Ewerbeck V, Suda AJ, Carstens C. Current short- and long-term results of lumbar disc replacement : update 2008. Chirurg 2008 Oct;79(10):937-943. [doi: 10.1007/s00104-008-1519-5] [Medline: 18818895]
6. Li XC, Zhong CF, Deng GB, Liang RW, Huang CM. Full-endoscopic procedures versus traditional discectomy surgery for discectomy: a systematic review and meta-analysis of current global clinical trials. Pain Physician 2016 Mar;19(3):103-118 [FREE Full text] [Medline: 27008284]
7. Mayer TG, Vanharanta H, Gatchel RJ, Mooney V, Barnes D, Judge L, et al. Comparison of CT scan muscle measurements and isokinetic trunk strength in postoperative patients. Spine (Phila Pa 1976) 1989 Jan;14(1):33-36. [Medline: 2913665]
8. Hides JA, Richardson CA, Jull GA. Multifidus muscle recovery is not automatic after resolution of acute, first-episode low back pain. Spine (Phila Pa 1976) 1996 Dec 1;21(23):2763-2769. [Medline: 8979323]
9. Abbott AD, Tyni-Lenné R, Hedlund R. Early rehabilitation targeting cognition, behavior, and motor function after lumbar fusion: a randomized controlled trial. Spine (Phila Pa 1976) 2010 Apr 15;35(8):848-857. [doi: 10.1097/BRS.0b013e3181d1049f] [Medline: 20354468]
10. Oosterhuis T, Costa LO, Maher CG, de Vet HC, van Tulder MW, Ostelo RW. Rehabilitation after lumbar disc surgery. Cochrane Database Syst Rev 2014 Mar 14(3):CD003007. [doi: 10.1002/14651858.CD003007.pub3] [Medline: 24627325]
11. McGregor AH, Doré CJ, Morris TP, Morris S, Jamrozik K. ISSLS prize winner: Function After Spinal Treatment, Exercise, and Rehabilitation (FASTER): a factorial randomized trial to determine whether the functional outcome of spinal surgery can be improved. Spine (Phila Pa 1976) 2011 Oct 1;36(21):1711-1720. [doi: 10.1097/BRS.0b013e318214e3e6] [Medline: 21378603]
12. Christensen FB, Laurberg I, Bünger CE. Importance of the back-café concept to rehabilitation after lumbar spinal fusion: a randomized clinical study with a 2-year follow-up. Spine (Phila Pa 1976) 2003 Dec 1;28(23):2561-2569. [doi: 10.1097/01.BRS.0000097890.96524.A1] [Medline: 14652472]
13. Hu R, Dong S, Zhao Y, Hu H, Li Z. Assessing potential spatial accessibility of health services in rural China: a case study of Donghai County. Int J Equity Health 2013 May 20;12:35 [FREE Full text] [doi: 10.1186/1475-9276-12-35] [Medline: 23688278]
14. Quinn CC, Gruber-Baldini AL, Shardell M, Weed K, Clough SS, Peeples M, et al. Mobile diabetes intervention study: testing a personalized treatment/behavioral communication intervention for blood glucose control. Contemp Clin Trials 2009 Jul;30(4):334-346. [doi: 10.1016/j.cct.2009.02.004] [Medline: 19250979]
15. Quinn CC, Shardell MD, Terrin ML, Barr EA, Ballew SH, Gruber-Baldini AL. Cluster-randomized trial of a mobile phone personalized behavioral intervention for blood glucose control. Diabetes Care 2011 Sep;34(9):1934-1942 [FREE Full text] [doi: 10.2337/dc11-0366] [Medline: 21788632]
16. Webb TL, Joseph J, Yardley L, Michie S. Using the internet to promote health behavior change: a systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. J Med Internet Res 2010;12(1):e4 [FREE Full text] [doi: 10.2196/jmir.1376] [Medline: 20164043]

17. Carter MC, Burley VJ, Cade JE. Weight loss associated with different patterns of self-monitoring using the mobile phone app My Meal Mate. JMIR Mhealth Uhealth 2017 Feb 2;5(2):e8 [FREE Full text] [doi: 10.2196/mhealth.4520] [Medline: 28153814]
18. Direito A, Jiang Y, Whittaker R, Maddison R. Apps for IMproving FITness and increasing physical activity among young people: the AIMFIT pragmatic randomized controlled trial. J Med Internet Res 2015;17(8):e210 [FREE Full text] [doi: 10.2196/jmir.4568] [Medline: 26316499]
19. Tinschert P, Jakob R, Barata F, Kramer JN, Kowatsch T. The potential of mobile apps for improving asthma self-management: a review of publicly available and well-adopted asthma apps. JMIR Mhealth Uhealth 2017 Aug 2;5(8):e113 [FREE Full text] [doi: 10.2196/mhealth.7177] [Medline: 28768606]
20. Whitehead L, Seaton P. The effectiveness of self-management mobile phone and tablet apps in long-term condition management: a systematic review. J Med Internet Res 2016;18(5):e97 [FREE Full text] [doi: 10.2196/jmir.4883] [Medline: 27185295]
21. Rothgangel A, Braun S, Smeets R, Beurskens A. Design and development of a telerehabilitation platform for patients with phantom limb pain: a user-centered approach. JMIR Rehabil Assist Technol 2017 Feb 15;4(1):e2 [FREE Full text] [doi: 10.2196/rehab.6761] [Medline: 28582249]
22. Archer KR, Devin CJ, Vanston SW, Koyama T, Phillips SE, Mathis SL, et al. Cognitive-behavioral-based physical therapy for patients with chronic pain undergoing lumbar spine surgery: a randomized controlled trial. J Pain 2016 Jan;17(1):76-89 [FREE Full text] [doi: 10.1016/j.jpain.2015.09.013] [Medline: 26476267]
23. Hawker GA, Mian S, Kendzerska T, French M. Measures of adult pain: Visual Analog Scale for Pain (VAS Pain), Numeric Rating Scale for Pain (NRS Pain), McGill Pain Questionnaire (MPQ), Short-Form McGill Pain Questionnaire (SF-MPQ), Chronic Pain Grade Scale (CPGS), Short Form-36 Bodily Pain Scale (SF-36 BPS), and Measure of Intermittent and Constant Osteoarthritis Pain (ICOAP). Arthritis Care Res (Hoboken) 2011 Nov;63(Suppl 11):S240-S252 [FREE Full text] [doi: 10.1002/acr.20543] [Medline: 22588748]
24. Holm I, Friis A, Storheim K, Brox JI. Measuring self-reported functional status and pain in patients with chronic low back pain by postal questionnaires: a reliability study. Spine (Phila Pa 1976) 2003 Apr 15;28(8):828-833. [Medline: 12698128]
25. EuroQol Group. EuroQol--a new facility for the measurement of health-related quality of life. Health Policy 1990 Dec;16(3):199-208. [Medline: 10109801]
26. Ware Jr JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care 1992 Jun;30(6):473-483. [Medline: 1593914]
27. Kim BJ, Kim T, Ahn J, Cho H, Kim D, Yoon B. Manipulative rehabilitation applied soon after lumbar disc surgery improves late post-operative functional disability: A preliminary 2-year follow-up study. J Back Musculoskelet Rehabil 2017 Sep 22;30(5):999-1004. [doi: 10.3233/BMR-169546] [Medline: 28505954]
28. van Erp RM, Jelsma J, Huijnen IP, Lundberg M, Willems PC, Smeets RJ. Spinal surgeons' opinions on pre- and postoperative rehabilitation in patients undergoing lumbar spinal fusion surgery: a survey-based study in the Netherlands and Sweden. Spine (Phila Pa 1976) 2018 May 15;43(10):713-719. [doi: 10.1097/BRS.0000000000002406] [Medline: 28885297]
29. Madera M, Brady J, Deily S, McGinty T, Moroz L, Singh D, for the Seton Spine Rehabilitation Study Group. The role of physical therapy and rehabilitation after lumbar fusion surgery for degenerative disease: a systematic review. J Neurosurg Spine 2017 Jun;26(6):694-704. [doi: 10.3171/2016.10.SPINE16627] [Medline: 28291412]
30. Nedungadi P, Jayakumar A, Raman R. Personalized health monitoring system for managing well-being in rural areas. J Med Syst 2017 Dec 14;42(1):22. [doi: 10.1007/s10916-017-0854-9] [Medline: 29242996]
31. Argent R, Daly A, Caulfield B. Patient involvement with home-based exercise programs: can connected health interventions influence adherence? JMIR Mhealth Uhealth 2018 Mar 1;6(3):e47 [FREE Full text] [doi: 10.2196/mhealth.8518] [Medline: 29496655]
32. McGregor AH, Henley A, Morris TP, Doré CJ, FASTER team. An evaluation of a postoperative rehabilitation program after spinal surgery and its impact on outcome. Spine (Phila Pa 1976) 2012 Apr 1;37(7):E417-E422. [doi: 10.1097/BRS.0b013e31823b00b2] [Medline: 22024898]
33. Johnson RE, Jones GT, Wiles NJ, Chaddock C, Potter RG, Roberts C, et al. Active exercise, education, and cognitive behavioral therapy for persistent disabling low back pain: a randomized controlled trial. Spine (Phila Pa 1976) 2007 Jul 1;32(15):1578-1585. [doi: 10.1097/BRS.0b013e318074f890] [Medline: 17621203]
34. Bollen JC, Dean SG, Siegert RJ, Howe TE, Goodwin VA. A systematic review of measures of self-reported adherence to unsupervised home-based rehabilitation exercise programmes, and their psychometric properties. BMJ Open 2014 Jun 27;4(6):e005044 [FREE Full text] [doi: 10.1136/bmjopen-2014-005044] [Medline: 24972606]
35. Caulfield BM, Donnelly SC. What is Connected Health and why will it change your practice? QJM 2013 Aug;106(8):703-707. [doi: 10.1093/qjmed/hct114] [Medline: 23676416]
36. Wang XQ, Zheng JJ, Yu ZW, Bi X, Lou SJ, Liu J, et al. A meta-analysis of core stability exercise versus general exercise for chronic low back pain. PLoS One 2012;7(12):e52082 [FREE Full text] [doi: 10.1371/journal.pone.0052082] [Medline: 23284879]

37. Walker UA, Mueller RB, Jaeger VK, Theiler R, Forster A, Dufner P, et al. Disease activity dynamics in rheumatoid arthritis: patients' self-assessment of disease activity via WebApp. Rheumatology (Oxford) 2017 Dec 1;56(10):1707-1712 [FREE Full text] [doi: 10.1093/rheumatology/kex229] [Medline: 28957553]
38. Schnall R, Mosley JP, Iribarren SJ, Bakken S, Carballo-Diéguez A, Brown IW. Comparison of a user-centered design, self-management app to existing mHealth apps for persons living with HIV. JMIR Mhealth Uhealth 2015;3(3):e91 [FREE Full text] [doi: 10.2196/mhealth.4882] [Medline: 26385783]
39. Bendixen RM, Fairman AD, Karavolis M, Sullivan C, Parmanto B. A user-centered approach: understanding client and caregiver needs and preferences in the development of mHealth apps for self-management. JMIR Mhealth Uhealth 2017 Sep 26;5(9):e141 [FREE Full text] [doi: 10.2196/mhealth.7136] [Medline: 28951378]
40. Liu WT, Wang CH, Lin HC, Lin SM, Lee KY, Lo YL, et al. Efficacy of a cell phone-based exercise programme for COPD. Eur Respir J 2008 Sep;32(3):651-659 [FREE Full text] [doi: 10.1183/09031936.00104407] [Medline: 18508824]
41. Lambert TE, Harvey LA, Avdalis C, Chen LW, Jeyalingam S, Pratt CA, et al. An app with remote support achieves better adherence to home exercise programs than paper handouts in people with musculoskeletal conditions: a randomised trial. J Physiother 2017 Jul;63(3):161-167 [FREE Full text] [doi: 10.1016/j.jphys.2017.05.015] [Medline: 28662834]
42. van der Weegen S, Verwey R, Spreeuwenberg M, Tange H, van der Weijden T, de Witte L. The development of a mobile monitoring and feedback tool to stimulate physical activity of people with a chronic disease in primary care: a user-centered design. JMIR Mhealth Uhealth 2013 Jul 2;1(2):e8 [FREE Full text] [doi: 10.2196/mhealth.2526] [Medline: 25099556]
43. Seto E, Leonard KJ, Cafazzo JA, Barnsley J, Masino C, Ross HJ. Mobile phone-based telemonitoring for heart failure management: a randomized controlled trial. J Med Internet Res 2012;14(1):e31 [FREE Full text] [doi: 10.2196/jmir.1909] [Medline: 22356799]
44. Logan AG, McIsaac WJ, Tisler A, Irvine MJ, Saunders A, Dunai A, et al. Mobile phone-based remote patient monitoring system for management of hypertension in diabetic patients. Am J Hypertens 2007 Sep;20(9):942-948. [doi: 10.1016/j.amjhyper.2007.03.020] [Medline: 17765133]
45. Harris LT, Tufano J, Le T, Rees C, Lewis GA, Evert AB, et al. Designing mobile support for glycemic control in patients with diabetes. J Biomed Inform 2010 Oct;43(5 Suppl):S37-S40 [FREE Full text] [doi: 10.1016/j.jbi.2010.05.004] [Medline: 20937484]
46. Matheve T, Brumagne S, Demoulin C, Timmermans A. Sensor-based postural feedback is more effective than conventional feedback to improve lumbopelvic movement control in patients with chronic low back pain: a randomised controlled trial. J Neuroeng Rehabil 2018 Sep 26;15(1):85 [FREE Full text] [doi: 10.1186/s12984-018-0423-6] [Medline: 30253807]
47. Hellum C, Johnsen LG, Storheim K, Nygaard OP, Brox JI, Rossvoll I, Norwegian Spine Study Group. Surgery with disc prosthesis versus rehabilitation in patients with low back pain and degenerative disc: two year follow-up of randomised study. Br Med J 2011 May 19;342:d2786 [FREE Full text] [doi: 10.1136/bmj.d2786] [Medline: 21596740]
48. Hampel P, Gemp S, Mohr B, Schulze J, Tlach L. [Long-term effects of a cognitive-behavioral intervention on pain coping among inpatient orthopedic rehabilitation of chronic low back pain and depressive symptoms]. Psychother Psychosom Med Psychol 2014 Nov;64(11):439-447. [doi: 10.1055/s-0034-1370913] [Medline: 24838435]
49. Tlach L, Hampel P. Long-term effects of a cognitive-behavioral training program for the management of depressive symptoms among patients in orthopedic inpatient rehabilitation of chronic low back pain: a 2-year follow-up. Eur Spine J 2011 Dec;20(12):2143-2151 [FREE Full text] [doi: 10.1007/s00586-011-1810-x] [Medline: 21516327]
50. Kempf C, Keddad K, Jolivet-Landreau I, Bertin P. [Data recording using a personal digital assistant. Experience of prospective survey on pain in 3196 patients]. Presse Med 2005 Mar 12;34(5):343-347. [Medline: 15859566]
51. 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 [FREE Full text] [Medline: 12803354]
52. Sun T, West N, Ansermino JM, Montgomery CJ, Myers D, Dunsmuir D, et al. A smartphone version of the Faces Pain Scale-Revised and the Color Analog Scale for postoperative pain assessment in children. Paediatr Anaesth 2015 Dec;25(12):1264-1273. [doi: 10.1111/pan.12790] [Medline: 26507916]
53. Pawar SG, Ramani PS, Prasad A, Dhar A, Babhulkar SS, Bahurupi YA. Software version of Roland Morris Disability questionnaire for outcome assessment in low back pain. Neurol Res 2017 Apr;39(4):292-297. [doi: 10.1080/01616412.2017.1297555] [Medline: 28337948]
54. Zhang J, Sun L, Liu Y, Wang H, Sun N, Zhang P. Mobile device-based electronic data capture system used in a clinical randomized controlled trial: advantages and challenges. J Med Internet Res 2017 Mar 8;19(3):e66 [FREE Full text] [doi: 10.2196/jmir.6978] [Medline: 28274907]
55. Skonnord T, Steen F, Skjeie H, Fetveit A, Brekke M, Klovning A. Survey email scheduling and monitoring in eRCTs (SESAMe): a digital tool to improve data collection in randomized controlled clinical trials. J Med Internet Res 2016 Dec 22;18(11):e311 [FREE Full text] [doi: 10.2196/jmir.6560] [Medline: 27876689]

Abbreviations

EH: eHealth program
eHealth: electronic health

EQ-5D: EuroQol 5-Dimension health questionnaire
HC: highest compliance
IQR: interquartile range
LBP: low back pain
ODI: Oswestry Disability Index
SF-36: 36-item Short-Form Health Survey
SF-36 GH: General health for 36-item Short-Form Health Survey
SF-36 PF: Physical functioning for 36-item Short-Form Health Survey
UC: usual care treatment
VAS: visual analog scale

Edited by G Eysenbach; submitted 23.02.18; peer-reviewed by J Zhang, M Tian; comments to author 22.07.18; revised version received 01.11.18; accepted 08.12.18; published 20.02.19.

Please cite as:
Hou J, Yang R, Yang Y, Tang Y, Deng H, Chen Z, Wu Y, Shen H
The Effectiveness and Safety of Utilizing Mobile Phone–Based Programs for Rehabilitation After Lumbar Spinal Surgery: Multicenter, Prospective Randomized Controlled Trial
JMIR Mhealth Uhealth 2019;7(2):e10201
URL: http://mhealth.jmir.org/2019/2/e10201/
doi:10.2196/10201
PMID:30785406

Original Paper

Accuracy of Samsung Gear S Smartwatch for Activity Recognition: Validation Study

Anis Davoudi[1], MSc; Amal Asiri Wanigatunga[2], PhD; Matin Kheirkhahan[3], PhD; Duane Benjamin Corbett[4], PhD; Tonatiuh Mendoza[5], MSc; Manoj Battula[3], MSc; Sanjay Ranka[3], PhD; Roger Benton Fillingim[6], PhD; Todd Matthew Manini[4], PhD; Parisa Rashidi[1], PhD

[1]Department of Biomedical Engineering, University of Florida, Gainesville, FL, United States

[2]Department of Epidemiology, University of Florida, Gainesville, FL, United States

[3]Department of Computer and Information Science and Engineering, University of Florida, Gainesville, FL, United States

[4]Department of Aging and Geriatric Research, University of Florida, Gainesville, FL, United States

[5]Department of Health Outcomes and Biomedical Informatics, University of Florida, Gainesville, FL, United States

[6]Department of Community Dentistry and Behavioral Science, University of Florida, Gainesville, FL, United States

Corresponding Author:
Parisa Rashidi, PhD
Department of Biomedical Engineering
University of Florida
1064 Center Drive, NEB 459
Gainesville, FL, 32611
United States
Phone: 1 352 392 5469
Email: parisa.rashidi@ufl.edu

Abstract

Background: Wearable accelerometers have greatly improved measurement of physical activity, and the increasing popularity of smartwatches with inherent acceleration data collection suggest their potential use in the physical activity research domain; however, their use needs to be validated.

Objective: This study aimed to assess the validity of accelerometer data collected from a Samsung Gear S smartwatch (SGS) compared with an ActiGraph GT3X+ (GT3X+) activity monitor. The study aims were to (1) assess SGS validity using a mechanical shaker; (2) assess SGS validity using a treadmill running test; and (3) compare individual activity recognition, location of major body movement detection, activity intensity detection, locomotion recognition, and metabolic equivalent scores (METs) estimation between the SGS and GT3X+.

Methods: To validate and compare the SGS accelerometer data with GT3X+ data, we collected data simultaneously from both devices during highly controlled, mechanically simulated, and less-controlled natural wear conditions. First, SGS and GT3X+ data were simultaneously collected from a mechanical shaker and an individual ambulating on a treadmill. Pearson correlation was calculated for mechanical shaker and treadmill experiments. Finally, SGS and GT3X+ data were simultaneously collected during 15 common daily activities performed by 40 participants (n=12 males, mean age 55.15 [SD 17.8] years). A total of 15 frequency- and time-domain features were extracted from SGS and GT3X+ data. We used these features for training machine learning models on 6 tasks: (1) individual activity recognition, (2) activity intensity detection, (3) locomotion recognition, (4) sedentary activity detection, (5) major body movement location detection, and (6) METs estimation. The classification models included random forest, support vector machines, neural networks, and decision trees. The results were compared between devices. We evaluated the effect of different feature extraction window lengths on model accuracy as defined by the percentage of correct classifications. In addition to these classification tasks, we also used the extracted features for METs estimation.

Results: The results were compared between devices. Accelerometer data from SGS were highly correlated with the accelerometer data from GT3X+ for all 3 axes, with a correlation ≥.89 for both the shaker test and treadmill test and ≥.70 for all daily activities, except for computer work. Our results for the classification of activity intensity levels, locomotion, sedentary, major body movement location, and individual activity recognition showed overall accuracies of 0.87, 1.00, 0.98, 0.85, and 0.64, respectively. The results were not significantly different between the SGS and GT3X+. Random forest model was the best model for METs estimation (root mean squared error of .71 and r-squared value of .50).

Conclusions: Our results suggest that a commercial brand smartwatch can be used in lieu of validated research grade activity monitors for individual activity recognition, major body movement location detection, activity intensity detection, and locomotion detection tasks.

(JMIR Mhealth Uhealth 2019;7(2):e11270) doi:10.2196/11270

KEYWORDS

actigraphy; activity recognition; machine learning; metabolic equivalent; physical activity

Introduction

Wearable accelerometers have greatly improved the objective measurement of physical activity over the past 20 years [1]. They have enabled the detection and tracking of activity intensity and patterns (eg, bouts) in the population and during intervention studies [2-5]. Most research uses accelerometers that have been specifically designed to record accelerations to quantify the amount of time spent in performing activities of different intensities, which is important for understanding the health benefits of physical activity. However, there has been a rapid growth of smartwatches that collect accelerations for both usability purposes (eg, screen orientation) and for tracking activity patterns. In fact, the smartwatch market is expected to grow at an annual rate of 18% through 2021 [6,7], and this allows an unprecedented opportunity to evaluate activity patterns without using a dedicated research device.

Compared with dedicated devices, smartwatches contain some conventional sensors such as heart rate sensors, Global Positioning Systems, and ultraviolet exposure, but also more novel utilities such as a speaker, microphone, and Global System for Mobile Communications data plan for communications. These additional sensors and utilities open new opportunities coupling companion measures along with activity patterns that can be continuously uploaded through wireless networks. Their multitasking platform and increasing popularity make smartwatches an ideal tool for researchers to monitor physical activity in real time without requiring users to wear any additional, dedicated device. For smartwatches to be acceptable in research, the accelerometer needs to be validated and data need to be compared with existing research-grade monitors.

In this study, we validate the triaxial accelerometer in the Samsung Gear S smartwatch (SGS) that currently makes up 16% of the smartwatch market [8]. This is necessary to ensure data from this device is acceptable for objectively measuring time spent in performing activities of different intensities and for recognizing physical activity type. First, the SGS underwent validation on a mechanical shaker table, and raw data were compared against an Actigraph GT9X. Next, a series of experiments were performed on a treadmill and during common daily activities. High-resolution raw accelerometer data were used to extract frequency- and time-domain features that are used to train and test classification models for activity recognition tasks. We compared the accuracies of the SGS and Actigraph GT3X+ (GT3X+) for assessing (1) activity intensity level, (2) locomotion versus nonlocomotion, (3) location of major body movement during the activity, (4) sedentary versus nonsedentary, and (5) individual activity recognition classification. Moreover, a portable metabolic unit was used to record metabolic equivalent scores (METs) to estimate activity intensity and used to further validate the SGS against the Actigraph. We hypothesized that the SGS would test valid and be accurate at assessing activity intensity and recognize activity type as compared with GT3X+.

Methods

Data Collection

We collected data in 3 different experiment setups: (1) a shaker table, (2) treadmill walking, and (3) 15 daily activities. In all experiments, we collected data simultaneously from both SGS [9] and GT3X+. Device characteristics are compared and described in Table 1 [10,11]. All study procedures were approved by the University of Florida Institutional Review Board. All participants provided written informed consent before participation in the study. First, we collected data using a mechanical 1-dimensional shaker table by applying acceleration to both devices for 3 min for 7 different speeds (0.5 Hz, 1 Hz, 1.5 Hz, 2 Hz, 2.5 Hz, 3 Hz, and 3.5 Hz), repeated for each axis. SGS monitor was positioned on top of the GT3X to ascertain that they experience the same accelerations. Then, we collected data during 6 speeds of treadmill walking, where 1 participant wore both devices on the right wrist and ambulated at 6 different speeds (1, 2, 3, 4, 5, and 6 mph), 3 min each at each speed. Data were collected at a frequency of 10 Hz for the SGS and 100 Hz for GT3X+ for both the shaker table and treadmill test.

In the third experiment, participants wore both devices on their right wrist while performing several daily activities, as listed in Table 2 [12]. Expiratory gas was collected during each activity using a portable, chest-worn, indirect calorimeter (Cosmed K4b2; COSMED USA). Energy expenditure was estimated using oxygen uptake (VO_2=milliliter $min^{-1}kg^{-1}$) at a steady state, generally beginning at 3 min after starting the activity. Oxygen consumption was subsequently converted to METs, a value that is often used to gauge the intensity of an activity relative to a reference resting value (ie, 3.0 METs=3 times the equivalent of resting oxygen consumption). METs were calculated as the oxygen consumption per minute relative to body mass (ml/min/kg) divided by a resting value of 3.5 ml/min/kg [12,13]. We linked the resulting MET value for each task to the average of the extracted features for the task.

Table 1. Technical specifications of ActiGraph GT3X+ and Samsung Gear S smartwatch.

Characteristics	ActiGraph GT3X+	Samsung Gear S smartwatch
Dimensions	4.6 cm x 3.3 cm x 1.5 cm	5.8 cm x 4.0 cm x 1.2 cm
Weight	19 gm	67 gm
Sampling rate	100 Hz	100 Hz
Dynamic range	±8G	±2G
Memory	4 GB	4 GB

Table 2. Characteristics of each activity. Accelerometer data were collected from 40 participants.

Activity	Major body movement location	Intensity	Locomotion	Duration (min)	Participants (n)
Computer work	Upper	Sedentary	No	8	11
Ironing	Upper	Light	No	8	12
Yoga	Total	Light	No	8	8
Shopping	Total	Light	No	8	13
Laundry	Upper	Light	No	8	9
Washing windows	Upper	Moderate	No	8	15
Home maintenance	Upper	Moderate	No	8	14
Replacing bed sheet	Upper	Moderate	No	8	13
Mopping	Upper	Moderate	No	8	13
Trash removal	Total	Moderate	No	8	13
Heavy lifting	Total	Moderate	No	8	12
Leisure walk	Lower	Moderate	Yes	8	9
Rapid walk	Lower	Moderate	Yes	8	9
Walk at RPE[a] 1	Lower	Moderate	Yes	5	14
Walk at RPE 5	Lower	Moderate	Yes	5	13

[a]RPE: ratings of perceived exertion.

Textbox 1. Inclusion criteria for the study.

1. Community-dwelling adults aged 20 years or older
2. Willingness to undergo all testing procedures
3. Weight stable for at least 3 months (±5 lbs)
4. English speaking

A total of 40 community-dwelling adults aged between 20 and 85 years (n=12 males, mean age 55.2 [SD 17.8] years, mean body mass index 26.8 [SD 6.2] kg/m^2), participated in this study. These 40 participants belonged to a subset of a larger study whose primary aim was to assess the age effect on energy expenditure during activities common to daily life in the United States [12]. The inclusion criteria are provided in Textbox 1. For a complete list of exclusion criteria refer to [12].

The activities in our study are common daily activities. They included ironing, yoga, shopping, laundry washing, computer work, washing windows, home maintenance, replacing bedsheet, mopping, trash removal, heavy lifting, and walking at 2 different ratings of perceived exertion (RPE), as well as leisure and rapid walking [12]. Activity instructions and the setup of the experiments are described in more detail in the study by Corbett et al [12]. Participants performed activities based on instructions, in a laboratory setting, each for 8 min, except for the walking activities performed at different RPE [14,15]. Activities were designed to be repeated to achieve an accumulated duration of 8 min.

Analysis

To validate data from the SGS against data from the GT3X+, we calculated the root mean square (RMS) value of the accelerometer data over 1 second, for all 3 axes, for each device. RMS is calculated using Equation 1 (Figure 1).

Here, N is the number of data points and x_{ij} refers to a single data point i from axis j. In this step, we calculated the RMS of

1-second windows of the data in each axis. Then, we computed the correlation between the data from the 2 devices along each axis. This process was performed for the shaker table test data, treadmill test data, and daily activity data. We used the middle 2 min of 3-min shaker table data, the middle 2 min of 3-min treadmill data, and the middle 5 min of the 8-min daily activity tests. The purpose of this selection was to exclude the boundary start and end segments of activities.

We extracted 15 time- and frequency-domain features found in the current literature (Table 3) [16,17]. Here, vector magnitude is defined as in Equation 2 (Figure 2), where x, y, and z are accelerations in the Cartesian coordinate system. In addition to the features previously suggested [16], our feature set also included (1) kurtosis, which is a descriptor of the shape of the distribution of the acceleration data in each window; (2) skewness, which measures the asymmetry of the distribution of acceleration data in the window; and (3) entropy, which helps in discriminating between activities with similar power spectral density but different movement patterns [17].

The length of the window of data used for statistical feature calculation is also important [18,19]. Previous individual activity recognition studies have used different window lengths, ranging from 0.1 seconds to 128 seconds [16,20-24]. In this study, we do not use window lengths smaller than 1 second because we are using frequency-domain features. We calculated the statistical features for 6 commonly used window lengths (1-, 2-, 4-, 8-, 15-, and 16-second lengths), to choose the best window length for our classification tasks. These window lengths were chosen in accordance with previous literature from the above list, limited to between windows smaller than 16 seconds. We did not go above the 16-second length as it would drastically reduce the number of sample points and will potentially include different actions in a single window of an activity. We used the overall accuracy metric to select the best model. The features calculated for the chosen window length are used in the prediction models.

Classification tasks included (1) detection of the location of major body movement, (2) detection of activity intensity level, (3) detection of locomotion, (4) detection of sedentary activity, and (5) individual activity recognition for the 15 daily activities.

Figure 1. Root mean square (RMS) shows the arithmetic mean of the squares of the accelerometer values.

$$RMS_j = \sqrt{\frac{\sum_{i=0}^{N} x_{ij}^2}{N}}$$

Table 3. Description of the features extracted from the raw data.

Feature	Description
Mean of vector magnitude (MVM)	Sample MVM in the window
Standard deviation of vector magnitude (SDVM)	SDVM in the window
Percentage of the power of the vector magnitude (VM) that is in the range of 0.6-2.5 Hz	Sum of moduli corresponding to frequency in this range/sum of moduli of all frequencies
Dominant frequency (DF) of VM	Frequency corresponding to the largest modulus
Fraction of power in VM at DF	Modulus of the dominant frequency divided by sum of moduli at each frequency
Mean angle of acceleration relative to vertical on the device	Sample mean of the angle between x-axis and VM in the window
SD of the angle of acceleration relative to vertical on the device	Sample SD of the angles in the window
Covariance	Covariance of the VM in the window
Skewness	Skewness of the VM in the window
Kurtosis	Kurtosis of the VM in the window
Entropy	Entropy of the VM in the window
Coefficient of variation	SD of VM in the window divided by the mean, multiplied by 100
Corr(x,y)	Correlation between x-axis and y-axis
Corr(y,z)	Correlation between y-axis and z-axis
Corr(x,z)	Correlation between x-axis and z-axis

Figure 2. Vector magnitude (VM) here is defined as Euclidean norm of the vector from the origin to the point shown by x, y, and z.

$$VM = \sqrt{x^2 + y^2 + z^2}$$

Figure 3. Gini impurity measures the likelihood of incorrect classification of a randomly selected instance of a set, if it were randomly classified according to the distribution of class labels from the set.

$$G = \sum_{i=1}^{n_c} p_i(1 - p_i)$$

The activities in the study include both simple and complex activities [25]. Simple activities, which include walking and computer work in our dataset, consist of repeating a single action to complete the activity. Complex activities, which include ironing, yoga, shopping, laundry washing, washing windows, home maintenance, replacing bedsheet, mopping, trash removal, and heavy lifting, are composed of several simple actions. For example, trash removal activity consists of sorting the trash, finding smaller trash cans, picking up the trash bags, and taking all the trash to a predetermined location. The heterogeneity of complex activities can significantly degrade the performance of an individual activity recognition classifier, as opposed to recognition of simple individual activities. In this study, we report the overall accuracy for simple activities, complex activities, and all activities as well as the balanced accuracy of each activity. Balanced accuracy is calculated as the arithmetic mean of sensitivity and specificity.

For the classification tasks, we used decision tree, random forest, support vector machines, and neural networks models. These models can flexibly represent various relationships between the features and the outcome and have been previously used for activity type classification tasks [16].

We report the ranking of the features' importance in the random forest model. The random forest model calculated the importance of the features using the decrease in the Gini index. The Gini impurity index is calculated using Equation 3 (Figure 3). Every time a split is made on a variable, the Gini impurity index for the 2 descendent nodes is less than the Gini impurity index of the parent node. The decrease in the Gini index for each variable is calculated by adding up the Gini decreases for the variables over all trees in the forest.

Here, n_c is the number of classes in the outcome and p_i is the ratio of the class i. For the MET estimation regression task, the feature importance values are calculated based on the total decrease in node impurities (difference between the residual sum of squares before and after splitting on the variable), averaged over all trees in the random forest model.

For both classification and regression tasks, we used nested cross-validation for evaluating our models. In each fold, we divided the data into 20% test data and 80% development data. The development data was further divided into 20% validation data and 80% training data. Validation data are used for tuning the parameters of the models and removing collinear features. Data partitioning was based on each feature extracted from the window length, rather than by partitioning based on participants, as individual participants did not have all the tasks, and separating based on participants might lead to the absence of some of the classes in some classification tasks, for example, individual activity recognition classification task. All statistical analyses were performed in R (version 3.1.3) [26].

Results

We performed 3 different experiments to evaluate the validity of SGS in comparison with the GT3X+: (1) shaker table, (2) treadmill, and (3) daily activities. Figure 4 shows the 1-second RMS of the raw acceleration data collected from both devices along the 3 x-, y-, and z-axes for the shaker table test and treadmill test. Correlation values for the 1-second RMS for all 3 axes were high for both the shaker table test and treadmill test (x-corr=.97, y-corr=.97, and z-corr=.95 for the shaker table and x-corr=.98, y-corr=.89, and z-corr=.93 for the treadmill test). Acceleration data from the 3 axes are highly correlated, except for small shifts in amplitude (Figure 5). Correlation values between SGS data and GT3X+ for 3 axes are high in daily activities as well (x-corr>.70, y-corr>.70, and z-corr>.71, except for computer work; Table 4). Figure 6 shows the similarity in acceleration measurement for walking activity as an example of simple activities, and Figure 7 shows the similarity in acceleration measurement for mopping as an example of complex activities, along the 3 x-, y-, and z-axes. Figure 6 shows the repeated single actions for a simple activity, and Figure 7 shows the varying actions occurring during a complex activity.

To test the effect of window length, we repeated our classification tasks with features extracted using varying window length (1, 2, 4, 8, 15, and 16 seconds). We initially observed that random forest has the highest overall accuracy. We evaluated the effect of window length used for extracting the features on the performance of the random forest model. Models trained on the features extracted based on larger window length had better performance (Table 5) [18]. We used features extracted from 16-seconds windows for the remainder of the paper.

Figure 8 shows the performance of our different models in terms of the overall accuracy, simple individual activity recognition accuracy, and complex individual activity recognition accuracy. Random forest has the best performance across all 3 activity recognition tasks. Although the random forest model is technically a collection of decision trees, it is designed to correct the overfitting of decision trees. The random forest model generally works better than support vector machines in multi-class classification problems, but the difference is much smaller for binary classification tasks. Figure 9 shows how SGS and GT3X+ perform in our classification tasks. Random forest model's performance in terms of balanced accuracy of each activity shows that the devices can detect the simple activities—computer work and walking activities—better (Figure 10). Multimedia Appendix 1 gives the normalized confusion matrix of the detected labels versus actual labels of the activities as classified by the random forest model.

Figure 4. The 1-second root mean square of the acceleration data from Samsung Gear S smartwatch and ActiGraph GT3X+ for each axis for the shaker table.

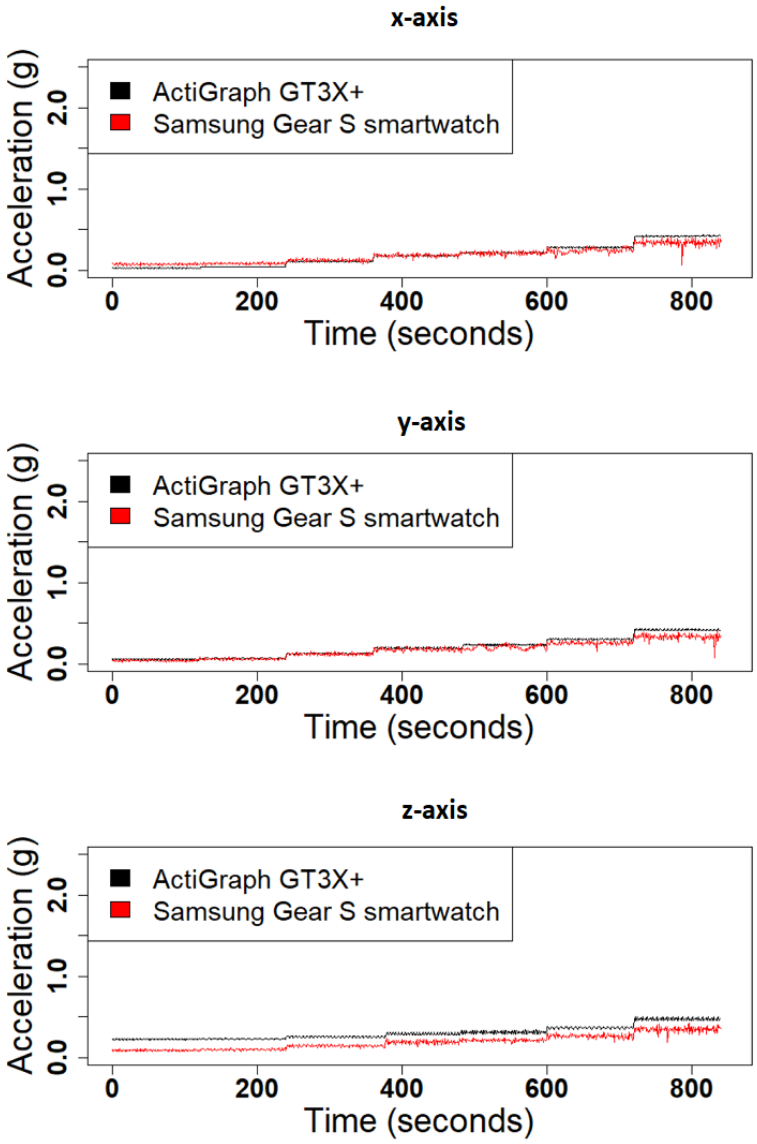


Figure 5. The 1-second root mean square of the acceleration data from Samsung Gear S smartwatch and ActiGraph GT3X+ for each axis for the treadmill test.

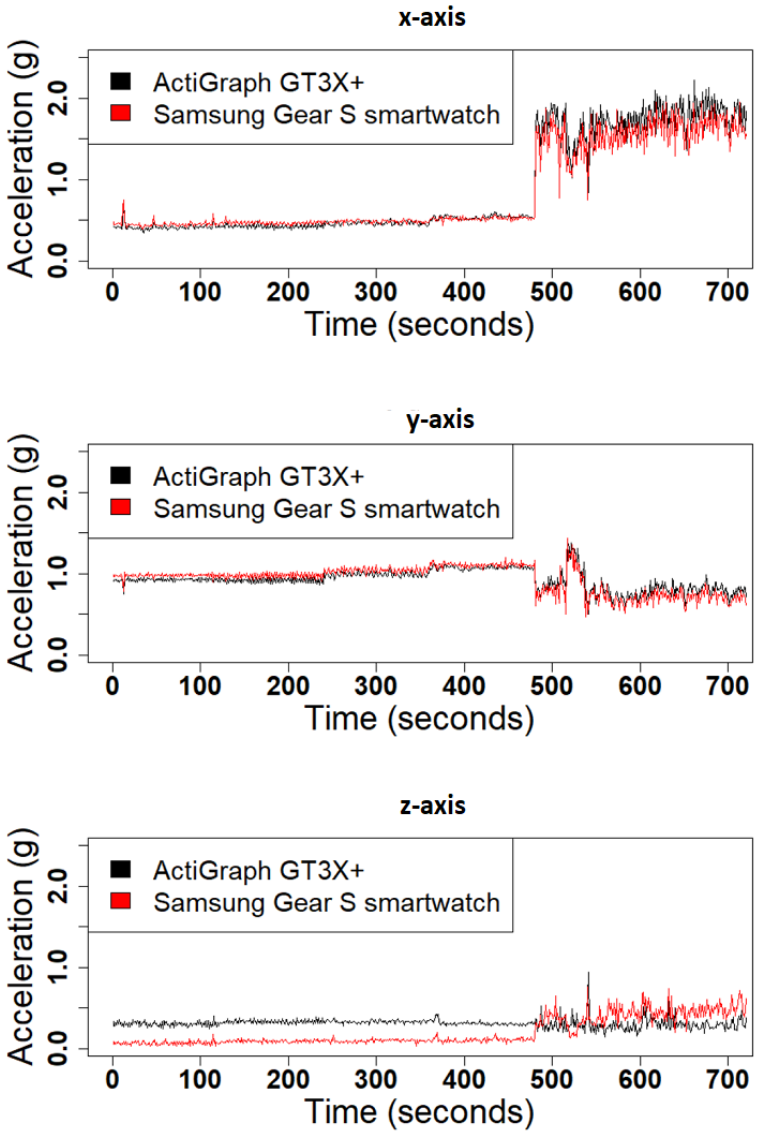


Table 4. Correlations of 1-second root mean square for all 3 axes of the Samsung Gear S smartwatch and ActiGraph GT3X+ by activity. All correlation values were statistically significant.

Task: correlation	x-axis	y-axis	z-axis
Simple activities			
Computer work	.83	.55	.63
Leisure walk	.88	.86	.85
Rapid walk	.78	.70	.76
Walk at RPE[a] 1	.79	.92	.87
Walk at RPE 5	.70	.75	.71
Complex activities			
Shopping	.89	.84	.86
Mopping	.94	.89	.75
Home maintenance	.91	.84	.81
Washing windows	.90	.80	.82
Heavy lifting	.94	.90	.71
Ironing	.83	.81	.78
Replacing bedsheet	.87	.88	.78
Yoga	.93	.88	.90
Trash removal	.92	.92	.84
Laundry washing	.89	.85	.79

[a]RPE: ratings of perceived exertion.

Figure 6. Acceleration data from leisure walking (simple activity) for all 3 axes.

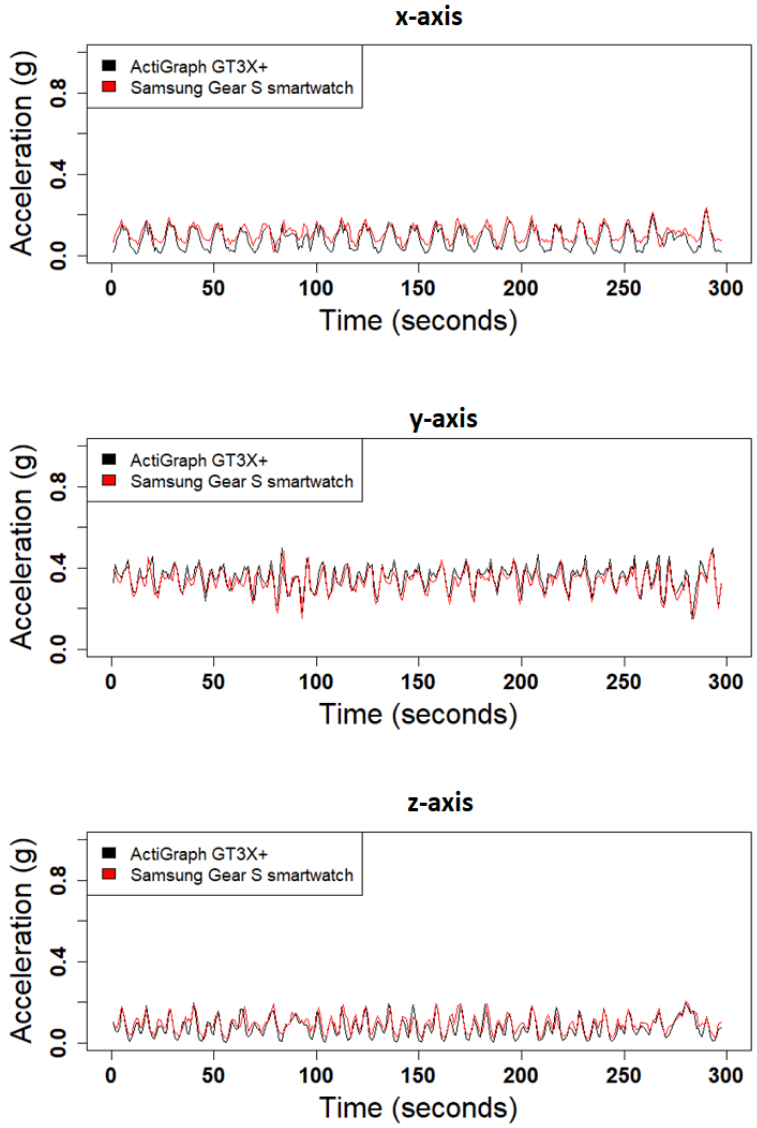


Figure 7. Acceleration data from mopping (complex activity) for all 3 axes.

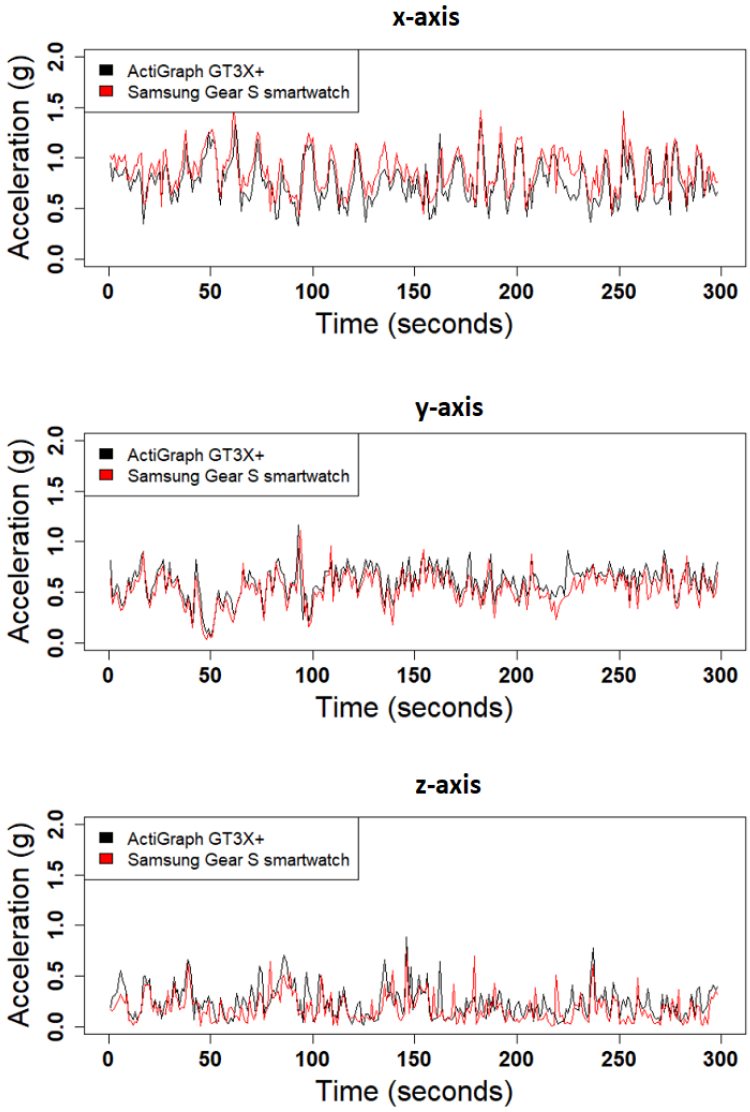


Table 5. Effect of window length for feature extraction on micro-averaged accuracy using random forest model for individual activity recognition, locomotion detection, sedentary activity detection, activity intensity level classification, and major body movement location detection. The best performance for each classification task is presented in italics.

Classification task: window length	Classification accuracy (seconds)					
	1	2	4	8	15	16
Individual activity recognition	.45	.47	.56	.60	.64	.64
Locomotion	1	1	1	1	1	1
Sedentary	.96	.97	.97	.98	.98	.98
Intensity	.80	.81	.82	.84	.87	.87
Major body movement location	.78	.80	.81	.83	.85	.85

Figure 8. Comparison of the performance of different classifiers for activity recognition task, in terms of overall accuracy for simple tasks, complex tasks, and total set of tasks. TREE: decision tree model, NNET: neural network model, RF: random forest model, and SVM: support vector machines model.

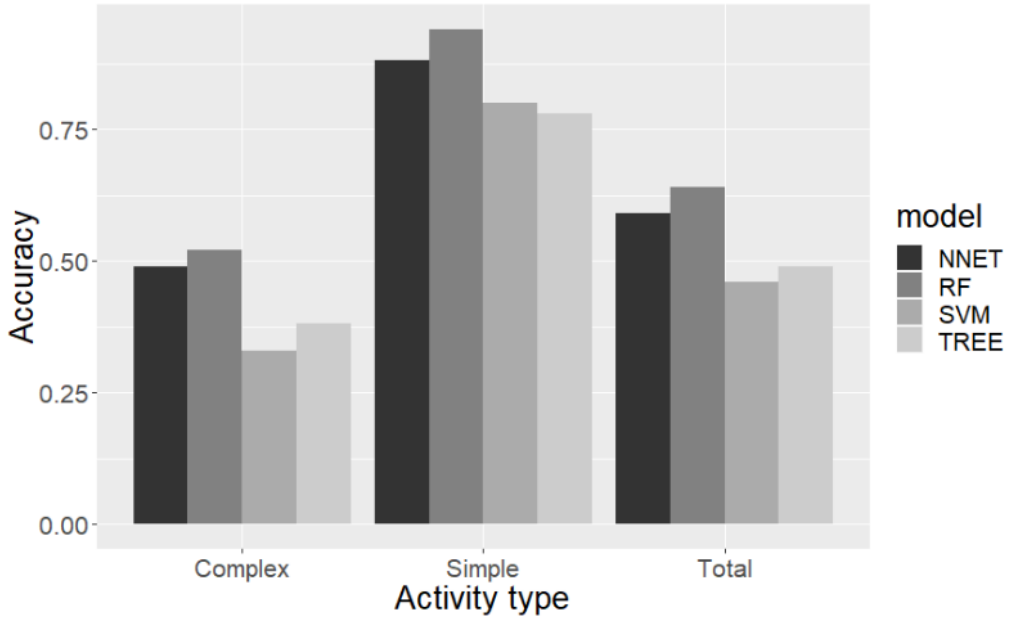
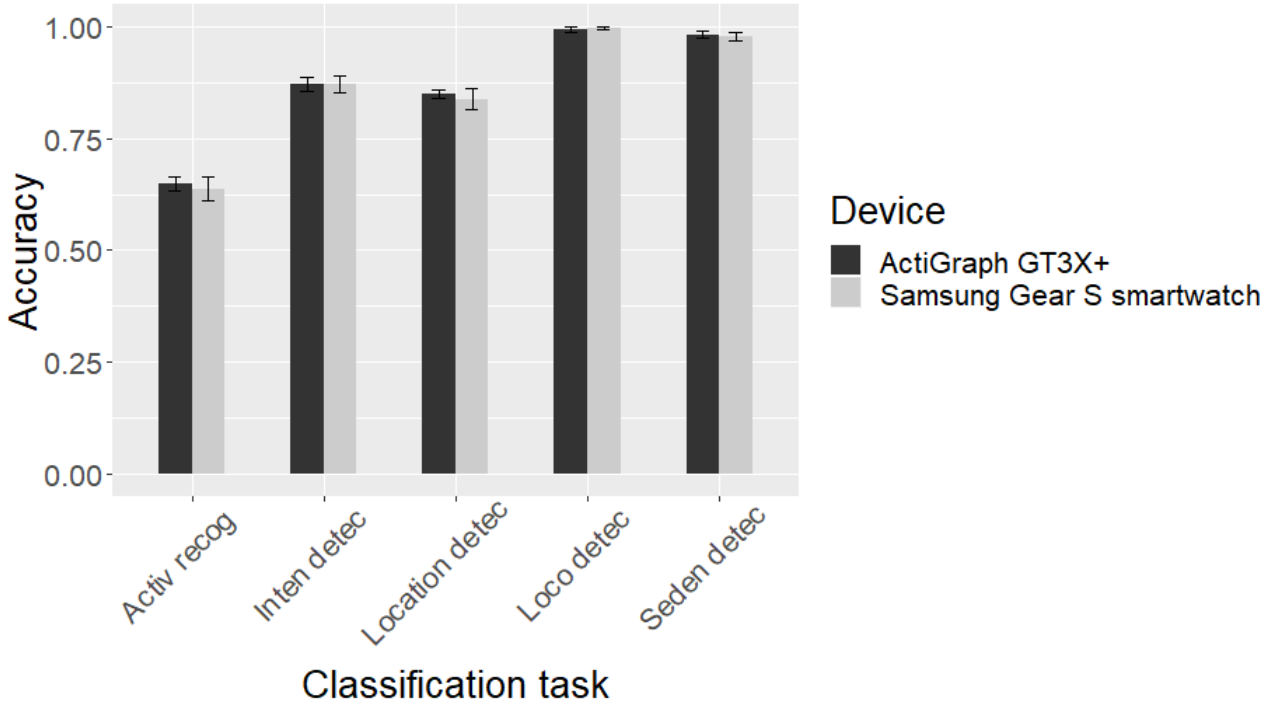


Figure 9. Accuracies of the 5-fold cross-validation classification tasks performed on Samsung Gear S smartwatch data and ActiGraph GT3X+ data. We used the best model and best window length (random forest model and 16-second window length feature extraction). Activ recog: Activity recognition; Inten detec: Activity intensity level detection; Location detec: Major body movement location detection; Loco detec: Locomotion detection; Seden detec: Sedentary activity detection.



In the next step, we used our classifiers (decision trees, random forest, support vector machines, and neural networks) on the extracted features to classify the activities based on their intensity, sedentary status, locomotion status, and location of major body movement. The overall accuracies of the models are compared for each classification task in Figure 11. The confusion matrices of the model with the highest accuracy-random forest model for each task are given in Tables 6 to 9. Random forest has the best performance in all tasks based on overall accuracy. Figure 12 shows the ranking of features used in the random forest model for all classification and regression tasks.

Figure 10. Balanced accuracy for each activity using the random forest model and 16-seconds window for extraction of features. RPE: ratings of perceived exertion.

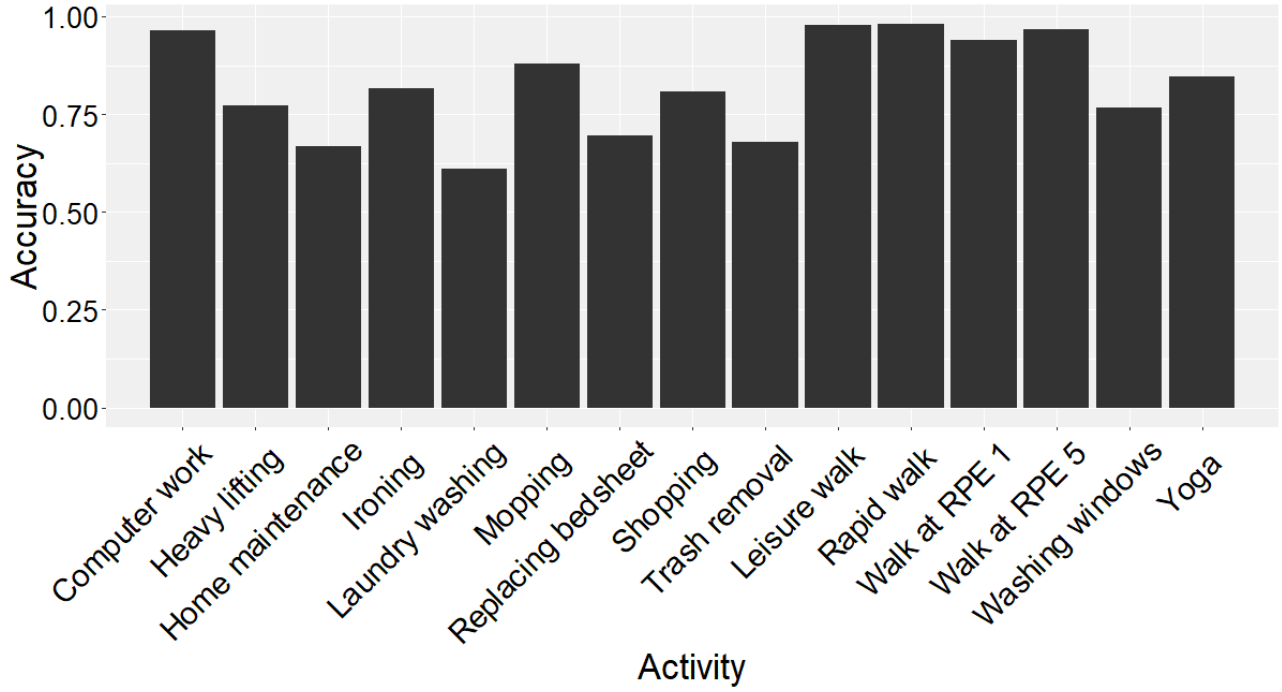


Figure 11. Performance of 4 classifier models in activity type classification tasks in terms of accuracy. NNET: neural networks model; RF: random forest model; SVM: support vector machines model, TREE: decision trees model.

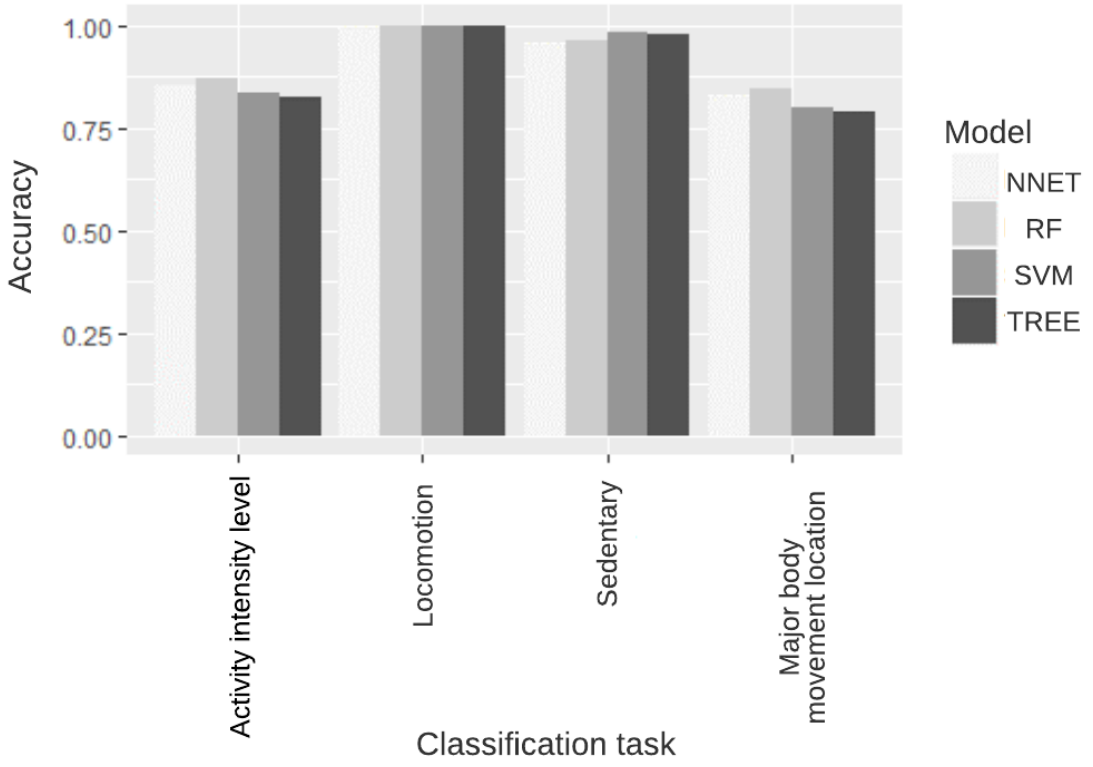


Table 6. Normalized confusion matrix showing the percentage of correctly classified instances for the random forest model used for activity type classification tasks based on intensity level. Each column shows the actual activity and rows represent the predicted labels.

Activity	Sedentary	Light	Moderate
Sedentary	0.87	0.02	0
Light	0.12	0.65	0.04
Moderate	0.01	0.33	0.95

Table 7. Normalized confusion matrix showing the percentage of correctly classified instances for the random forest model used for activity type classification tasks based on locomotion. Each column shows the actual activity and rows represent the predicted labels.

Activity	Locomotion	Nonlocomotion
Locomotion	0.98	0
Nonlocomotion	0.02	1

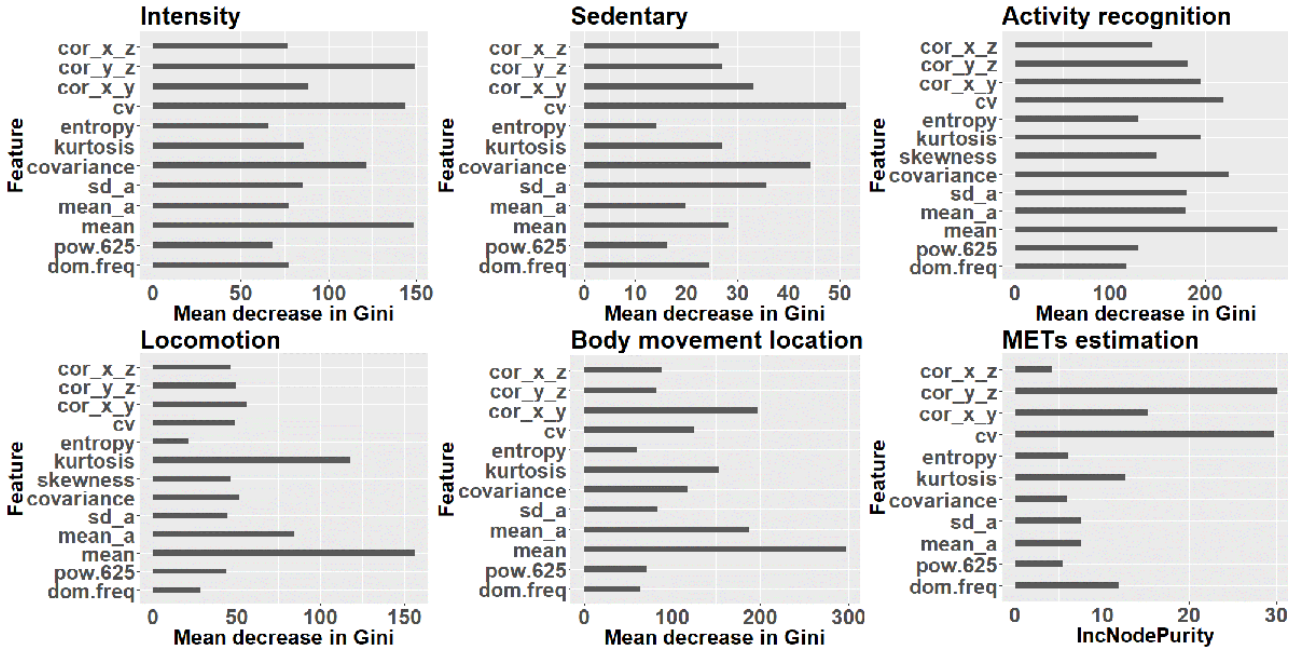
Table 8. Normalized confusion matrix showing percentage of correctly classified instances for the random forest model used for activity type classification tasks based on major body movement location. Each column shows the actual activity and rows represent the predicted labels.

Activity	Lower	Total	Upper
Lower	1	0	0
Total	0	0.61	0.11
Upper	0	0.39	0.89

Table 9. Normalized confusion matrix showing percentage of correctly classified instances for the random forest model used for activity type classification tasks based on sedentary. Each column shows the actual activity and rows represent the predicted labels.

Activity	Sedentary	Nonsedentary
Sedentary	0.78	0.01
Nonsedentary	0.22	0.99

Figure 12. Importance of features in terms of the mean decrease in the Gini index in each classification task and in terms of the increase in node purity for the metabolic equivalent scores estimation regression task, as reported by the random forest model and 16-second window length for feature extraction (note the different range for the x-axis). METs: metabolic equivalent scores; IncNodePurity: Increase in node purity; cor_x_z: correlation between x-axis and z-axis; cor_y_z: correlation between y-axis and z-axis; cor_x_y: correlation between x-axis and y-axis; cv: covariance of the vector magnitude; sd_a: SD of angle; mean_a: mean of angle; pow.625: percentage of the power of the vector magnitude that is in the range of 0.6-2.5 Hz; dom.freq: dominant frequency of vector magnitude.



Using random forest, support vector machines, neural network, and decision trees models for the regression task of METs estimation shows that the random forest model also has the best performance for the METs estimation. Random forest, decision tree, neural networks, support vector machines models had RMS error values of .71, .77, .99, and .77, respectively. Their r-squared values—which is the coefficient of determination defined as the proportion of the variance in the output that is predictable from the input features—were .50, .40, .01, and .41, respectively.

Discussion

Principal Findings

The goal of this study was to validate the SGS accelerometer data against the GT3X+ accelerometer data using several comparisons. First, the accelerometer data from SGS and GT3X+ were compared using a mechanical shaker at different speeds. Second, data were collected and compared from the 2 devices worn by a participant ambulating on a treadmill at different speeds. Third, data collected during activities common in daily life were collected and compared in all 3 axes. We also compared the performance of activity recognition and classification models using accelerometer data recorded using SGS against using accelerometer data recorded using GT3X+.

Accelerometer data from SGS and GT3X+ have high correlations along all 3 axes during shaker tests. However, the correlations were slightly lower when performing daily activities. There are several scenarios where these 2 devices might show different acceleration values. The lower correlation values during the daily activities might partially be caused by nonalignment and movement of the 2 devices during the activities. When the 2 devices are not perfectly aligned, their axes are not pointing in the same direction. This may cause the movement to be dispersed across multiple axes with varying ratios for each device. Another factor that might result in a lower correlation between the 2 devices is the relative slowness of movements during activities such as computer work or yoga, where even small differences are comparatively more significant.

In our study, all 4 models performed well for activity type classification tasks. Random forest had the best performance for all classification tasks. This difference was greater for the individual activity recognition task. Neural networks model's performance was similar to the random forest model, which were superior to support vector machines, which is more suitable for binary classification, and decision tree model, which tends to overfit [27]. Mean vector magnitude was the most important feature in our tasks, whereas the ranking of other features understandably varied among the different classification tasks. We also used the random forest, decision trees, support vector machines, and neural network models for METs estimation. Random forest had the best performance in this task as well, but other models, except for the neural network, had similar performances.

In this work, we experimented with 1 second, 2 seconds, 4 seconds, 8 seconds, 15 seconds, and 16 seconds nonoverlapping time windows. Larger window lengths resulted in better classification accuracy (Table 5). This may be because of smaller window length capturing shorter duration of actions that are common among different activities. A window length of 16 seconds resulted in high accuracies for activity type classifications such as locomotion versus nonlocomotion. However, it did not lead to high accuracy for individual activity recognition. The lower overall accuracy of classifiers for individual activity recognition can be attributed to the heterogeneous nature of complex activities that are composed of multiple simpler actions under the same label. The fact that the same label here corresponds to different actions deteriorates the performance of the classifier. To remedy this problem for future models, one can separately label every single action in a complex activity.

We observed a slight shift in acceleration amplitude during the mechanical shaker tests between the 2 devices. For our classification purposes, this shift would not affect the models because the relative change in acceleration appeared to be preserved in the SGS. Therefore, it is not expected to impact the classification results in practice. However, a comparison between devices is cautioned until corrections can be made to equate absolute thresholds of accelerations. Overall, our results show that the performances of the SGS and GT3X in various activity recognition tasks and METs estimation are similar and can potentially be used interchangeably between studies. However, the raw data are not interchangeable because of the slight shift explained, and thus, any threshold derived for 1 device needs to be validated for the other device to be compared.

Limitations

There are limitations that need to be acknowledged. First, inter-device reliability was not tested, given the use of a commercial device. Another limitation is that SGS was evaluated in standard settings (eg, shaker table and laboratory activities) and thus may not be applicable to free-living conditions. In this study, we compared the performances of the machine learning models trained on data collected by the 2 devices in laboratory settings and structured activities. However, free-living activity recognition tasks are more complicated because of factors such as a temporal overlap between the activities, similar action units in several activities, activity fragmentation, and the interpersonal and intrapersonal variation in activities as well as variation in wear locations. Future works can focus on comparing the performance of SGS-based models with GT3X-based models in free-living activity recognition tasks. Such efforts would need to implement methods for activity label determination, such as body-worn camera recordings that allow for later labeling of the activities performed. In addition, the tasks that were evaluated are not representative of all the tasks that an individual may perform. Although the activities constitute a wide range of movements, the results reported in this study are limited to the activities tested.

Conclusions

In this study, we showed that data collected from a commercial brand smartwatch performed similarly to a research-grade accelerometer to detect a variety of simple and complex activity types. The comparable performance of models relying on SGS

and GT3X+ data for activity recognition and energy expenditure estimation verifies the validity of the SGS for research purposes.

Acknowledgments

This work is supported by grants from the National Institutes of Health (NIH AG042525), the Data Science and Applied Technology Core of the Claude D Pepper Older Americans Independence Center at the University of Florida (P30 AG028740-11), and K07AG046371. AAW was supported by the National Institutes of Health/National Institute on Aging (NIH/NIA T32AG000247 and NIH/NIA P30AG021334). Support for DBC was provided from the NIH/NIA 1T32AG049673-01.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Normalized confusion matrix for individual activity recognition task, using the random forest model and 16-second window for extraction of features. Each column shows the actual activity and rows represent the predicted labels. Italicized diagonal elements show the percentage of accurately classified points for each activity.

[DOCX File, 17KB - mhealth_v7i2e11270_app1.docx]

References

1. Troiano RP, Berrigan D, Dodd KW, Mâsse LC, Tilert T, McDowell M. Physical activity in the United States measured by accelerometer. Med Sci Sports Exerc 2008 Jan;40(1):181-188. [doi: 10.1249/mss.0b013e31815a51b3] [Medline: 18091006]
2. Troiano RP. A timely meeting: objective measurement of physical activity. Med Sci Sports Exerc 2005 Nov;37(11 Suppl):S487-S489. [doi: 10.1249/01.mss.0000185473.32846.c3] [Medline: 16294111]
3. Troiano RP, McClain JJ, Brychta RJ, Chen KY. Evolution of accelerometer methods for physical activity research. Br J Sports Med 2014 Jul;48(13):1019-1023 [FREE Full text] [doi: 10.1136/bjsports-2014-093546] [Medline: 24782483]
4. Napolitano MA, Borradaile KE, Lewis BA, Whiteley JA, Longval JL, Parisi AF, et al. Accelerometer use in a physical activity intervention trial. Contemp Clin Trials 2010 Nov;31(6):514-523 [FREE Full text] [doi: 10.1016/j.cct.2010.08.004] [Medline: 20723619]
5. 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]
6. Beaver L. Business Insider. 2016 Sep 27. The smartwatch report: forecasts, adoption trends, and why the market isn't living up to the hype URL: http://www.businessinsider.com/smartwatch-and-wearables-research-forecasts-trends-market-use-cases-2016-9 [accessed 2018-06-11] [WebCite Cache ID 705zlC4O2]
7. Smith A. Pew Research Center. 2017 Jan 12. Record shares of Americans now own smartphones, have home broadband URL: http://www.pewresearch.org/fact-tank/2017/01/12/evolution-of-technology/ [accessed 2018-06-11] [WebCite Cache ID 705zaWXNJ]
8. Business Wire. 2016 Jul 21. Worldwide Smartwatch Market Experiences Its First Decline as Shipments Fall 32% in the Second Quarter of 2016, According to IDC URL: https://www.businesswire.com/news/home/20160721005647/en/Worldwide-Smartwatch-Market-Experiences-Decline-Shipments-Fall [accessed 2019-01-10] [WebCite Cache ID 705ydyTHu]
9. Nair S, Kheirkhahan M, Davoudi A, Rashidi P, Wanigatunga AA, Corbett DB, et al. ROAMM: a software infrastructure for real-time monitoring of personal health. 2016 Presented at: 2016 IEEE 18th International Conference on e-Health Networking, Applications and Services; September 14-16, 2016; Munich, Germany p. 14-16. [doi: 10.1109/HealthCom.2016.7749479]
10. Mortazavi B, Nemati E, VanderWall K, Flores-Rodriguez HG, Cai JY, Lucier J, et al. Can smartwatches replace smartphones for posture tracking? Sensors 2015 Oct 22;15(10):26783-26800 [FREE Full text] [doi: 10.3390/s151026783] [Medline: 26506354]
11. Samsung. Samsung Gear S URL: https://www.samsung.com/hk_en/wearables/samsung-gear-s/ [accessed 2018-08-10] [WebCite Cache ID 71ZegtkOV]
12. Corbett DB, Wanigatunga AA, Valiani V, Handberg EM, Buford TW, Brumback B, et al. Metabolic costs of daily activity in older adults (Chores XL) study: design and methods. Contemp Clin Trials Commun 2017 Jun;6:1-8 [FREE Full text] [doi: 10.1016/j.conctc.2017.02.003] [Medline: 28649668]
13. McLaughlin JE, King GA, Howley ET, Bassett DR, Ainsworth BE. Validation of the COSMED K4 b2 portable metabolic system. Int J Sports Med 2001 May;22(4):280-284. [doi: 10.1055/s-2001-13816] [Medline: 11414671]

14. Borg GA. Psychophysical bases of perceived exertion. Med Sci Sports Exerc 1982;14(5):377-381. [doi: 10.1249/00005768-198205000-00012] [Medline: 7154893]
15. Dawes HN, Barker KL, Cockburn J, Roach N, Scott O, Wade D. Borg's rating of perceived exertion scales: do the verbal anchors mean the same for different clinical groups? Arch Phys Med Rehabil 2005 May;86(5):912-916. [doi: 10.1016/j.apmr.2004.10.043] [Medline: 15895336]
16. Staudenmayer J, He S, Hickey A, Sasaki J, Freedson P. Methods to estimate aspects of physical activity and sedentary behavior from high-frequency wrist accelerometer measurements. J Appl Physiol (1985) 2015 Aug 15;119(4):396-403 [FREE Full text] [doi: 10.1152/japplphysiol.00026.2015] [Medline: 26112238]
17. Attal F, Mohammed S, Dedabrishvili M, Chamroukhi F, Oukhellou L, Amirat Y. Physical human activity recognition using wearable sensors. Sensors (Basel) 2015 Dec 11;15(12):31314-31338 [FREE Full text] [doi: 10.3390/s151229858] [Medline: 26690450]
18. Bashir SA, Doolan DC, Petrovski A. The Impact of Feature Vector Length on Activity Recognition Accuracy on Mobile Phone. In: Proceedings of the World Congress on Engineering 2015 Volume I. 2015 Presented at: World Congress on Engineering 2015; July 1-3, 2015; London, UK.
19. Banos O, Galvez JM, Damas M, Pomares H, Rojas I. Window size impact in human activity recognition. Sensors 2014 Apr 9;14(4):6474-6499 [FREE Full text] [doi: 10.3390/s140406474] [Medline: 24721766]
20. Huynh T, Schiele B. Analyzing features for activity recognition. In: Proceedings of the 2005 Joint Conference on Smart Objects and Ambient Intelligence: Innovative Context-Aware Services: Usages and Technologies. 2005 Presented at: sOc-EUSAI'05; October 12-14, 2005; Grenoble, France p. 159-163. [doi: 10.1145/1107548.1107591]
21. Stikic M, Huynh T, Van Laerhoven K, Schiele B. ADL recognition based on the combination of RFID and accelerometer sensing. 2008 Presented at: 2008 Second International Conference on Pervasive Computing Technologies for Healthcare; January 30-February 1, 2008; Tampere, Finland p. 258-263. [doi: 10.1109/PCTHEALTH.2008.4571084]
22. Pirttikangas S, Fujinami K, Nakajima T. Feature Selection Activity Recognition from Wearable Sensors. In: Proceedings of the Third International Conference on Ubiquitous Computing Systems. 2006 Presented at: UCS'06; October 11-13, 2006; Seoul, Korea p. 516-527. [doi: 10.1007/11890348_39]
23. Mannini A, Intille SS, Rosenberger M, Sabatini AM, Haskell W. Activity recognition using a single accelerometer placed at the wrist or ankle. Med Sci Sports Exerc 2013 Nov;45(11):2193-2203 [FREE Full text] [doi: 10.1249/MSS.0b013e31829736d6] [Medline: 23604069]
24. Krause A, Siewiorek DP, Smailagic A, Farringdon J. Unsupervised, dynamic identification of physiological and activity context in wearable computing. In: Proceedings of the 7th IEEE International Symposium on Wearable Computers. 2003 Presented at: ISWC'03; October 21-23, 2003; White Plains, NY, USA p. 88. [doi: 10.1109/ISWC.2003.1241398]
25. Dernbach S, Das B, Krishnan NC, Thomas BL, Cook DJ. Simple and Complex Activity Recognition through Smart Phones. In: Proceedings of the 2012 Eighth International Conference on Intelligent Environments. 2012 Presented at: IE'12; June 26-29, 2012; Guanajuato, Mexico p. 214-221. [doi: 10.1109/IE.2012.39]
26. R Development Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2014.
27. Bramer M. Avoiding overfitting of decision trees. In: Principles of Data Mining. London, UK: Springer International Publishing; 2013:121-136.

Abbreviations

GT3X+: ActiGraph GT3X+
MET: metabolic equivalent score
RMS: root mean square
RPE: ratings of perceived exertion
SGS: Samsung Gear S smartwatch

Edited by G Eysenbach; submitted 11.06.18; peer-reviewed by J Sasaki, S Barbieri; comments to author 13.07.18; revised version received 08.11.18; accepted 09.12.18; published 06.02.19.

Please cite as:
Davoudi A, Wanigatunga AA, Kheirkhahan M, Corbett DB, Mendoza T, Battula M, Ranka S, Fillingim RB, Manini TM, Rashidi P
Accuracy of Samsung Gear S Smartwatch for Activity Recognition: Validation Study
JMIR Mhealth Uhealth 2019;7(2):e11270
URL: http://mhealth.jmir.org/2019/2/e11270/
doi:10.2196/11270
PMID:30724739

Original Paper

Applying Multivariate Segmentation Methods to Human Activity Recognition From Wearable Sensors' Data

Kenan Li[1], PhD; Rima Habre[1], PhD; Huiyu Deng[1], MSc; Robert Urman[1], PhD; John Morrison[1], MSc; Frank D Gilliland[1], PhD; José Luis Ambite[2], PhD; Dimitris Stripelis[2], MSc; Yao-Yi Chiang[3], PhD; Yijun Lin[3], MSc; Alex AT Bui[4], PhD; Christine King[5], PhD; Anahita Hosseini[6], MSc; Eleanne Van Vliet[1], PhD; Majid Sarrafzadeh[6], PhD; Sandrah P Eckel[1], PhD

[1]Department of Preventive Medicine, Keck School of Medicine of University of Southern California, Los Angeles, CA, United States
[2]Information Sciences Institute, University of Southern California, Los Angeles, CA, United States
[3]Spatial Sciences Institute, University of Southern California, Los Angeles, CA, United States
[4]Department of Radiological Sciences, University of California Los Angeles, Los Angeles, CA, United States
[5]Department of Biomedical Engineering, University of California, Irvine, Irvine, CA, United States
[6]Department of Computer Science, University of California Los Angeles, Los Angeles, CA, United States

Corresponding Author:
Kenan Li, PhD
Department of Preventive Medicine
Keck School of Medicine of University of Southern California
Soto Building Room 202-09
2001 North Soto Street
Los Angeles, CA, 90089
United States
Phone: 1 2256102559
Email: kenanl@usc.edu

Abstract

Background: Time-resolved quantification of physical activity can contribute to both personalized medicine and epidemiological research studies, for example, managing and identifying triggers of asthma exacerbations. A growing number of reportedly accurate machine learning algorithms for human activity recognition (HAR) have been developed using data from wearable devices (eg, smartwatch and smartphone). However, many HAR algorithms depend on fixed-size sampling windows that may poorly adapt to real-world conditions in which activity bouts are of unequal duration. A small sliding window can produce noisy predictions under stable conditions, whereas a large sliding window may miss brief bursts of intense activity.

Objective: We aimed to create an HAR framework adapted to variable duration activity bouts by (1) detecting the change points of activity bouts in a multivariate time series and (2) predicting activity for each homogeneous window defined by these change points.

Methods: We applied standard fixed-width sliding windows (4-6 different sizes) or greedy Gaussian segmentation (GGS) to identify break points in filtered triaxial accelerometer and gyroscope data. After standard feature engineering, we applied an Xgboost model to predict physical activity within each window and then converted windowed predictions to instantaneous predictions to facilitate comparison across segmentation methods. We applied these methods in 2 datasets: the *human activity recognition using smartphones* (*HARuS*) dataset where a total of 30 adults performed activities of approximately equal duration (approximately 20 seconds each) while wearing a waist-worn smartphone, and the Biomedical REAl-Time Health Evaluation for Pediatric Asthma (*BREATHE*) dataset where a total of 14 children performed 6 activities for approximately 10 min each while wearing a smartwatch. To mimic a real-world scenario, we generated artificial unequal activity bout durations in the BREATHE data by randomly subdividing each activity bout into 10 segments and randomly concatenating the 60 activity bouts. Each dataset was divided into ~90% training and ~10% holdout testing.

Results: In the HARuS data, GGS produced the least noisy predictions of 6 physical activities and had the second highest accuracy rate of 91.06% (the highest accuracy rate was 91.79% for the sliding window of size 0.8 second). In the BREATHE data, GGS again produced the least noisy predictions and had the highest accuracy rate of 79.4% of predictions for 6 physical activities.

Conclusions: In a scenario with variable duration activity bouts, GGS multivariate segmentation produced *smart-sized* windows with more stable predictions and a higher accuracy rate than traditional fixed-size sliding window approaches. Overall, accuracy was good in both datasets but, as expected, it was slightly lower in the more real-world study using wrist-worn smartwatches in children (BREATHE) than in the more tightly controlled study using waist-worn smartphones in adults (HARuS). We implemented GGS in an offline setting, but it could be adapted for real-time prediction with streaming data.

(JMIR Mhealth Uhealth 2019;7(2):e11201) doi:10.2196/11201

KEYWORDS

machine learning; physical activity; smartphone; statistical data analysis wearable devices

Introduction

Background

Time-resolved quantification of physical activity is important because physical activity is linked with human health. Physical activity has direct health benefits, and the American College of Sports Medicine and the Centers for Disease Control and Prevention [1] publish physical activity guidelines to promote and maintain public health (eg, children should do at least 60 min of physical activity per day). Physical activity also has indirect effects on health by modifying exposures of pollutants. The National Human Activity Pattern Survey [2] found that human activity patterns play a key role in explaining variation in pollutant exposures—by impacting the timing, location, and degree of exposures—and related health outcomes. It follows that high-resolution time-resolved monitoring of human activity may have clinical and research applications. Not only could a person's moderate-to-vigorous activity (or inactivity) be logged to quantify typical spatio-temporal patterns but deviations from the typical routine could also be identified as possible targets for intervention. The widespread use of wearable smartphones and smartwatches, together with advances in communication, computation, and sensing capabilities, makes real-time human activity recognition (HAR) possible by providing remote data acquisition and on-device processing.

Indeed, wearable sensors and mobile devices are being increasingly used in studies assessing physical activity, sleep, mobility, medication adherence, and a variety of other areas [3]. Our study is motivated by the "Pediatric Research using Integrated Sensor Monitoring Systems" (PRISMS) program—launched in 2015 by the National Institute of Biomedical Imaging and Bioengineering—to develop a sensor-based, integrated health monitoring system for studying pediatric asthma. Asthma is a heterogeneous, multifactorial disease that is one of the most common causes of emergency hospital visits in children [4]. Important risk factors for asthma exacerbation include allergen and air pollutant exposures and viral infection [4], but physical activity also plays an important role in asthma incidence [5], acute symptoms [6], and long-term control [7,8]. In a framework such as PRISMS, HAR may facilitate the management of asthma and the identification of triggers of exacerbation.

Windowing in Human Activity Recognition Modeling Approaches

Data for HAR are increasingly collected using wearable sensors (eg, accelerometers and gyroscopes) that permit continuous, real-time monitoring [9-13]. Most HAR studies summarize accelerometer and gyroscope data streams—as well as the resulting instantaneous activity predictions—using a time-based windowing approach. The reasons for this are two-fold. First, the typical duration of human activities is significantly longer than the sensors' sampling rate (eg, 10-50 Hz). Second, raw data from an accelerometer or gyroscope are highly variable, noisy, and oscillatory, so instantaneous raw values may provide insufficient information to differentiate the associated activity. The size of the window is constrained by the sensor sampling frequency and is an important parameter that affects the accuracy of the HAR prediction, the computational loads of the algorithm, and the energy consumption on the wearable device. When selecting the size of a fixed-size window, there is a trade-off between being too short (captures fine details and produces noisy predictions) and being too long (misses short-duration activity bouts and produces more stable predictions). In a platform such as PRISMS where researchers might want to tailor context-sensitive interactions with study participants (eg, triggering a notification or survey) based on physical activity patterns, windows that are too short could generate frequent interactions with users, leading to notification fatigue and reduced compliance. Longer windows could perform well at certain times of the day when activities are fairly constant over long periods (eg, sedentary classroom time) but poorly during periods of high variability (eg, gym class and getting ready for school). A variable-sized sampling window approach with data-driven break points (at times when the activities may change) has the potential to improve HAR and improve the usability of platforms involving HAR.

Time Series Segmentation

Fixed-size sliding windows are 1 type of a larger class of segmentation methods in time series analysis. Segmentation methods divide a time series into segments having similar characteristics. Most segmentation algorithms can be framed in several ways: (1) producing the best representation using only a given number of segments, (2) producing the best representation such that the maximum error for any segment does not exceed the given threshold, or (3) producing the best representation such that the combined error of all segments is less than the given threshold [14]. Multivariate segmentation methods segment multidimensional signals. Multivariate segmentation has been studied in several contexts using various approaches (each with different assumptions), including Bayesian change point detection [15], hypothesis testing [16], mixture models, hidden Markov models [17], and convex segmentation [18]. For this study, we selected a multivariate

segmentation algorithm called greedy Gaussian segmentation (GGS) [19], which is based on maximizing the likelihood of the data for a fixed number of segments. GGS assumes that in each segment, the mean and covariance are constant and independent of the means and covariances in all other segments. GGS is a scalable greedy algorithm and is applicable to solve much larger problems (in terms of vector dimension and time series length) than many of other above methods.

In this paper, we provide background on the GGS algorithm and perform a novel application of GGS to offline HAR, comparing GGS with the standard fixed-size sliding window approach. We use data from 2 HAR studies with different prescribed activity durations and different sensor wear modalities (waist-worn sensor and wrist-worn sensor). After processing the data using either segmentation approach, we used standard feature engineering and machine learning methods to predict activities and compared the accuracy of the 2 different segmentation approaches.

Methods

Data

The *human activity recognition using smartphones* (HARuS) dataset consists of 61 experiments conducted by 30 volunteers aged 19 to 48 years [20]. Triaxial accelerometery and gyroscope data were collected at 50 Hz by a waist-worn smartphone (Samsung Galaxy S II). Each experiment was about 7 min long. In each experiment, the HARuS protocol scripted 12 ambulation activities, including 6 basic activities (each approximately 20 seconds in duration) and 6 postural transition activities (stand-to-sit, sit-to-stand, sit-to-lie, lie-to-sit, stand-to-lie, and lie-to-stand). The 6 basic activities include 3 static postures (standing, sitting, and lying) and 3 dynamic activities (walking, walking downstairs, and walking upstairs). The raw data were directly acquired from the smartphone readings, and the activities were labeled by manual review of video recordings of each experiment. To be consistent with previous studies [11], we only modeled the 6 basic activities and deleted the 6 types of postural transition activity bouts and all unlabeled sessions, all of which were of relatively short duration and unlikely, for example, to be strongly associated with asthma exacerbation in studies using PRISMS [5]. The dataset was divided into the first 55 experiments for training (2 experiments each for 26 people and 3 experiments for 1 participant) and 6 experiments (2 experiments each for 3 people) for holdout testing. The 6 raw signals of experiment 1 are plotted in Multimedia Appendix 1.

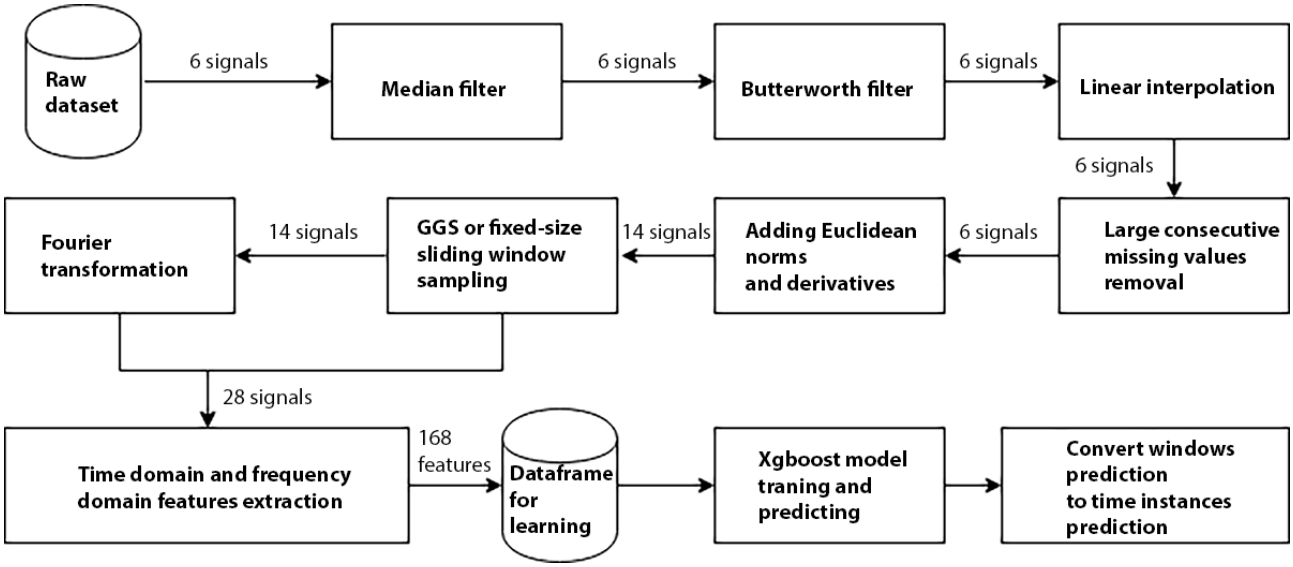
The Los Angeles PRISMS Center *BREATHE* dataset [21-23] was collected on 16 participants, aged 5 to 15 years, using the BREATHE Kit, an informatics platform designed to monitor multiple exposures, behaviors, and activities in context to identify personal triggers and predict the risk of pediatric asthma exacerbations in real time. Triaxial accelerometry and gyroscope data were collected at 10 Hz using a wrist-worn Motorola Moto 360 Sport smartwatch. Participants performed each of the 5 activities (standing, sitting, lying, walking, and walking on stairs) for 10 min and running for 5 min (to minimize discomfort). Unlike the HARuS dataset, participants were permitted to perform natural movements (especially free arm movement such as sitting while typing or using a smartphone) during each activity. The raw data were acquired as the end product of a data pipeline (from smartwatch to the BREATHE app on the smartphone via Bluetooth and then securely uploaded to the BREATHE servers wirelessly and in real times). For the BREATHE dataset, we modeled all 6 scripted activities: standing, sitting, lying, walking, walking on stairs (labels did not differentiate up and down stairs), and running. We used experiments from 14 of the 16 participants as 2 participants had substantial quantities of missing data. In the BREATHE dataset, data were saved as separate files for each activity, for each participant. To evaluate whether GGS segmentation improves prediction under a scenario of variable activity bout durations, we generated artificial activity data files for each participant by (1) randomly dividing his or her activity sessions (each about 10 min long) into 10 subsessions; then (2) randomly shuffling all subsessions (60 in total); and finally (3) concatenating all 60 subsessions into 1 data file, potentially resulting in fewer than 60 distinct activity bouts if bouts with identical activities are located next to each other. Hence, we produced 14 artificial activity files with artificial unequal activity bout durations, one for each of the 14 participants. The artificial dataset was divided into the first 12 participants for training and the last 2 participants for holdout testing. The 6 raw signals of experiment 1 are plotted in Multimedia Appendix 1.

Workflow

Figure 1 provides an overview of our workflow. For both datasets, the raw data were first preprocessed by applying a median filter (kernel size=3) to remove outliers. Afterwards, a Butterworth [24] filter was used to remove artifacts and baseline wandering noise associated with the data acquisition process (eg, the constant force of gravity or shaking the device). Specifically, a third-order low-pass Butterworth filter was applied separately to each triaxial component (x, y, and z of the accelerometer and gyroscope). A power spectral density (PSD) was calculated and used to choose the cut-off frequency, over which the sensor signals were attenuated. PSD is a metric that estimates the distribution of power over frequency, and it has been widely implemented to evaluate filters of high-frequency with baseline-wandering noise [25].

Subsequently, the data streams were temporally aligned. The sampling frequency observed in practice can be a result of practical constraints (eg, battery saving and restricted access by the software stack in mobile device's operating systems). Thus, observed data can be sampled irregularly, with mismatch between the 2 sensors. In the HARuS dataset, there were no mismatched time stamps (ie, only existing for 1 sensor) when we concatenated accelerometer and gyroscope readings according to their time stamps. However, the BREATHE dataset contained considerable mismatching, and both the accelerometer and the gyroscope were not perfectly collected at 10 Hz. To align the 2 sensor readings, we first downscale sampled the raw data at 50 Hz to round their time stamps to the nearest 50 Hz sampling point, and then we applied a linear interpolation method.

Figure 1. The workflow of the human activity recognition framework. GGS: greedy Gaussian segmentation.



Specifically, we added (as necessary) records for all 50 Hz time stamps to both sensor data files and linearly interpolated missing sensor readings (approximately 80% because of the downscaling) based on the left 5 adjacent nonmissing values and the right 5 adjacent nonmissing values. In addition to the missing values caused by the mismatching time stamps, there was also a number of longer periods with missing values in the BREATHE dataset. After aligning the 2 sensors, we truncated time periods with more than 10 seconds of consecutive missing values.

Data transformation was used to augment the original data (6 signals from 2 triaxial sensors) with additional transformed signals. Statistical features were later extracted from both the raw and transformed signals. Specifically, 8 new signals were generated: 6 derivatives with respect to time (1 for each of the 6 original signals) and 2 Euclidean norms (1 for the x-, y-, and z-axis of each sensor). Hence, a total of 14 signals were available (6 original measured signals and 8 new calculated signals).

Time windows were generated using 2 approaches. First, multivariate segmentation on the 6 original signals produced windows of varying sizes, with break points selected using training data to reflect changes of the means and covariances of the raw signals (a detailed description follows). Second, for comparison, we created various sizes of nonoverlapping fixed-length sliding windows (4 sizes for HARuS dataset: 0.2 second, 0.8 second, 3 seconds, and 8 seconds; 6 sizes for BREATHE dataset: 0.2 second, 0.8 second, 3 seconds, 8 seconds, 12 seconds, and 40 seconds). Window sizes were chosen to include approximately the window size in the original HARuS study (2.56 seconds) [20] and to reflect a wide enough range to include the optimum window size for both datasets.

Within each set of windows, we extracted statistical features for input into a machine learning model. These statistical features were either based on time domain (the original time-based windows) or frequency domain (Fourier transformation of the original time-based windows). For each set of windows, we calculated a total of 168 features: 6 statistics (arithmetic mean, SD, median absolute deviation, minimum, maximum, and entropy) on 14 signals and on both the time and frequency domains (6 x 14 x 2=168).

Multivariate Segmentation

A brief description of GGS [19] is as follows. Consider a multivariate time series consisting of T time instants x_1, x_2,..., $x_T \in R^m$, where m is the number of features (ie, m=6 in our study). The time series need not be uniformly sampled in real time (see note in Discussion on the independence assumption). Given K break points b_1,..., $b_K \in (1,..., T)$ between a starting point $b_0 = 1$ and an end point $b_{K+1} = T$, we assume that $x_t \sim MVN(\mu_{bi}, \Sigma_{bi})\ \forall\ t \in (b_i,..., b_{i+1})\ \forall\ i \in [0, K]$ and are independent samples, where μ_{bi} and Σ_{bi} denote the mean vector and covariance matrix of the multivariate normal distribution within the interval of (b_i,.., b_{i+1}). A GGS can be learned on the multivariate time series by fitting a greedy algorithm to maximize the covariance-regularized log-likelihood.

In Figure 2 equation a, where l (b, μ, Σ) denotes the log-likelihood before regularization, b denotes the vector of break points, μ denotes [μ_{b0},..., μ_{bK}], Σ denotes [Σ_{b0},..., Σ_{bK}], and $\lambda \geq 0$ is an *a priori* specified hyperparameter that controls the amount of regularization [19]. The greedy heuristic algorithm follows a top-down subroutines of adding a new break point with the largest increase of $\Phi(b,\mu,\Sigma)$ at each step until K, and then in a bottom-up way adjusts the positions of all break points until no change of any 1 break point increases $\Phi(b,\mu,\Sigma)$. A curve of the covariance-regularized log-likelihood versus K can be used to select K for a given dataset.

Figure 2. Equations.

$$\Phi(b, \mu, \Sigma) = l(b, \mu, \Sigma) - \lambda \sum_{i=0}^{K} tr\left(\Sigma_{b_i}^{-1}\right) \quad \text{(a)}$$

$$\hat{y}_i = \sum_{m=1}^{M} f_m(x_i) \quad \text{(b)}$$

$$J^m = \sum_{i=1}^{n} j\left(\hat{y}_i^{m-1} + f_m(x_i),\ y_i\right) + \sum_m \Omega(f_m) \quad \text{(c)}$$

$$\hat{y}_i^{m-1}$$

Gradient Boosted Trees Classification

To achieve high accuracy using a scalable method, we predicted activity classes using Xgboost [26], an implementation of a tree-based boosting widely used in machine learning challenges. For a given dataset (D) with *n* observations and *p* features (ie, *p*=168 in our analysis), $D=\{(x_i \in R^p, y_i \in R)\} \ \forall \ i \in [1, n]$, Xgboost ensembles M trees denoted f_m to predict the output y_i.

The model is trained in a greedy, additive manner starting from m=1 (Figure 2, equation b). Let $\hat{y}_i^{m-1}$ be the prediction of y_i at the $(m-1)^{th}$ iteration. We add f_m to minimize the following objective (J^m) until the satisfying convergence between the prediction and the ground truth, where j is a predefined differentiable convex loss function that measures the difference between the current prediction and the ground truth and Ω is a predefined regularization term that penalizes the complexity of the model to prevent overfitting:

Xgboost has features that can outperform other implementations of tree-based boosting (eg, boosted trees in scikit-learn and generalized boosted regression model in R) such as (1) using an exact (or approximate, for large datasets) greedy algorithm to enumerate over all possible splits to find the best solution, (2) alleviating slow-downs using a cache-aware prefetching algorithm, and (3) enabling out-of-core computation by dividing the data into multiple blocks, each stored on disk, to use machine's maximum resources (see Figure 2, equation c).

For the HARuS and BREATHE datasets, we tuned and implemented an Xgboost model with *m*=200 trees and learning rate=0.1 (more specifications in Tables 1 and 2) using *p*=168 features calculated on each segment (from fixed-sized windows or GGS) of the training data. Segment-specific predictions for the testing data were translated into instantaneous predictions to facilitate comparison across segmentation approaches. Final evaluations of accuracy were based on instantaneous predictions.

Table 1. Confusion matrix of instantaneous predictions using greedy Gaussian segmentation from the 6 test experiments in the Human Activity Recognition using Smartphones dataset.

True categories	Xgboost[a] predicted categories						Recall (%)	Precision (%)
	W[b]	WU[c]	WD[d]	ST[e]	STD[f]	LY[g]		
W	11238	0	935	0	0	0	100	92.32
WU	0	11070	1297	0	12	0	99.61	89.43
WD	0	0	11659	0	40	0	83.36	99.66
ST	0	0	0	11037	2798	0	85.14	79.78
STD	0	24	96	1926	12546	0	81.49	85.98
LY	0	19	0	0	0	15266	100	99.88

[a]Xgboost specification: base_score=0.5, booster="gbtree," colsample_bylevel=1, colsample_bytree=1, gamma=0, learning_rate=0.1, max_delta_step=0, max_depth=2, min_child_weight=1, missing=None, n_estimators=200, n_jobs=1, nthread=None, objective="multi:softprob," random_state=0, reg_alpha=0, reg_lambda=1, scale_pos_weight=1, seed=None, silent=True, subsample=1. Overall accuracy: 91.06%.

[b]W: walking.

[c]WU: walking upstairs.

[d]WD: walking downstairs.

[e]ST: sitting.

[f]STD: standing.

[g]LY: laying.

Table 2. Confusion matrix of instantaneous predictions using greedy Gaussian segmentation from the 2 test experiments in the BREATHE dataset.

True categories	Xgboost[a] Predicted categories						Recall (%)	Precision (%)
	L[b]	R[c]	ST[d]	STR[e]	STD[f]	WK[g]		
L	38874	166	6920	830	3938	0	76.63	68.76
R	1587	31593	0	12402	791	11693	54.41	82.54
S	12483	0	38596	864	8030	154	64.19	72.65
STR	559	6505	1929	46751	2320	6156	72.80	71.17
STD	887	0	5127	0	54300	77	89.91	72.05
WK	2146	12	555	4846	5976	52455	79.49	74.37

[a]Xgboost specification: base_score=0.5, booster="gbtree," colsample_bylevel=1, colsample_bytree=1, gamma=0, learning_rate=0.1, max_delta_step=0, max_depth=3, min_child_weight=1, missing=None, n_estimators=200, n_jobs=1, nthread=None, objective="multi:softprob," random_state=0, reg_alpha=0, reg_lambda=1, scale_pos_weight=1, seed=None, silent=True, subsample=1. Overall accuracy: 79.4%.

[b]L: lie.

[c]R: run.

[d]S: sit.

[e]STR: stair.

[f]STD: stand.

[g]WK: walk.

Results

Human Activity Recognition Using Smartphones Dataset

The PSD curves to determine the cut-off frequency of the Butterworth filter are displayed in Figure 3. All 6 PSD curves taper to 0 at higher frequencies, with largest values in the lower frequency range from 0 Hz to 5 Hz. There is little baseline wandering noise in high frequencies (>10 Hz). For consistency with previous studies [11], we chose 20 Hz as the cut-off frequency.

For GGS in the HARuS training data, the total covariance-regularized log-likelihood elevated rapidly as K increased from 0 to an inflection point around 16, and then even less rapidly (Figure 4). To favor more detailed segmentation results and allow for some incorrectly identified break points, especially during noisy periods and the transitory periods, we conservatively selected 50 break points.

Figure 3. Triaxial (x, y, and z) power spectra density curves of accelerometer (top row) and gyroscope (bottom row) of the human activity recognition using smartphones training dataset. ACC: accelerometer; Gyro: gyroscope; PSD: power spectral density.

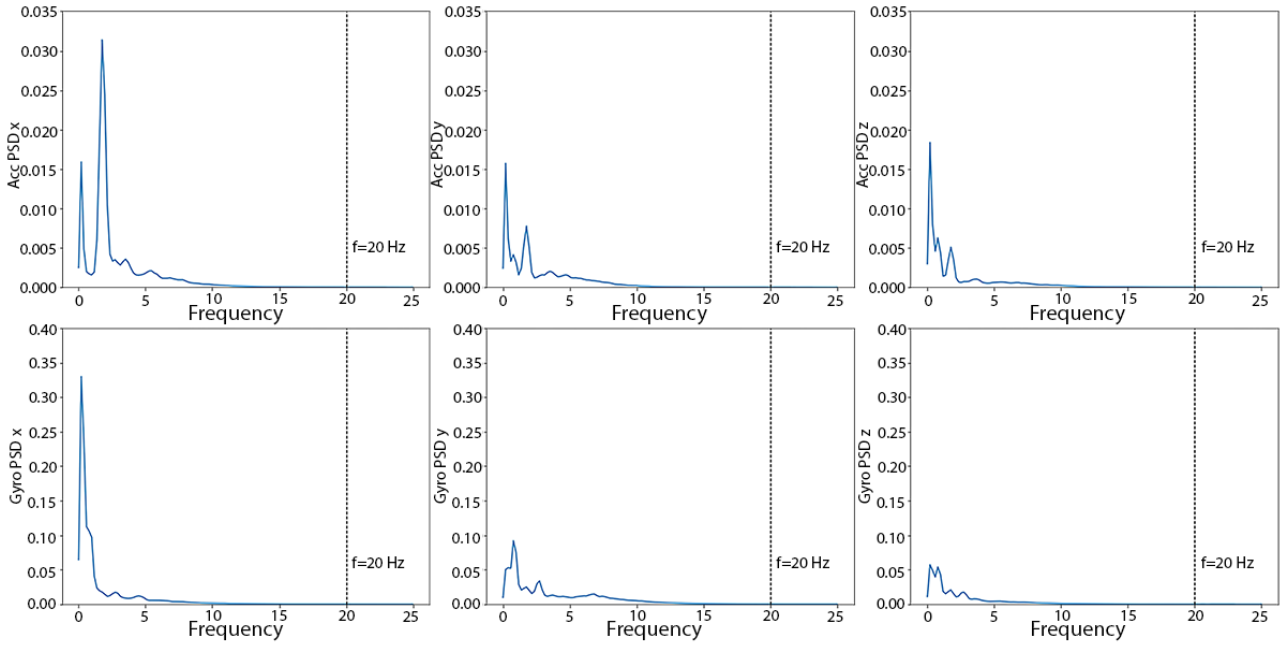


Figure 4. Total covariance-regularized log-likelihood curve of the human activity recognition using smartphones training dataset.

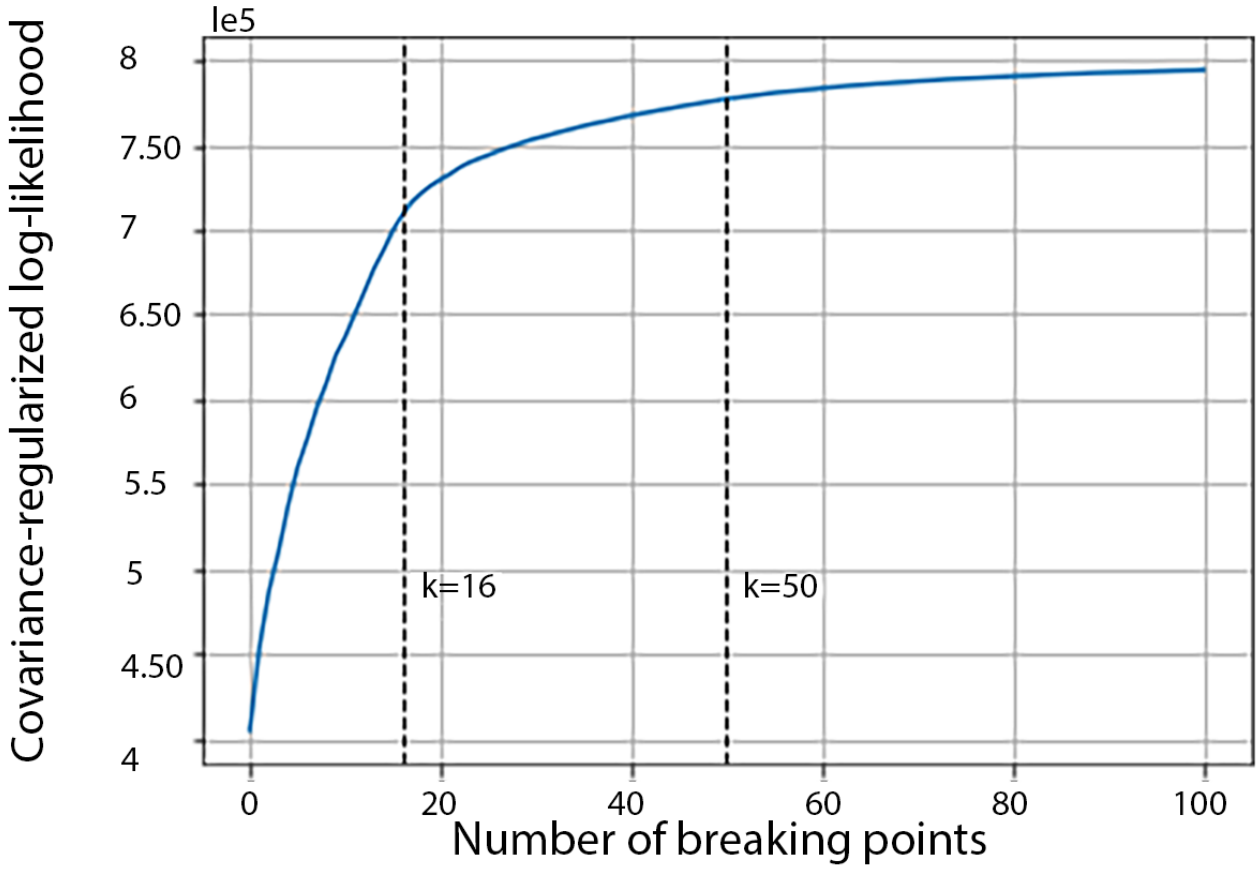
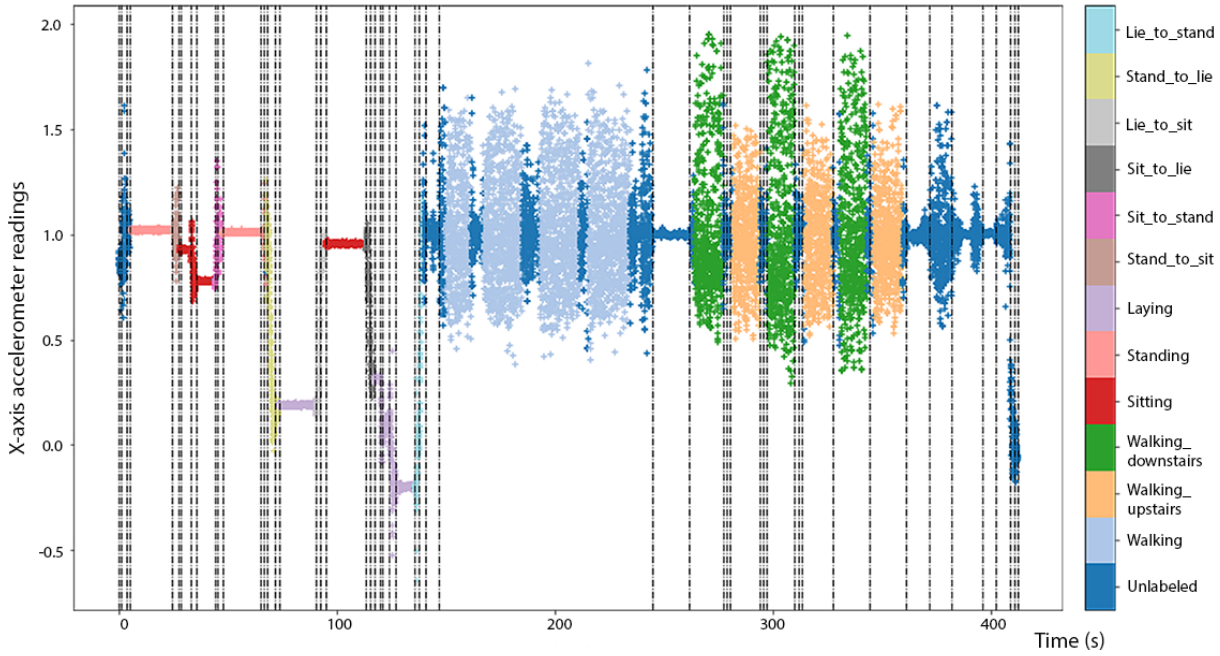


Figure 5. Multivariate segmentation break points (K=50) displayed using vertical dashed lines on the time series of x-axis accelerometer readings from experiment 1 in the human activity recognition using smartphones training dataset.



As shown in Figure 5 for experiment 1, the 13 bouts of the 6 nontransitory activities were generally well separated by the 50 break points. For this experiment, the first bout of sitting and the second bout of laying were both relatively noisy, and erroneous break points were created within these sessions.

We trained an Xgboost model (Figure 6), a support vector machine (SVM) model using a radial basis function kernel and a random forest model using the segmented data. The instantaneous accuracy rate of the Xgboost model using GGS in the 6 holdout experiments was 91.06% (Table 1). This result is higher than the 89.3% accuracy reported in the original HARuS study on the same set of 6 activities [11], and it also should be noted that their accuracy was calculated using sliding window predictions and not instantaneous predictions. Had we calculated accuracy using segment-level predictions, our accuracy would have been 95.96%. When activities were misclassified, they tended to be misclassified as other similar energy activities (Table 1). For example, sitting was most frequently misclassified as standing. The results of the SVM model and the random forest model are summarized in Multimedia Appendix 1.

In comparison, the instantaneous accuracy of Xgboost models fitted using fixed-width sliding windows was highest for the 0.8-second window (91.79%), as shown in Figure 7. This *optimal* window size is smaller than the one used in the original HARuS paper (2.56 seconds) [20]. As might be expected from experiments designed to have equally sized activity bouts, the 0.8-second fixed-size sliding window accuracy was slightly higher than that from GGS (91.06%). In the HARuS data, predictions were relatively stable, with some additional variability for the smallest size sliding windows (Figure 8). The 3 most important features from Xgboost using GGS were the segment-specific mean, minimum of the x-axis of the accelerometer, and the mean of the x-axis of the gyroscope (Figure 9).

Figure 6. Instantaneous predictions using greedy Gaussian segmentation (top row) and ground truth (bottom row) from the 6 test experiments in the human activity recognition using smartphones dataset.

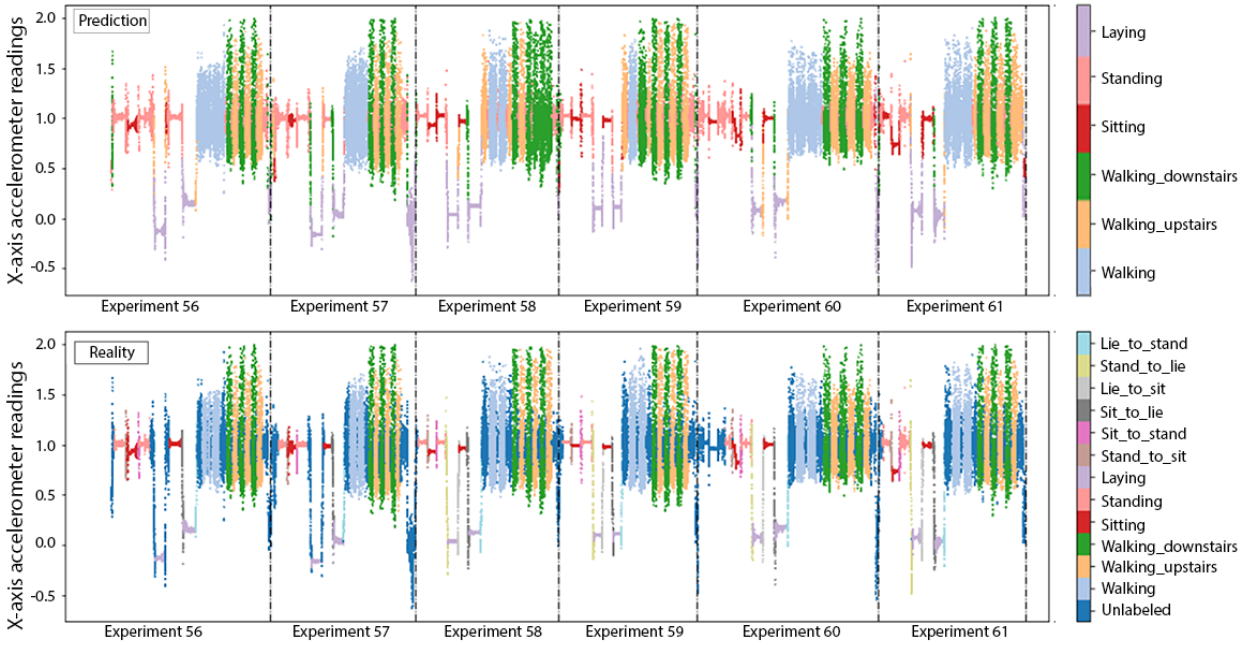


Figure 7. Accuracy of instantaneous predictions using 4 different fixed-size sliding windows (SWs) in the 6 test experiments in the human activity recognition using smartphones dataset. The horizontal dashed line represents the accuracy using greedy Gaussian segmentation. GGS: greedy Gaussian segmentation.

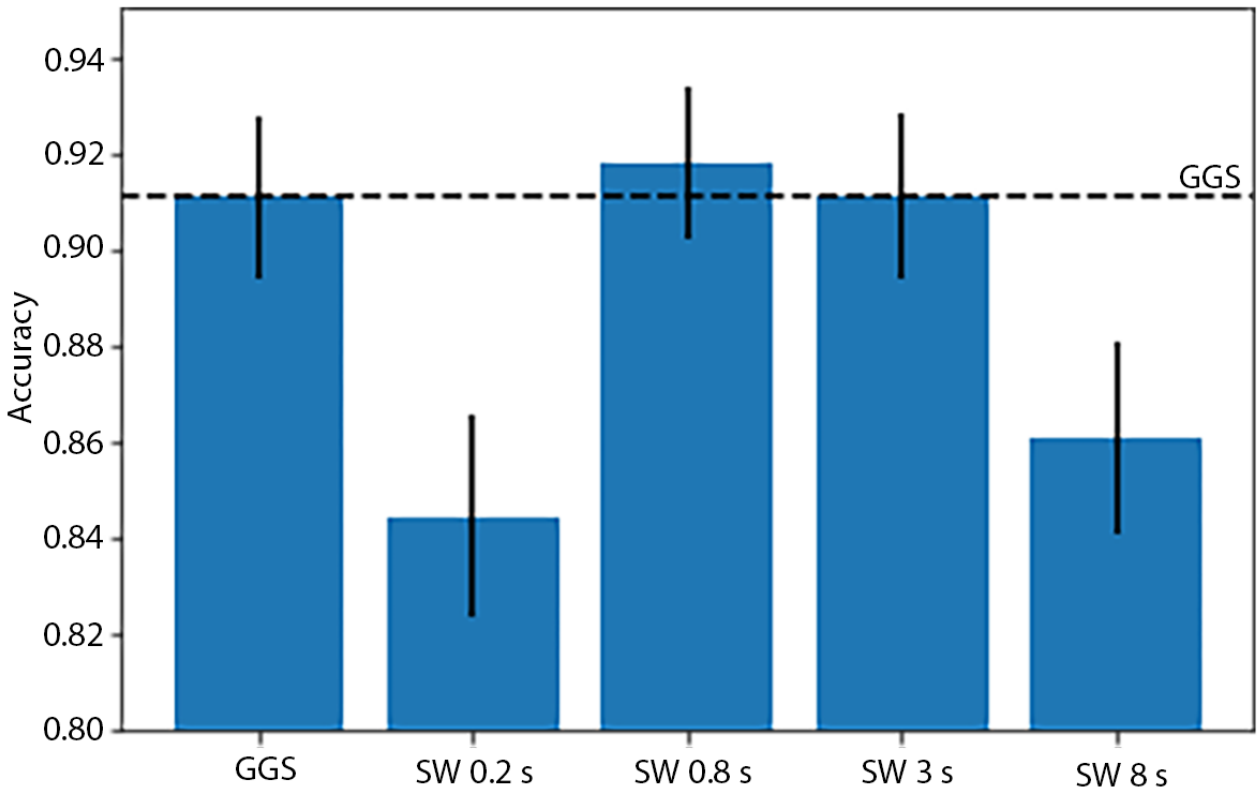


Figure 8. Predictions using 4 different fixed-sized sliding windows (SWs) and greedy Gaussian segmentation, as well as the ground truth for the 6 test experiments in the human activity recognition using smartphones dataset. GGS: greedy Gaussian segmentation; SW: sliding window.

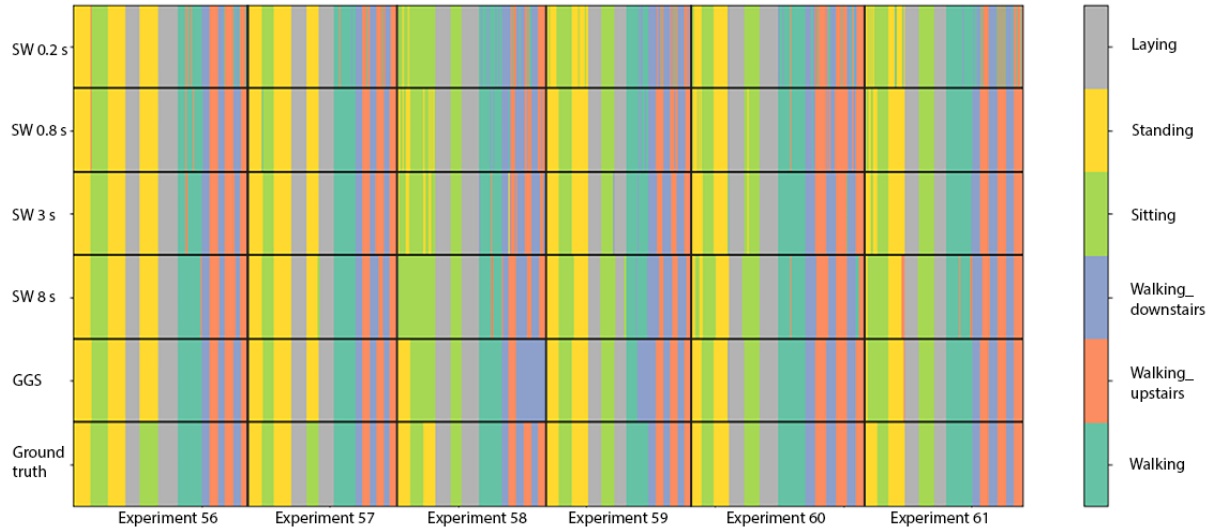
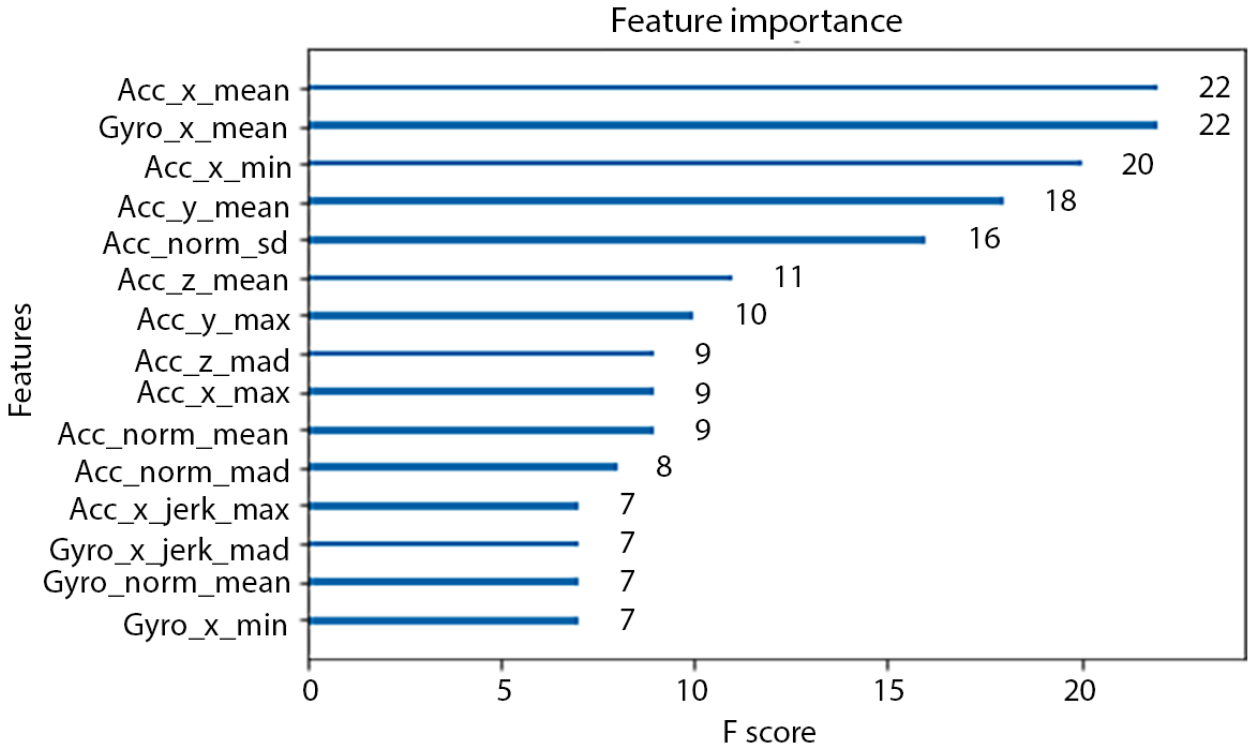


Figure 9. Importance of the top 15 features from Xgboost using greedy Gaussian segmentation from the human activity recognition using smartphones dataset. Abbreviations in the feature names are standard deviation (std), minimum (min), maximum (max), mean absolute deviation (mad), Euclidean magnitude (norm), and derivative (jerk). The operators in the names should be read in the order of from the right to the left. For example, acc_x_jerk_max means the maximum value of the derivative values on the x-axis of the accelerometer sensor. Acc: accelerometer; Gyro: gyroscope.



BREATHE Dataset

On the basis of the PSD plots of the training data (Figure 10), we again chose 20 Hz as the cut-off frequency for the Butterworth filter. The gyroscope energies are in the same scale as the HARuS dataset; however, the accelerometer readings have much larger amplitudes, which makes the curves look smoother in the range of approximately 5 Hz. The zoom-in windows in the accelerometer's 3 subplots show the variations of the PSD curves in the range from 2.5 Hz to 7.5 Hz on a similar scale to that used in the PSD plots for the HARuS data.

The covariance-regularized log-likelihood curve for the 12 training experiments in the BREATHE dataset (Figure 11) had one inflection point at approximately K=60 but no clear second inflection point (through K=300) as we had observed in the HARuS dataset. Interestingly, there were, by design, approximately 60 activity bouts in each BREATHE experiment, demonstrating that GGS again identified the number of different activity bouts. We arbitrarily chose K=100 break points for multivariate segmentation as it was a round number larger than the most obvious inflection point. From Figure 12, it appears that 100 was an adequate number of break points. A choice of

60 break points would have been inadequate to segment approximately 60 bouts as some noisier bouts were erroneously partitioned into multiple segments.

Similar to the HARuS dataset, 3 models were trained: Xgboost, SVM, and random forestAs evident from Figure 13, the predictive accuracy for certain activities varied across participants (eg, the accuracy for running was 71.5% for the participant in experiment 13 and 74.4% for the participant in experiment 14). Similar to the HARuS results, most misclassified records were shuffled either within the active group (walk, stair, and run) or the inactive group (sit, lie, and stand). If the activities had been grouped into active or inactive, the instantaneous accuracy rate would have been 95.0%. The results of the SVM model and the random forest model are shown in Multimedia Appendix 1. The instantaneous accuracy rate of the Xgboost model using GGS was 79.4% (Table 2 and Figure 14).

The accuracies of Xgboost from the 4 smallest fixed-size sliding windows (the same sizes as used in the HARuS dataset) increased monotonously. To achieve the reverse U-shape curve indicating that we obtained the optimum window size, we included 2 additional window sizes. The highest accuracy was achieved for the 8-second window (72.7%) as shown in Figure 13. As expected in this dataset with activity bouts of unequal duration, the *smart-sized* GGS segmentation (79.4% accuracy) considerably outperformed the fixed-size sliding windows. Not only was GGS more accurate but it also produced considerably less noisy predictions as shown in Figure 15. The 2 most important features from Xgboost using GGS were segment specific: mean z-axis and the minimum norm of the triaxial accelerometer signal (Figure 16).

Figure 10. Triaxial (x, y, and z) power spectra density curves of accelerometer (upper 3 subplots) and gyroscope meter (lower 3 subplots) of the BREATHE training dataset. ACC: accelerometer; Gyro: gyroscope; PSD: power spectral density.

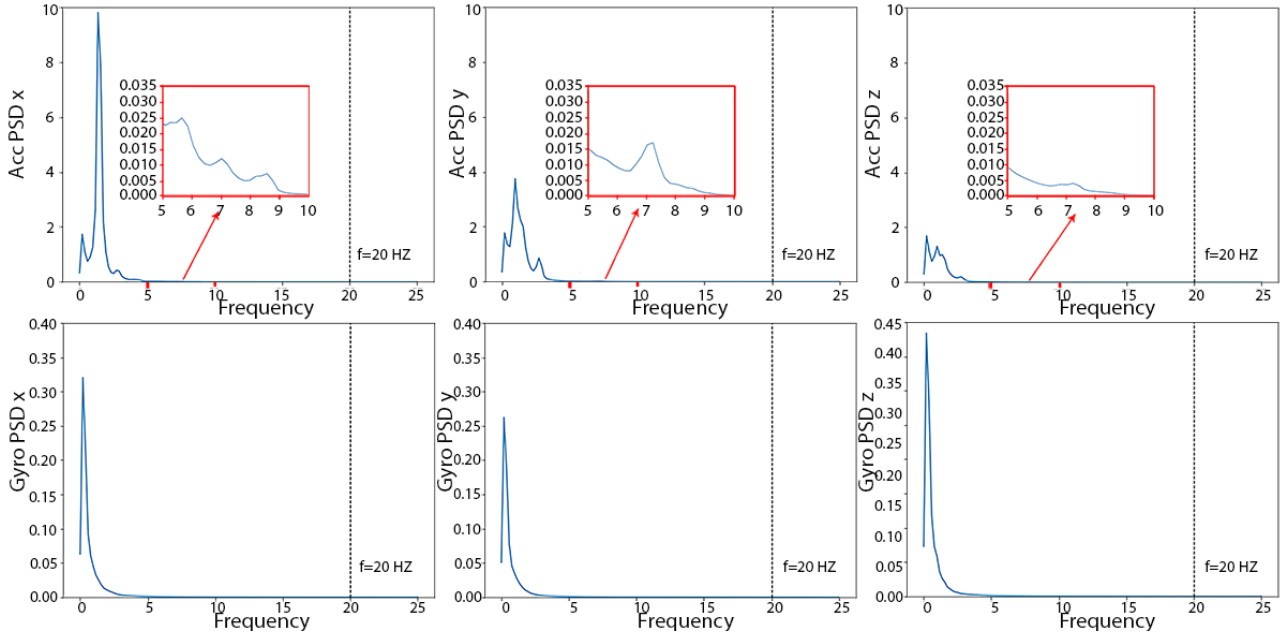


Figure 11. Total covariance-regularized log-likelihood curve of the BREATHE training dataset.

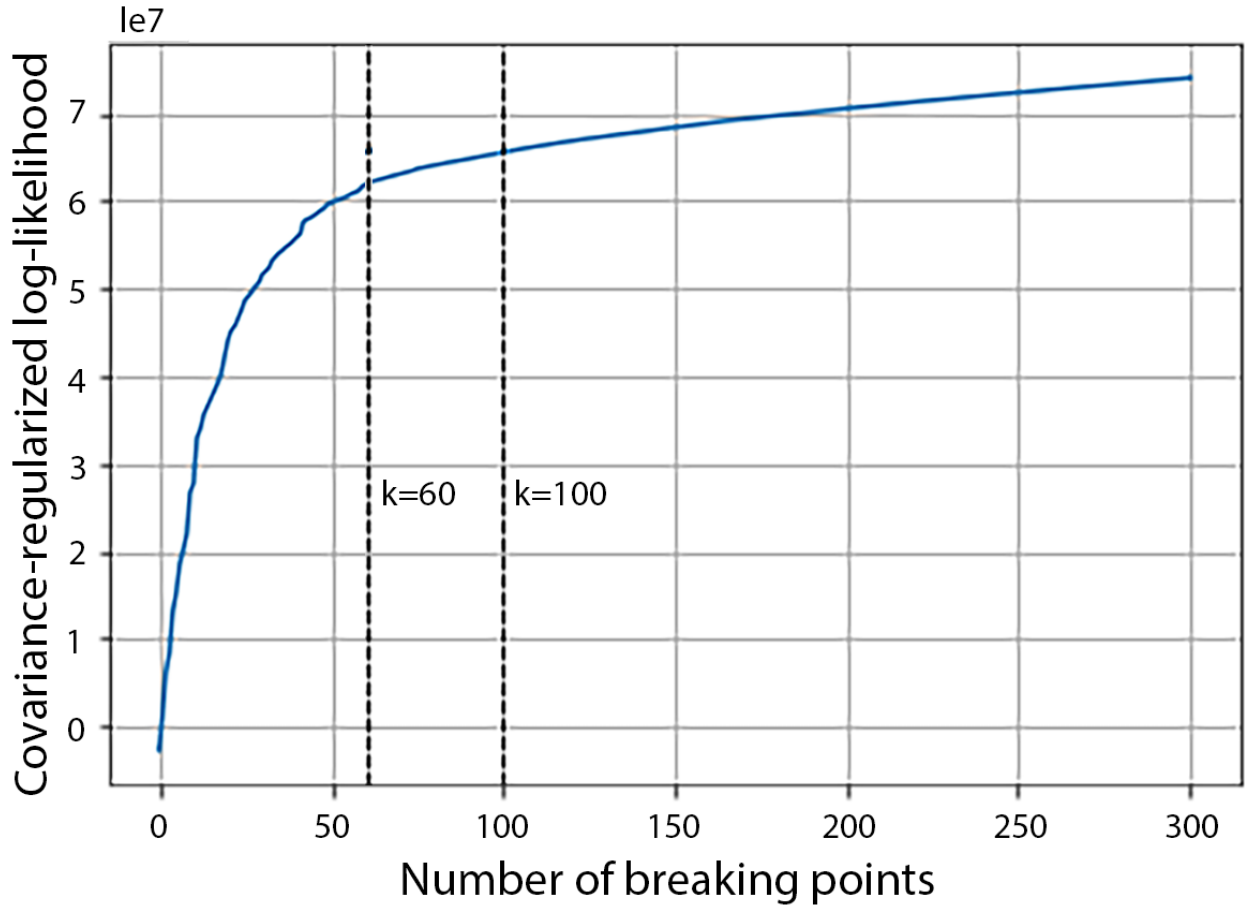


Figure 12. Multivariate segmentation break points (K=100) displayed using vertical dashed lines on the time series of x-axis accelerometer readings from experiment 1 in the BREATHE training dataset.

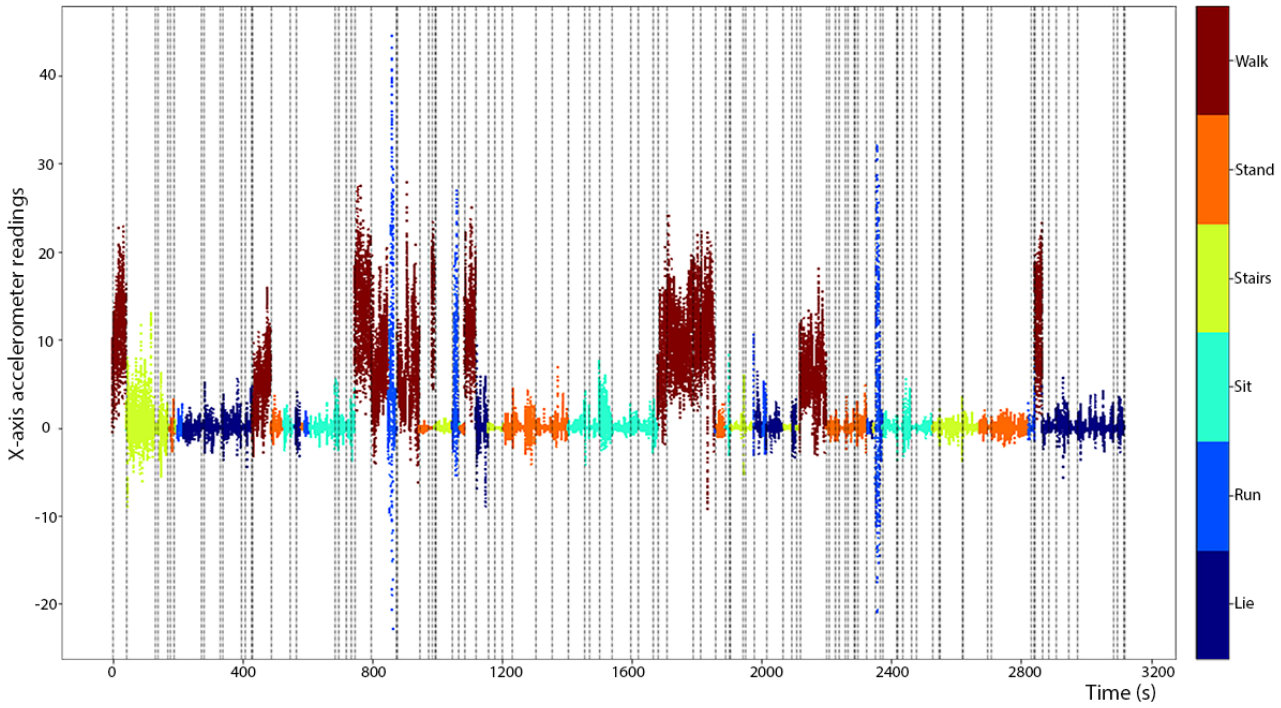


Figure 13. Instantaneous predictions using greedy Gaussian segmentation (top) and ground truth (bottom) from the 2 test experiments (13 and 14) in the BREATHE dataset.

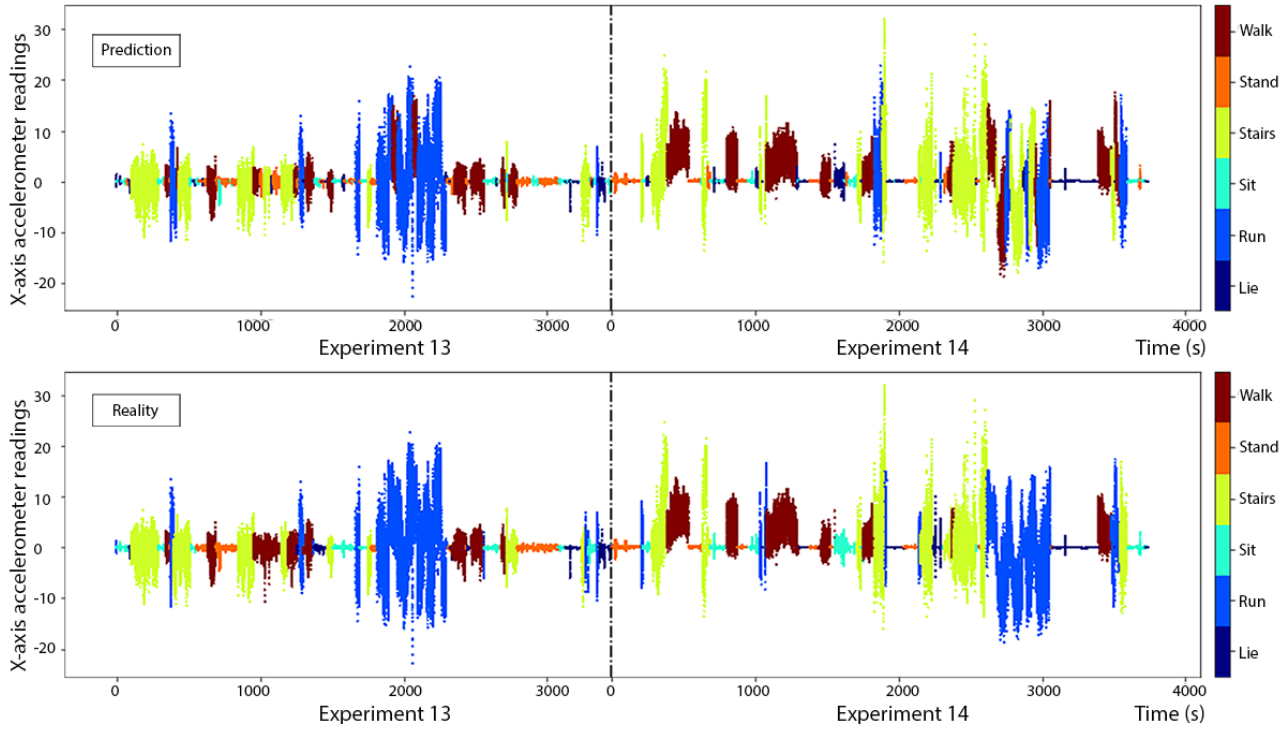


Figure 14. Accuracy of instantaneous predictions from Xgboost using 6 different fixed-size sliding windows (SWs) in the 2 test experiments in the BREATHE dataset. The horizontal dashed line represents the accuracy from Xgboost with greedy Gaussian segmentation. SW: sliding window; GGS: greedy Gaussian segmentation.

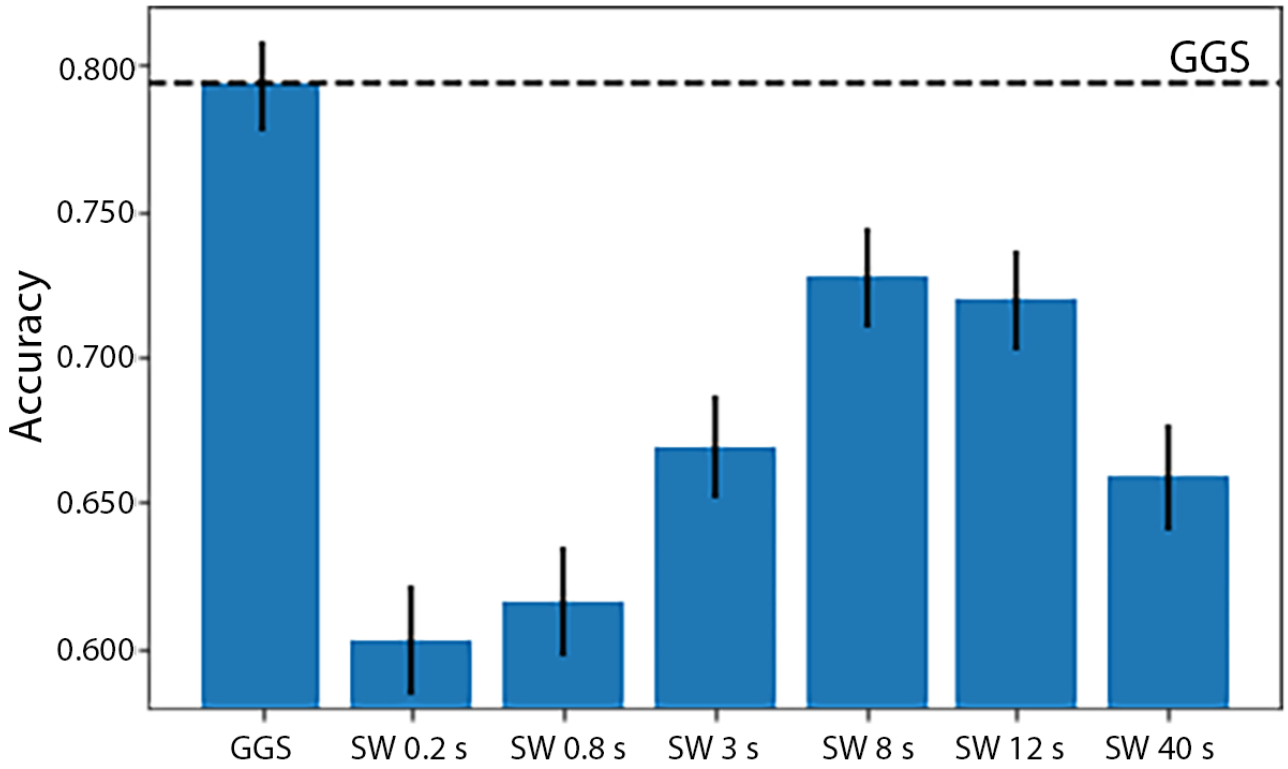


Figure 15. Predictions from Xgboost using 6 different fixed-sized sliding windows (SWs) and greedy Gaussian segmentation as well as the ground truth for experiment 13 of the BREATHE test data. SW: sliding window; GGS: greedy Gaussian segmentation.

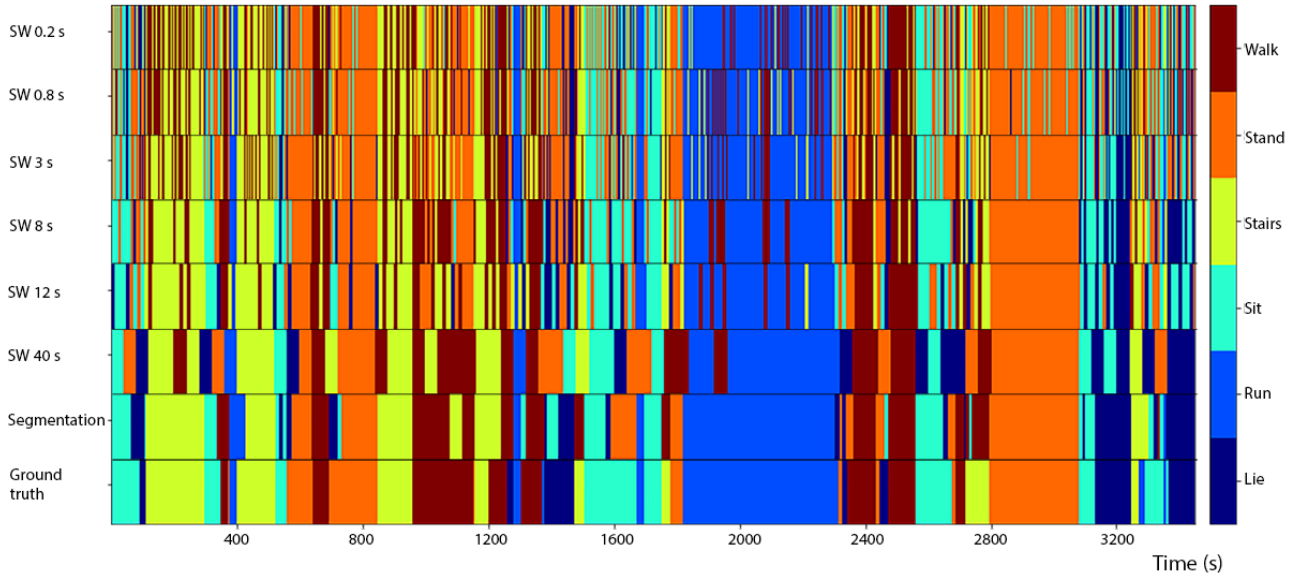
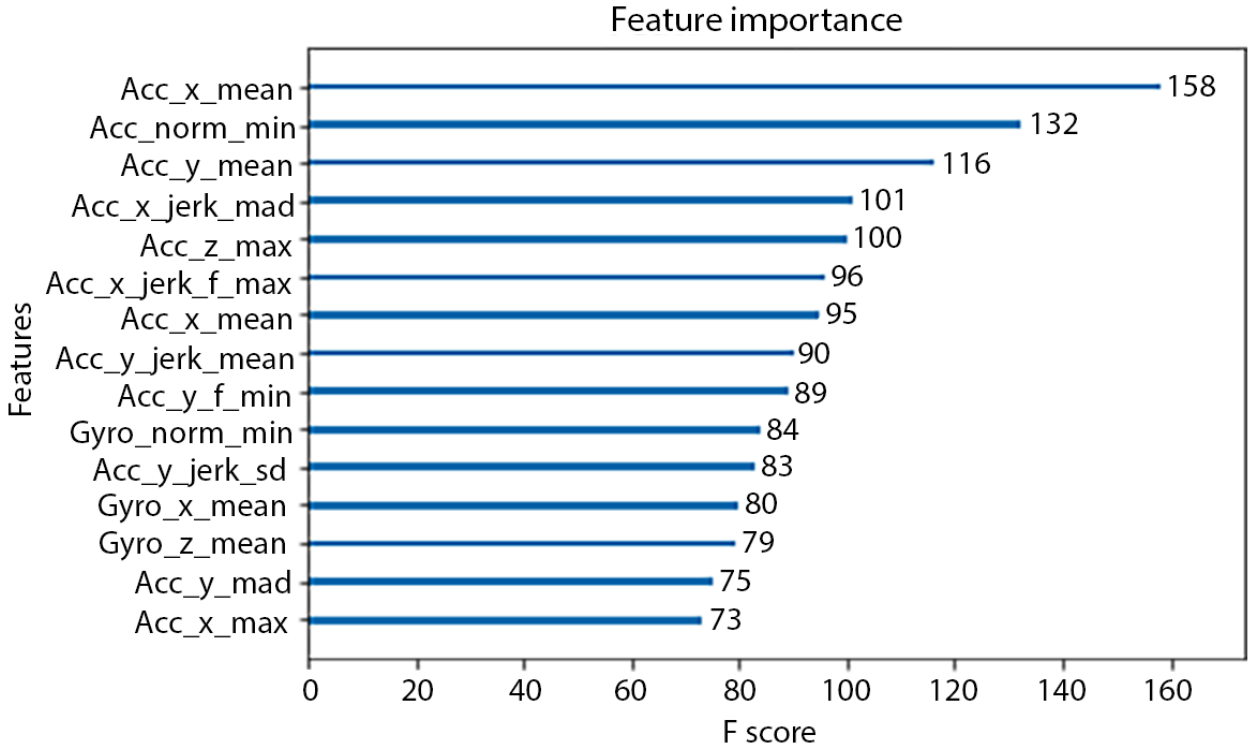


Figure 16. Importance of the top 15 features from Xgboost using greedy Gaussian segmentation from the BREATHE dataset. Abbreviations in the feature names are SD, minimum (min), maximum (max), mean absolute deviation (mad), Euclidean magnitude (norm), and derivative (jerk). The operators in the names should be read in the order of from the right to the left. For example, acc_x_jerk_max means the maximum value of the derivative values on the x-axis of the accelerometer sensor.



Discussion

Summary of Findings

We found that Xgboost using GGS outperformed Xgboost using fixed-size sliding windows in a dataset with unequal activity bout durations (BREATHE), by producing more accurate and considerably more stable predictions. When implemented in a platform such as PRISMS, GGS should be able to identify short bursts of activity while still producing relatively smooth predictions. Identification of short activity bouts is particularly important for appropriately quantifying vigorous activity in children [27]. Noisy predictions from fixed-size sliding windows might need to be smoothed by pooling (ie, majority vote) for improved face validity of reported activity classifications and to avoid triggering excessive user notifications. Note that we presented our results using instantaneous predictions—to allow for comparisons across segmentation methods—that resulted

in slightly lower accuracy than previous studies presenting segment-level predictions. In practice, segment-level predictions are typically used.

Major differences between the HARuS and BREATHE datasets included not only activity bout duration (equal vs unequal), participant ages (adults vs children), and experimental protocol (tightly proscribed activities vs activities allowing for more natural movements) but also how the sensors were worn. This difference in wear location is likely the cause of the differences between the most important features in the Xgboost models. The axes of a device (smartwatch or smartphone) are typically labeled as x, denoting the side-to-side dimension; y, denoting the forward and backward dimension; and z, denoting the up and down dimension. Incorporating these axes with the wearing position of the 2 datasets, forward movement would correspond to signal along the x-axis for HARuS participants and the z-axis (slightly deviated to x-axis) for BREATHE participants. For both datasets, the most important features appeared to be related to forward motion (x-axis for the HARuS data and z-axis or combination of axes, ie, the norm for the BREATHE data) and the direction perpendicular to this motion (eg, mean values of the y-axis of the accelerometer, acc_y_mean, which had the third highest score in the HARuS data and the fourth highest score in the BREATHE data).

Limitations

In this study, the models were trained by *clip-independent* method. Time dependency is more obvious in datasets with temporal context, and many researches applied hidden Markov model (HMM) to such datasets as motion videos or images [28], body makers [29], and so on. For pure waist- or wrist-worn accelerometer or gyroscope meter, the signals do not have the strong time dependency as those temporal context data. Second, to compare the *time-dependent* methods, HMM should be tested with other analogic methods such as long short-term memory (LSTM), but not GGS. GGS is a way to clip the data such as the *fixed-length* sliding window. We can either apply *clip-independent* method as in this study or HMM or LSTM to test the time dependency among those clips.

The major weaknesses of the GGS approach are computational load and space requirements. To deploy GGS on streaming data, we would need to maintain a much larger cache memory of the latest received streaming data in comparison with the traditional fixed-length sliding window methods. GGS also requires time series of continuous features. However, sensor data (such as accelerometer and gyroscope) are typically quantitative, so this requirement is reasonable. Furthermore, missing values need to be either removed or interpolated. As for scalability, GGS has a runtime complexity of $O(KTn^3)$ in the normal mode and $O(Tn^3)$ in a *warm start* mode, in which the algorithm directly starts with a random set of K breaking points. Fixed-size sliding window approaches have better runtime complexity of O(n). Thus, the greedy heuristics needs to be improved in our future study. However, as the number of segments (K) is generally much smaller than the optimum number of fixed-size windows, GGS could largely save computational loads in the subsequent feature engineering, especially when tremendous feature to be extracted. Statistically, the GGS algorithm assumes that the multivariate time series can be described as independent samples from a multivariate Gaussian distribution within each segment. Time series data typically display autocorrelation, which would violate the independence assumption, especially when breaking points were not enough to separate the autocorrelated parts into different segments.

Conclusions

Identification of the break points that signify changes in physical activity plays an important role in quantifying HAR. In platforms such as PRISMS, HAR can be used not only to quantify the total duration of time in, for example, light, moderate, or vigorous activity but also to trigger user notifications or alerts or provide real-time feedback on activity. Our GGS-based approach shows great potential in variable activity bout duration scenarios and produces fewer variable predictions that should minimize unnecessary interactions with the user. However, computational and implementation limitations exist. Interesting future work will be focused on deploying GGS in real-time data streams and, more generally, finding heterogeneous segments when introducing additional sensor signals measured at different frequencies and on different scales (eg, sensors for physiological signals such as heart rate).

Acknowledgments

The authors would like to thank the participants of both studies and all the staff who made data collection possible, particularly Lisa Valencia. This study was supported by the National Institute of Biomedical Imaging and Bioengineering (grants U24EB021996, U54EB022002) and the National Institute of Environmental Health Sciences Center (grant 5P30ES07048).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Supplemental tables and figures.

[PDF File (Adobe PDF File), 1MB - mhealth_v7i2e11201_app1.pdf]

References

1. Haskell WL, Lee IM, Pate RR, Powell KE, Blair SN, Franklin BA, American College of Sports Medicine, American Heart Association. Physical activity and public health: updated recommendation for adults from the American College of Sports Medicine and the American Heart Association. Circulation 2007 Aug 28;116(9):1081-1093. [doi: 10.1161/CIRCULATIONAHA.107.185649] [Medline: 17671237]
2. Klepeis NE, Nelson WC, Ott WR, Robinson JP, Tsang AM, Switzer P, et al. The National Human Activity Pattern Survey (NHAPS): a resource for assessing exposure to environmental pollutants. J Expo Anal Environ Epidemiol 2001;11:231-252. [doi: 10.1038/sj.jea.7500165] [Medline: 11477521]
3. Perry B, Herrington W, Goldsack JC, Grandinetti CA, Vasisht KP, Landray MJ, et al. Use of mobile devices to measure outcomes in clinical research, 2010-2016: a systematic literature review. Digit Biomark 2018;2(1):11-30. [doi: 10.1159/000486347] [Medline: 29938250]
4. Murray CS, Poletti G, Kebadze T, Morris J, Woodcock A, Johnston SL, et al. Study of modifiable risk factors for asthma exacerbations: virus infection and allergen exposure increase the risk of asthma hospital admissions in children. Thorax 2006 May;61(5):376-382 [FREE Full text] [doi: 10.1136/thx.2005.042523] [Medline: 16384881]
5. Lucas SR, Platts-Mills TA. Physical activity and exercise in asthma: relevance to etiology and treatment. J Allergy Clin Immunol 2005 May;115(5):928-934. [doi: 10.1016/j.jaci.2005.01.033] [Medline: 15867847]
6. Cochrane LM, Clark CJ. Benefits and problems of a physical training programme for asthmatic patients. Thorax 1990 May;45(5):345-351 [FREE Full text] [Medline: 2116678]
7. Eijkemans M, Mommers M, Draaisma JM, Thijs C, Prins MH. Physical activity and asthma: a systematic review and meta-analysis. PLoS One 2012;7(12):e50775. [doi: 10.1371/journal.pone.0050775] [Medline: 23284646]
8. Robinson DM, Egglestone DM, Hill PM, Rea HH, Richards GN, Robinson SM. Effects of a physical conditioning programme on asthmatic patients. N Z Med J 1992 Jul 08;105(937):253-256. [Medline: 1620508]
9. Khan AM, Lee YK, Kim TS. Accelerometer signal-based human activity recognition using augmented autoregressive model coefficients and artificial neural nets. 2008 Aug 20 Presented at: 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society; August 20-25, 2008; Vancouver, BC, Canada p. 5173-5175. [doi: 10.1109/IEMBS.2008.4650379]
10. Hache G, Lemaire ED, Baddour N. Wearable mobility monitoring using a multimedia smartphone platform. IEEE Trans Instrum Meas 2011;60(9):3153-3161. [doi: 10.1109/TIM.2011.2122490]
11. Anguita D, Ghio A, Oneto L, Parra X, Reyes-Ortiz J. Human Activity Recognition on Smartphones Using a Multiclass Hardware-Friendly Support Vector Machine. 2012 Presented at: International Workshop on Ambient Assisted Living 2012; December 3-5, 2012; Vitoria-Gasteiz, Spain p. 216-223 URL: https://www.icephd.org/sites/default/files/IWAAL2012.pdf
12. Lara OD, Labrador MA. A survey on human activity recognition using wearable sensors. IEEE Commun Surv Tutor 2013;15(3):1192-1209. [doi: 10.1109/SURV.2012.110112.00192]
13. Torres-Huitzil C, Alvarez-Landero A. Accelerometer-Based Human Activity Recognition in Smartphones for Healthcare Services. In: Adibi S, editor. Mobile Health - A Technology Road Map. Cham, Switzerland: Springer International Publishing; 2015:147-169.
14. Keogh E, Chu S, Hart D, Pazzani M. Segmenting time series: a survey and novel approach. In: Last M, Kandel A, Bunke H, editors. Data Mining in Time Series Databases. Singapore: World Scientific Publishing Co Pte Ltd; 2004:1-21.
15. Booth NB, Smith AF. A Bayesian approach to retrospective identification of change-points. J Econom 1982;19(1):7-22. [doi: 10.1016/0304-4076(82)90048-3]
16. Galeano P, Wied D. Multiple break detection in the correlation structure of random variables. Comput Stat Data Anal 2014;76:262-282. [doi: 10.1016/j.csda.2013.02.031]
17. Nystrup P, Madsen H, Lindstrom E. Long memory of financial time series and hidden Markov models with time-varying parameters. J Forecast 2017;36(8):989-1002. [doi: 10.1002/for.2447]
18. Katz I, Crammer K. arXiv e-prints. Outlier-Robust Convex Segmentation URL: http://adsabs.harvard.edu/abs/2014arXiv1411.4503K [accessed 2018-12-03] [WebCite Cache ID 73utxwLSR]
19. Hallac D, Nystrup P, Boyd S. arXiv e-prints. Greedy Gaussian Segmentation of Multivariate Time Series URL: http://adsabs.harvard.edu/abs/2016arXiv161007435H [accessed 2018-12-03] [WebCite Cache ID 73uu9z7Al]
20. Anguita D, Ghio A, Oneto L, Parra X, Reyes-Ortiz JL. A Public Domain Dataset for Human Activity Recognition using Smartphones. In: ESANN 2013 proceedings. 2013 Presented at: 21st European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning; April 24-26, 2013; Belgium URL: https://www.elen.ucl.ac.be/Proceedings/esann/esannpdf/es2013-84.pdf
21. Hosseini A, Fazeli S, Vliet EV, Valencia L, Habre R, Sarrafzadeh M, et al. Children Activity Recognition: Challenges and Strategies. In: Conf Proc IEEE Eng Med Biol Soc. 2018 Jul Presented at: Annual International Conference of the IEEE Engineering in Medicine and Biology Society; July 17-21, 2018; Honolulu p. 4331-4334. [doi: 10.1109/EMBC.2018.8513320]
22. Hosseini A, Buonocore CM, Hashemzadeh S, Hojaiji H, Kalantarian H, Sideris C, et al. Feasibility of a secure wireless sensing smartwatch application for the self-management of pediatric asthma. Sensors (Basel) 2017;17(8):E1780. [doi: 10.3390/s17081780] [Medline: 28771168]

23. Hosseini A, Buonocore CM, Hashemzadeh S, Hojaiji H, Kalantarian H, Sideris C, et al. HIPAA compliant wireless sensing smartwatch application for the self-management of pediatric asthma. 2016 Jun 14 Presented at: 13th International Conference on Wearable and Implantable Body Sensor Networks (BSN); June 14-17, 2016; San Francisco p. 49-54. [doi: 10.1109/BSN.2016.7516231]
24. Selesnick IW, Burrus CS. Generalized digital Butterworth filter design. IEEE Trans Signal Process 1998;46(6):1688-1694. [doi: 10.1109/78.678493]
25. Kaur M, Singh B. Comparison of different approaches for removal of baseline wander from ECG signal. In: Proceedings of the International Conference & Workshop on Emerging Trends in Technology. 2011 Presented at: ICWET'11; February 25-26, 2011; Mumbai, Maharashtra, India. [doi: 10.1145/1980022.1980307]
26. Chen T, Guesrin C. arXiv e-prints. 2016. XGBoost: A Scalable Tree Boosting System URL: http://adsabs.harvard.edu/abs/2016arXiv160302754C [accessed 2018-12-01] [WebCite Cache ID 73uuGAveZ]
27. Rowlands A, Powell SM, Humphries R, Eston RG. The effect of accelerometer epoch on physical activity output measures. J Exerc Sci Fit 2006;4(1):52-58 [FREE Full text]
28. Qu Z, Lu T, Liu X, Wu Q, Wang M. A new method for human action recognition: Discrete HMM with improved LBG algorithm. 2015 Sep 25 Presented at: 9th International Conference on Anti-counterfeiting, Security, and Identification; September 25-27, 2015; Xiamen, China. [doi: 10.1109/ICASID.2015.7405672]
29. Gedat E, Fechner P, Fiebelkorn R, Vandenhouten R. Human action recognition with hidden Markov models and neural network derived poses. 2017 Sep 14 Presented at: 15th International Symposium on Intelligent Systems and Informatics; September 14-16, 2017; Subotica, Serbia. [doi: 10.1109/SISY.2017.8080544]

Abbreviations

GGS: greedy Gaussian segmentation
HAR: human activity recognition
HARuS: human activity recognition using smartphones
HMM: hidden Markov model
LSTM: long short-term memory
PRISMS: Pediatric Research using Integrated Sensor Monitoring Systems
PSD: power spectral density
SVM: support vector machine

Edited by M Focsa; submitted 01.06.18; peer-reviewed by M Albert, S Krishnan; comments to author 20.07.18; revised version received 30.09.18; accepted 14.11.18; published 07.02.19.

Please cite as:
Li K, Habre R, Deng H, Urman R, Morrison J, Gilliland FD, Ambite JL, Stripelis D, Chiang YY, Lin Y, Bui AAT, King C, Hosseini A, Vliet EV, Sarrafzadeh M, Eckel SP
Applying Multivariate Segmentation Methods to Human Activity Recognition From Wearable Sensors' Data
JMIR Mhealth Uhealth 2019;7(2):e11201
URL: http://mhealth.jmir.org/2019/2/e11201/
doi:10.2196/11201
PMID:30730297

Original Paper

Breast Cancer Survivors' Experiences With an Activity Tracker Integrated Into a Supervised Exercise Program: Qualitative Study

Hoi San Wu[1], BSc; Roxanne Gal[2], MSc; Niek C van Sleeuwen[3], BSc; Aarnout C Brombacher[3], BSc, MSc, PhD; Wijnand A IJsselsteijn[1], PhD; Anne M May[2], PhD; Evelyn M Monninkhof[2], PhD

[1]Human-Technology Interaction Group, Department of Industrial Engineering & Innovation Sciences, Eindhoven University of Technology, Eindhoven, Netherlands

[2]Julius Center of Health Sciences and Primary Care, University Medical Center Utrecht, University of Utrecht, Utrecht, Netherlands

[3]Systematic Change Group, Department of Industrial Design, Eindhoven University of Technology, Eindhoven, Netherlands

Corresponding Author:
Roxanne Gal, MSc
Julius Center of Health Sciences and Primary Care
University Medical Center Utrecht, University of Utrecht
STR 6.131
PO Box 85500
Utrecht, 3508 GA
Netherlands
Phone: 31 887569624
Email: R.Gal@umcutrecht.nl

Abstract

Background: There is growing evidence that physical activity is related to a better prognosis after a breast cancer diagnosis, whereas sedentary behavior is associated with worse outcomes. It is therefore important to stimulate physical activity and reduce sedentary time among patients with breast cancer. Activity trackers offer a new opportunity for interventions directed at stimulating physical activity behavior change.

Objective: This study aimed to explore the experience of patients with breast cancer who used an activity tracker in addition to a supervised exercise intervention in the randomized UMBRELLA Fit trial.

Methods: A total of 10 patients with breast cancer who completed cancer treatment participated in semistructured in-depth interviews about their experience with and suggestions for improvements for the Jawbone UP2 activity tracker.

Results: The activity tracker motivated women to be physically active and created more awareness of their (sedentary) lifestyles. The women indicated that the automatically generated advice (received via the Jawbone UP app) lacked individualization and was not applicable to their personal situations (ie, having been treated for cancer). Furthermore, women felt that the daily step goal was one-dimensional, and they preferred to incorporate other physical activity goals. The activity tracker's inability to measure strength exercises was a noted shortcoming. Finally, women valued personal feedback about the activity tracker from the physiotherapist.

Conclusions: Wearing an activity tracker raised lifestyle awareness in patients with breast cancer. The women also reported additional needs not addressed by the system. Potential improvements include a more realistic total daily physical activity representation, personalized advice, and personalized goals.

(JMIR Mhealth Uhealth 2019;7(2):e10820) doi:10.2196/10820

KEYWORDS

breast cancer; activity trackers; physical activity; sedentary behavior; qualitative research

Introduction

Breast cancer is the most common cancer type among women worldwide [1]. In 2015, 14,551 new cases were diagnosed in the Netherlands. The overall 5-year survival rate is now 87%, and this rate is still increasing because of better treatment and earlier detection [2-4]. Breast cancer treatment comes with short- and long-term side effects. During and after treatment, patients often complain of fatigue, reduced fitness, and impaired quality of life [5].

A reduction in physical activity levels and increase in sedentary time is also seen as side effects [6,7]. However, previous exercise interventions have produced physical and psychological health benefits in patients with breast cancer, such as lower fatigue levels, increased physical fitness, and improved emotional well-being and quality of life [8]. In addition, evidence suggests that physically active patients with breast cancer have a lower risk of recurrence of the disease and mortality [9]. Thus, it is important to stimulate physical activity and reduce sedentary time among patients with breast cancer.

Physical activity trackers are a popular tool used in health interventions. They can stimulate people to be more physically active and less sedentary as they provide insights into physical activity patterns, resulting in greater feelings of empowerment to set and stick to health goals [10,11]. The use of activity trackers is in line with the motto of the Quantified Self movement [12]: "self-knowledge through numbers." This movement reflects the fact that people increasingly integrate technology into their lives to gather personally relevant information. The underlying assumption about acquiring objective and quantitative data about one's behavior is that it leads to more accurate self-knowledge, which in turn empowers users to improve themselves. A recent literature review demonstrated that there is mounting evidence for this "self-improvement hypothesis," in which users gain usable insights from self-tracking data [13]. However, to date, few studies have used activity trackers in the rehabilitation of patients with cancer but are increasingly used for patients with cancer [14,15]. The first few studies explored the acceptability of activity trackers as a tool to stimulate physical activity and reduce sitting time and showed promising results [16-18]. However, more insight into the desires and needs of patients with breast cancer is needed to integrate an activity tracker in health care to increase physical activity and to achieve behavioral change.

The aim of this study was to explore experiences of patients with breast cancer with an activity tracker and its usage while participating in the intervention arm of the randomized controlled UMBRELLA Fit trial [19]. We used data collected from in-depth interviews to analyze the gap between the current and desired situations regarding the system's information feedback and the user interface from the patient's point of view.

Methods

Participants and Intervention

This study is embedded in the UMBRELLA Fit trial [19] which investigates the effects of an exercise intervention on the short- (6 months) and long-term (24 months) quality of life of inactive patients with breast cancer after primary treatment completion. The exercise intervention is a 12-week program consisting of supervised sessions (strength and endurance training) at a physiotherapy center twice a week. In addition, patients are encouraged to develop an active lifestyle on the 5 other days, defined as being moderate to highly physically active for at least 30 min a day, and reduce sedentary time. To help them achieve these goals, patients developed personal physical activity goals in consultation with a physiotherapist, they kept an activity log, and the patients were given an activity tracker. The path to achieving these activity goals was discussed every 2 weeks with the physiotherapist and when needed, goals were adapted.

In total, 60 women with breast cancer completed the UMBRELLA Fit exercise intervention. At the time this qualitative study was performed, 13 women had completed the exercise intervention and were asked to participate in this qualitative study. Of these, 10 women agreed to this study and were approached by telephone. The women were at least 12 months post diagnosis and had completed their primary breast cancer treatment (except hormonal treatment). On an average, the women were interviewed 90 days after completion of the trial. The UMBRELLA Fit study was approved by the Medical Ethics Committee of the University Medical Center Utrecht (UMCU).

The Activity Tracker

Initial requirements for the activity tracker were that it should possess the ability to track physical activity and inactivity, as well as the ability to synchronize this information to a database. In addition, the activity tracker should use an alert to remind the wearer of sedentary behavior that exceeded a certain amount of time.

The Jawbone UP2 (Jawbone, San Francisco, California, USA) met these requirements and was, therefore, used. This activity tracker is a wristband worn on the nondominant arm. It tracks steps, activity, and sleep and connects with a smartphone or tablet app. The Jawbone UP2 has been shown to have good reliability and validity in measuring the daily step count [20-22]. Using the accompanying "UP" app, users can monitor their step count, physical activity data, and calories burned. It is also possible to manually log workouts, track moods, and food intake, and to set certain goals for themselves. The UP app also incorporates a Smart Coach that provides personal, informative, and motivational messages and challenges based on what is measured. For example, *You are close to maintaining your 7-day, […] step average! Another […] steps, or a […] minute walk, will take you there,* and *Even if you don't walk to work, you can still find moments to step. Hop off the bus one stop early. Park your car at the far end of the lot. It all adds up.*

In the UMBRELLA Fit study, the Jawbone UP2 was especially used to signal sedentary behavior by using the idle alert as an inactivity reminder: the Jawbone UP2 vibrated when the women were inactive for at least 45 min. Also, the daily step goal in the app was set at 10,000 steps as a guideline. The study did not focus on sleep, mood, or dietary intake.

Before the intervention, all women received the Jawbone UP2, downloaded the UP app on their smartphone if they had one, and set up an account. They were instructed to wear the Jawbone UP2 the whole day and to synchronize the activity tracker daily with the app.

Interviews

The interviews were semistructured. We used a standard list of open questions in 5 categories to gain more insight into the

users' experiences with an activity tracker and its usage, added to a supervised exercise intervention (Textbox 1).

Textbox 1. Overview of the interview questions.

1. Jawbone UP2 activity tracker
   - How did you experience the Jawbone UP2 (added value, shortcomings)?
   - Was it clear how to use the Jawbone UP2?
   - Did the Jawbone UP2 meet your expectations?
2. Idle alert
   - How was your experience with the idle alert?
   - Did it create awareness about your sedentary behavior?
   - What did you think of the 45-minute time span that triggered the idle alert?
3. UP app and Smart Coach
   - How was your experience with the UP app (added value, shortcomings)?
   - Was it clear how to use the app?
   - Which function(s) did you use?
   - Did the app contribute to creating awareness about your inactivity and activity?
4. Exercise intervention
   - What is your opinion about the addition of the Jawbone UP2 to the exercise intervention?
   - Did wearing the Jawbone UP2 activity tracker and using the UP app change your behavior?
5. Suggestions or remarks
   - Do you have any additional suggestions or remarks regarding the Jawbone UP2?

Face-to-face interviews were performed by 2 researchers (HSW and NCvS), with the exception of 1 patient who used the activity tracker but not the app and was interviewed by phone. The interviews were conducted at a location of the patients' choice: in the hospital, in a lunchroom, or in their living room at home. All interviews were recorded using a smartphone. The interviews were transcribed verbatim and analyzed by 2 researchers (HSW and NCvS) following the guidelines of thematic analysis [23].

Measures

Daily step count during the intervention period was measured with the activity tracker and synchronized to the app. After the intervention period, data from the activity tracker were downloaded from the Jawbone website. Information on usage of the activity tracker was obtained through logs in which women registered the daily wearing time. Before and after the exercise intervention, patients performed a maximal cardiopulmonary exercise test on a cycle ergometer to determine peak oxygen uptake (VO_{2peak}). Finally, the level of physical activity was measured before and after the exercise intervention with the Short QUestionnaire to ASsess Health-enhancing physical activity (SQUASH) [24].

Results

Participants

After 10 interviews, data saturation was reached. The inter-rater reliability was calculated for 2 interviews using NVivo 11 (QRS International Pty Ltd, Melbourne, Australia). The Cohen kappa coefficient was 0.62, which is acceptable. The age of the women ranged from 33 to 64 years (median 57 [SD 8.8]).

Adherence to Wearing the Activity Tracker

Jawbone UP2 data were available from 8 of the 10 women. Woman no. 7 stopped participating after 3 weeks because of dissatisfaction with different elements of the study, including the Smart Coach. Similarly, woman no. 8 had no smartphone and, therefore, data synchronization was not possible.

Overall adherence to wearing the activity tracker ranged from 35% to 99% of the days, with 5 of the 8 women having an adherence of 89% or more (Table 1). The 3 other women wore the activity tracker on 35%, 42%, and 67% of the days, respectively. The reasons for not wearing the activity tracker were problems with charging, the tracker broke down, holiday, and flu. Woman no. 6 wore the activity tracker for less than 10 days and did not register compliance consequently. She explained that she lost interest in the training sessions of the study and, therefore, she felt that she did not need to wear the activity tracker anymore.

Table 1. Percentage of days with a daily step count above 10,000 and adherence to wearing the Jawbone UP2 during the intervention period.

Participant	Days worn, n (%)[a]	Reasons for not wearing
1	84 (89)	Forgot to wear (4 d[b]); did not wear (last 6 d)
2	77 (91)	Charger did not work (6 d)
3	40 (42)	Forgot to wear (4 d); on holiday (16 d); lost Jawbone UP2 (last 37 d)
4	64 (67)	Forgot to wear (3 d); band broke (last 28 d)
5	87 (99)	Perfect adherence
6	41 (35)	Did not wear (first 6 d); on holiday (9 d); flu (9 d); overall low adherence (51 d)
7[c]	—[d]	Dropout
8[e]	—	No Jawbone UP2 data available (no smartphone)
9	70 (95)	Forgot to wear after charging (3 d)
10	102 (97)	Did not wear on 3 d

[a]The intervention period was extended for some participants due to planned vacations, physical symptoms, family issues and time constraints.

[b]d: days.

[c]Woman no. 7 stopped participating in the trial after 3 weeks.

[d]Not applicable.

[e]Woman no. 8 used the Jawbone UP2, but had no smartphone and, therefore, synchronizing the data was not possible.

Change in Daily Step Count, Physical Activity Level, and Peak Oxygen Uptake

Overall, the mean number of steps per day was 8403 (SD 1994; Multimedia Appendix 1; Figure 1), with the highest step count at week 6 with an average 10,123 steps per day. On an average, patients reached on 30% (SD 20) of the days a step count of more than 10,000 steps. Moreover, 3 of the 8 patients had a mean daily step count of around 10,000, defined as the daily step goal, and 4 patients had an average daily step count between 7000 and 8000.

Mean change from baseline to postintervention in self-reported total physical activity level (including commuting activities, walking, cycling, and sport activities extracted from SQUASH) was plus 79 min per week (SD 123; Table 2), and mean change in VO_{2peak} was plus 1 ml/kg/min (SD 2).

Figure 1. Mean number of steps per day as measured with the Jawbone UP2.

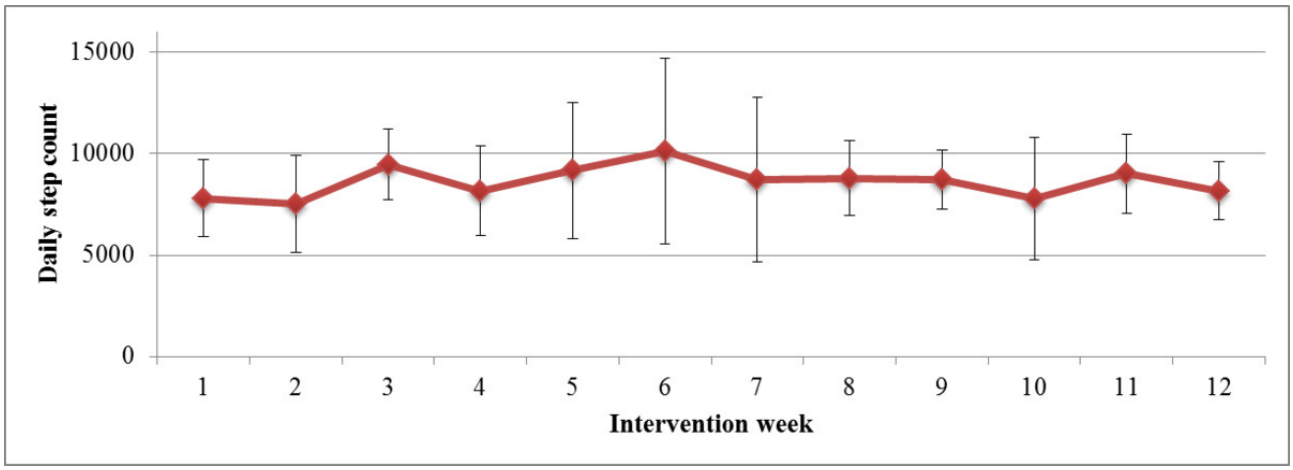


Table 2. Physical activity level and peak oxygen uptake (VO_{2peak}) at baseline and change during the intervention period.

Participant	Self-reported total minutes of activity; baseline, minutes per week[a]	Self-reported total minutes of activity; change, minutes per week[a]	VO_{2peak}; baseline, ml/kg/min	VO_{2peak}; change, ml/kg/min
1	60	+90	22	+2
2	360	−210	30	−1
3	330	+110	29	+3
4	30	+90	21	0
5	0	0	23	−2
6	0	+135	19	+4
7[b]	330	—[c]	29	—
8	0	+180	19	+3
9	160	+200	29	+1
10	0	+120	25	+1
Total, mean (SD)	127 (155)	+79 (123)	25 (4)	+1 (2)
Median	45	+110	24	+1

[a]Total minutes of activity including commuting activities, walking, cycling, and sport activities extracted from SQUASH.

[b]Woman no. 7 stopped participating in the trial after 3 weeks.

[c]Not applicable.

Textbox 2. Overview of the themes and subthemes.

1. The use of an activity tracker and accompanying app raises lifestyle awareness.
   - Activity tracker functions motivate (especially goals and idle alert).
   - More awareness of lifestyle: more physically active and less sedentary time.
2. Patients need personalized advice.
   - Generated advice (Smart Coach) is not applicable to personal situations.
   - Lack of personal advice.
3. Patients need a more realistic total daily physical activity representation.
   - Step goal was too one-dimensional.
   - Prefer the possibility of tracking other physical activity goals.
4. Patients need more integration between the intervention components of the study.
   - No feedback about the activity tracker from the physiotherapist.
   - The activity tracker does not adequately measure fitness activities during supervised exercise sessions.

Experiences With the Activity Tracker and Accompanying App

Following the thematic analysis, 4 themes emerged from the data (Textbox 2).

Theme 1: The Activity Tracker and Accompanying App Raises Lifestyle Awareness

All the women emphasized the important role the activity tracker played as a motivator and tool for gaining insights into their physical activity lifestyles. Woman no. 9 mentioned:

> *[I] became more aware, the switch is flipped: sometimes I sat for too long a period of time, so I should start moving again. So, it gave me insights.*

The main motivating feature of the activity tracker and app was the daily step goal. Almost all the women did something extra to achieve this goal, as woman no. 5 illustrates:

> *It became an obsession to reach the 10,000 steps.*

As an example of this motivation, woman no. 9 said:

> *You've got a goal of a certain number of steps a day and when you've almost reached it, you think: "I'll walk the dog tonight."*

Another motivator for raising awareness was the idle alert function. This function not only motivated the women to start moving before or at the alert, but also made them aware that they should avoid prolonged sitting periods. Woman no. 5 mentioned:

> *I tried to avoid the [idle alert] notification.*

This illustrates how the alert raised awareness to avoid long periods of sedentary time. Woman no. 10 also kept track of her sedentary time so she would move again before the alert vibrated. Woman no. 9 explained how the alert made her conscious of her sedentary behavior:

> *Oh, it [the activity tracker] is vibrating again, I have to move.*

Although the women took different approaches to dealing with the alert, it helped them interrupt their sedentary time, thus raising awareness about their lifestyle.

At the moment of the interview, most of the women were still more aware of their lifestyles, even though they no longer wore the activity tracker for some time. The activity tracker was an important component of the intervention in developing this awareness. Woman no. 9 explained:

> *It's something that can help you when you want to live healthier [...] It's a very good tool to start with. I'd say it increases your awareness and it works very well.*

When the women no longer had the activity trackers, they actively tried to maintain a more active and less sedentary lifestyle. Woman no. 8 gave an example:

> *Now, I more often think I have to get up after an hour, I have to walk.*

Woman no. 9 also pointed out:

> *The funny thing is that I still do it. If a colleague asks: "Shall I get you coffee?" I think: "No, I'll get it myself." You become aware that you do not move so much. That's what you discover in those three months [of the trial].*

Comparing her life before and after the trial, woman no. 5 said:

> *Previously, it was maybe 10 minutes, walking the dog, but now, I walk the dog 20 [minutes] or half an hour more or so, a few times a day.*

Theme 2: Patients Need Personalized Advice

Almost all the women emphasized a preference for more personalized advice than they received from the Smart Coach. They only viewed directly applicable messages as motivating. Personalized messages about the women's daily progress toward their individual goals were the most motivating, as exemplified by woman no. 5:

> *When, for example, in the evening, the device told me: "Well you've done a good job, but you still need a certain number of steps, something like 900 or 1,000 steps. Go for a short walk or something else." And then I went again, or I took the stairs ten times up and down, or I walked an extra time in my garden.*

In addition to the motivational aspect, the women enjoyed the personalization of these messages. The longer a woman used the activity tracker, the more personalized messages were generated, and the more she enjoyed them. Woman no. 9 said:

> *Yes, the Smart Coach, yes, very amusing. The longer you wear it, it figured out how you score, what you like [...]. Then it indicated things that made you think: "It gives me that little bit extra. Yes, I really like it."*

However, the Smart Coach did not always deliver personalized advice. Woman no. 7 explained that the Smart Coach was programmed with standard advice, focusing on healthy people:

> *If you don't reach the targets, this Smart Coach will not ask empathetically why it not works [limitations caused by cancer and its treatment]. The only thing he [the Smart Coach] says is "Don't be a fool and do it. [...] That gives a negative impression."*

She found this very annoying as her disease and its treatment had limited her physical condition and ability to achieve those goals. Nonpersonalized advice was either ignored or experienced as irritating, as woman no. 5 emphasized:

> *Sometimes it was advice, which I found useful, but sometimes I skipped or ignored it.*

Woman no. 9 pointed out that a certain capacity for placing the advice in perspective was needed to cope with the Smart Coach's messages that did not fit her personal situation. Some women even lowered their goals in the app, so the Smart Coach would not complain; this made them feel better.

Theme 3: Patients Need a More Realistic Total Daily Physical Activity Representation

Many women indicated that a disadvantage of the activity tracker was that it only registered steps; physical activities such as resistance exercise and cycling were not registered correctly. The women were trying to develop an active lifestyle of being moderate to highly physically active for at least 30 min a day, but the tracker merely measured their steps and not all types of physical activity. Woman no. 2 said:

> *Yes, those other activities apart from the steps weren't registered.*

Women no. 7 also pointed out that the activity tracker did not track activities other than walking:

> *This Smart Coach [...] doesn't add up activities, it only counts steps. For example, I cycled for three hours but I only reached 6000 steps. [...] At that point, it doesn't say: "You already cycled for three hours." Instead, it says: "You need 18 more minutes walking the stairs.*

The women mentioned 3 activities that they found important and thought should have been captured by the activity tracker and incorporated into their total physical activity representation. First, almost all of them indicated that the fitness workouts at the physiotherapist were inadequately measured. Second, all the women indicated that the bicycle was important as a means of physically active transportation. They were bothered that it was not adequately captured. Finally, swimming was not

registered by the activity tracker because it was not water resistant. Woman no. 3 mentioned:

> *I also found it a pity [...] because at that time I swam an hour and a half every week and I would have preferred to be able to wear it [the activity tracker].*

The women were not only bothered that their other activities were not tracked by the activity tracker, but also that they were not specifically shown in the app, as woman no. 1 described while discussing the disadvantages:

> *And I also can't find my fitness sessions in the app [...] you can't find cycling and [lifting] weights [in the app].*

Woman no. 5 also shared her frustration about this:

> *When I went to the physiotherapist for an hour, I had to register it [...] but it didn't show up [the registered activity].*

Theme 4: Patients Need More Integration Between the Intervention Components of the Study

Almost all the women indicated that they felt a lack of connection between the use of the activity tracker and the accompanying app, and their physiotherapy sessions. Woman no. 3 said:

> *I did nothing with the bracelet at the physiotherapist sessions and neither did they.*

The 2 components were closely related for the women, as the activity tracker measured activity outside their physiotherapy appointments, which was their checkup and feedback moments regarding their physical condition. Woman no. 3 explained:

> *I thought of it as a feedback moment [...] I'd have some questions about the activity tracker, but I couldn't ask them to the physiotherapist. It surprised me.*

Woman no. 1 noticed:

> *There were supervised sessions in the UMBRELLA Fit program, but the UP [app] was not integrated in what the physiotherapist did [during the sessions].*

As the physiotherapy sessions were an important aspect of the trial, the women expected the activity tracker to register activities during these sessions. Woman no. 3 said:

> *It'd be amazing if everything I did regarding physical activities was registered, including workouts at the physiotherapy sessions.*

Discussion

Principal Findings

A physically active lifestyle is known to be important for patients recovering from breast cancer, as it may ameliorate the negative side effects of treatment and help to limit some of the comorbidities (eg, heart disease, diabetes, or other cancers) [8,9,25]. This qualitative study investigated the experiences of patients with breast cancer who used an activity tracker and the accompanying smartphone app aimed at increasing physical activity and decreasing sedentary time. The activity tracker was added to a supervised exercise intervention for women who had completed their breast cancer treatment. Aside from logistical reasons for not using the activity tracker, adherence was high. Daily step count, self-reported physical activity level, and VO_{2peak} of participants slightly increased following the exercise intervention. As the activity tracker was part of the exercise intervention, it was not possible to draw conclusions on the isolated effect of the activity tracker and associated app, but a recent meta-analysis indicated that exercise interventions comprising activity trackers and smartphone apps were more effective than exercise interventions without activity trackers and smartphone apps [26].

From our interviews with 10 participants, 4 important themes emerged: (1) the use of an activity tracker and accompanying app raises lifestyle awareness, (2) patients need personalized advice, (3) patients need a more realistic total daily physical activity representation, and (4) patients need more integration between the intervention components of the study.

From the interviews, we found that the activity tracker and accompanying app functioned as a motivational tool and created more awareness of physical activity behavior and sedentary behavior. The women indicated that the daily step goal defined in the accompanying app was the most important motivator. Similar results were found in a comparable study of patients with breast cancer by Nguyen et al [17]. They found that goals presented in an achievable and easy-to-understand indicator of physical activity (eg, step goal presented in numbers) worked well. It helped the patients with breast cancer to be more aware of their physical activity levels and to incorporate physical activity into their daily routines. This is also supported by Wang et al [27], who reported that, in general, using a physical activity app facilitated more exercise. Furthermore, feedback concerning the progress toward their goal made the women more determined to stick to their daily activity goal which was also in line with Nelson et al [10].

The next theme that emerged from the interviews was a need for more tailored and personalized advice. The women indicated that the Smart Coach messages which were applicable to their activity levels were motivating and enjoyable. The advice given in a positive way encouraged the women to be more engaged in an active lifestyle. These kinds of positive encouragements were comparable with the rise in motivation to reach goals, when rewards were provided [28]. However, some of the messages generated by the Smart Coach did not fit the women's situations and were tailored to activity levels of the general population. As all the women had undergone cancer treatment and experienced disease- and treatment-related side effects as fatigue and a decreased physical fitness level, it was sometimes frustrating to read messages generated by a system that did not consider the side effects of the treatment. Our results suggested that using generic commercial trackers and their associated software for purposes of rehabilitation support may require additional design adaptations and references that are specific to the person and the limitations imposed by the disease and treatment. Women also mentioned the important role the physiotherapist played as a physical activity expert who was

very helpful in the process of becoming fit. Therefore, combining the physiotherapist's feedback with the immediate, quantitative feedback from the tracker in the app was seen as a potential improvement. In addition to messages, the physiotherapist could give personal advice about matters such as the women's goals and possible adjustments to them.

Other studies support this suggestion of combining activity tracker data with feedback from a medical professional [25,29,30]. Including input from a health professional could make interventions more effective. Nguyen et al [17] also suggested that support from peers with similar conditions may be helpful both in terms of practical advice (eg, what is a realistic goal to strive for) as well as psychosocial support, making the intervention more enjoyable and motivating. This is in line with recent insights related to personal tracking, which increasingly acknowledge the importance of data sharing and social interactions around data [31]. The third theme emerging from the interviews was the need for a more realistic total daily physical activity representation. As the activity tracker was not capable of accurately measuring forms of physical activity besides steps, these activities were also not part of the daily amount of physical activity shown in the app. The most important inadequately captured activity was cycling, which is a common means of transportation in the Netherlands. Dutch women cycle around 2.3 kilometers every day [32]. As cycling is so prominent in the Netherlands and contributes significantly to an active lifestyle, not registering this activity gives an inaccurate impression of a person's physical activity level. This may not be a problem in cultures where cycling is not that prominent. In addition, the women indicated that swimming and fitness sessions were important physical activities.

As the representation of total physical activity level is inaccurate, it is more difficult to monitor their path to a physically active lifestyle. This might lead to frustration and reduce motivation to continue using the activity tracker. Rosenburg et al [18] also reported inaccurate measuring of other activities apart from walking as a barrier to use an activity tracker in men with prostate cancer. Thus, additional design requirements for activity trackers in this context include a greater diversity in automated activity registration, as well as water resistance. This allows women to swim and bath with the tracker and have a more accurate activity representation that includes more activities. Activity trackers with these features are now widely available and should be considered for future studies.

The women also liked the feedback system (ie, Smart Coach) in the accompanying app. An opportunity for improvement could be to use an activity tracker with an integrated screen that gives immediate feedback, thus eliminating the need to go online to inspect the information in the app. However, this would be more expensive than the current system.

The last theme that emerged from the interviews was the need for more integration between the health professional and the activity tracker. As discussed above, a health professional can add value to the patient's intervention experience when provided with data from the activity tracker (eg, the health professional could help the patient determine an appropriate higher or lower step goal and give her specific tips for reaching the goal). In addition, the health professional could fulfill the role of the "Smart Coach" in which the positive qualitative feedback could be adjusted to individual patients.

Strengths and Limitations

This study had several strengths. First, it was a field study where participants used the activity tracker during a 12-week intervention program. This allowed them to get beyond first impressions and realistically use and evaluate the activity tracker in context. Second, through in-depth interviews, we gained insight into the use of, and experiences with, the activity tracker, the accompanying app, and its functions (eg, the idle alert and the Smart Coach) in contrast to reviews focused on the use of an activity tracker alone. This gave insight into the drivers and barriers to acceptance and into the use of activity trackers for patients with breast cancer. These interviews helped us identify several important requirements and improvements for next-generation trackers to be used in the context of patient recovery. Third, as this study was part of a larger medically supervised intervention, we were able to explore some of the strengths of combining activity tracker feedback with tailored and personalized feedback from medical professionals. This pertains to the larger issue of patient empowerment through self-tracking and the changing roles of patients and doctors, as digital technology increasingly democratizes through digital technology. Finally, because most activity trackers and smartphone apps have comparable functions (eg, step counting, goal setting and tracking, a daily report, an idle alert, and sleep tracking), our results also apply to other activity trackers apart from the Jawbone UP2.

A study of this kind also has limitations. First, the particularities of the activity tracker we used may not generalize to all trackers, and new technological developments in sensing, data processing, and interface design will allow current and new-generation trackers to overcome some of the issues we identified in this study. Second, although many of our insights are likely to hold true for populations other than recovering patients with breast cancer, our results also illustrate that the particularities of a disease and the effects (or side effects) of its treatment are important. This supports a general argument favoring personalized approaches through adaptive digital interfaces and personal coaching over a one-size-fits-all approach. By necessity, this also limits the generalizability of our findings to other specific patient populations. Finally, even though our study participants used the tracker for several weeks, it remains to be seen what the long-term health effects could be of using an activity tracker in comparison with other fitness-promoting programs. Also, the key factors contributing to long-term use and effectiveness of activity trackers are a topic of interest for future investigation.

Conclusions

This study explored experiences with an activity tracker and its usage in an exercise intervention among inactive patients with breast cancer. The interviews showed that an activity tracker raises awareness of a physically active lifestyle and sedentary behavior. Our study also showed the potential of using a wearable activity tracker to improve supportive care after primary treatment for breast cancer. However, there is a need

for a more realistic representation of the total daily physical activity, a more personalized advice that is tailored to their current situation after breast cancer treatment, and a better integration of the activity tracker into clinical practice. To optimize the use of an activity tracker in clinical practice, we suggest to base personalized advice on references that are specific to the patient population and to integrate the use of activity trackers and smartphone apps in rehabilitation programs, which requires more intensive guidance of a health care professional on usage and goal setting.

Acknowledgments

The authors would like to thank the patients for their participation, and the study team of UMBRELLA at the department of Radiotherapy, Imaging Division, and Julius Centre for Health Sciences and Primary Care of the UMCU for running the cohort.

This work was supported by a Veni grant from the Netherlands Organisation for Health Research and Development (ZonMw 016.156.050). For the UMBRELLA cohort, funding was provided by the department of Radiotherapy, Imaging Division of the UMCU, and the department of Clinical Epidemiology, Julius Center for Health Sciences and Primary Care, UMCU.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Mean number of steps per day per week during the intervention period.

[PDF File (Adobe PDF File), 259KB - mhealth_v7i2e10820_app1.pdf]

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2018 Nov;68(6):394-424 [FREE Full text] [doi: 10.3322/caac.21492] [Medline: 30207593]
2. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. Int J Cancer 2015 Mar 1;136(5):E359-E386 [FREE Full text] [doi: 10.1002/ijc.29210] [Medline: 25220842]
3. [Integral Cancer Center Netherlands]. [Dutch Cancer Registry] URL: http://www.cijfersoverkanker.nl [accessed 2017-10-10] [WebCite Cache ID 75XhUX70B]
4. Vondeling GT, Menezes GL, Dvortsin EP, Jansman FG, Konings IR, Postma MJ, et al. Burden of early, advanced and metastatic breast cancer in the Netherlands. BMC Cancer 2018 Dec 7;18(1):262 [FREE Full text] [doi: 10.1186/s12885-018-4158-3] [Medline: 29514651]
5. Servaes P, Verhagen C, Bleijenberg G. Fatigue in cancer patients during and after treatment: prevalence, correlates and interventions. Eur J Cancer 2002 Jan;38(1):27-43. [doi: 10.1016/S0959-8049(01)00332-X] [Medline: 11750837]
6. Phillips SM, Dodd KW, Steeves J, McClain J, Alfano CM, McAuley E. Physical activity and sedentary behavior in breast cancer survivors: new insight into activity patterns and potential intervention targets. Gynecol Oncol 2015 Aug;138(2):398-404 [FREE Full text] [doi: 10.1016/j.ygyno.2015.05.026] [Medline: 26026737]
7. Sabiston CM, Brunet J, Vallance JK, Meterissian S. Prospective examination of objectively assessed physical activity and sedentary time after breast cancer treatment: sitting on the crest of the teachable moment. Cancer Epidemiol Biomarkers Prev 2014 Jul;23(7):1324-1330 [FREE Full text] [doi: 10.1158/1055-9965.EPI-13-1179] [Medline: 24753546]
8. Fong DY, Ho JW, Hui BP, Lee AM, Macfarlane DJ, Leung SS, et al. Physical activity for cancer survivors: meta-analysis of randomised controlled trials. Br Med J 2012 Jan 31;344:e70. [doi: 10.1136/bmj.e70]
9. McTiernan A, Irwin M, Vongruenigen V. Weight, physical activity, diet, and prognosis in breast and gynecologic cancers. J Clin Oncol 2010 Sep 10;28(26):4074-4080 [FREE Full text] [doi: 10.1200/JCO.2010.27.9752] [Medline: 20644095]
10. Nelson EC, Verhagen T, Noordzij ML. Health empowerment through activity trackers: an empirical smart wristband study. Comput Human Behav 2016 Sep;62:364-374. [doi: 10.1016/j.chb.2016.03.065]
11. Poirier J, Bennett WL, Jerome GJ, Shah NG, Lazo M, Yeh H, et al. Effectiveness of an activity tracker- and internet-based adaptive walking program for adults: a randomized controlled trial. J Med Internet Res 2016;18(2):e34 [FREE Full text] [doi: 10.2196/jmir.5295] [Medline: 26860434]
12. Lee VR. What's Happening in the " Quantified Self " Movement? In: ICLS 2014 Proceedings. 2014 Presented at: International Conference of the Learning Sciences; June 23-27, 2014; Boulder, Colorado p. 1032-1036. [doi: 10.13140/2.1.1132.1126]
13. Kersten-van Dijk ET, Westerink JH, Beute F, IJsselsteijn WA. Personal informatics, self-insight, and behavior change: a critical review of current literature. Hum Comput Interact 2017 Jan 19;32(5-6):268-296. [doi: 10.1080/07370024.2016.1276456]

14. Hartman SJ, Nelson SH, Weiner LS. Patterns of Fitbit use and activity levels throughout a physical activity intervention: exploratory analysis from a randomized controlled trial. JMIR Mhealth Uhealth 2018 Feb 5;6(2):e29 [FREE Full text] [doi: 10.2196/mhealth.8503] [Medline: 29402761]
15. Valle CG, Deal AM, Tate DF. Preventing weight gain in African American breast cancer survivors using smart scales and activity trackers: a randomized controlled pilot study. J Cancer Surviv 2017 Feb;11(1):133-148. [doi: 10.1007/s11764-016-0571-2] [Medline: 27631874]
16. Lloyd GR, Oza S, Kozey-Keadle S, Pellegrini CA, Conroy DE, Penedo FJ, et al. Breast cancer survivors' beliefs and preferences regarding technology-supported sedentary behavior reduction interventions. AIMS Public Health 2016;3(3):592-614 [FREE Full text] [doi: 10.3934/publichealth.2016.3.592] [Medline: 29057279]
17. Nguyen NH, Hadgraft NT, Moore MM, Rosenberg DE, Lynch C, Reeves MM, et al. A qualitative evaluation of breast cancer survivors' acceptance of and preferences for consumer wearable technology activity trackers. Support Care Cancer 2017 Dec;25(11):3375-3384. [doi: 10.1007/s00520-017-3756-y] [Medline: 28540402]
18. Rosenberg D, Kadokura E, Bouldin E, Miyawaki C, Higano C, Hartzler A. Acceptability of Fitbit for physical activity tracking within clinical care among men with prostate cancer. 2016 Presented at: AMIA Annual Symposium; November 12-16, 2016; Chicago p. 1050-1059.
19. Gal R, Monninkhof EM, Groenwold RH, van Gils CH, van den Bongard DH, Peeters PH, et al. The effects of exercise on the quality of life of patients with breast cancer (the UMBRELLA Fit study): study protocol for a randomized controlled trial. Trials 2017 Oct 27;18(1):504 [FREE Full text] [doi: 10.1186/s13063-017-2252-5] [Medline: 29078800]
20. Ferguson S, Davis D, Browne J, Taylor J. Examining the validity and reliability of Antonovsky's sense of coherence scale in a population of pregnant Australian women. Eval Health Prof 2015 Jun;38(2):280-289. [doi: 10.1177/0163278715578558] [Medline: 25825376]
21. Gomersall SR, Ng N, Burton NW, Pavey TG, Gilson ND, Brown WJ. Estimating physical activity and sedentary behavior in a free-living context: a pragmatic comparison of consumer-based activity trackers and actigraph accelerometry. J Med Internet Res 2016 Sep 7;18(9):e239 [FREE Full text] [doi: 10.2196/jmir.5531] [Medline: 27604226]
22. Kooiman TJ, Dontje ML, Sprenger SR, Krijnen WP, van der Schans CP, de Groot M. Reliability and validity of ten consumer activity trackers. BMC Sports Sci Med Rehabil 2015;7:24 [FREE Full text] [doi: 10.1186/s13102-015-0018-5] [Medline: 26464801]
23. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
24. Wendel-Vos G, Schuit AJ, Saris WH, Kromhout D. Reproducibility and relative validity of the short questionnaire to assess health-enhancing physical activity. J Clin Epidemiol 2003 Dec;56(12):1163-1169. [doi: 10.1016/S0895-4356(03)00220-8] [Medline: 14680666]
25. Crowgey T, Peters KB, Hornsby WE, Lane A, McSherry F, Herndon JE, et al. Relationship between exercise behavior, cardiorespiratory fitness, and cognitive function in early breast cancer patients treated with doxorubicin-containing chemotherapy: a pilot study. Appl Physiol Nutr Metab 2014 Jun;39(6):724-729 [FREE Full text] [doi: 10.1139/apnm-2013-0380] [Medline: 24869976]
26. Gal R, May AM, van Overmeeren EJ, Simons M, Monninkhof EM. The effect of physical activity interventions comprising wearables and smartphone applications on physical activity: a systematic review and meta-analysis. Sports Med Open 2018 Sep 03;4(1):42 [FREE Full text] [doi: 10.1186/s40798-018-0157-9] [Medline: 30178072]
27. Wang Q, Egelandsdal B, Amdam GV, Almli VL, Oostindjer M. Diet and physical activity apps: perceived effectiveness by app users. JMIR Mhealth Uhealth 2016;4(2):e33 [FREE Full text] [doi: 10.2196/mhealth.5114] [Medline: 27056639]
28. Owen N, Sugiyama T, Eakin EE, Gardiner PA, Tremblay MS, Sallis JF. Adults' sedentary behavior determinants and interventions. Am J Prev Med 2011 Aug;41(2):189-196. [doi: 10.1016/j.amepre.2011.05.013] [Medline: 21767727]
29. Glasgow RE, Eakin EG, Fisher EB, Bacak SJ, Brownson RC. Physician advice and support for physical activity: results from a national survey. Am J Prev Med 2001 Oct;21(3):189-196. [doi: 10.1016/s0749-3797(01)00350-6] [Medline: 11567839]
30. Phillips SM, Conroy DE, Keadle SK, Pellegrini CA, Lloyd GR, Penedo FJ, et al. Breast cancer survivors' preferences for technology-supported exercise interventions. Support Care Cancer 2017 Dec;25(10):3243-3252 [FREE Full text] [doi: 10.1007/s00520-017-3735-3] [Medline: 28470368]
31. Kersten-van Dijk ET, IJsselsteijn W. Design beyond the numbersharing, comparing, storytelling and the need for a Quantified Us. Int Des Archit J 2016;2016(29):121-135.
32. CBS [Centraal Bureau voor de Statistiek] StatLine. [People mobility in the Netherlands; personal characteristics and modes of transport, region]. Netherlands: CBS StatLine URL: https://opendata.cbs.nl/statline/#/CBS/nl/dataset/83499ned/table?dl=111E1

Abbreviations

SQUASH: Short QUestionnaire to ASsess Health-enhancing physical activity
UMCU: University Medical Center Utrecht

VO2peak: peak oxygen uptake

Edited by G Eysenbach; submitted 20.04.18; peer-reviewed by G Jerome, B Arnoldussen, B Lynch; comments to author 07.08.18; revised version received 11.11.18; accepted 11.11.18; published 21.02.19.

Please cite as:
Wu HS, Gal R, van Sleeuwen NC, Brombacher AC, IJsselsteijn WA, May AM, Monninkhof EM
Breast Cancer Survivors' Experiences With an Activity Tracker Integrated Into a Supervised Exercise Program: Qualitative Study
JMIR Mhealth Uhealth 2019;7(2):e10820
URL: https://mhealth.jmir.org/2019/2/e10820/
doi:10.2196/10820
PMID:30789349

Corrigenda and Addenda

Multimedia Appendix Correction: mHealth Supportive Care Intervention for Parents of Children With Acute Lymphoblastic Leukemia: Quasi-Experimental Pre- and Postdesign Study

Jingting Wang[1], PhD; Doris Howell[2,3], PhD; Nanping Shen[4], PhD; Zhaohui Geng[1], PhD; Fulei Wu[1], PhD; Min Shen[5], BA; Xiaoyan Zhang[6], BA; Anwei Xie[7], MSc; Lin Wang[8], PhD; Changrong Yuan[9], PhD, FAAN

[1]School of Nursing, Second Military Medical University, Shanghai, China
[2]Department of Supportive Care, Princess Margaret Cancer Center, Toronto, ON, Canada
[3]Faculty of Nursing, University of Toronto, Toronto, ON, Canada
[4]Department of Nursing, Shanghai Children's Medical Center, Shanghai, China
[5]Department of Hematology, Children's Hospital of Soochow University, Suzhou, China
[6]Department of Hematology and Oncology, Shanghai Children's Medical Center, Shanghai, China
[7]Department of Infectious Disease, Children's Hospital of Soochow University, Suzhou, China
[8]United States Census Bureau, Washington, DC, United States
[9]School of Nursing, Fudan University, Shanghai, China

Corresponding Author:
Changrong Yuan, PhD, FAAN
School of Nursing
Fudan University
Room 403, Building 8
305 Fenglin Road
Shanghai, 200032
China
Phone: 86 02164431803
Fax: 86 02164431803
Email: yuancr@fudan.edu.cn

Related Article:

Correction of: http://mhealth.jmir.org/2018/11/e195/

(JMIR Mhealth Uhealth 2019;7(2):e13159) doi:10.2196/13159

The authors of mHealth Supportive Care Intervention for Parents of Children With Acute Lymphoblastic Leukemia: Quasi-Experimental Pre- and Postdesign Study (JMIR Mhealth Uhealth 2018;6(11):e195) noticed problems with all three Multimedia Appendices which were published alongside the article as supplementary files. The link for Multimedia Appendix 1 was broken and would not allow the file to be downloaded, while the contents of Multimedia Appendices 2 and 3 were incorrect. All three Multimedia Appendices have been replaced with the appropriate files and the links have been confirmed as functional.

The correction will appear in the online version of the paper on the JMIR website on February 20, 2019, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article also has been resubmitted to those repositories.

Multimedia Appendix 1

Participants' baseline sociodemographic characteristics stratified by the study group.

[PDF File (Adobe PDF File), 134KB - mhealth_v7i2e13159_app1.pdf]

Multimedia Appendix 2

Comparison of outcomes' changes in the intervention group (n=49), in the observation group (n=43), and D value between the 2 groups.

[PDF File (Adobe PDF File), 93KB - mhealth_v7i2e13159_app2.pdf]

Multimedia Appendix 3

The qualitative results from the interviews in the intervention group (n=11).

[PDF File (Adobe PDF File), 97KB - mhealth_v7i2e13159_app3.pdf]

Edited by G Eysenbach; submitted 16.12.18; this is a non–peer-reviewed article;accepted 20.12.18; published 20.02.19.

Please cite as:
Wang J, Howell D, Shen N, Geng Z, Wu F, Shen M, Zhang X, Xie A, Wang L, Yuan C
Multimedia Appendix Correction: mHealth Supportive Care Intervention for Parents of Children With Acute Lymphoblastic Leukemia: Quasi-Experimental Pre- and Postdesign Study
JMIR Mhealth Uhealth 2019;7(2):e13159
URL: http://mhealth.jmir.org/2019/2/e13159/
doi:10.2196/13159
PMID:30785403

Original Paper

Middle-Aged Men With HIV Have Diminished Accelerometry-Based Activity Profiles Despite Similar Lab-Measured Gait Speed: Pilot Study

Timothy M Hale[1,2], PhD; Viola Guardigni[2,3,4], MD; Eva Roitmann[5], MSc; Matthieu Vegreville[5]; Brooke Brawley[4], MPA; Erin Woodbury[4], MS; Thomas W Storer[2,4], PhD; Paul E Sax[2,4], MD; Monty Montano[2,4], PhD

[1]Partners HealthCare, Pivot Labs, Boston, MA, United States

[2]Harvard Medical School, Boston, MA, United States

[3]S.Orsola-Malpighi Hospital, Department of Medical and Surgical Sciences-DIMEC, University of Bologna, Bologna, Italy

[4]Brigham & Women's Hospital, Boston, MA, United States

[5]Withings, Paris, France

Corresponding Author:
Monty Montano, PhD
Brigham & Women's Hospital
221 Longwood Avenue
Boston, MA, 02115
United States
Phone: 1 6175259046
Email: mmontano@bwh.harvard.edu

Abstract

Background: People aging with HIV are living with increased risk for functional decline compared with uninfected adults of the same age. Early preclinical changes in biomarkers in middle-aged individuals at risk for mobility and functional decline are needed.

Objective: This pilot study aims to compare measures of free-living activity with lab-based measures. In addition, we aim to examine differences in the activity level and patterns by HIV status.

Methods: Forty-six men (23 HIV+, 23 HIV−) currently in the MATCH (Muscle and Aging Treated Chronic HIV) cohort study wore a consumer-grade wristband accelerometer continuously for 3 weeks. We used free-living activity to calculate the gait speed and time spent at different activity intensities. Accelerometer data were compared with lab-based gait speed using the 6-minute walk test (6-MWT). Plasma biomarkers were measured and biobehavioral questionnaires were administered.

Results: HIV+ men more often lived alone (P=.02), reported more pain (P=.02), and fatigue (P=.048). In addition, HIV+ men had lower blood CD4/CD8 ratios (P<.001) and higher Veterans Aging Cohort Study Index scores (P=.04) and T-cell activation (P<.001) but did not differ in levels of inflammation (P=.30) or testosterone (P=.83). For all participants, accelerometer-based gait speed was significantly lower than the lab-based 6-MWT gait speed (P<.001). Moreover, accelerometer-based gait speed was significantly lower in HIV+ participants (P=.04) despite the absence of differences in the lab-based 6-MWT (P=.39). HIV+ participants spent more time in the lowest quartile of activity compared with uninfected (P=.01), who spent more time in the middle quartiles of activity (P=.02).

Conclusions: Accelerometer-based assessment of gait speed and activity patterns are lower for asymptomatic men living with HIV compared with uninfected controls and may be useful as preclinical digital biomarkers that precede differences captured in lab-based measures.

(JMIR Mhealth Uhealth 2019;7(2):e11190) doi:10.2196/11190

KEYWORDS

aging; digital biomarker; gait speed; HIV

Introduction

Survival for people aging with HIV (PAWH) is now approaching the lifespan of uninfected individuals, in part, owing to the success of effective antiretroviral therapy [1]. In developed countries, more than half of the PAWH are over 50 years old and may be at greater risk for age-related comorbidities, including a decline in functional capacity [2]. Why these age-related comorbidities more often occur at younger ages in PAWH compared with uninfected (HIV−) individuals is presently unclear [3].

The functional decline in PAWH increases their risk for premature frailty and early mortality [4,5]. In multiple studies, slower walking speed has been consistently associated with a decline in mobility and performance of activities of daily living [6-8], and faster decline in walking speed is associated with increased mortality [9]. An emerging challenge is that while current assessments of functional capacity and mobility limitations in PAWH aged <65 years have been effective in identifying impairment [5,10], these assessment tools were historically developed in the context of geriatric population studies of adults aged >65 years [7] and may not always be suitable for assessing middle-aged individuals without overt mobility limitations [11]. Therefore, there is an unmet need for additional diagnostic tools sensitive to preclinical risk factors for the functional decline that occurs in middle-aged (ie, 50-65 years) persons [12]. One such tool might be physical activity monitoring. People living with HIV tend to be sedentary and less fit [13,14]. In this study, we posited that the objective assessment of habitual physical activity might provide a more sensitive way to identify preclinical changes in middle-aged HIV+ men.

Wearable accelerometers are now frequently used in research studies to measure free-living physical activity and are useful in providing an objective measurement of low-intensity activity [15,16]. The use of accelerometry in the function assessment of PAWH is limited [17-21]. The use of consumer-grade accelerometers has become cost-effective [22]. Although technical and methodological limits remain [16,22], many consumer-grade accelerometers display strong validity when compared with lab-based measurements and research-grade accelerometers [23,24].

This study included participants in the MATCH (Muscle and Aging in Treated Chronic HIV) cohort of middle-aged men and women living with HIV and uninfected control participants of similar age [25]. Overall, HIV+ participants in the MATCH cohort are relatively healthy but displayed increased levels of inflammation and immune activation compared with their uninfected counterparts and displayed modest but significant subclinical deficits in lab-based physical function [25]. However, when men and women were analyzed separately, we noticed that performance in the lab-based 6-minute walk test (6-MWT) in men did not reach significance despite reporting greater fatigue [25]. We hypothesized that measuring physical activity in daily life using consumer-grade accelerometers might provide additional information, potentially detecting subclinical deficits that do not as yet reach significance in the lab-based assessment of gait speed. Therefore, this pilot study aims to measure and compare walking speed and activity patterns between HIV+ and HIV− men in the MATCH cohort using a consumer-grade accelerometer. Accelerometer data were used to measure steps taken per minute and to estimate the gait speed and classify patterns of activity.

Methods

Study Population

Participants in this study (MPACT: MATCH Physical ACtivity Tracker) consisted of HIV+ and HIV− men aged 50-65 years who were enrolled in the MATCH study [25] and who consented to this study. The MATCH cohort is a longitudinal observational study of middle-aged people living with HIV on effective ART, along with age-matched uninfected controls, all living in the Boston metropolitan area in Massachusetts [25]. At the first visit, the study protocol and procedures were explained and consent obtained. After consent, demographic information was collected, and questionnaires were administered to assess health-related quality of life (HRQoL; Patient Reported Outcome Measurements Information System [PROMIS-29]), disability (Pepper Assessment Tool for Disability [PAT-D]), and depression (Center for Epidemiological Studies Depression [CES-D]). All study procedures were approved by the Partners Human Research Committee institutional review board.

Questionnaires

Patient Reported Outcome Measurements Information System-29

PROMIS-29 assesses general HRQoL on 7 dimensions, including depression, anxiety, physical function, pain, fatigue, sleep, and social engagement with each of the 7 domains having 4 questions and a single question assessing pain intensity [26]. For example, in the pain domain, a sample question is, "*How much did pain interfere with your day to day activities?*" Responses include a range of 5 choices (eg, not at all, very much). The pain intensity is assessed by, *In the past 7 days, how would you rate your pain on average?* Responses range from 0 "No pain" to 10 "Worst imaginable pain." Scores are negatively associated with HRQoL.

Pepper Assessment Tool for Disability

The PAT-D is a 19-item questionnaire used to assess whether functional limitations impact disability across 3 domains—mobility, activities of daily living, and instrumental activities of daily living [27]. Responses are recorded on a 5-point Likert scale. An example question is, "*How much difficulty walking one block?*" Responses include a range of 7 choices (eg, no difficulty, a lot of difficulty).

Center for Epidemiological Studies Depression

The CES-D scale measures clinical depression [28] and uses a 20-item score to assess clinically significant depression [29]. An example question is, "*How often over the past week were you depressed?*" Responses include a range of 4 choices (eg, rarely, most or all the time).

Functional Assessment of Chronic Illness Therapy-Fatigue

Fatigue was assessed using the 13-item Functional Assessment of Chronic Illness Therapy-Fatigue (FACIT-F) subscale [30]. An example question is, "*How often do you feel fatigued?*" Responses include a range of 5 choices (eg, not at all, very much).

Veterans Aging Cohort Study Index Score and Biomarkers

The Veterans Aging Cohort Study Index calculation includes age, CD4 (cluster of differentiation 4 T cells), HIV RNA, Hb, creatinine, FIB-4, and HCV [31,32]. Blood biomarkers (CD4 T-cell count, CD4/CD8 T-cell ratio), an inflammatory composite score of circulating biomarkers for inflammation (ie, sCD163, sCD14, C-reactive protein, and interleukin-6 expression) and a T-cell immune activation composite score (ie, CD8, CD38, and HLA-DR expression) were evaluated in this study sample similar to what is described for the larger cohort [25]. Total testosterone (TT) and serum hormone-binding globulin (SHBG) were measured in the Brigham Research Assay Core facility using liquid chromatographic-tandem mass spectrometry (LC-MS/MS). SHBG was measured using radioimmunoassay. Free testosterone (FT) was calculated from TT and SHBG [33].

Accelerometer Data Collection and Measures

The accelerometer used in this study was the Withings Pulse Ox, a consumer-grade triaxial accelerometer with published validity for step counts when compared with research-grade accelerometers [23,24]. Participants were instructed to wear the accelerometer on their nondominant wrist 24 hours per day, 7 days per week for 3 consecutive weeks. Each individual in this study returned to the study center weekly for study personnel to recharge the Pulse Ox battery and upload data using a randomly assigned and anonymous study account. Data were collected in 1-minute intervals when any step activity was detected by the accelerometer, including low-intensity levels of activity that are not typically analyzed or accessible using consumer-grade trackers. In 13 cases, the tracker was lost or malfunctioned during the study period and was replaced to continue with one additional week of data collection. In 10 cases, data were collected for <3 weeks, primarily because of participants failure to wear the device. In total, 3 weeks of data were collected from 36 participants, 2 weeks from 6 participants, and 1 week of data from 4 participants.

In addition to step counts, accelerometer data were used to calculate gait speed, metabolic equivalents (METs), and relative activity intensity based on each participant's activity distribution throughout the day. The mean gait speed in meters/second was based on the number of steps per minute and a formula using participant's height in meters: (steps×height×0.414)/60) [34]. METs were calculated using a formula developed by Withings: 1+1/2×speed (km/h)+0.086×speed (km/h)×slope (in %) (Eva Roitmann, Withings, email communication; March 10, 2017 and March 28, 2018).

The standardized 6-MWT [11] was administered at the study site by the same trained and experienced exercise physiologist as in this study. A comparison test was conducted between both measures, in which participants wore an activity tracker (Pulse Ox) to determine the gait speed using the formula indicated above. The simultaneous measurement of gait speed using the lab-based assessment 6-MWT [calculated by the distance walked (m) in 6 minutes] and accelerometry indicated that the gait speed differed on average by <2%, similar to results in prior studies ([23,24] and Multimedia Appendix 1).

Accelerometer-derived patterns of physical activity were defined as the percent of the time a participant spends at one of 3 levels (ie, low, moderate, or intense) of their maximum observed intensity or gait speed. Two measures were created, a weekly measure using the relative maximum intensity observed for each participant during each week, and a summary measure that used the maximum observed intensity for a participant across 3 weeks of activity. The total activity for each individual was divided into 3 groups, with low activity defined as the lowest quartile (ie, 0%-25%, Q1), moderate activity as the middle 2 quartiles (ie, 25%-75%, Q2+Q3), and vigorous (intense) activity as the highest quartile (ie, 75%, Q4) for each participant.

Statistical Analysis

Continuous and categorical variables were expressed as means (SD) or percentages. Differences in demographics, plasma biomarkers, and self-reported measures between HIV+ and HIV− men were analyzed by logistic regression and Student's *t* tests. Differences by HIV status in accelerometer-derived measures of gait speed, METs, and activity patterns were hypothesized to be lower in HIV+ men and, therefore, conducted using a one-sided *t* test. One-sided *t* tests were used to analyze differences between groups (HIV+ vs HIV−) in lab-based physical function assessments of gait speed (6-MWT) and METs collected within the 16 months prior to this substudy [25]. Whether the HIV status and other markers were predictive of physical activity (ie, as gait speed assessment in the lab, free-living gait speed based on activity trackers, and lowest quartile of activity based on activity trackers) was assessed using univariate and multivariate linear regression. To evaluate what factors might be significant predictors of gait speed, the following variables were evaluated in univariate analysis: body mass index (BMI), HIV status, testosterone levels, and CD4/CD8 ratios. Multivariate models were used to examine BMI and HIV status as independent predictors of activity, and the significance of these predictors after adjusting for race, smoking, and living alone. An alpha of <.05 was used to indicate statistical significance.

Results

Characteristics of Study Participants

Forty-six men participating in the MATCH cohort [25] consented to participate in this substudy. Tables 1 and 2 present the characteristics of the study participants. Participants were aged between 50 and 65 years and included 23 asymptomatic, generally healthy HIV+ participants on effective antiretroviral therapy with undetectable viremia (<50 copies/mL) and 23 HIV− (confirmed by HIV serology) participants, all living in the Boston area. HIV+ participants were more often nonwhite (P=.001) and more often lived alone (P=.02). We observed no significant differences in education or working status. BMI and

smoking behavior approach statistical significance with HIV+ men having a lower BMI compared with HIV− men (*P*=.05) and elevated smoking behavior (*P*=.51). Blood biomarker profiles indicated that when compared with HIV−, HIV+ participants had significantly reduced CD4/CD8 ratios (*P*<.001), higher Veterans Aging Cohort Study scores (*P*=.04), and elevated immune activation (*P*<.001) but did not differ in levels of inflammation (*P*=.30), TT (*P*=.83), FT (*P*=.72), or SHBG (*P*=.29). Biobehavioral questionnaires were administered to obtain self-reported measures and indicated that when compared with uninfected HIV− participants, HIV+ participants reported more fatigue (FACIT-F: *P*=.048) and pain intensity (PROMIS-29: *P*=.02); however, there were no self-reported differences in alcohol abuse (*P*=.79), depression (CES-D: *P*=.74), or disability (PAT-D: *P*=.39).

Table 1. The baseline characteristics of the study participants (continuous variables).

Continuous variables	HIV+ (n=23), mean (SD)	HIV− (n=23), mean (SD)	*P* value[a]
Age (years)	58.1 (3.8)	59.2 (4.5)	.38
Body mass index (kg/m^2)	26.9 (3.8)	29.9 (5.8)	.05
CD4/CD8 ratio	0.99 (0.65)	2.80 (1.56)	<.001
Veterans Aging Cohort Study Index score	22.39 (8.48)	17.70 (6.49)	.04
Inflammation composite score[b]	1.09 (1.00)	0.78 (0.95)	.30
T-cell immune activation composite score[c]	1.43 (1.12)	0.09 (0.29)	<.001
Total testosterone[d]	2.82 (0.53)	2.78 (0.48)	.83
Free testosterone[d]	2.58 (0.56)	2.63 (0.43)	.72
Serum hormone-binding globulin[d]	4.00 (0.51)	3.84 (0.53)	.29
Fatigue (Functional Assessment of Chronic Illness Therapy-Fatigue)	41.91 (9.69)	46.61 (5.17)	.048
Pain intensity (Patient Reported Outcome Measurements Information System)[e]	3.22 (2.70)	1.52 (1.75)	.02
Disability (Pepper Assessment Tool for Disability)	1.38 (0.56)	1.26 (0.42)	.39

[a]*P* values were determined by 2-sided *t* tests.

[b]Composite score of sCD163, sCD14, CRP, and interleukin-6 expression.

[c]Composite score of CD8, CD38, and HLA-DR expression.

[d]Values are natural log-transformed.

[e]Scale 0 to 10, in the past 7 days.

Table 2. The baseline characteristics of the study participants (categorical variables).

Categorical variables	HIV+ (n=23), n (%)	HIV− (n=23), n (%)	*P* value[a]	Odds ratio (95% CI)
White	8 (35)	22 (96)	.001	0.02 (.00-.21)
Living alone	17 (77)[b]	10 (43)	.02	4.42 (1.21-16.12)
College degree	9 (41)[b]	13 (57)	.30	0.53 (.16-1.74)
Working	9 (39)	10 (43)	.77	0.84 (.26-2.71)
≥100 cigarettes lifetime	16 (73)[b]	10 (43)	.05	3.47 (.99-.12.09)
Alcohol abuse[c]	5 (23)[b]	6 (26)	.79	0.83 (.21-3.26)
Self-reported depression (Center for Epidemiological Studies Depression)[d]	7 (30)	6 (26)	.74	1.24 (.34-4.49)

[a]*P* values were determined by logistic regression with HIV status as the outcome.

[b]n=22.

[c]Defined as ≥3 drinks per day.

[d]CES-D score ≥16.

Physical Function Assessment

Laboratory-based physical function assessment was recently described in MATCH participants [25], and those data were used to evaluate the subset of individuals in this study. In the subset of individuals in this study, no significant differences were detected between HIV+ and HIV− men in gait speed using the 6-MWT (*P*=.39), nor in predicted METs (*P*=.15; Table 3) [13]. However, all participants, both HIV+ and HIV−, displayed a significantly reduced accelerometer-based (aka, free-living or volitional) gait speed when compared with the lab-based 6-MWT measurements (0.44 [SD 0.13] m/s vs 1.52 [SD 0.27] m/s; *P*<.001). Furthermore, HIV+ participants displayed significantly slower free-living gait speeds (0.41 [SD 0.11] m/s vs 0.48 [SD 0.14] m/s; *P*=.04) and METs (1.03 [SD 0.31] vs 1.24 [SD 0.42]; *P*=.03) compared with HIV− participants under free-living conditions. The MET difference is equivalent to about 300 Kcal/day (based on 16 hours of waking time).

Personalized Physical Activity Distribution

To evaluate patterns of daily activity, HIV+ participants were compared with HIV− participants based on the time spent as a percentage of the total activity for each participant. As shown in Figure 1, HIV+ participants spent significantly more of their time in the lowest quartile of activity compared with HIV− participants, during both the first week (*P*=.01) and all 3 weeks combined (*P*=.01; Table 3; Figure 1). By contrast, HIV− participants spent more of their time in the middle quartiles of activity compared with HIV+ participants, in the first week (*P*=.04) and again in all 3 weeks combined (*P*=.02; Table 3; Figure 1).

Tables 4 and 5 present the results from linear regression models. In univariate linear regression analyses, the BMI was the only significant predictor of lab-based gait speed (*P*=.04), whereas HIV status (*P*=.02), total and FT (*P*=.04), and the CD4/CD8 ratio (*P*=.02) were significant predictors of accelerometer-based metrics. In a multivariate regression model, the BMI and HIV status were independent predictors of accelerometer but not lab-based gait speed (Model 1). However, when adjusted for race, smoking, and living alone in Model 2, only BMI was an independent predictor of the accelerometer and lab-based gait speed, likely because of the small sample size of the study.

Table 3. Lab-based and accelerometer-based measurements of physical function and activity patterns.

Measurements	HIV+ (n=23)	HIV− (n=23)	*P* value[a]
Lab-based assessments, mean (SD)			
6-minute walk test gait speed (m/s)	1.51 (0.23)	1.53 (0.31)	.39
Predicted metabolic equivalents[b]	2.93 (0.30)	3.04 (0.41)	.15
Accelerometer assessments, mean (SD)			
Gait speed (m/s)[c]	0.41 (0.11)	0.48 (0.14)	.04
Predicted metabolic equivalents[c]	1.03 (0.31)	1.24 (0.42)	.03
Activity patterns[d], n (%)			
Week 1: Q1 (Low, %)	67 (12)	58 (11)	.01
Week 1: Q2-Q3 (Moderate, %)	22 (7)	26 (10)	.04
Week 1: Q4 (Intense, %)	11 (9)	15 (11)	.11
Week 1,2,3: Q1 (Low, %)	68 (13)	60 (10)	.01
Week 1,2,3: Q2-Q3 (Moderate, %)	23 (7)	28 (10)	.02
Week 1,2,3: Q4 (Intense, %)	9 (9)	12 (12)	.19

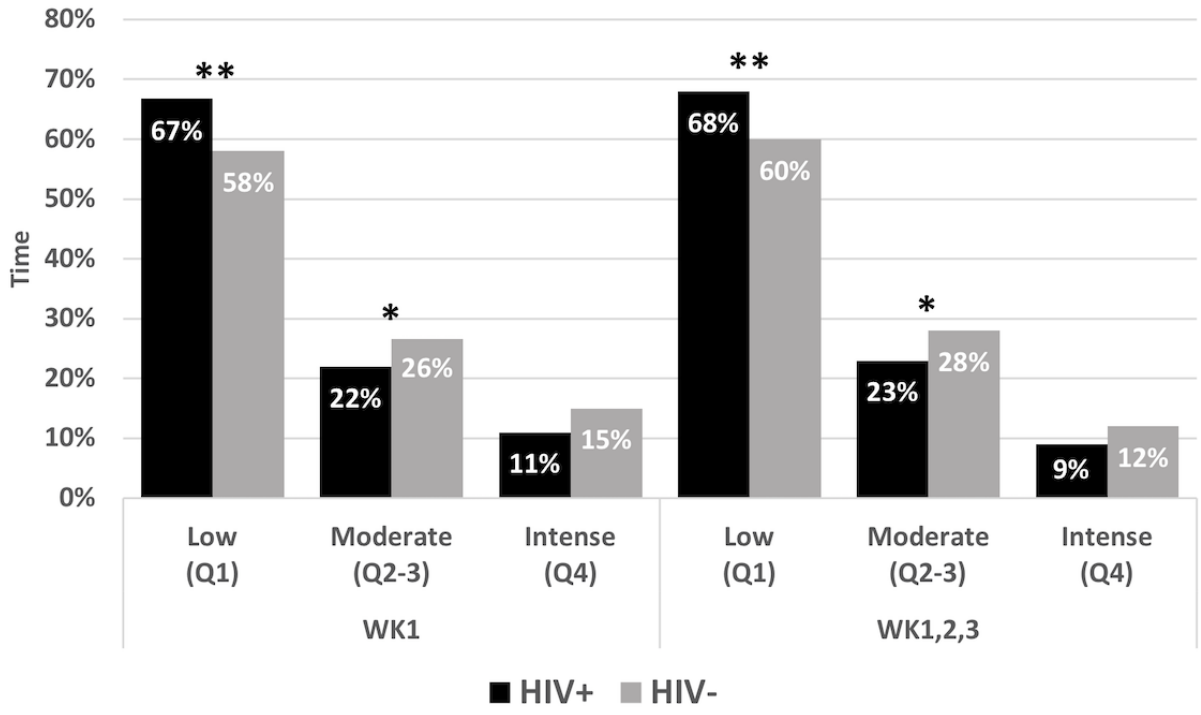
[a]*P* values were determined by one-sided *t* tests.

[a]Adjusted for height.

[c]Based on the following formula calculation for gait speed: ([steps×height×0.414]/60) [34]] and METs: 1+1/2×speed (km/h)+0.086×speed (km/h) ×slope (in %).

[d]Percent time spent in lowest (Q1), moderate (Q2-Q3), intense (Q4) quartiles of activity, as assessed using the maximum intensity from each week or from the overall period (21 days).

Figure 1. Comparison of free-living activity patterns between men living with HIV (HIV+) and uninfected men (HIV–), represented as mean percent time spent at different activity levels.



$^{*}P < 0.05$, $^{**}P < 0.01$

Table 4. Univariate linear regression models for activity outcomes.

Predictors[a]	Lab-based outcomes		Accelerometer-based outcomes			
	Gait speed		Gait speed		Low activity	
	Coefficient (95% CI)	*P* value	Coefficient (95% CI)	*P* value	Coefficient (95% CI)	*P* value
HIV status	−0.01 (−0.11 to 0.10)	.91	−0.15 (−0.34 to 0.03)	.10	0.16 (0.02 to 0.29)	.02
Body mass index	−0.32 (−0.63 to −0.02)	.04	−0.44 (−0.98 to 0.097)	.11	−0.06 (−0.50 to 0.38)	.78
Total testosterone	0.06 (−0.05 to 0.18)	.26	0.21 (0.02 to 0.40)	.04	−0.02 (−0.19 to 0.14)	.77
Free testosterone	0.05 (−0.06 to 0.17)	.34	0.21 (0.01 to 0.40)	.04	−0.05 (−0.22 to 0.13)	.60
CD4/CD8 ratio	0.01 (−0.06 to 0.08)	.81	0.09 (−0.03 to 0.20)	.14	−0.10 (−0.18 to −0.02)	.02

[a]All variables were natural log-transformed before analysis.

Table 5. Multivariate linear regression models for activity outcomes.

Predictors[a]	Lab-based outcomes		Accelerometer-based outcomes			
	Gait speed		Gait speed		Low activity	
	Coefficient (95% CI)	*P* value	Coefficient (95% CI)	*P* value	Coefficient (95% CI)	*P* value
Model 1						
HIV status	0.04 (−0.15 to 0.07)	.47	−0.20 (−0.39 to −0.02)	.03	−0.15 (0.01 to 0.29)	.04
Body mass index	−0.35 (−0.67 to −0.36)	.03	−0.59 (−1.13 to −0.06)	.03	0.01 (−0.41 to 0.43)	.96
Model 2[b]						
HIV status[b]	0.12 (−0.17 to 0.26)	.08	−0.10 (−0.37 to 0.16)	.44	0.14 (−0.06 to 0.33)	.17
Body mass index	−0.32 (−0.61 to −0.02)	.04	−0.65 (−1.22 to −0.07)	.03	0.02 (−0.06 to 0.33)	.94

[a]All variables were natural log-transformed before analysis.

[b]Adjusted for race (ie, white), smoking, and living alone.

Discussion

Principal Findings

An emerging challenge in managing PAWH is identifying preclinical sentinel biomarkers to identify changes in the risk for functional decline [35,36]. While consistently associated with adverse health outcomes [6-8], historically, physical function assessments, such as slow walking speed, were designed and validated in older geriatric uninfected populations [7] and may be less sensitive to preclinical features of functional decline in healthier middle-aged individuals. There is, therefore, a need for additional tools for preclinical detection of decline in mobility to prevent the development of disability later in life [12].

In our previously published study of middle-aged HIV+ and HIV− men and women, we observed modest but significant deficits in gait speed, predicted METs, and lab-based stair climb power in HIV+ compared with HIV− men and women [25]. However, when men from that study were analyzed as a subset, despite significantly reduced METs and trend deficits in stair climb power, the differences in gait speed, though lower on average did not reach significance, raising the speculation in this study that lab-based assessment may not fully capture subclinical deficits in mobility emerging in this population. The gait speed obtained during the 6-MWT done under standardized, lab-based conditions did not differentiate performance between HIV+ and HIV− individuals and were not different from published reference values for healthy men of the same age as those in this study [37]. The short duration and controlled conditions under which the 6-MWT is performed may result in faster gait speeds that may not be representative of average and variable daily physical activity and, possibly because of the current reserve capacity of participants, may allow them to compensate for this short-duration assessment [38,39]. In contrast, our measures of physical activity through accelerometry were able to differentiate between groups (P=.04), suggesting that this might be a useful preclinical measure of average daily physical activity in this population. As the habitual level of physical activity is associated with physical performance ability (greater physical activity, better physical function), monitoring physical activity may provide a useful marker for identifying preclinical signs of functional decline. The reduced volitional activity observed in this study may reflect diminished physiological energy reserve prior to overt signs of functional deficit.

Prior studies have evaluated PAWH with more advanced inflammation and functional deficits than the current status of participants in this study [5,10]. While HIV+ participants in this study did not differ in a composite score for inflammation (Table 1), they did have elevated levels of the monocyte activation marker sCD163 (data not shown) and did display higher levels of an immune activation composite score and a lower CD4/CD8 ratio, collectively signaling increased risk for multiple comorbid conditions [40], including functional impairment [41]. HIV+ participants reported more pain and increased fatigue (Table 1), consistent with a recent study that showed that pain and exhaustion were predictors of falls in middle-aged HIV+ adults and that functional impairment was associated with elevated immune activation as well as inflammation [4,41].

Conditions, such as depression and pain, which often are identified in PAWH, are known to be associated with the development of frailty [42]. This study shows that men living with HIV were less active than uninfected men when assessed by accelerometers and were also more likely to live alone and experience chronic pain. Social isolation (in this study measured as living alone; Table 1) results in a reduced social network of support and can be a barrier to physical activity that may accelerate the risk for functional decline. Moreover, smoking, common among individuals with HIV [43] may disincentivize activity and exercise because of its negative impact on cardiovascular and pulmonary health [44]. Such modifiable social and behavioral risk factors will need to be further investigated to develop preclinical intervention strategies to prevent disability later in life. Physical activity is an efficacious strategy for the prevention of many chronic conditions, including metabolic, cardiovascular, pulmonary, and musculoskeletal [45], and found to be beneficial for people living with HIV [46-49]. Therefore, a better assessment of activity can improve intervention and monitoring strategies.

A novel aspect of this study was to develop person-specific activity profiles enabled by accelerometry monitoring; this

approach may provide individuals with a reference activity pattern that can be used to capture longitudinal changes that signal the development of declines in their reserve that is not yet clinically apparent [50]. Notably, in univariate regression analysis, lab-based gait speed was sensitive to BMI as a predictor, whereas accelerometer-based metrics were sensitive to HIV status, testosterone levels, and CD4/CD8 ratios as predictors (Tables 4 and 5). Despite the small sample size, in a multivariate regression model, BMI and HIV status were independent predictors of accelerometer but not lab-based gait speed (Table 4). However, when adjusted for race, smoking, and living alone, only BMI was an independent predictor of the accelerometer and lab-based gait speed (Table 4), likely because of the small sample size of the study. Collectively, these data suggest that accelerometer- and lab-based physical activity may be linked to physiological conditions that underlie physical activity and may be useful in stratifying individuals into risk groups for future functional limitations.

Limitations

This study has multiple limitations. The sample size is small and will need confirmation in a larger study. Also, laboratory assessments were performed an average of 16 months prior to activity tracker data collection. However, participants in this study are clinically stable, with an overall accelerometry-based gait speed (0.44 m/s) similar to gait speeds in an independent study sample of over 800 men (0.46 m/s) using Withings trackers in the same age range and sampled randomly across the East Coast (Multimedia Appendix 2). Additional limitations include potential differences in activity because of seasonal variation, physical autonomy, and access to urban greenspaces [51].

Conclusions

In summary, this study identifies a significantly reduced gait speed and activity pattern in otherwise asymptomatic middle-aged men living with HIV compared with those without HIV, in the absence of detectable differences in physical performance assessed in a laboratory setting. HIV+ participants reported more fatigue and pain, which when coupled with the observed reduced activity may signal a state of preclinical risk for functional decline. Free-living accelerometry may provide a useful biometric tool for monitoring the efficacy of future interventions focused on reducing decline in the physical function in PAWH.

Acknowledgments

This work was supported by grants from the National Institute of Allergy and Infectious Diseases (R01 AI108541 to MM, and the Harvard University Center for AIDS Research 5P30AI060354-12 subaward to TMH), the National Institute of Aging (R21 AG055415 to MM, the Boston Claude D. Pepper Older Americans Independence Center P30 AG031679 subproject 6777 to MM, and the R24AG044325 subaward to TMH).

We thank the MATCH study participants, the Boston Pepper Center, and the Harvard University Center for AIDS Research. We also thank the Withings for their gift supporting this research.

Authors' Contributions

ER assisted in extracting accelerometer data and creating measures. MV and ER provided input on the methods section of this manuscript.

Conflicts of Interest

MM and ER are employed by Withings. The other authors have no other conflicts of interest to declare.

Multimedia Appendix 1

Comparison of the Withings Pulse Ox step count and estimates of gait speed to results from the 6-minute walk test.

[PDF File (Adobe PDF File), 31KB - mhealth_v7i2e11190_app1.pdf]

Multimedia Appendix 2

Comparison between the mean free-living (volitional) gait speed measured in the MPACT substudy to reference data.

[PDF File (Adobe PDF File), 38KB - mhealth_v7i2e11190_app2.pdf]

References

1. Wada N, Jacobson LP, Cohen M, French A, Phair J, Muñoz A. Cause-specific life expectancies after 35 years of age for human immunodeficiency syndrome-infected and human immunodeficiency syndrome-negative individuals followed simultaneously in long-term cohort studies, 1984-2008. Am J Epidemiol 2013 Jan 15;177(2):116-125 [FREE Full text] [doi: 10.1093/aje/kws321] [Medline: 23287403]
2. Hawkins KL, Brown TT, Margolick JB, Erlandson KM. Geriatric syndromes: new frontiers in HIV and sarcopenia. AIDS 2017 Dec 01;31 Suppl 2:S137-S146. [doi: 10.1097/QAD.0000000000001444] [Medline: 28471944]

3. Deeks SG. HIV infection, inflammation, immunosenescence, and aging. Annu Rev Med 2011;62:141-155 [FREE Full text] [doi: 10.1146/annurev-med-042909-093756] [Medline: 21090961]
4. Erlandson KM, Allshouse AA, Jankowski CM, Duong S, MaWhinney S, Kohrt WM, et al. Risk factors for falls in HIV-infected persons. J Acquir Immune Defic Syndr 2012 Dec 01;61(4):484-489 [FREE Full text] [doi: 10.1097/QAI.0b013e3182716e38] [Medline: 23143526]
5. Greene M, Covinsky K, Astemborski J, Piggott DA, Brown T, Leng S, et al. The relationship of physical performance with HIV disease and mortality. AIDS 2014 Nov 28;28(18):2711-2719 [FREE Full text] [doi: 10.1097/QAD.0000000000000507] [Medline: 25493597]
6. Guralnik JM, Ferrucci L, Pieper CF, Leveille SG, Markides KS, Ostir GV, et al. Lower extremity function and subsequent disability: consistency across studies, predictive models, and value of gait speed alone compared with the short physical performance battery. J Gerontol A Biol Sci Med Sci 2000 Apr;55(4):M221-M231. [Medline: 10811152]
7. 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]
8. Studenski S, Perera S, Patel K, Rosano C, Faulkner K, Inzitari M, et al. Gait speed and survival in older adults. JAMA 2011 Jan 5;305(1):50-58 [FREE Full text] [doi: 10.1001/jama.2010.1923] [Medline: 21205966]
9. Sabia S, Dumurgier J, Tavernier B, Head J, Tzourio C, Elbaz A. Change in fast walking speed preceding death: results from a prospective longitudinal cohort study. J Gerontol A Biol Sci Med Sci 2014 Mar;69(3):354-362 [FREE Full text] [doi: 10.1093/gerona/glt114] [Medline: 23913931]
10. Schrack JA, Althoff KN, Jacobson LP, Erlandson KM, Jamieson BD, Koletar SL, et al. Accelerated Longitudinal Gait Speed Decline in HIV-Infected Older Men. JAIDS Journal of Acquired Immune Deficiency Syndromes 2015;70(4):370-376. [doi: 10.1097/QAI.0000000000000731] [Medline: 26102450]
11. LeBrasseur NK, Bhasin S, Miciek R, Storer TW. Tests of muscle strength and physical function: reliability and discrimination of performance in younger and older men and older men with mobility limitations. J Am Geriatr Soc 2008 Nov;56(11):2118-2123 [FREE Full text] [doi: 10.1111/j.1532-5415.2008.01953.x] [Medline: 18811607]
12. Ferrucci L, Cooper R, Shardell M, Simonsick EM, Schrack JA, Kuh D. Age-Related Change in Mobility: Perspectives From Life Course Epidemiology and Geroscience. J Gerontol A Biol Sci Med Sci 2016 Dec;71(9):1184-1194 [FREE Full text] [doi: 10.1093/gerona/glw043] [Medline: 26975983]
13. Sullivan K, Shikuma CM, Chow D, Cornelius E, Romine RK, Lindsey RA, et al. Aerobic fitness levels and validation of a non exercise VO2max prediction equation for HIV-infected patients on HAART. HIV Clin Trials 2014;15(2):69-77 [FREE Full text] [doi: 10.1310/hct1502-69] [Medline: 24710921]
14. Vancampfort D, Mugisha J, De Hert M, Probst M, Stubbs B. Sedentary Behavior in People Living With HIV: A Systematic Review and Meta-Analysis. J Phys Act Health 2017 Jul;14(7):571-577. [doi: 10.1123/jpah.2016-0507] [Medline: 28290767]
15. Jacquemin C, Servy H, Molto A, Sellam J, Foltz V, Gandjbakhch F, et al. Physical Activity Assessment Using an Activity Tracker in Patients with Rheumatoid Arthritis and Axial Spondyloarthritis: Prospective Observational Study. JMIR Mhealth Uhealth 2018 Jan 02;6(1):e1 [FREE Full text] [doi: 10.2196/mhealth.7948] [Medline: 29295810]
16. Schrack JA, Cooper R, Koster A, Shiroma EJ, Murabito JM, Rejeski WJ, et al. Assessing Daily Physical Activity in Older Adults: Unraveling the Complexity of Monitors, Measures, and Methods. J Gerontol A Biol Sci Med Sci 2016 Dec;71(8):1039-1048 [FREE Full text] [doi: 10.1093/gerona/glw026] [Medline: 26957472]
17. Guimarães MP, Ferriolli E, Pfrimer K, Navarro AM. Doubly Labeled Water Method and Accelerometer for the Measurement of Energy Expenditure in Human Immunodeficiency Virus-Infected Patients. Ann Nutr Metab 2017;70(1):66-73. [doi: 10.1159/000458766] [Medline: 28278503]
18. Rehm KE, Konkle-Parker D. Physical activity levels and perceived benefits and barriers to physical activity in HIV-infected women living in the deep south of the United States. AIDS Care 2016 Dec;28(9):1205-1210 [FREE Full text] [doi: 10.1080/09540121.2016.1164802] [Medline: 27023306]
19. Henry BL, Moore DJ. Preliminary Findings Describing Participant Experience With iSTEP, an mHealth Intervention to Increase Physical Activity and Improve Neurocognitive Function in People Living With HIV. J Assoc Nurses AIDS Care 2016;27(4):495-511. [doi: 10.1016/j.jana.2016.01.001] [Medline: 26847379]
20. Ramírez-Marrero FA, Santana-Bagur JL, Joyner MJ, Rodríguez-Zayas J, Frontera W. Metabolic syndrome in relation to cardiorespiratory fitness, active and sedentary behavior in HIV+ Hispanics with and without lipodystrophy. P R Health Sci J 2014 Dec;33(4):163-169 [FREE Full text] [Medline: 25563033]
21. Fillipas S, Cicuttini F, Holland AE, Cherry CL. The international physical activity questionnaire overestimates moderate and vigorous physical activity in HIV-infected individuals compared with accelerometry. J Assoc Nurses AIDS Care 2010;21(2):173-181. [doi: 10.1016/j.jana.2009.11.003] [Medline: 20116301]
22. Chiauzzi E, Rodarte C, DasMahapatra P. Patient-centered activity monitoring in the self-management of chronic health conditions. BMC Med 2015 Apr;13:77 [FREE Full text] [doi: 10.1186/s12916-015-0319-2] [Medline: 25889598]
23. Ferguson T, Rowlands AV, Olds T, Maher C. The validity of consumer-level, activity monitors in healthy adults worn in free-living conditions: a cross-sectional study. Int J Behav Nutr Phys Act 2015 Mar 27;12:42 [FREE Full text] [doi: 10.1186/s12966-015-0201-9] [Medline: 25890168]

24. Kooiman TJM, Dontje ML, Sprenger SR, Krijnen WP, van der Schans CP, de Groot M. Reliability and validity of ten consumer activity trackers. BMC Sports Sci Med Rehabil 2015 Oct;7:24 [FREE Full text] [doi: 10.1186/s13102-015-0018-5] [Medline: 26464801]
25. Tran T, Guardigni V, Pencina KM, Amato AA, Floyd M, Brawley B, et al. Atypical Skeletal Muscle Profiles in HIV+ Asymptomatic Middle-Aged Adults. Clin Infect Dis 2017 Dec 25. [doi: 10.1093/cid/cix1121] [Medline: 29293942]
26. Cella D, Riley W, Stone A, Rothrock N, Reeve B, Yount S, et al. The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005-2008. J Clin Epidemiol 2010 Nov;63(11):1179-1194 [FREE Full text] [doi: 10.1016/j.jclinepi.2010.04.011] [Medline: 20685078]
27. Rejeski WJ, Ip EH, Marsh AP, Miller ME, Farmer DF. Measuring disability in older adults: the International Classification System of Functioning, Disability and Health (ICF) framework. Geriatr Gerontol Int 2008 Mar;8(1):48-54 [FREE Full text] [doi: 10.1111/j.1447-0594.2008.00446.x] [Medline: 18713189]
28. Radloff LS. The CES-D Scale: A Self-Report Depression Scale for Research in the General Population. Applied Psychological Measurement 1977 Jun 01;1(3):385-401. [doi: 10.1177/014662167700100306]
29. Lewinsohn PM, Seeley JR, Roberts RE, Allen NB. Center for Epidemiologic Studies Depression Scale (CES-D) as a screening instrument for depression among community-residing older adults. Psychology and Aging 1997 Nov;12(2):277-287. [doi: 10.1037/0882-7974.12.2.277]
30. Smith E, Lai J, Cella D. Building a measure of fatigue: the functional assessment of Chronic Illness Therapy Fatigue Scale. PM R 2010 May;2(5):359-363. [doi: 10.1016/j.pmrj.2010.04.017] [Medline: 20656617]
31. Justice AC, Modur SP, Tate JP, Althoff KN, Jacobson LP, Gebo KA, NA-ACCORDVACS Project Teams. Predictive accuracy of the Veterans Aging Cohort Study index for mortality with HIV infection: a North American cross cohort analysis. J Acquir Immune Defic Syndr 2013 Feb 01;62(2):149-163 [FREE Full text] [doi: 10.1097/QAI.0b013e31827df36c] [Medline: 23187941]
32. Tate JP, Justice AC, Hughes MD, Bonnet F, Reiss P, Mocroft A, et al. An internationally generalizable risk index for mortality after one year of antiretroviral therapy. AIDS 2013 Feb 20;27(4):563-572 [FREE Full text] [doi: 10.1097/QAD.0b013e32835b8c7f] [Medline: 23095314]
33. Zakharov MN, Bhasin S, Travison TG, Xue R, Ulloor J, Vasan RS, et al. A multi-step, dynamic allosteric model of testosterone's binding to sex hormone binding globulin. Mol Cell Endocrinol 2015 Jan 05;399:190-200. [doi: 10.1016/j.mce.2014.09.001] [Medline: 25240469]
34. Murray MP, Drought AB, Kory RC. Walking Patterns of Normal Men. The Journal of Bone & Joint Surgery 1964;46(2):335-360.
35. Deeks SG, Lewin SR, Havlir DV. The end of AIDS: HIV infection as a chronic disease. Lancet 2013 Nov 02;382(9903):1525-1533 [FREE Full text] [doi: 10.1016/S0140-6736(13)61809-7] [Medline: 24152939]
36. Smit M, Brinkman K, Geerlings S, Smit C, Thyagarajan K, Sighem AV, ATHENA observational cohort. Future challenges for clinical care of an ageing population infected with HIV: a modelling study. Lancet Infect Dis 2015 Jul;15(7):810-818 [FREE Full text] [doi: 10.1016/S1473-3099(15)00056-0] [Medline: 26070969]
37. Enright PL, Sherrill DL. Reference equations for the six-minute walk in healthy adults. Am J Respir Crit Care Med 1998 Nov;158(5 Pt 1):1384-1387. [doi: 10.1164/ajrccm.158.5.9710086] [Medline: 9817683]
38. Cooper R, Bann D, Wloch EG, Adams JE, Kuh D. "Skeletal muscle function deficit" in a nationally representative British birth cohort in early old age. J Gerontol A Biol Sci Med Sci 2015 May;70(5):604-607 [FREE Full text] [doi: 10.1093/gerona/glu214] [Medline: 25431354]
39. Shumway-Cook A, Guralnik JM, Phillips CL, Coppin AK, Ciol MA, Bandinelli S, et al. Age-associated declines in complex walking task performance: the Walking InCHIANTI toolkit. J Am Geriatr Soc 2007 Jan;55(1):58-65 [FREE Full text] [doi: 10.1111/j.1532-5415.2006.00962.x] [Medline: 17233686]
40. Serrano-Villar S, Moreno S, Fuentes-Ferrer M, Sánchez-Marcos C, Avila M, Sainz T, et al. The CD4:CD8 ratio is associated with markers of age-associated disease in virally suppressed HIV-infected patients with immunological recovery. HIV Med 2014 Jan;15(1):40-49 [FREE Full text] [doi: 10.1111/hiv.12081] [Medline: 24007533]
41. Erlandson KM, Allshouse AA, Jankowski CM, Lee EJ, Rufner KM, Palmer BE, et al. Association of functional impairment with inflammation and immune activation in HIV type 1-infected adults receiving effective antiretroviral therapy. J Infect Dis 2013 Jul 15;208(2):249-259 [FREE Full text] [doi: 10.1093/infdis/jit147] [Medline: 23559466]
42. Onen NF, Agbebi A, Shacham E, Stamm KE, Onen AR, Overton ET. Frailty among HIV-infected persons in an urban outpatient care setting. J Infect 2009 Nov;59(5):346-352. [doi: 10.1016/j.jinf.2009.08.008] [Medline: 19706308]
43. Jacquet J, Peyriere H, Makinson A, Peries M, Nagot N, Donnadieu-Rigole H, MesConsos Study group. Psychoactive substances, alcohol and tobacco consumption in HIV-infected outpatients. AIDS 2018 Jun 01;32(9):1165-1171. [doi: 10.1097/QAD.0000000000001830] [Medline: 29683847]
44. Kamholz SL. Pulmonary and cardiovascular consequences of smoking. Clin Occup Environ Med 2006;5(1):157-71, x. [doi: 10.1016/j.coem.2005.10.001] [Medline: 16446260]
45. Pedersen BK, Saltin B. Exercise as medicine - evidence for prescribing exercise as therapy in 26 different chronic diseases. Scand J Med Sci Sports 2015 Dec;25 Suppl 3:1-72. [doi: 10.1111/sms.12581] [Medline: 26606383]

46. Gomes Neto M, Conceição CS, Oliveira Carvalho V, Brites C. Effects of Combined Aerobic and Resistance Exercise on Exercise Capacity, Muscle Strength and Quality of Life in HIV-Infected Patients: A Systematic Review and Meta-Analysis. PLoS One 2015 Sep;10(9):e0138066 [FREE Full text] [doi: 10.1371/journal.pone.0138066] [Medline: 26378794]
47. Hand GA, Lyerly GW, Jaggers JR, Dudgeon WD. Impact of Aerobic and Resistance Exercise on the Health of HIV-Infected Persons. Am J Lifestyle Med 2009 Nov 01;3(6):489-499 [FREE Full text] [doi: 10.1177/1559827609342198] [Medline: 20508736]
48. Neff HA, Kellar-Guenther Y, Jankowski CM, Worthington C, McCandless SA, Jones J, et al. Turning disability into ability: barriers and facilitators to initiating and maintaining exercise among older men living with HIV. AIDS Care 2018 Jul 03:1-5. [doi: 10.1080/09540121.2018.1493186] [Medline: 29968493]
49. Quigley A, O'Brien K, Parker R, MacKay-Lyons M. Exercise and cognitive function in people living with HIV: a scoping review. Disabil Rehabil 2018 Jan 29:1-12. [doi: 10.1080/09638288.2018.1432079] [Medline: 29376434]
50. Goldhahn J. Need for Digital Biomarkers in Musculoskeletal Trials. Digit Biomark 2017 Aug 17;1(1):82-86. [doi: 10.1159/000479753]
51. Taylor L, Hochuli D. Defining greenspace: Multiple uses across multiple disciplines. Landscape and Urban Planning 2017 Feb;158:25-38 [FREE Full text] [doi: 10.1016/j.landurbplan.2016.09.024]

Abbreviations

6-MWT: 6 minutes walking test
BMI: body mass index
CES-D: Center for Epidemiological Studies Depression
FT: free testosterone
FACIT-F: Functional Assessment of Chronic Illness Therapy-Fatigue
HRQoL: health-related quality of life
MATCH: Muscle Function and Aging in Chronic HIV Infection
MET: metabolic equivalent
PAT-D: Pepper Assessment Tool for Disability
PAWH: People aging with HIV
PROMIS: Patient Reported Outcome Measurements Information System
SHBG: serum hormone-binding globulin

Edited by G Eysenbach; submitted 02.06.18; peer-reviewed by NA Ismail, A Algarin, C Adewole; comments to author 03.10.18; revised version received 22.10.18; accepted 26.10.18; published 01.02.19.

Please cite as:
Hale TM, Guardigni V, Roitmann E, Vegreville M, Brawley B, Woodbury E, Storer TW, Sax PE, Montano M
Middle-Aged Men With HIV Have Diminished Accelerometry-Based Activity Profiles Despite Similar Lab-Measured Gait Speed: Pilot Study
JMIR Mhealth Uhealth 2019;7(2):e11190
URL: http://mhealth.jmir.org/2019/2/e11190/
doi:10.2196/11190
PMID:30707104

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org