Journal of Medical Internet Research

Journal Impact Factor (JIF) (2023): 5.8
Volume 18 (2016), Issue 4   ISSN 1438-8871   Editor in Chief: Gunther Eysenbach, MD, MPH

Contents

Original Papers

Cardiac Patients' Walking Activity Determined by a Step Counter in Cardiac Telerehabilitation: Data From the Intervention Arm of a Randomized Controlled Trial (e69)

Charlotte Thorup, John Hansen, Mette Grønkjær, Jan Andreasen, Gitte Nielsen, Erik Sørensen, Birthe Dinesen. 3

Efficacy and External Validity of Electronic and Mobile Phone-Based Interventions Promoting Vegetable Intake in Young Adults: Systematic Review and Meta-Analysis (e58)

Monica Nour, Juliana Chen, Margaret Allman-Farinelli. 16

Effectiveness of a Web-Based Computer-Tailored Multiple-Lifestyle Intervention for People Interested in Reducing their Cardiovascular Risk: A Randomized Controlled Trial (e78)

Vera Storm, Julia Dörenkämper, Dominique Reinwand, Julian Wienert, Hein De Vries, Sonia Lippke. 35

Enhancement of Self-Monitoring in a Web-Based Weight Loss Program by Extra Individualized Feedback and Reminders: Randomized Trial (e82)

Melinda Hutchesson, Chor Tan, Philip Morgan, Robin Callister, Clare Collins. 50

The Impact of Automated Brief Messages Promoting Lifestyle Changes Delivered Via Mobile Devices to People with Type 2 Diabetes: A Systematic Literature Review and Meta-Analysis of Controlled Trials (e86)

Carukshi Arambepola, Ignacio Ricci-Cabello, Pavithra Manikavasagam, Nia Roberts, David French, Andrew Farmer. 61

Evaluation of QuitNow Men: An Online, Men-Centered Smoking Cessation Intervention (e83)

Joan Bottorff, John Oliffe, Gayl Sarbit, Paul Sharp, Cristina Caperchione, Leanne Currie, Jonathan Schmid, Martha Mackay, Sean Stolp. . . .
7 3

Effectiveness of Social Media Interventions for People With Schizophrenia: A Systematic Review and Meta-Analysis (e92)

Maritta Välimäki, Christina Athanasopoulou, Mari Lahti, Clive Adams. 83

An Internet-Based Physical Activity Intervention to Improve Quality of Life of Inactive Older Adults: A Randomized Controlled Trial (e74)

Karen Broekhuizen, Jelle de Gelder, Carolien Wijsman, Liselotte Wijsman, Rudi Westendorp, Evert Verhagen, Pieternella Slagboom, Anton de Craen, Willem van Mechelen, Diana van Heemst, Frans van der Ouderaa, Simon Mooijaart. 97

Interprofessional Communication of Clinicians Using a Mobile Phone App: A Randomized Crossover Trial Using Simulated Patients (e79)

Bhavesh Patel, Maximilian Johnston, Natalie Cookson, Dominic King, Sonal Arora, Ara Darzi. 108

Comparison Between a Self-Administered and Supervised Version of a Web-Based Cognitive Test Battery: Results From the NutriNet-Santé Cohort Study (e68)
Karen Assmann, Marion Bailet, Amandine Lecoffre, Pilar Galan, Serge Hercberg, Hélène Amieva, Emmanuelle Kesse-Guyot. 121

Awareness and Use of the After-Visit Summary Through a Patient Portal: Evaluation of Patient Characteristics and an Application of the Theory of Planned Behavior (e77)
Srinivas Emani, Michael Healey, David Ting, Stuart Lipsitz, Harley Ramelson, Vladimir Suric, David Bates. 134

Is Participatory Design Associated with the Effectiveness of Serious Digital Games for Healthy Lifestyle Promotion? A Meta-Analysis (e94)
Ann DeSmet, Debbe Thompson, Tom Baranowski, Antonio Palmeira, Maïté Verloigne, Ilse De Bourdeaudhuij. 147

What is the Profile of Individuals Joining the KNEEguru Online Health Community? A Cross-Sectional Mixed-Methods Study (e84)
Philip Bright, Karen Hambly, Sandra Tamakloe. 166

The Pros and Cons of Getting Engaged in an Online Social Community Embedded Within Digital Cognitive Behavioral Therapy for Insomnia: Survey Among Users (e88)
Neil Coulson, Richard Smedley, Sophie Bostock, Simon Kyle, Rosie Gollancz, Annemarie Luik, Peter Hames, Colin Espie. 178

Crowdsourcing and the Accuracy of Online Information Regarding Weight Gain in Pregnancy: A Descriptive Study (e81)
Tammy Chang, Bianca Verma, Trevor Shull, Michelle Moniz, Lauren Kohatsu, Melissa Plegue, Kevyn Collins-Thompson. 187

Patient Perspectives on Sharing Anonymized Personal Health Data Using a Digital System for Dynamic Consent and Research Feedback: A Qualitative Study (e66)
Karen Spencer, Caroline Sanders, Edgar Whitley, David Lund, Jane Kaye, William Dixon. 196

Opportunities for Web-based Drug Repositioning: Searching for Potential Antihypertensive Agents with Hypotension Adverse Events (e76)
Kejian Wang, Mei Wan, Rui-Sheng Wang, Zuquan Weng. 207

Answers to Health Questions: Internet Search Results Versus Online Health Community Responses (e95)
Shaheen Kanthawala, Amber Vermeesch, Barbara Given, Jina Huh. 217

A Web-Based Computer-Tailored Alcohol Prevention Program for Adolescents: Cost-Effectiveness and Intersectoral Costs and Benefits (e93)
Ruben Drost, Aggie Paulus, Astrid Jander, Liesbeth Mercken, Hein de Vries, Dirk Ruwaard, Silvia Evers. 230

A Patient-Centered Framework for Evaluating Digital Maturity of Health Services: A Systematic Review (e75)
Kelsey Flott, Ryan Callahan, Ara Darzi, Erik Mayer. 253

Original Paper

Cardiac Patients' Walking Activity Determined by a Step Counter in Cardiac Telerehabilitation: Data From the Intervention Arm of a Randomized Controlled Trial

Charlotte Thorup[1,2], MSN; John Hansen[3], PhD; Mette Grønkjær[4], PhD; Jan Jesper Andreasen[1,2], PhD; Gitte Nielsen[5], PhD; Erik Elgaard Sørensen[2,4], PhD; Birthe Irene Dinesen[6], PhD

[1]Department of Cardiothoracic Surgery, Aalborg University Hospital, Aalborg, Denmark

[2]Department of Clinical Medicine, Aalborg University, Aalborg, Denmark

[3]Laboratory for Cardio technology, Medical Informatics Group, Department of Health Science and Technology, Faculty of Medicine, Aalborg University, Aalborg, Denmark

[4]Clinical Nursing Research Unit, Aalborg University Hospital, Aalborg, Denmark

[5]Department of Cardiology, Vendsyssel Hospital, Hjoerring, Denmark

[6]Laboratory of Assistive Technologies - Telehealth and Telerehabilitation, SMI, Department of Health Science and Technology, Faculty of Medicine, Aalborg University, Aalborg, Denmark

Corresponding Author:
Charlotte Thorup, MSN
Department of Cardiothoracic Surgery
Aalborg University Hospital
Søndre Skovvej 5, 313
Aalborg, 9000
Denmark
Phone: 45 20729950
Fax: 45 78453075
Email: cbt@rn.dk

Abstract

Background: Walking represents a large part of daily physical activity. It reduces both overall and cardiovascular diseases and mortality and is suitable for cardiac patients. A step counter measures walking activity and might be a motivational tool to increase and maintain physical activity. There is a lack of knowledge about both cardiac patients' adherence to step counter use in a cardiac telerehabilitation program and how many steps cardiac patients walk up to 1 year after a cardiac event.

Objective: The purpose of this substudy was to explore cardiac patients' walking activity. The walking activity was analyzed in relation to duration of pedometer use to determine correlations between walking activity, demographics, and medical and rehabilitation data.

Methods: A total of 64 patients from a randomized controlled telerehabilitation trial (Teledi@log) from Aalborg University Hospital and Hjoerring Hospital, Denmark, from December 2012 to March 2014 were included in this study. Inclusion criteria were patients hospitalized with acute coronary syndrome, heart failure, and coronary artery bypass grafting or valve surgery. In Teledi@log, the patients received telerehabilitation technology and selected one of three telerehabilitation settings: a call center, a community health care center, or a hospital. Monitoring of steps continued for 12 months and a step counter (Fitbit Zip) was used to monitor daily steps.

Results: Cardiac patients walked a mean 5899 (SD 3274) steps per day, increasing from mean 5191 (SD 3198) steps per day in the first week to mean 7890 (SD 2629) steps per day after 1 year. Adherence to step counter use lasted for a mean 160 (SD 100) days. The patients who walked significantly more were younger (*P*=.01) and continued to use the pedometer for a longer period (*P*=.04). Furthermore, less physically active patients weighed more. There were no significant differences in mean steps per day for patients in the three rehabilitation settings or in the disease groups.

Conclusions: This study indicates that cardiac telerehabilitation at a call center can support walking activity just as effectively as telerehabilitation at either a hospital or a health care center. In this study, the patients tended to walk fewer steps per day than cardiac patients in comparable studies, but our study may represent a more realistic picture of walking activity due to the continuation of step counter use. Qualitative studies on patients' behavior and motivation regarding step counter use are needed to shed light on adherence to and motivation to use step counters.

Trial Registration: ClinicalTrails.gov NCT01752192; https://clinicaltrials.gov/ct2/show/NCT01752192 (Archived by WebCite at http://www.webcitation.org/6fgigfUyV)

(J Med Internet Res 2016;18(4):e69) doi:10.2196/jmir.5191

KEYWORDS

heart disease; rehabilitation; step counters; physical activity; telerehabilitation

Introduction

Cardiac diseases are the main causes of death and account for 13% to 15% of all deaths worldwide [1] and 24.8% of all deaths in Europe [2]. Cardiac rehabilitation aims to improve cardiac patients' functional capacity, recovery, psychosocial well-being, and health-related quality of life through a multidisciplinary intervention consisting of physical activity support, exercise training, diet and weight counseling, psychosocial coping, and management of the disease [2]. Cardiac rehabilitation is vital for recovery after cardiac disease, yet cardiac rehabilitation has poor compliance and adherence [2,3]. Home-based cardiac rehabilitation programs, such as cardiac telerehabilitation, have been introduced to increase access, participation, and adherence. Cardiac telerehabilitation is defined as cardiac rehabilitation that uses information and communication technology to improve health and lifestyle by monitoring and communicating through interactive tools. Cardiac telerehabilitation has proven to be just as effective in decreasing morbidity and mortality as center- and hospital-based cardiac rehabilitation programs [4-6]. Furthermore, cardiac telerehabilitation has the potential to reach citizens who live long distances from rehabilitation centers [4-12]. Physical activity decreases cardiovascular mortality and morbidity [13-20]. Walking is a simple physical activity that forms a large part of daily physical activity in both sedentary and active individuals, reducing both overall and cardiovascular disease mortality [21]. As a result, walking strategies need to be included in cardiac rehabilitation. Walking is suitable for cardiac patients because it is safe and feasible for almost all patients [22-24]. Step counters are recognized tools to count steps and measure walking activity, and they represent an important motivational tool to measure and increase adherence to physical activity [17,25,26]. Studies of cardiac patients' use of step counters show an increase of physical activity with their use [18,21,25,27-31]. Cardiac patients who receive walking interventions have higher levels of walking activity compared to control groups, and their walking activity increases over 1 year [16-18].

People are considered physically active when they perform more than 30 minutes of moderate to intense activity per day (on most days of the week) [32]. Researchers agree that approximately 7000-10,000 steps per day is equivalent to 30 minutes per day of moderate to intense physical activity. More than 10,000 steps per day is considered highly active [15,17,24,32]. There is a lack of knowledge about cardiac patients' adherence to, and use of, step counters during rehabilitation. We also lack knowledge regarding how many steps cardiac patients walk during the period up to 1 year after a cardiac event [16,17]. A 1-year follow-up is considered relevant because sustained behavioral changes require a long observation period [33]. In an attempt to increase patients' physical activity, it is important to identify realistic and appropriate goals [21,29,31,34]. Six studies have been identified that explore the amount of walking activity by patients with cardiac disease [15-18,21,35] and these studies show diversity in results. Further studies on cardiac patients' walking activity and use of step counters are needed.

This study is part of a larger Danish research project, Teledi@log, in which a cardiac telerehabilitation program has been developed and tested as a randomized controlled trial for patients with heart diseases. This paper focuses explicitly on cardiac patients' walking activity. The walking activity will be analyzed in relation to the duration of step counter use to determine correlations between walking activity, demographics, and medical and rehabilitation data.

Methods

Cardiac Telerehabilitation

The patients selected were participants from a randomized controlled cardiac telerehabilitation study, Teledi@log (ClinicalTrails.gov NCT01752192), and were included from December 2012 to March 2014. The Teledi@log project was approved by the Danish Ethical Committee (N-20120051). The general objective of Teledi@log was to tailor cardiac telerehabilitation based on the patient's individual needs. The telerehabilitation program lasted for 3 months. Patients in the intervention group were provided with a step counter, a scale, a sphygmomanometer, and a tablet. The tablet contained a tailored personal health record (PHR) for health information and communication between the patient and health professionals. The patients measured blood pressure, pulse, and weight twice a week and number of steps recorded daily on a step counter. Data were transmitted wirelessly from the devices to the PHR. Based on the patient's individual condition, the rehabilitation nurse created a tailored rehabilitation plan for each patient, containing an activity plan with goals for daily steps. This was done in accordance with European Association of Cardiovascular Prevention and Rehabilitation recommendations [3] and in collaboration with the patient before discharge from the hospital. The plan was displayed in the PHR. Both the patient and the health professionals at the hospital and health care center had access to and communicated via the patient's PHR. All patients had personal goals for daily steps in the PHR. In addition to access to health information, the step counter was the only telerehabilitation technology that the patients retained after 3 months, allowing them to continue monitoring steps for

12 months. The telerehabilitation technology provided the patients with insights into their own walking activity, enabling them to monitor and tailor their own activity plans. All patients were assigned a personal nurse attached to either the health care center or the hospital. The control group received traditional cardiac rehabilitation. In Denmark, Danish national guidelines specify that cardiac patients can be offered either cardiac rehabilitation at a hospital or at a health care center [36]. To match this in the intervention group, cardiac patients in the Teledi@log trial selected one of three rehabilitation settings: call center and telerehabilitation, individualized cardiac telerehabilitation at a community health care center, or individualized cardiac telerehabilitation at the hospital (Textbox 1).

Textbox 1. Settings for cardiac telerehabilitation.

1. Call center (contact person: cardiac nurse)

- A cardiac nurse from the hospital was in charge of the patient's rehabilitation and all rehabilitation activities were provided through the personal health record and in collaboration with the rehabilitation nurse in accordance with European Association of Cardiovascular Prevention and Rehabilitation, individualized activities were planned (patient decided what activities to follow)
- Follow-up time was based on individual needs
- Self-monitoring
- Step counter

2. Health care center (contact person: cardiac rehabilitation nurse)

- A rehabilitation nurse from the health care center was in charge of the patient's rehabilitation and all rehabilitation activities were provided through the personal health record and in collaboration with the rehabilitation nurse in accordance with recommendations from European Association of Cardiovascular Prevention and Rehabilitation; furthermore, the rehabilitation consisted of individual and group sessions once or twice a week for 12 weeks
- Group consultation and exercise sessions took place together with other cardiac patients
- Follow-up time was based on individual needs
- Group exercise
- Self-monitoring
- Step counter

3. Hospital (contact person: cardiac nurse)

- A cardiac nurse from the hospital was in charge of the patient's rehabilitation and all rehabilitation activities were provided through the personal health record and in collaboration with the rehabilitation nurse, in accordance with recommendations from European Association of Cardiovascular Prevention and Rehabilitation; furthermore, the rehabilitation consisted of individual and group sessions once or twice a week for 12 weeks
- Consultation and exercise sessions took place in a group with other cardiac patients
- Follow-up time was based on individual needs
- Group exercise
- Self-monitoring
- Step counter

Participants and Recruitment

A computer-based block randomization in groups of 10 was performed. In total, 151 cardiac patients participated in Teledi@log. Of these, 72 were in the intervention group who received telerehabilitation technologies, including the step counter. Eight of the patients in the intervention group dropped out during the 1-year study period (four died, one had severe progression of illness, one was unreachable at follow-up, and two could not cope with or finish the project), leaving 64 patients for this step counter substudy. Figure 1 shows the CONSORT flow diagram of Teledi@log, with the 64 patients in this study shown as the intervention group.

Patients were recruited from Aalborg University Hospital, Aalborg, Denmark, and Vendsyssel Hospital, Hjoerring, Denmark. The inclusion criteria were patients hospitalized with acute coronary syndrome (ACS), heart failure (ejection fraction <40%), coronary artery bypass graft (CABG), or valve replacement/mitral valve repair. Patients were excluded in cases of pregnancy or breastfeeding, or if they did not speak Danish. Telerehabilitation nurses at the participating hospitals reviewed patients' charts for eligibility and eligible patients were approached. Those who agreed to participate signed an informed consent form. Demographic data were registered by the nurse. Furthermore, the patients were instructed in how to use the telerehabilitation technologies and measurement started immediately after discharge. Fourteen days after inclusion, each patient was visited by a research assistant to ensure that the patients were using the telerehabilitation technologies correctly.

Figure 1. Teledi@log CONSORT flow diagram.

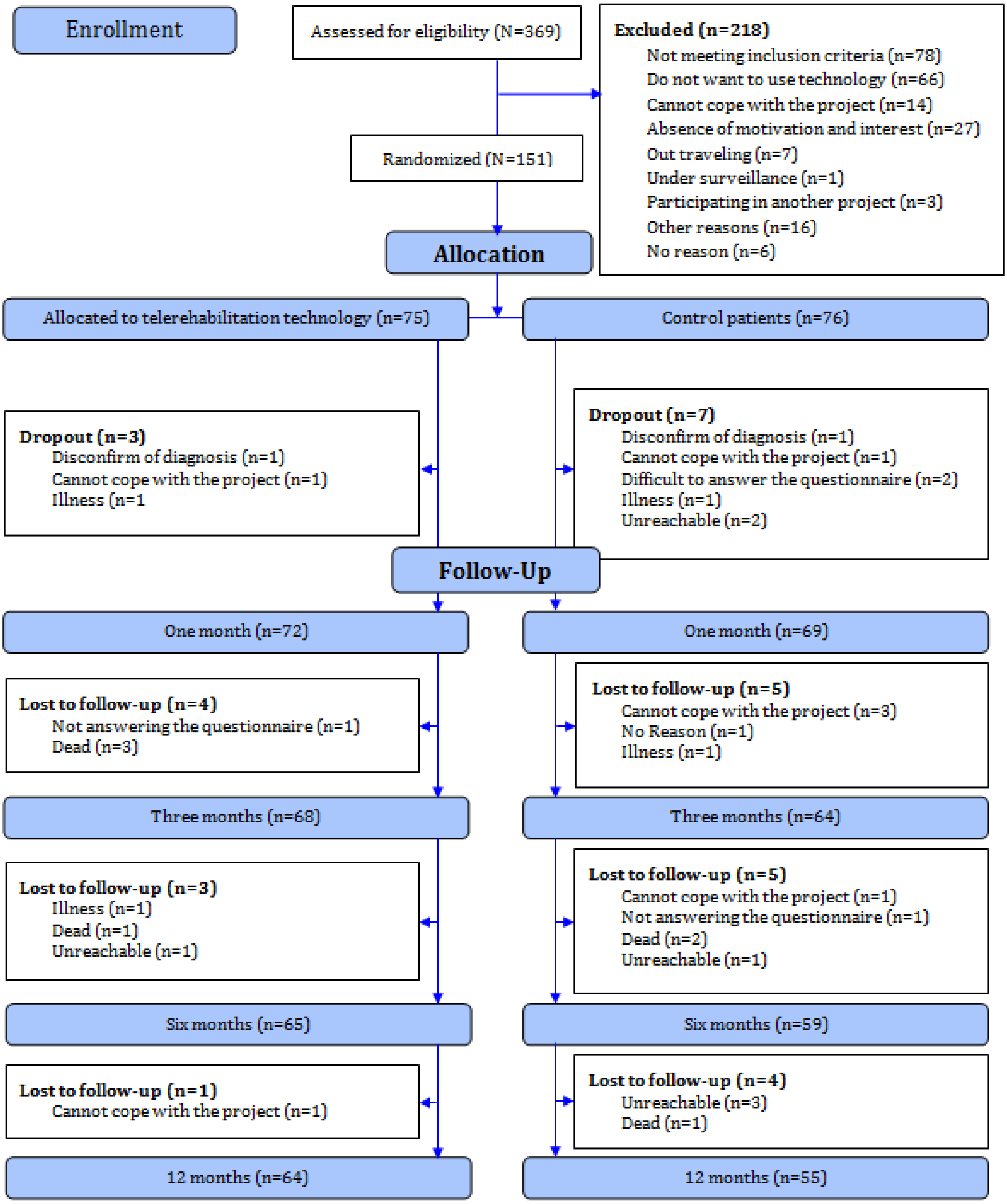
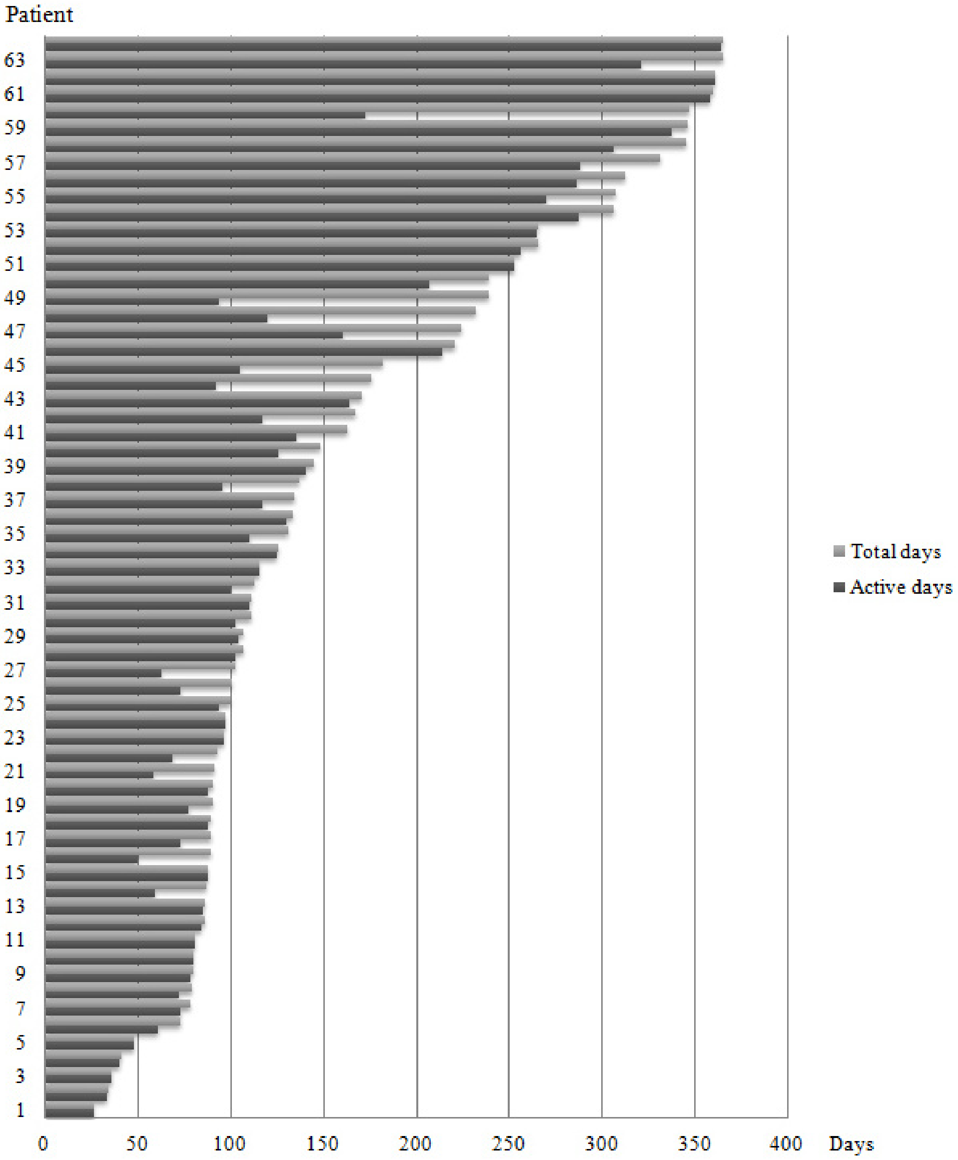


Figure 2. Duration of Fitbit use. Total days (gray) in relation to active days (black) of use for each patient.



Step Measurements

The number of steps was assessed using the Fitbit Zip step counter (Fitbit Inc, San Francisco, CA, USA) [37].

Patients were asked to fasten the Fitbit at the breast pocket or hip during all waking hours (except for bathing and swimming) for at least 3 months after hospital discharge and for a period of up to 1 year. Before being used, the Fitbit was programmed with the patient’s date of birth, gender, weight, and height to ensure accuracy of the step counts obtained. Fitbit Zip uses a 3-axis accelerometer and converts accelerometer data into step data. Step data were continually visible on the Fitbit’s display; each day at midnight, the steps per day were downloaded to the patient’s PHR. For statistical analysis, each patient’s step data were downloaded on a secure database at intervals of 1 minute for 365 days from the day of inclusion. Bäck et al [35] proposed a graduated step index to describe activity in cardiac patients: (1) <3000 steps per day (low activity/sedentary), (2) 3000-9999 steps per day (medium activity), and (3) ≥10,000 steps per day (high activity). This classification was used in the presentation of data.

Statistical Analysis

Means with standard deviation are presented for continuous variables and proportions (%) for categorical variables. Mean body mass index (BMI) was measured as the difference between baseline BMI and BMI at 3 months divided by 2.

The grand mean for every patient was measured as the mean of all the patients’ active days. *Active days* were defined as days with more than 100 steps per day. The low cut-off point of ≥100 steps per day was set to capture very low activity and still avoid failure measurement when the Fitbit was moved around and not worn. Activity of less than 100 steps per day was considered to indicate moving the Fitbit around but not wearing it. Furthermore, steps per day was also measured as a weekly mean at day 7 and at 1 month, 3 months, 6 months, 9 months, and 12 months. Patients who stopped using the Fitbit were called “nonusers” from the time they stopped.

Duration of use was measured as total days, counted from the starting day to the final day of Fitbit use. The final day was determined as the last day of ≥4 consecutive active days of Fitbit recordings despite any temporary break in use. If the patients had less than four active days of consecutive Fitbit recordings, the use was considered terminated.

The association between Fitbit groups and baseline characteristics was tested using 1-way ANOVA for continuous data values and the Fisher exact test for categorical data. In the case of significant difference, a post hoc Bonferroni test was carried out. For nonnormally distributed continuous data, a Kruskal-Wallis test was done (days and total days) and a *t* test was performed for gender-sorted grand mean of steps walked. To test association between mean numbers of steps walked and termination of step counter use, a repeated measure logistic regression analysis was carried out. All tests were considered statistically significant if *P*<.05. MATLAB release 2014b (MathWorks, Natick, MA, USA) and STATA version 13.1 (StataCorp, College Station, TX, USA) were used for statistical analyses.

Results

Patient Characteristics

Patient baseline characteristics are shown in Table 1. Of the 64 patients in our sample, 14 (22%) were classified as low active with a mean of <3000 steps per day and 14% (9/64) were highly active, walking ≥10,000 steps per day. The remaining 41 patients (64%) were medium active, walking between 3000 and 9999 steps per day. The mean age of the entire sample was 62.8 years (range 35-88 years). There was a significant difference in age between the activity groups (*P*=.01). Patients in the low activity group were significantly older than patients in both the medium activity group (*P*=.03) and the high activity group (*P*=.02). The mean ages in the three activity groups of low, medium, and high were 70.7 (SD 10.7), 61.1 (SD 11.4), and 58.2 (SD 8.3) years, respectively. Males represented 51 of 64 (80%) participants. Even though the less active patients' mean BMI was higher, this was not significant. Almost half (48%, 31/64) of the patients had a primary diagnosis of ACS and 21 of 64 patients (33%) were treated with surgery (CABG or valve replacement/mitral valve replacement). Eight patients (11%) had heart failure and five patients (8%) had both heart failure and ACS. For cardiac telerehabilitation, 29 of 64 patients (45%) chose the health care center and 23 of 64 (36%) chose the hospital. The remaining 12 patients (19%) chose the call center.

Table 1. Patient baseline characteristics.

Characteristic[a]	All patients	Activity level (steps/day)			*P* [b]
		Low (<2999)	Medium (3000-9999)	High (≥10,000)	
Demographic variables					
Participants, n (%)	64 (100)	14 (22)	41 (64)	9 (14)	
Age (years), mean (SD)	62.8 (11.5)	70.7 (10.7)	61.1 (11.4)	58.2 (8.3)	.01[c]
Sex, n (%)					.29
Male	51 (80)	13 (20)	30 (47)	8 (13)	
Female	13 (20)	1 (2)	11 (17)	1 (2)	
BMI (kg/m^2), mean (SD)	28 (5.1)	29.7 (5.1)	27.7 (5.4)	27.0 (4.2)	.38
Primary diagnosis or treatment, n (%)					
ACS	33 (48)	6 (9)	20 (31)	7 (11)	
Surgery[d]	18 (33)	6 (9)	12 (19)	0 (0)	
Heart failure	8 (11)	1 (2)	5 (8)	2 (3)	
ACS & heart failure	5 (8)	1 (2)	4 (6)	0 (0)	
Cardiac telerehabilitation, n (%)					.08
Health care center	29 (45)	7 (11)	20 (31)	2 (3)	
Hospital	23 (36)	3 (5)	17 (27)	3 (5)	
Call center	12 (19)	4 (6)	4 (6)	4 (6)	

[a] ACS: acute coronary syndrome; BMI: body mass index.

[b] *P* value for comparison of all three activity groups (low, medium, and high).

[c] Post hoc Bonferroni corrected values: low versus medium activity groups (*P*=.03), low versus high activity groups (*P*=.02), and medium versus high activity groups (*P*<.99).

[d] Surgery includes valve replacement, mitral valve repair, and coronary artery bypass grafting.

Duration of Step Counter Use

Two patients used the Fitbit for a total of 365 days each; the overall mean total days was 160 (SD 100, range 26-365 days). Patients in the low, medium, and high activity groups used the pedometer for a mean of 109 (SD 56), 168 (SD 103), and 208 (SD 112) days, respectively. There was a significant difference between both the low and high activity groups in total days of Fitbit use (*P*=.01). Active days comprised 139 (SD 93) of 160 (87%) total days. There was a significant difference between both the low and medium activity groups (*P*=.01) and the low and high activity groups (*P*=.003) in active days of Fitbit use (Table 2).

Table 2. Duration of step counter use and mean daily steps.

Step counter use	All patients	Activity level (steps/day)			*P*
		Low (<2999)	Medium (3000-9999)	High (≥10,000)	
Duration of use					
Total days, mean (SD)	160 (100)	109 (56)	168 (103)	208 (112)	.04[a,b]
Active days, mean (SD)	139 (93)	79 (26)	148 (97)	189 (102)	.006[a,c]
Active days/total days, %	87	72	88	91	
Walking activity (steps/day), mean (SD)					
Grand mean	5899 (3151)	1996 (716)	6016 (1784)	11,439 (440)	
Gender					.82[a]
Male	5853 (3274)	2064 (696)	6008 (1785)	11,430 (469)	
Female	6078 (2725)	1105 (0)	6037 (1869)	11,501 (0)	
Week					.004[d]
Week 1 (7 days)	5191 (3198)	1578 (500)	5366 (2306)	9611 (2995)	
Week 4 (30 days)	6362 (3834)	1807 (780)	6305 (2536)	12,697 (1678)	
Week 13 (90 days)	6186 (3013)	2304 (1189)	6073 (2271)	10,637 (1095)	
Week 26 (180 days)	6794 (3518)	808 (0)	6506 (3569)	9011 (884)	
Week 39 (270 days)	8235 (4220)	0 (0)	5960 (2461)	12,784 (3127)	
Week 52 (365 days)	7890 (2629)	0 (0)	7426 (2730)	9050 (2811)	
Primary diagnose or treatment					.12[a]
ACS[e]	6549 (3149)	2194 (712)	6127 (1558)	11,588 (354)	
Surgery[f]	4781 (3023)	1879 (653)	6058 (845)	0 (0)	
Heart failure	7340 (3190)	2636 (0)	5949 (2605)	10,916 (306)	
ACS & heart failure	4505 (163)	865 (0)	5415 (845)	0 (0)	
Cardiac telerehabilitation					.34[a]
Health care center	5324 (2579)	2197 (585)	5779 (1243)	11,737 (333)	
Hospital	6128 (3084)	2082 (980)	5950 (2242)	11,187 (518)	
Call center	6847 (4353)	1578 (751)	7485 (1608)	11,478 (418)	

[a] *P* value for comparison of all three activity groups (low, medium, and high).

[b] Post hoc 2-sample Wilcoxon rank-sum test values: low versus medium activity groups (*P*=.05), low versus high activity groups (*P*=.01), and medium versus high activity groups (*P*=.22).

[c] Post hoc 2-sample Wilcoxon rank-sum values: low versus medium activity groups (*P*=.01), low versus high activity groups (*P*=.003), and medium versus high activity groups (*P*=.21).

[d] *P* value for correlation between termination of step counter use and mean steps walked at the specified days.

[e] ACS: acute coronary syndrome.

[f] Surgery includes valve replacement, mitral valve repair, and coronary artery bypass grafting.

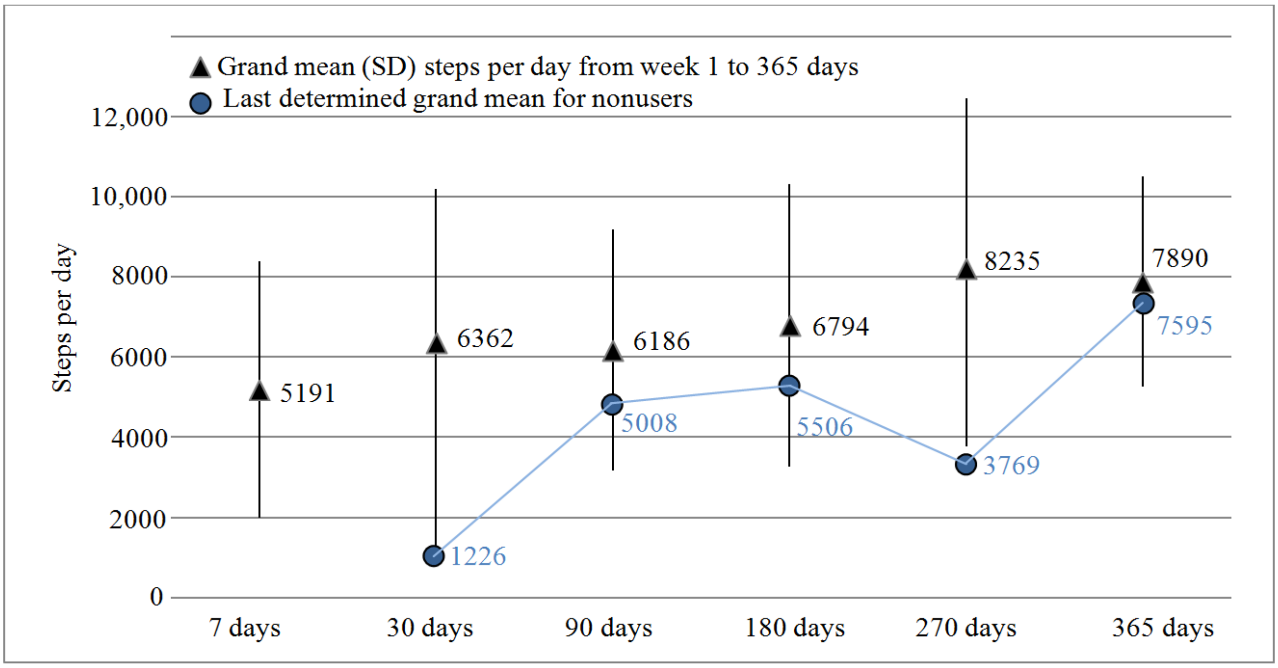


Figure 3. Grand mean steps per day and standard deviation (length of whiskers) at different days for both users and nonusers.

Step Measurement

The grand mean for all patients for all active days was 5899 (SD 3151) steps per day. The grand mean steps per day were as follows: day 7: mean 5191, SD 3198; day 30: mean 6362, SD 3834; day 90: mean 6186, SD 3013; day 180: mean 6794, SD 3518; day 270: mean 8235, SD 4220; and day 365: mean 7890, SD 2629 steps per day indicating an increase in walking activity over time (Table 2). Knowing that the increase in walking activity could be a result of the low activity patients' termination of step counter use, the week mean of nonusers (the last determined) was calculated (Figure 3) together with the increase in patients' grand mean. In addition, a linear regression revealed a significant relationship between termination of step counter use and low week mean steps at different weeks (*P*=.004) (Table 2). Despite the slight increase in nonusers' weekly means, it cannot be ruled out that the increase in steps per week over the year might be due to the dropping out of those patients with low walking activity.

No significant association was found between gender and mean steps per day (*P*=.82) in the different activity groups. There were no significant differences in mean steps per day between the four treatment groups (*P*=.12). Patients who choose the call center for cardiac telerehabilitation had the highest mean steps per day (mean 6847, SD 4353) and patients using the health care center for cardiac telerehabilitation had the lowest (mean 5324, SD 2579 steps/day). There was no significant relation between choice of rehabilitation setting and mean steps per day (*P*=.34) (Table 2 and Figure 4).

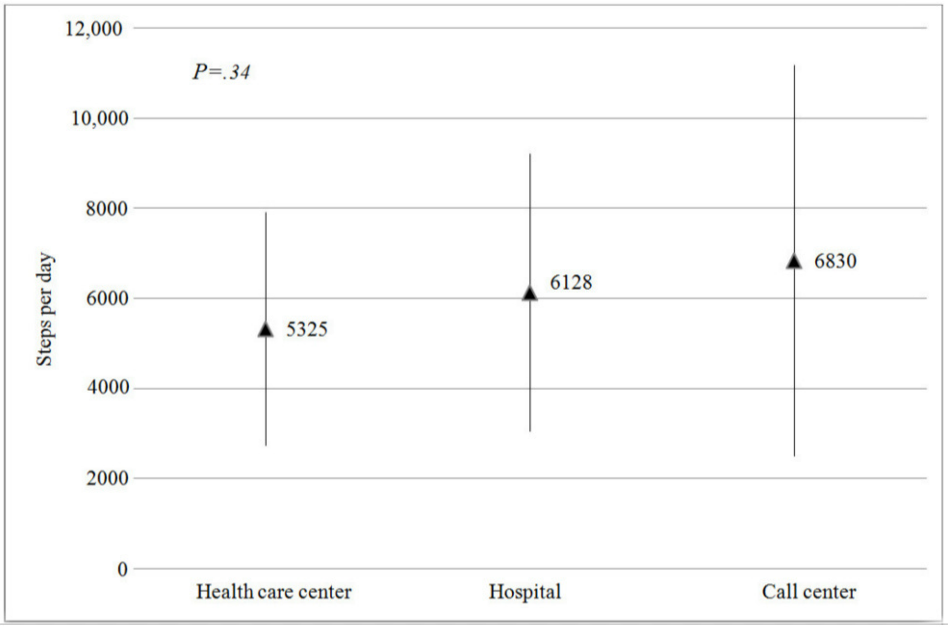


Figure 4. Mean steps per day and standard deviation (length of whiskers) at the three cardiac telerehabilitation settings.

Discussion

Principal Results

We found a significant correlation in age in the step activity groups. The patients used the step counter for a mean of 160 (SD 100) days and there was a significantly lower duration of step counter use in the low activity group than in the medium and high activity groups. The overall mean number of daily steps was 5899 (SD 3151), increasing slightly from mean 5191 (SD 3198) steps per day at the first week to mean 7890 (SD 2629) steps per day after 1 year. Nevertheless, this increase might be due to dropout of patients with low activity. Furthermore, the less physically active patients weighed more. There were no significant differences in patients' treatment or rehabilitation in the three activity groups.

The patients were given the opportunity to wear the Fitbit for up to 365 days after the cardiac event; the mean total days were 160 (SD 100), of which 87% (139/160) were active days. Inactive days might be due to forgetfulness or a decision not to wear the Fitbit every day. A closer look at the data revealed that some patients had 7 to 14 consecutive days without step recordings, which might indicate holiday trips with no Fitbit use. Qualitative studies of the patients' behavior and motivation in relation to Fitbit wearing are needed to shed light on these issues.

We can compare our results with a study by Izawa et al [18] of cardiac patients who walked a mean of 8609 steps per day 1 month after discharge. In our study, the mean at 30 days was 6362 steps, indicating that our patients were less active at 1 month than those in Izawa et al's study. The same pattern was observed throughout the year. The mean for the patients in our study was up to 25% fewer steps per day than cardiac patients in comparable studies [15,16,18,21,35,38,39]. The main difference in the studies was the continuity of step counter use. In all studies except ours, the step counter was given to the patient for the first month or for 1 week before the time of measurement (at 1, 3, 6, 9, or 12 months), whereas patients in our study retained the Fitbit for 365 days. This may explain the discrepancy between the studies. There is reason to believe that our findings represent a more realistic picture of walking activity because patients might change their behavior to what is expected on the days of measurement, meaning that they may walk more than usual. This effect might have been eliminated due to the continuous wearing of the Fitbit in our study. None of the previously mentioned studies used the Fitbit Zip, which may prevent adequate comparison of step results. Accuracy studies on Fitbit Zip in healthy adults have revealed satisfactory step measurements in free-living physical activity [40,41], but slow walking speed seems to provide inaccurate step measurements [42]. Older people (>70 years) [43] and heart failure patients [44] walk at a slow speed, which might hamper the Fitbit Zip's capability to measure steps accurately. Studies of Fitbit Zip's accuracy when used by cardiac patients are needed.

The explanation for the significant association between long-term use of the Fitbit and high step activity was not identified in this study. In line with other studies [18,21,25,27-31], this might indicate that Fitbit users were encouraged to increase their walking activity. The telerehabilitation setting also provided the patients with goals for daily steps and an opportunity to monitor and follow their own walking activity. In another of Teledi@log's substudies, Thorup et al [45] found that self-monitoring of steps provided a conscious awareness of walking activity due to the immediate feedback on step activity. In our study, males represented 51 of 64 (80%) participants, almost the same as in comparable studies [12,35,39,46]. The correlation between high BMI and low steps per day has been seen in other studies [13,15,24] and may be considered a health problem for cardiac patients [3]. The significant negative relation between increasing age and steps per day is also evident in other studies [21,25,32,46]. In our study, 22% of the patients were classified as low active (mean <3000 steps/day) and 14% were highly active (walking ≥10,000 steps/day). The remaining 64% were medium active and walked between 3000 and 9999 steps per day. The recommendation of 10,000 daily steps to achieve health benefits appears to be a reasonable estimate of daily activity for healthy adults, but this goal may be too ambitious for people with cardiac disease [17,26,46]. Research suggests that a target of approximately 7000 [35] to 7500 [17] daily steps might reduce waist circumference, BMI, and cardiovascular disease risk factors in patients with coronary artery disease [17,35]. Bearing in mind the dose-response relationship between physical activity and health status [46-48], 7500 daily steps may not be sufficient to reach optimal health status in cardiac patients [17], yet it could be the starting point to improving their physical activity levels. However, experts do not agree on the number of steps needed per day for cardiac patients to provide better health in secondary prevention of cardiac disease. Similar to other studies [13,15,46], our study found no significant association between gender and mean steps per day. Patients who choose cardiac telerehabilitation through the call center did not follow group exercises, whereas patients who chose cardiac rehabilitation at either the health care center or the hospital did follow group exercises. Despite the lack of group exercise, they had the highest mean steps per day (mean 6847, SD 4353) compared to the other cardiac telerehabilitation settings (although the difference was not significant). This might indicate that cardiac telerehabilitation at a call center can support walking activity just as effectively as cardiac telerehabilitation at hospitals and health care centers. In line with this, Thorup et al [45] found that the Fitbit led to self-monitoring, which then led to independence of standardized rehabilitation programs.

Limitations

Despite the strength of our study, in that the patients' steps were monitored continuously for 1 year, the results must be viewed cautiously due to the low number of participants. Although Fitbits provide an objective measurement of physical activity, they are not designed to capture different modes of physical activity (eg, cycling and swimming). Furthermore, Fitbits are considered less valid during slow walking and in obese patients [49,50], and there is reason to believe that cardiac patients might have a slower walking pace [24,44]. The Fitbit was also part of a larger telerehabilitation program; therefore, other factors might have influenced the patients' activity.

Conclusion

This study has demonstrated that cardiac patients in the Teledi@log program walked a mean of 5899 steps per day in the year after a cardiac event, increasing from a mean 5191 steps per day at the first week to a mean 7890 steps per day after 1 year. In this study, the patients tended to walk less than cardiac patients in comparable studies, but there is reason to believe that this study represents a more realistic picture of walking activity due to the continuation of Fitbit use (mean 160 days). The patients who walked more tended to be of younger age, had a lower BMI, and continued using the Fitbit for a longer period. There were no significant differences in mean steps per day for patients based on their type of treatment or rehabilitation setting. Qualitative studies on the patients' behavior and motivation regarding Fitbit use are needed to shed light on adherence to and motivation to use the Fitbit.

Acknowledgments

The research was funded by Eir Research and Business Park, The European Regional Development Fund, UNIK Partnership, Danish Council of Research and Innovation, Departments of Cardiothoracic Surgery and Clinical Nursing Research Unit, Aalborg University Hospital, Aalborg University, and the Lundbeck Foundation (FP25/2013). The authors wish to thank the research team behind Teledi@log for support on data collection and analysis throughout the research and writing of this paper.

Conflicts of Interest

None declared.

References

1. World Health Organization. 2014 May. The top 10 causes of death URL: http://www.who.int/mediacentre/factsheets/fs310/en/ [accessed 2015-09-09] [WebCite Cache ID 6bPtI3I4T]
2. Anderson L, Taylor RS. Cardiac rehabilitation for people with heart disease: an overview of Cochrane systematic reviews. Cochrane Database Syst Rev 2014;12:CD011273. [doi: 10.1002/14651858.CD011273.pub2] [Medline: 25503364]
3. Piepoli MF, Corrà U, Benzer W, Bjarnason-Wehrens B, Dendale P, Gaita D, Cardiac Rehabilitation Section of the European Association of Cardiovascular PreventionRehabilitation. Secondary prevention through cardiac rehabilitation: from knowledge to implementation. A position paper from the Cardiac Rehabilitation Section of the European Association of Cardiovascular Prevention and Rehabilitation. Eur J Cardiovasc Prev Rehabil 2010 Feb;17(1):1-17. [doi: 10.1097/HJR.0b013e3283313592] [Medline: 19952757]
4. Beatty AL, Fukuoka Y, Whooley MA. Using mobile technology for cardiac rehabilitation: a review and framework for development and evaluation. J Am Heart Assoc 2013;2(6):e000568 [FREE Full text] [doi: 10.1161/JAHA.113.000568] [Medline: 24185949]
5. Scalvini S, Zanelli E, Comini L, Dalla TM, Troise G, Febo O, et al. Home-based versus in-hospital cardiac rehabilitation after cardiac surgery: a nonrandomized controlled study. Phys Ther 2013 Aug;93(8):1073-1083 [FREE Full text] [doi: 10.2522/ptj.20120212] [Medline: 23599353]
6. Varnfield M, Karunanithi MK, Särelä A, Garcia E, Fairfull A, Oldenburg BF, et al. Uptake of a technology-assisted home-care cardiac rehabilitation program. Med J Aust 2011 Feb 21;194(4):S15-S19. [Medline: 21401482]
7. Frederix I, Hansen D, Coninx K, Vandervoort P, Vandijck D, Hens N, et al. Medium-term effectiveness of a comprehensive Internet-based and patient-specific telerehabilitation program with text messaging support for cardiac patients: randomized controlled trial. J Med Internet Res 2015;17(7):e185 [FREE Full text] [doi: 10.2196/jmir.4799] [Medline: 26206311]
8. Frederix I, Vanhees L, Dendale P, Goetschalckx K. A review of telerehabilitation for cardiac patients. J Telemed Telecare 2015 Jan;21(1):45-53. [doi: 10.1177/1357633X14562732] [Medline: 25475219]
9. Munro J, Angus N, Leslie SJ. Patient focused Internet-based approaches to cardiovascular rehabilitation--a systematic review. J Telemed Telecare 2013 Sep;19(6):347-353. [doi: 10.1177/1357633X13501763] [Medline: 24163299]
10. Taylor RS, Dalal H, Jolly K, Zawada A, Dean SG, Cowie A, et al. Home-based versus centre-based cardiac rehabilitation. Cochrane Database Syst Rev 2015;8:CD007130. [doi: 10.1002/14651858.CD007130.pub3] [Medline: 26282071]
11. Huang K, Liu W, He D, Huang B, Xiao D, Peng Y, et al. Telehealth interventions versus center-based cardiac rehabilitation of coronary artery disease: a systematic review and meta-analysis. Eur J Prev Cardiol 2014 Dec 8:959-971. [doi: 10.1177/2047487314561168] [Medline: 25488550]
12. Varnfield M, Karunanithi M, Lee C, Honeyman E, Arnold D, Ding H, et al. Smartphone-based home care model improved use of cardiac rehabilitation in postmyocardial infarction patients: results from a randomised controlled trial. Heart 2014 Nov;100(22):1770-1779 [FREE Full text] [doi: 10.1136/heartjnl-2014-305783] [Medline: 24973083]
13. Blanchard CM, Giacomantonio N, Lyons R, Cyr C, Rhodes RE, Reid RD, et al. Examining the steps-per-day trajectories of cardiac rehabilitation patients: a latent class growth analysis perspective. J Cardiopulm Rehabil Prev 2014;34(2):106-113. [doi: 10.1097/HCR.0000000000000031] [Medline: 24326902]
14. Taylor RS, Sagar VA, Davies EJ, Briscoe S, Coats AJ, Dalal H, et al. Exercise-based rehabilitation for heart failure. Cochrane Database Syst Rev 2014;4:CD003331. [doi: 10.1002/14651858.CD003331.pub4] [Medline: 24771460]

15. Brändström Y, Brink E, Grankvist G, Alsén P, Herlitz J, Karlson BW. Physical activity six months after a myocardial infarction. Int J Nurs Pract 2009 Jun;15(3):191-197. [doi: 10.1111/j.1440-172X.2009.01744.x] [Medline: 19531077]
16. Houle J, Doyon O, Vadeboncoeur N, Turbide G, Diaz A, Poirier P. Innovative program to increase physical activity following an acute coronary syndrome: randomized controlled trial. Patient Educ Couns 2011 Dec;85(3):e237-e244. [doi: 10.1016/j.pec.2011.03.018] [Medline: 21546203]
17. Houle J, Valera B, Gaudet-Savard T, Auclair A, Poirier P. Daily steps threshold to improve cardiovascular disease risk factors during the year after an acute coronary syndrome. J Cardiopulm Rehabil Prev 2013;33(6):406-410. [doi: 10.1097/HCR.0000000000000021] [Medline: 24104407]
18. Izawa KP, Watanabe S, Oka K, Hiraki K, Morio Y, Kasahara Y, et al. Relation between physical activity and exercise capacity of ≥5 metabolic equivalents in middle- and older-aged patients with chronic heart failure. Disabil Rehabil 2012;34(23):2018-2024. [doi: 10.3109/09638288.2012.667502] [Medline: 22458333]
19. Lawler PR, Filion KB, Eisenberg MJ. Efficacy of exercise-based cardiac rehabilitation post-myocardial infarction: a systematic review and meta-analysis of randomized controlled trials. Am Heart J 2011 Oct;162(4):571-584. [doi: 10.1016/j.ahj.2011.07.017] [Medline: 21982647]
20. Heran BS, Chen JM, Ebrahim S, Moxham T, Oldridge N, Rees K, et al. Exercise-based cardiac rehabilitation for coronary heart disease. Cochrane Database Syst Rev 2011(7):CD001800 [FREE Full text] [doi: 10.1002/14651858.CD001800.pub2] [Medline: 21735386]
21. Savage PD, Ades PA. Pedometer step counts predict cardiac risk factors at entry to cardiac rehabilitation. J Cardiopulm Rehabil Prev 2008;28(6):370-377. [doi: 10.1097/HCR.0b013e31818c3b6d] [Medline: 19008690]
22. Worringham C, Rojek A, Stewart I. Development and feasibility of a smartphone, ECG and GPS based system for remotely monitoring exercise in cardiac rehabilitation. PLoS One 2011;6(2):e14669 [FREE Full text] [doi: 10.1371/journal.pone.0014669] [Medline: 21347403]
23. Goodrich DE, Larkin AR, Lowery JC, Holleman RG, Richardson CR. Adverse events among high-risk participants in a home-based walking study: a descriptive study. Int J Behav Nutr Phys Act 2007;4:20 [FREE Full text] [doi: 10.1186/1479-5868-4-20] [Medline: 17521443]
24. Tudor-Locke C, Craig CL, Aoyagi Y, Bell RC, Croteau KA, De Bourdeaudhuij I, et al. How many steps/day are enough? For older adults and special populations. Int J Behav Nutr Phys Act 2011;8:80 [FREE Full text] [doi: 10.1186/1479-5868-8-80] [Medline: 21798044]
25. Bravata DM, Smith-Spangler C, Sundaram V, Gienger AL, Lin N, Lewis R, et al. Using pedometers to increase physical activity and improve health: a systematic review. JAMA 2007 Nov 21;298(19):2296-2304. [doi: 10.1001/jama.298.19.2296] [Medline: 18029834]
26. Tudor-Locke C, Washington TL, Hart TL. Expected values for steps/day in special populations. Prev Med 2009 Aug;49(1):3-11. [doi: 10.1016/j.ypmed.2009.04.012] [Medline: 19409409]
27. Trinh L, Wilson R, Williams HM, Sum AJ, Naylor P. Physicians promoting physical activity using pedometers and community partnerships: a real world trial. Br J Sports Med 2012 Mar;46(4):284-290. [doi: 10.1136/bjsm.2009.069567] [Medline: 21282802]
28. Korkiakangas EE, Alahuhta MA, Husman PM, Keinänen-Kiukaanniemi S, Taanila AM, Laitinen JH. Pedometer use among adults at high risk of type 2 diabetes, Finland, 2007-2008. Prev Chronic Dis 2010 Mar;7(2):A37 [FREE Full text] [Medline: 20158965]
29. Vaes AW, Cheung A, Atakhorrami M, Groenen MT, Amft O, Franssen FM, et al. Effect of 'activity monitor-based' counseling on physical activity and health-related outcomes in patients with chronic diseases: A systematic review and meta-analysis. Ann Med 2013 Sep;45(5-6):397-412. [doi: 10.3109/07853890.2013.810891] [Medline: 23952917]
30. Kurti AN, Dallery J. Internet-based contingency management increases walking in sedentary adults. J Appl Behav Anal 2013;46(3):568-581. [doi: 10.1002/jaba.58] [Medline: 24114220]
31. Evangelista LS, Dracup K, Erickson V, McCarthy WJ, Hamilton MA, Fonarow GC. Validity of pedometers for measuring exercise adherence in heart failure patients. J Card Fail 2005 Jun;11(5):366-371. [Medline: 15948087]
32. Dontje ML, van der Wal MH, Stolk RP, Brügemann J, Jaarsma T, Wijtvliet PE, et al. Daily physical activity in stable heart failure patients. J Cardiovasc Nurs 2014;29(3):218-226. [doi: 10.1097/JCN.0b013e318283ba14] [Medline: 23416939]
33. 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]
34. Annegarn J, Spruit MA, Uszko-Lencer NH, Vanbelle S, Savelberg HH, Schols AM, et al. Objective physical activity assessment in patients with chronic organ failure: a validation study of a new single-unit activity monitor. Arch Phys Med Rehabil 2011 Nov;92(11):1852-1857. [doi: 10.1016/j.apmr.2011.06.021] [Medline: 22032218]
35. Bäck M, Cider A, Gillström J, Herlitz J. Physical activity in relation to cardiac risk markers in secondary prevention of coronary artery disease. Int J Cardiol 2013 Sep 20;168(1):478-483. [doi: 10.1016/j.ijcard.2012.09.117] [Medline: 23041099]
36. Sundhedsstyrelsen. National klinisk retningslinje for hjerterehabilitering. In: Sundhedsstyrelsen. KÃƒÆ’Ã‚Â¸benhavn S: Sundhedsstyrelsen; Feb 20, 2015:3-62.
37. Fitbit. URL: https://www.fitbit.com/uk/zip [accessed 2016-03-08] [WebCite Cache ID 6frbOneIj]

38. Devi R, Powell J, Singh S. A web-based program improves physical activity outcomes in a primary care angina population: randomized controlled trial. J Med Internet Res 2014;16(9):e186 [FREE Full text] [doi: 10.2196/jmir.3340] [Medline: 25217464]
39. Antypas K, Wangberg SC. An Internet- and mobile-based tailored intervention to enhance maintenance of physical activity after cardiac rehabilitation: short-term results of a randomized controlled trial. J Med Internet Res 2014;16(3):e77 [FREE Full text] [doi: 10.2196/jmir.3132] [Medline: 24618349]
40. Tully MA, McBride C, Heron L, Hunter RF. The validation of Fibit Zip™ physical activity monitor as a measure of free-living physical activity. BMC Res Notes 2014;7:952 [FREE Full text] [doi: 10.1186/1756-0500-7-952] [Medline: 25539733]
41. 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;12:42 [FREE Full text] [doi: 10.1186/s12966-015-0201-9] [Medline: 25890168]
42. Beevi FH, Miranda J, Pedersen CF, Wagner S. An evaluation of commercial pedometers for monitoring slow walking speed populations. Telemed J E Health 2015 Oct 9:e1. [doi: 10.1089/tmj.2015.0120] [Medline: 26451900]
43. Kuys SS, Peel NM, Klein K, Slater A, Hubbard RE. Gait speed in ambulant older people in long term care: a systematic review and meta-analysis. J Am Med Dir Assoc 2014 Mar;15(3):194-200. [doi: 10.1016/j.jamda.2013.10.015] [Medline: 24388775]
44. Pepera GK, Sandercock GR, Sloan R, Cleland JJ, Ingle L, Clark AL. Influence of step length on 6-minute walk test performance in patients with chronic heart failure. Physiotherapy 2012 Dec;98(4):325-329. [doi: 10.1016/j.physio.2011.08.005] [Medline: 23122439]
45. Thorup C, Groenkjaer M, Spindler H, Andreasen J, Hansen J, Dinesen B, et al. Pedometer use as motivation for physical activity in cardiac tele-rehabilitation. In: Int J Integr Care. 2015 Oct 20 Presented at: European Telemedicien Conference (ETC); Oct 21-22, 2015; Odense, Denmark URL: http://www.ijic.org/articles/abstract/2288/
46. Ayabe M, Brubaker PH, Dobrosielski D, Miller HS, Kiyonaga A, Shindo M, et al. Target step count for the secondary prevention of cardiovascular disease. Circ J 2008 Feb;72(2):299-303 [FREE Full text] [Medline: 18219170]
47. Garber CE, Blissmer B, Deschenes MR, Franklin BA, Lamonte MJ, Lee I, American College of Sports Medicine. American College of Sports Medicine position stand. Quantity and quality of exercise for developing and maintaining cardiorespiratory, musculoskeletal, and neuromotor fitness in apparently healthy adults: guidance for prescribing exercise. Med Sci Sports Exerc 2011 Jul;43(7):1334-1359. [doi: 10.1249/MSS.0b013e318213fefb] [Medline: 21694556]
48. Hills AP, Street SJ, Byrne NM. Physical activity and health: "what is old is new again". Adv Food Nutr Res 2015;75:77-95. [doi: 10.1016/bs.afnr.2015.06.001] [Medline: 26319905]
49. Jehn M, Schmidt-Trucksäess A, Schuster T, Hanssen H, Weis M, Halle M, et al. Accelerometer-based quantification of 6-minute walk test performance in patients with chronic heart failure: applicability in telemedicine. J Card Fail 2009 May;15(4):334-340. [doi: 10.1016/j.cardfail.2008.11.011] [Medline: 19398082]
50. Jehn M, Schmidt-Trucksäss A, Schuster T, Hanssen H, Halle M, Köhler F. Pedometer accuracy in patients with chronic heart failure. Int J Sports Med 2010 Mar;31(3):186-191. [doi: 10.1055/s-0029-1243641] [Medline: 20166006]

Abbreviations

ACS: acute coronary syndrome
BMI: body mass index
CABG: coronary artery bypass graft
PHR: personal health record

Edited by G Eysenbach; submitted 02.10.15; peer-reviewed by I Frederix, N Farre; comments to author 26.10.15; revised version received 26.12.15; accepted 04.01.16; published 04.04.16.

Please cite as:
Thorup C, Hansen J, Grønkjær M, Andreasen JJ, Nielsen G, Sørensen EE, Dinesen BI
Cardiac Patients' Walking Activity Determined by a Step Counter in Cardiac Telerehabilitation: Data From the Intervention Arm of a Randomized Controlled Trial
J Med Internet Res 2016;18(4):e69
URL: http://www.jmir.org/2016/4/e69/
doi:10.2196/jmir.5191
PMID:27044310

Original Paper

Efficacy and External Validity of Electronic and Mobile Phone-Based Interventions Promoting Vegetable Intake in Young Adults: Systematic Review and Meta-Analysis

Monica Nour[1], BSc, MNutDiet; Juliana Chen[1], BSc, MNutDiet; Margaret Allman-Farinelli[1], PhD, FDAA

[1] School of Life and Environmental Sciences, University of Sydney, Sydney, NSW, Australia

Corresponding Author:
Monica Nour, BSc, MNutDiet
School of Life and Environmental Sciences
University of Sydney
Level 4 East, Charles Perkins Centre D17
The University of Sydney
Sydney, NSW, 2006
Australia
Phone: 61 403094095
Fax: 61 86271605
Email: mnou2973@uni.sydney.edu.au

Abstract

Background: Young adults (18–35 years) remain among the lowest vegetable consumers in many western countries. The digital era offers opportunities to engage this age group in interventions in new and appealing ways.

Objective: This systematic review evaluated the efficacy and external validity of electronic (eHealth) and mobile phone (mHealth) -based interventions that promote vegetable intake in young adults.

Methods: We searched several electronic databases for studies published between 1990 and 2015, and 2 independent authors reviewed the quality and risk of bias of the eligible papers and extracted data for analyses. The primary outcome of interest was the change in vegetable intake postintervention. Where possible, we calculated effect sizes (Cohen *d* and 95% CIs) for comparison. A random effects model was applied to the data for meta-analysis. Reach and representativeness of participants, intervention implementation, and program maintenance were assessed to establish external validity. Published validation studies were consulted to determine the validity of tools used to measure intake. We applied the Grading of Recommendations Assessment, Development and Evaluation (GRADE) system to evaluate the overall quality of the body of evidence.

Results: Of the 14 studies that met the selection criteria, we included 12 in the meta-analysis. In the meta-analysis, 7 studies found positive effects postintervention for fruit and vegetable intake, Cohen *d* 0.14–0.56 (pooled effect size 0.22, 95% CI 0.11–0.33, I^2=68.5%, *P*=.002), and 4 recorded positive effects on vegetable intake alone, Cohen *d* 0.11–0.40 (pooled effect size 0.15, 95% CI 0.04–0.28, I^2=31.4%, *P*=.2). These findings should be interpreted with caution due to variability in intervention design and outcome measures. With the majority of outcomes documented as a change in combined fruit and vegetable intake, it was difficult to determine intervention effects on vegetable consumption specifically. Measurement of intake was most commonly by self-report, with 5 studies using nonvalidated tools. Longer-term follow-up was lacking from most studies (n=12). Risk of bias was high among the included studies, and the overall body of evidence was rated as low quality. The applicability of interventions to the broader young adult community was unclear due to poor description of external validity components.

Conclusions: Preliminary evidence suggests that eHealth and mHealth strategies may be effective in improving vegetable intake in young adults; whether these small effects have clinical or nutritional significance remains questionable. With studies predominantly reporting outcomes as fruit and vegetable intake combined, we suggest that interventions report vegetables separately. Furthermore, to confidently establish the efficacy of these strategies, better-quality interventions are needed for young adults, using valid measures of intake, with improved reporting on costs, sustainability and long-term effects of programs.

Trial registration: PROSPERO International Prospective Register of Systematic Reviews: CRD42015017763; http://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42015017763 (Archived by WebCite at http://www.webcitation.org/6fLhMgUP4)

(J Med Internet Res 2016;18(4):e58) doi:10.2196/jmir.5082

KEYWORDS

young adults; vegetable consumption; mHealth; eHealth; social marketing

Introduction

Poor fruit and vegetable intake contributes to 2.635 million deaths per year [1]. Consuming the recommended 600 g daily could reduce this global burden by 1.8% [1], with adequate fruit and vegetable intake linked to minimized adiposity, improved weight management [2], and reduced risk of heart disease and some cancers [1]. Despite several decades of government-led social marketing campaigns, alongside concerted effort by researchers and practitioners to facilitate behavior change, intake of vegetables remains suboptimal in many countries [3-6].

Australian young adults (18-34 years) are among the lowest consumers of vegetables, with only 4.7% consuming the recommended 5 or more servings a day [7]. During this transitional phase of life, young adults are developing self-determined food habits that will affect their future health. While the association between fruit and vegetable consumption and reduced chronic disease risk is well established in the literature [2,8-15], promoting these long-term health benefits, as is typically done in nationwide social marketing campaigns, does not appear to motivate young adults [16,17]. Young adults are typically less concerned about their future well-being and engage in more risky health behaviors [18]. Consequently, this population should be targeted separately in interventions.

Research in the area of digital interventions has revealed that electronic (eHealth) and mobile phone (mHealth) -based strategies are effective in promoting healthful behaviors [19-21]. eHealth and mHealth refer to the use of the Internet, mobile, or wireless devices to deliver health services and information to improve health outcomes or enhance health research [22,23]. Examples of eHealth and mHealth strategies include text messaging, email, mobile phone apps, phone calls, and websites. Young adults are among the highest users of mobile phones and wireless information sharing platforms [24], with 89% of 18- to 29-year-olds in the United States reporting use of social networking sites [25]. This offers an opportunity to engage young adults in interventions in new and appealing ways. Harnessing this technology to deliver social marketing and individually tailored programs could facilitate the widespread dissemination of interventions in an affordable, convenient, and age-appropriate manner.

Previous systematic reviews of fruit and vegetable consumption-promoting programs have identified that, while interventions produced some positive changes in knowledge and attitudes about the importance of fruit and vegetable consumption, there were only minor improvements in intake [26-28]. These interventions were typically delivered to adults and children, and targeted fruit and vegetable intake concurrently. To our knowledge, to date there is no published review investigating the efficacy and external validity of social marketing and eHealth and mHealth interventions on vegetable intake in young adults. With greater perceived barriers for the consumption of vegetables, poorer knowledge about vegetable servings [29], and just over half of the population already meeting the recommended 2 fruit servings a day [7], it is evident that increasing vegetable intake is a greater challenge. Thus, investigating the implications of interventions on vegetable intake alone will help us understand how we can better support and facilitate improved vegetable consumption.

When evaluating the efficacy of interventions, the accuracy of outcomes should be considered. This is dependent on the validity of intake measurement tools. To compare outcomes across studies, definitions of what constitutes a vegetable serving is also important. This is a source of confusion for the public and for researchers, with variations between countries [30]. In Australia, a serving of vegetables is approximately 75 g or half a cup of cooked vegetables [31], whereas in the United Kingdom a serving is equivalent to 80 g [32].

Furthermore, the specification of behavior change techniques used in interventions is essential to the process of revealing which strategies are effective in the target population and allowing replication of successful interventions [33]. A review of recent eHealth and mHealth interventions found that studies that incorporated a greater number of behavior change techniques had the largest effects [34]. Whether these effects can be generalized to the broader young adult population depends on external validity. Thus, evaluating the external validity of studies is as important as determining efficacy and will have implications for the translation of interventions into larger health promotion programs.

Therefore, in this review we aimed to (1) systematically examine the efficacy of social marketing, and electronic or mobile phone-based interventions in increasing vegetable intake in young adults, (2) assess the quality of the studies, including the validity of tools used to monitor changes in vegetable intake, and (3) review the adequacy of reporting of external validity components.

Methods

We used the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) framework [35] to develop the systematic review protocol, which has been published elsewhere [36]. During the review process, we replaced the quality-assessment tool specified in the original protocol with the Grading of Recommendations Assessment, Development and Evaluation (GRADE) system [37].

Search Strategy

We conducted the systematic literature search between April and August 2015 using the following electronic databases: ScienceDirect, MEDLINE, PyscINFO, Scopus, the Cochrane Library, CINAHL, Embase, and Web of Science. The last search was conducted on August 17, 2015, with no new relevant papers found. We excluded studies published before 1990, as email was not widely used before this period [38]. After hand searching reference lists of key reviews and included studies, as well as conducting a manual search of JMIR journals, we included other relevant studies.

We conducted 2 searches. The first used combinations, synonyms, and truncations of “online intervention,” “computer-assisted therapy,” “electronic mail,” “Internet,” “website,” “cell phones,” “young adult” or “adult,” “fruits,” and “vegetables.” While we were searching largely for eHealth and mHealth interventions, we used other relevant MEDLINE MeSH, such as “telemedicine,” to encompass the terms “mHealth,” “eHealth,” “telehealth,” and “mobile health.” Furthermore, although we were mainly interested in the efficacy of vegetable interventions, we extended the search terms to include “fruit,” as studies typically report on fruit and vegetables concurrently. Additionally, we used the term “adult” alongside “young adult” to broaden the search from 18- to 24-year-olds (the typical database definition of young adults) to 18- to 35-year-olds (based on the US National Institutes of Health cut-off for young adults) [39]. Table 1 shows the first search strategy used in the MEDLINE. The full search strategy is presented in Multimedia Appendix 1 (Tables S1 and S2).

Table 1. Electronic database search: MEDLINE (search 1: eHealth and mHealth interventions).

Search number	Search statement[a]	No. of citations retrieved
1	Online intervention.mp or Computer-assisted therapy.mp. or Therapy, Computer-Assisted/	5242
2	Internet/ or Website.mp	55,352
3	Cell phones.mp or Cell phones/	5040
4	Telemedicine/ or Cyber.mp	12,148
5	email.mp or Electronic mail/	5193
6	Adult/or Young adult/ or young adult*.mp	4,093,057
7	Fruit/ or Fruit*.mp	65,586
8	Vegetable*.mp or Vegetables/	39,576
9	1 or 2 or 3 or 4 or 5	77,751
10	7 or 8	87,363
11	6 and 9 and 10	120
12	Limit 11 to (English language and humans and yr = 1990-current)	120

[a] Modifiers are * (search term as major focus of articles), .mp (multiple purpose search including all fields: title, original title, abstract, subject heading, name of substance, and registry word fields), and / (valid controlled vocabulary term which has been searched in the subject headings field of the database).

We conducted separate database and Google searches to locate programs that used social marketing and mass media to increase fruit and vegetable intake in young adults. Search terms were “young adult,” “adults,” “fruits,” “vegetables,” “social marketing,” “social media,” and “mass media.” These studies were not limited by publication type and included gray literature, such as nonpublished evaluations of programs by organizations. Table 2 presents the second search strategy used in MEDLINE.

Table 2. Electronic database search: MEDLINE (search 2: social marketing and mass media interventions).

Search number	Search statement[a]	No. of citations retrieved
1	Adult/ or Young Adult/ or young adult*.mp.	4126,552
2	Fruit/ or fruit*.mp.	66,529
3	Vegetable*.mp. or Vegetables/	40,014
4	2 or 3	88,502
5	Social marketing.mp. or social marketing/	2976
6	Social media. mp or Mass Media/ or Social Media/	11,192
7	5 or 6	13,882
8	1 and 4 and 7	6
9	Limit 8 to (English language and humans and yr = 1990-current)	6

[a] Modifiers are * (search term as major focus of articles), .mp (multiple purpose search including all fields: title, original title, abstract, subject heading, name of substance, and registry word fields), and / (valid controlled vocabulary term which has been searched in the subject headings field of the database).

Eligibility Criteria

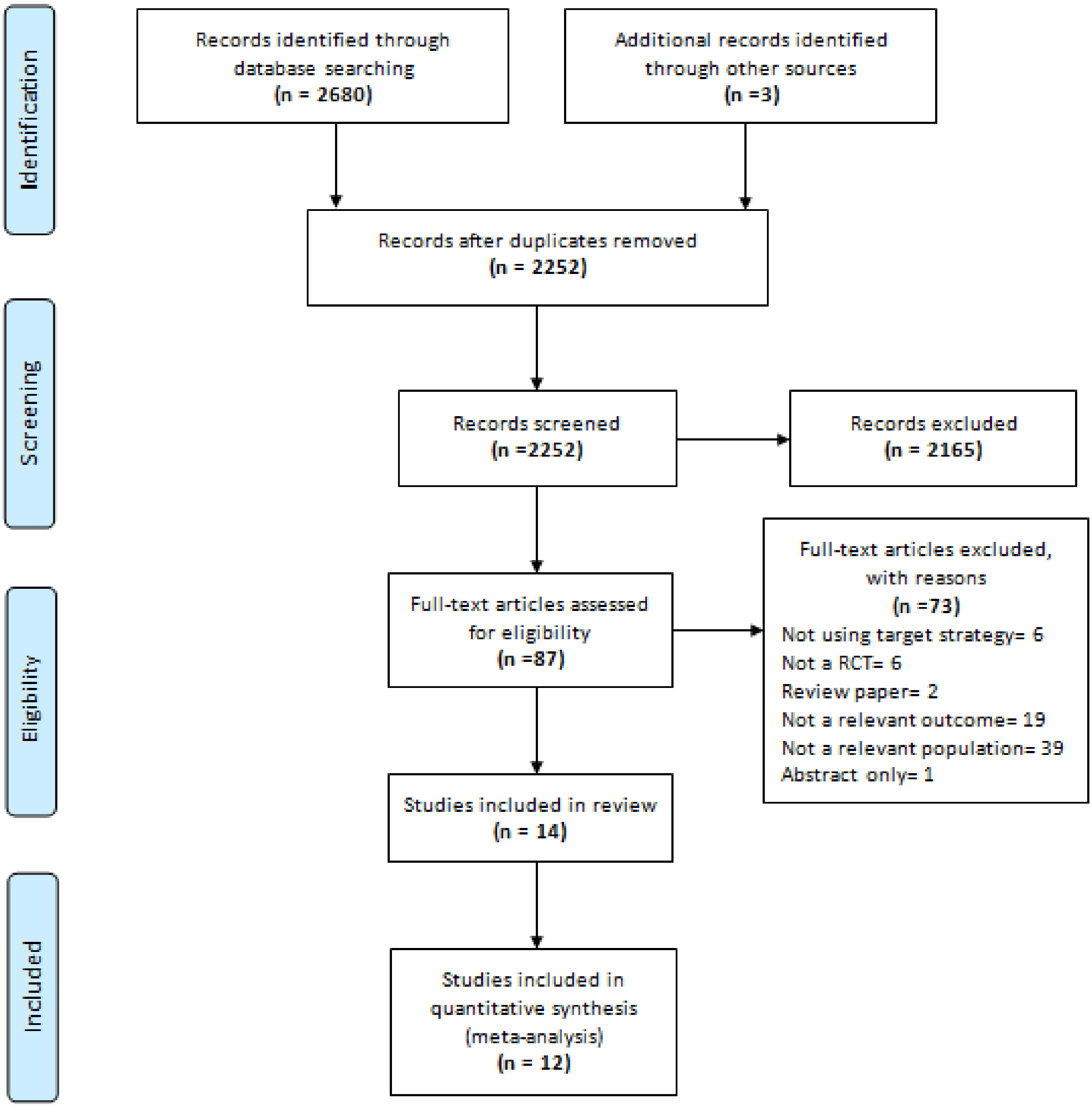
Criteria for inclusion of eHealth and mHealth interventions were as follows: (1) randomized controlled trials (RCTs) with a primary or secondary aim of increasing fruit and vegetable intake in young adults that (2) were targeted at young adults aged 18–35 years inclusive, (3) reported fruit and vegetable intake at baseline and follow-up, (4) involved healthy participants with no disease or illness that would affect the primary outcome or ability to modify fruit and vegetable intake, (5) were written in English, (6) were published after 1990, and (7) were limited to eHealth- and mHealth-based interventions, defined as studies using texting, email, mobile phone apps, phone calls, or websites to deliver the intervention.

Criteria for inclusion of social marketing and mass media interventions were identical to points (1) to (6) above, but were not limited by study design. Social marketing and mass media interventions were defined as those that used media advertising through the Internet, television, billboards, radio, or social media platforms such as Facebook.

Study Selection

We downloaded titles and abstracts of all retrieved studies to EndNote X6 citation management software (Thomson Reuters). Duplicates were removed, then titles and abstracts were reviewed by grouping papers into (1) those meeting selection criteria or (2) requiring further examination; or (3) they were excluded. Papers determined to be potentially relevant to the review were downloaded as full text and reviewed for eligibility by two assessors (MN, JC) and further categorized (Figure 1). We resolved discrepancies in assessors' results by discussion.

Figure 1. Flow diagram demonstrating the process of selecting the included studies of interventions promoting fruit and vegetable intake in young adults. Other sources included a Google search, a hand search of reference lists of relevant systematic reviews and included studies, and a manual search of JMIR journals.



Data Extraction Process

We created a data extraction table according to the principles of the PRISMA statement for reporting systematic reviews [35], with some additional elements included for completing the Cochrane Collaboration's risk of bias tool [40]. Once we had piloted the process on a random selection of 4 of the included studies, 2 independent reviewers extracted the following data in duplicate: study details (authors, year, country of publication, funding, and affiliations); participants (characteristics, setting, inclusion and exclusion criteria, attrition, and blinding); intervention and comparator details; duration; and the summary outcome measure (change in fruit and vegetable intake between baseline and follow-up for the intervention and control arms). We also extracted the name of the tool used to assess changes in fruit and vegetable intake, as well as citations of available validation studies.

Data Synthesis and Analysis

The primary outcome of interest was the change in vegetable intake postintervention. Where possible, for all study arms we recorded mean or median intakes (as servings, cups, frequency, or percentage consuming) pre- and postintervention. If vegetable intake was not reported separately, we documented the change in fruit and vegetable intake. We also noted the measures of error (SE or SD) and associated *P* values for change between groups over time. To determine the magnitude of intervention outcomes, we calculated effect sizes (Cohen *d* and 95% CIs) for studies that reported sufficient data (means, and measure of error or frequencies). Web-based calculators [41] based on Lipsey and Wilson's formulas [42] assisted with calculations. We assessed the magnitude of the effect sizes according to Cohen's categories, whereby an effect <0.2 is considered negligible, between 0.2 and 0.49 is small, 0.5-0.8 is medium, and >0.8 is large [43].

We also considered the clinical significance of outcomes. There is no consensus in the literature regarding what change in intake is considered clinically significant. However, several meta-analyses and longitudinal studies suggest a dose-response relationship, whereby an increase in vegetable intake by approximately 1 serving is protective for cardiovascular health (decreased risk of stroke and cardiovascular disease mortality by 11% and 4%, respectively) [44,45]. Furthermore, every 1-serving increase in vegetable intake has been associated with a 0.12 kg reduction in weight (95% CI -0.35 to -0.14) [46]. These studies define a serving of vegetables as approximately 1 cup of leafy vegetables or half a cup of cooked vegetables (frozen, fresh, or canned) in line with previous US and current Australian dietary guidelines [31,47].

To pool the outcomes for the meta-analysis, we grouped studies for which an effect size was calculated. We used STATA version 13 (StataCorp LP) to conduct the analyses using the metan, metabias, and metafunnel commands. A random effects model was applied. Publication bias was determined through Egger's statistical test for funnel plot asymmetry and visual inspection of the funnel plots of the Cohen *d* effect size (standardized mean difference), plotted against its standard error. The I^2 value for heterogeneity was calculated based on the Q statistic: [(Q statistic - df/Q statistic) × 100%]. Cochrane Collaboration guidelines [48] suggest that an I^2 for heterogeneity below 40% is considered low, and a value above 50% is considered substantial.

Quality Assessment

Risk of Bias Assessment

Using the Cochrane Collaboration's tool [40], we established risk of bias at the individual study level, based on the following study elements: selection of participants (random sequence generation and concealment of allocation methods); attrition (completeness of outcome data); detection (blinding of participants and personnel); and reporting (selective reporting of outcome measures). Two authors (MN and JC) independently evaluated each study for risk of bias and coded them as having low risk, high risk, or unclear risk. Any differences in judgment were clarified through discussion.

GRADE Assessment

The quality of the body of evidence was determined by 2 independent reviewers (MN, JC) using the GRADE system [37]. We considered 5 categories to ascribe a quality rating: limitations in study designs; consistency of results; directness of the evidence with regard to study populations, intervention design, and outcomes measured; precision of outcomes; and the presence of publication biases.

Rating Validity of Dietary Assessment Tools

We determined the validity of each tool used to measure changes in vegetable intake based on published literature demonstrating its accuracy [49-59]. The checklist of requirements by Nelson et al [60] was also consulted to qualitatively examine the effectiveness of reporting on measurement tools. This checklist assesses factors such as data-collection procedures (objective measure vs self-report), methods of quantifying portions, variety of foods captured, food composition databases used and whether checking procedures were applied.

Rating External Validity

We assessed the external validity of included studies based on the Green and Glasgow's criteria [61]. The assessment explored components under 3 sections: (1) reach and representativeness of participants, (2) intervention implementation and adaptation, and (3) program maintenance and institutionalization (sustainability of program implementation). Quantitative and qualitative data pertaining to these external validity components were extracted. We recorded specific data that were not reported as not reported, and if an assessment component did not apply to the particular study we reported it not applicable. Individual participation rate (%) was calculated as the percentage of eligible participants agreeing to participate. Attrition rate (%) was calculated as the percentage of participants who dropped out after randomization. Attrition was further grouped by intervention arm (treatment vs control). Extracted data were used to examine the number of studies adhering to the external validity components. The frequency and adequacy of reporting of these components were also examined and compared between studies.

Results

Study Selection

As the flow diagram in Figure 1 shows, we found 2680 studies through database searching and 3 additional studies through hand searching the references. We screened a total of 2252 papers by title and abstract. Of these, we assessed the full text of 87 studies. A total of 14 studies [62-75] met the selection criteria and were included in the review. See Multimedia Appendix 2 for the complete list of references excluded by full text with corresponding reasons. None of these studies used social marketing strategies or mass media to encourage vegetable consumption in young adults specifically. Therefore, the remaining results report the effectiveness and external validity of eHealth and mHealth interventions aimed at increasing fruit and vegetable intake in young adults. We included 12 studies in the meta-analyses. For the meta-analysis, we combined the reported results in 2 groups for comparability: fruit and vegetable (8 studies) and vegetable only (5 studies); 1 study contributed results for both groups [64].

Study Reach and Representativeness of Participants

Overall, 7984 healthy people participated in the eHealth and mHealth RCTs (see Multimedia Appendix 3, Table S3). There were, however, large discrepancies in the sample sizes. Only 3 of the 14 studies had recruited >500 participants at baseline [62-64], and 1 study had a sample size of <100 [64]. More than half of the interventions were conducted in the United States [62,64,67,69,71,73,75], 4 in Australia [65,65,67,72], 1 in New Zealand [74], and 1 in Malaysia [70]. The target audience was college or university students for the majority of the studies [62,63,67-75], and 3 studies reported their target audience to be young adults [64-66].

Recruitment methods were reported for 13 of the 14 studies, but limited details were provided. All but 2 studies recruited through the university or college setting [64,66]. Participants were recruited through undergraduate psychology courses in 2 studies [67,68], from random nonnutrition classes in 2 studies [69,70], and through advertisements and flyers posted on university grounds in 4 studies [62,63,65,71,72]. In 1 study a recruitment table was set up on campus [73], and another study invited patients attending the student university health service [74]. In 1 study [64] advertisements with a toll-free phone number were used, and the final study distributed letters of invitation through participating family doctors, along with electronic and print advertisements [66]. Of the included studies, 9 indicated their participation rate, with a mean of 78.0%. The inclusion criteria were detailed by 10 studies, all of which specified age (years) as one of their criteria. Demographic data were provided by most of the studies although not consistently. Baseline age (years) was reported in all but 1 study (Multimedia Appendix 3, Table S3), with a mean age of 20.8 years across the studies. The ethnicity of participants was reported to be >50% Caucasian or white in 7 studies. The percentage of female participants was reported by 13 studies, with women more commonly recruited than men (mean 69.8% female) (Multimedia Appendix 3, Table S3).

Intervention Implementation and Adaptation

Details of the intervention and comparator groups were provided in detail. All studies recruited an intervention and a control group (see Multimedia Appendix 4, Table S4), with 4 studies using multiple intervention and control arms [67,71,72,74]. A total of 6 studies provided no treatment to the control arm [67,68,70,71,74,75], 7 studies gave the comparator group general information not containing the intervention material [62,64-66,69,72,73], and 1 study provided the control group with the intervention material on completion of the follow-up assessment [63]. The duration of interventions and number of sessions were easily extrapolated from each study. The level of contact between researchers and participants ranged from one-off sessions (provision of feedback) to daily contact by email or text message (Multimedia Appendix 4, Table S4). The majority of the interventions used online education through learning platforms, websites, and emailing, with only 2 studies using apps [65,66] and 4 using text messaging [65,66,70,72]. No studies reported the use of social media platforms. The studies predominantly used goal setting for behavior change, with monitoring and feedback also commonly incorporated. For the majority of the interventions, the aim was to offset weight gain in young adulthood. Targeting improvements in fruit and vegetable intake was one such method used to address weight gain. While 1 study was designed to reduce health-risk behaviors in young adults [74], only 5 studies focused specifically on fruit and vegetable intake [64,68,69,72,75], and none targeted vegetables alone.

The reviewed studies varied in the detail provided regarding the behavior theories and techniques considered in the intervention design. The design of 5 studies was based on the transtheoretical model of behavior change, where the participants' stage of change determined the content received [63-66,75]. A total of 6 studies were theory or education based [62,63,67,70,71,73]. Social-cognitive theory informed 2 interventions [67,69]. Half of the reviewed studies applied the behavioral construct of self-efficacy in their intervention [62,64,69,70,71,73,75]. The study by Kypri and McAnally [74] did not report consideration of theoretical frameworks in their intervention design. The remaining 2 studies [68,72] were informed by the theory of planned behavior and the theory of habit formation (Multimedia Appendix 4, Table S4). All the studies that we reviewed intervened at the individual level. Only 2 studies were implemented outside of the university setting, thus limiting the generalizability of the interventions to the overall young adult population. Of these studies, one [64] was targeted at lower socioeconomic status young adults, while the other mainly captured young adults from higher socioeconomic areas [66].

The duration of the interventions (excluding postintervention follow-up) ranged from one-off contact to 6 months of treatment, with a mean of 10 weeks (Multimedia Appendix 4, Table S4). A total of 9 studies allocated a follow-up period [62-64,66,69,71-73], with a mean of 16 weeks. Adherence was most commonly documented as the number of sessions completed or the amount of materials viewed by participants (Multimedia Appendix 4, Table S4), but was not consistently reported across studies. The mean level of compliance among those reporting adherence was 85.4%.

Delivery expertise varied among the studies (Multimedia Appendix 4, Table S4). Research staff were more commonly reported to have conducted the interventions, with little specification of their qualifications and the number of research staff involved. Registered dietitians delivered 5 of the interventions [63,65,66,69,75]. Other expertise included a health promotion officer [71] and outreach educators [64].

Study Maintenance and Institutionalization

The rate of attrition was documented in all reviewed studies. At completion of the interventions the mean attrition rate was 19.6% (see Multimedia Appendix 5, Table S5). All but 4 studies [64,69,71,75] reported attrition for the control and intervention group separately, and 4 did not assess differences in characteristics between completers and noncompleters [65,70,71,73]. Only 2 studies looked at the long-term impacts of the study, by assessing outcomes at least 12 months following treatment [62,63]. Both of these studies found that the changes in fruit and vegetable intake were not maintained at follow-up (Multimedia Appendix 5, Table S5). The sustainability of program implementation was poorly reported, with only 1 study mentioning that results would be used to refine the intervention for trial in a broader young adult population using a larger sample size [66]. Finally, only 2 studies published a process evaluation documenting effective program elements [62,66].

Risk of Bias

We rated the majority of the studies reviewed as unclear to high risk because they did not perform intention-to-treat analyses, which introduced biases in the outcome data (attrition bias) [62,70-75] (see Multimedia Appendix 6, Table S6). We rated 2 studies high in a second domain (detection bias) [71,73]. The majority of the studies did not clarify their methods of blinding (n=8). Selection bias was mainly unclear within and across studies, with 5 studies not reporting the method of sequence generation in randomization [62,64,69,71,75] and only 2 studies specifying allocation concealment methods [66,74] (Multimedia Appendix 6, Table S6). While all of the studies reported results for prespecified outcomes, we could not completely rule out reporter bias across studies because only 5 RCTs published their original protocol [63,65,66,68,69] or provided details of their trial registration [66]. However, no selective reporting was apparent based on the methods within the reviewed manuscripts (both successful and unsuccessful outcomes recorded). Overall, the combined lack of clarity of the level of bias across studies raises concerns about the plausibility of the studies’ results.

GRADE Quality Rating

The reviewed interventions had several limitations in study design and did not address the research question directly, resulting in an overall low quality rating (Table 3).

Table 3. Overall assessment of quality in 14 studies (7984 participants in total) of promotion of fruit and vegetable intake using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) system.

Category	Rating with reasoning
Limitations	–2 quality levels due to very serious limitations
Consistency	No subtraction of levels, as inconsistency doesn’t affect confidence in results
Directness	–2 quality levels, as the population, outcomes, and study design are indirect
Precision	No subtraction of levels due to good precision
Publication bias	No subtraction of levels, as funnel plot symmetry suggests publication bias is unlikely
Overall quality	Low: our confidence in the effect estimate is limited

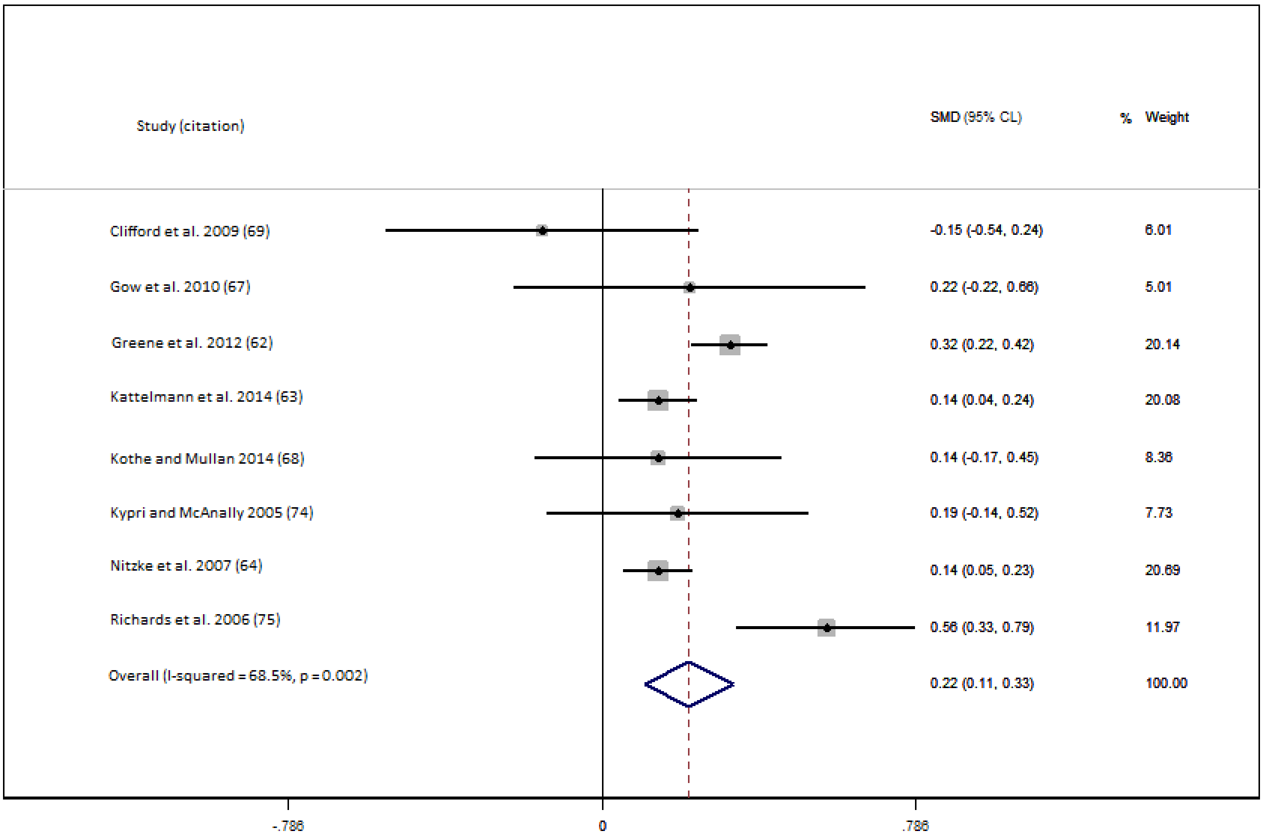
Study Limitations

All the included studies were RCTs. However, only 2 studies adequately concealed the difference between intervention arms [66,74]. In 1 study, the study design and purpose of randomization was explained to participants, preventing allocation concealment [71]. The remaining 11 studies did not clearly describe their method of concealment. Furthermore, 8 studies did not describe their method of blinding and 3 did not blind effectively [65,71,73]. Half of the included studies had a loss to follow-up of >20% [62-64,67,72,73,75] and did not conduct intention-to-treat analysis [62,70-74]. A total of 3 studies did not state methods for dealing with missing data or conducted analysis on completer populations [63,68,69]. Several studies used nonvalidated measures of intake, further limiting the quality of the body of evidence.

Consistency

The studies with effect sizes for change in fruit and vegetable intake yielded an I^2 statistic of 68.5% (P value for heterogeneity =.002), indicating that there may be considerable heterogeneity. However, a higher heterogeneity can be caused by small variations in point estimates from studies with larger sample sizes, as is evident in Figure 2. An I^2 of 31.4% (P value for heterogeneity =0.2) for studies reporting vegetable intake separately suggests low heterogeneity.

Figure 2. Forest plot of Cohen d effect size (standardized mean difference, SMD) for studies reporting change in fruit and vegetable intake combined. The diamond represents the overall effect size; the percentage weighting of each study toward the overall effect is indicated by the size of gray squares; and the 95% confidence limits are shown by horizontal lines. The overall intervention effect lies at the center of the larger clear diamond with right and left end points indicating the 95% confidence limits. Note: weights are from random effects analysis.



Directness

While comparisons between control and intervention arms were direct for the included interventions, variations in study design, populations, and outcome measures meant that the overall body of evidence was indirect. The population of included studies was predominantly college students. Only 2 interventions recruited beyond the university or college setting, but they were still not representative of the broader young adult population. This review allowed for the inclusion of studies that measured changes in intake as a secondary outcome. Consequently, several studies were weight management interventions targeting fruit and vegetable intake as a component of the program. Only 5 studies targeted fruit and vegetables specifically [64,68,69,72,75] and none targeted vegetables alone. Measures of fruit and vegetable intake also varied considerably. Thus, the overall evidence is an indirect representation of the impact of eHealth and mHealth on vegetable intake.

Precision

Only 6 of the 14 studies reported conducting power calculations [63,67-69,71,73]. However, these were mainly based on primary outcomes other than vegetable intake, such as change in nutrition knowledge or weight. Sample size varied from 51 to 2024 participants but yielded 7984 in total, which is considered sufficient.

Publication Bias

While we implemented a comprehensive search strategy to capture the gray literature, we may have missed unpublished studies (interventions with insignificant or negative findings) or those published in journals not indexed in major databases. The outcomes of statistical tests of publication bias (Egger’s test) were not reported, as these results are less accurate when based on fewer than 10 studies or when there is significant heterogeneity [48]. Visual inspection of funnel plots (Figures 3 and 4) indicated symmetry in the distribution of points around the mean effect size, suggesting that bias from missing studies is unlikely.

Figure 3. Funnel plot for risk of publication bias: intervention effect for fruit and vegetable intake represented by the standardized mean difference (SMD) plotted against the standard error, se(SMD). Dashed diagonal lines indicate the pseudo 95% confidence limits and scatter dots represent individual studies.

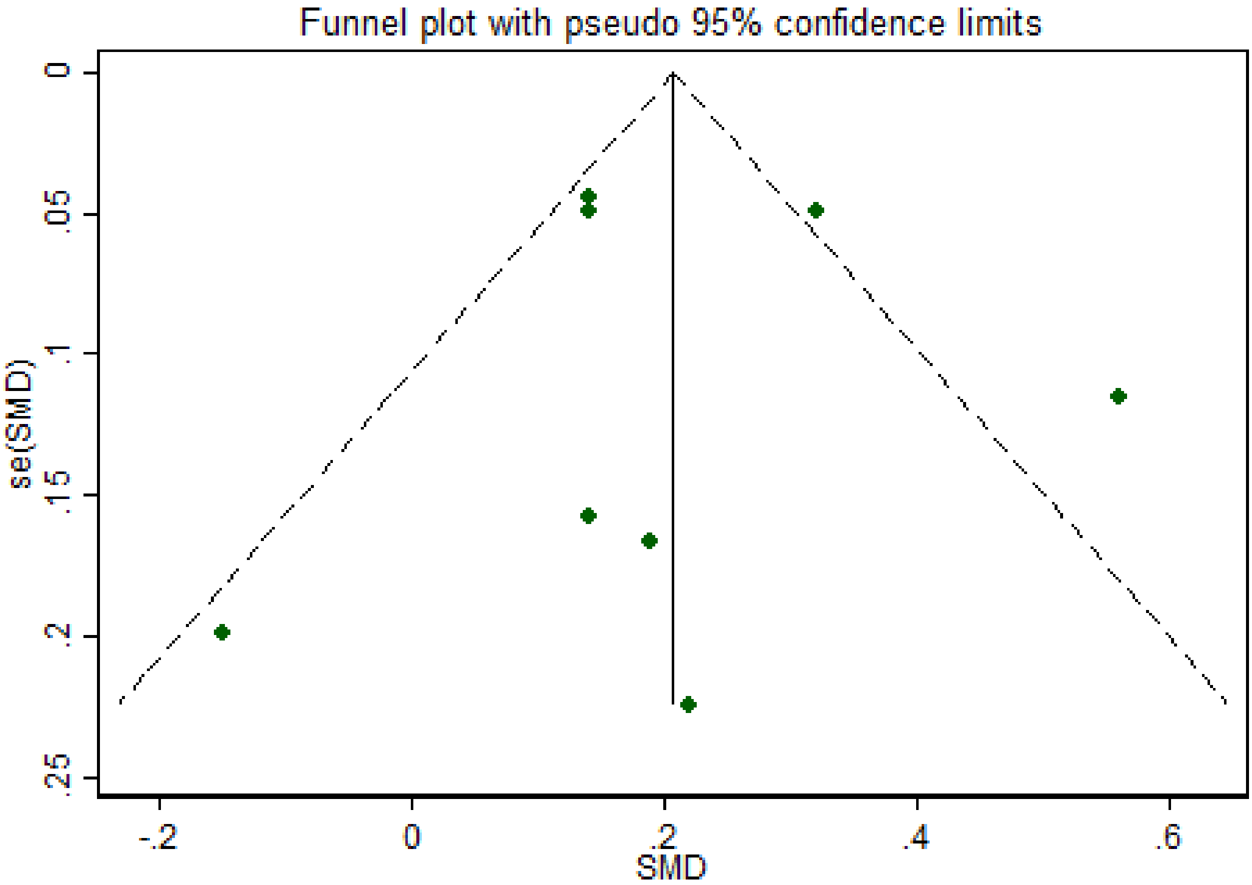
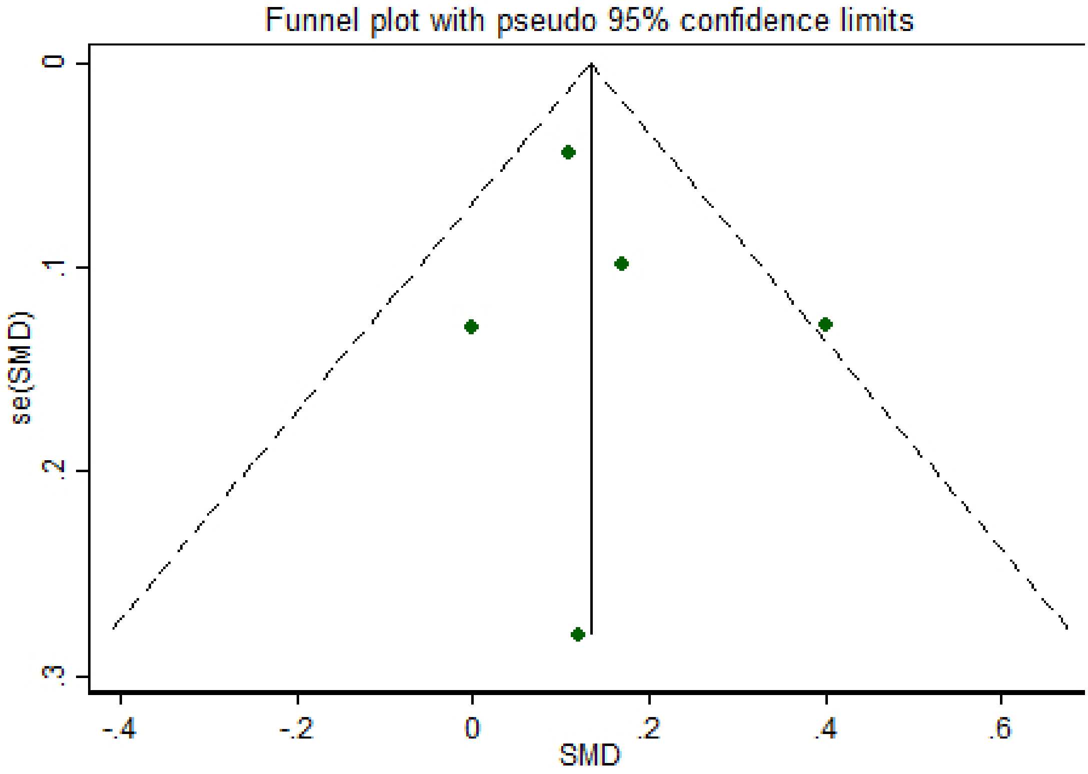


Figure 4. Funnel plot for risk of publication bias: intervention effect for vegetable intake represented by standardized mean difference (SMD) plotted against the standard error, se(SMD). Dashed diagonal lines indicate the pseudo 95% confidence limits and scatter dots represent individual studies.



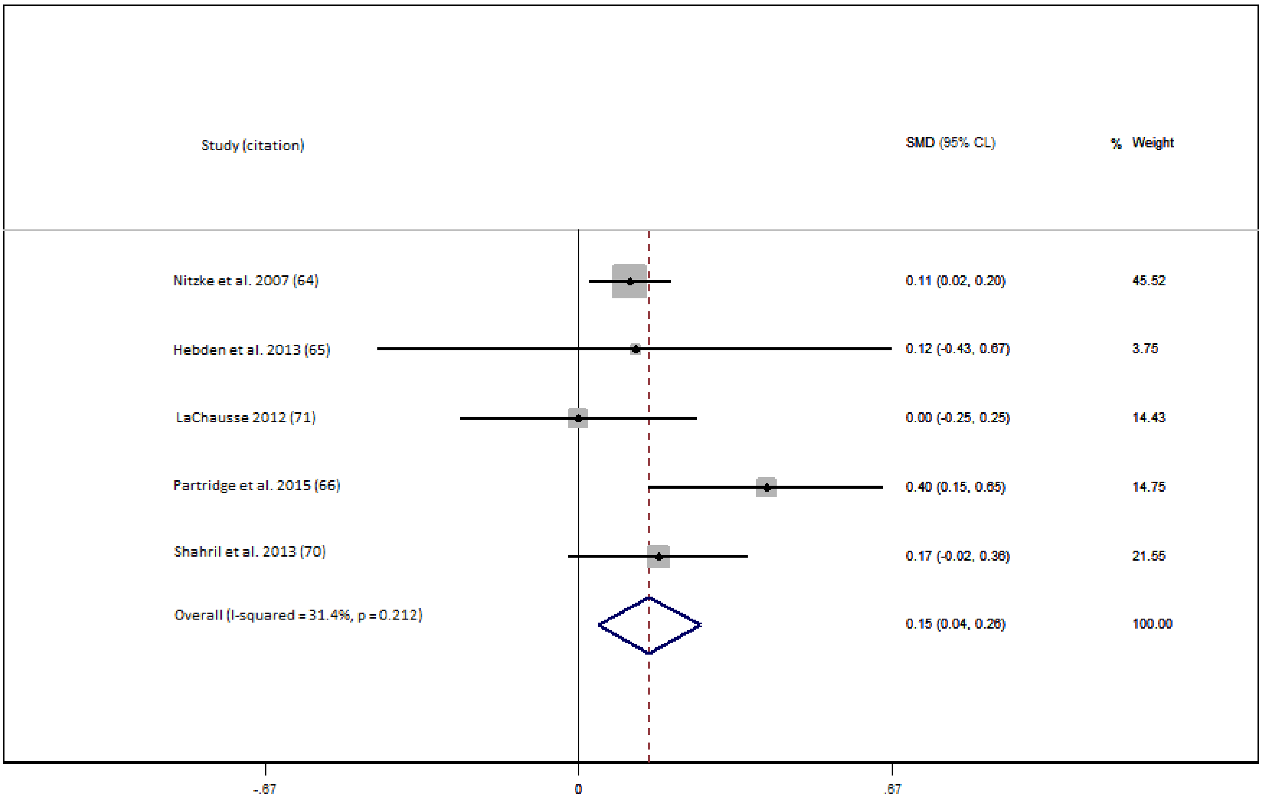
Efficacy of Interventions

Of the 14 reviewed studies, 9 provided results for fruit and vegetable intake, and we included 8 in the meta-analysis. Of these studies, 7 found positive effects postintervention [62-64,67,68,74,75] (Cohen *d* 0.14-0.56), 4 of which were statistically significant [62-64,75]. For all but 1 study [75], the magnitude of effect was small. In total, 2 studies also reported clinically significant improvements of ≥1 serving/day [62,75] (see Multimedia Appendix 7, Table S7). The pooled effect size for interventions reporting change in fruit and vegetable intake was 0.22 (95% CI 0.11 to 0.33), indicating a small positive effect of eHealth and mHealth interventions on fruit and vegetable intake. The 4 studies [62-64,75] with significant effects contributed 72.9% weighting (Figure 4). The I^2 was 68.5%, *P*=.002, suggesting considerable heterogeneity between these studies, and so findings should be interpreted with caution.

Of the 6 studies that assessed vegetable intake independently of fruit [64-66,70-72], we included 5 in the meta-analysis, 4 of which had positive effects on vegetable intake [64-66,70] (Cohen *d* 0.11-0.40). Two of these positive effects were statistically significant [64,66]. Increases in intake were <1 serving/day, with the exception of the results reported by Partridge et al [66] (Multimedia Appendix 7, Table S7). The pooled effect size for change in vegetable intake was negligible at 0.15 (95% CI 0.04 to 0.28; I^2=31.4%, *P*=.2) (Figure 5).

Studies that were more successful in improving fruit or vegetable intake provided participants with individually tailored advice and feedback based on their stage of change [64,66,75] and incorporated goal setting [62,66,75]. Of the studies producing clinically and statistically significant results for fruit or vegetable intake, or both [62,66,75], 1 used online theory education based on nondiet principles [62]. This intervention was designed according to 2 educational models, Carey and colleague’s system of instructional design [76] and Keller’s instructional motivational model [77]. Fruit and vegetable intake goals were set after completion of each weekly educational lesson, and self-evaluation of progress preceded the next weekly Web-based module. The study by Richards and colleagues [75] used motivational interviewing in combination with Web-based resources and emails. The resources were tailored to the participants’ stage of change, where precontemplators and contemplators were given reasons to and tips on how to eat more fruits and vegetables, as well as a goal-setting framework. Action and maintenance participants received emails with tips for maintaining consumption and trying new fruits and vegetables. Finally, the study by Partridge et al [66] combined multiple eHealth and mHealth strategies to support behavior change, with text messaging found to be the most popular, and the website and discussion boards the least popular, among participants. The text messages contained reminders and tips on how to achieve their individualized goal set during their phone counseling session with a dietitian and were based on the 10 processes of change (transtheoretical model). Participants could monitor their fruit and vegetable intake goals using a personalized app that also provided recipes and tips on how to increase their intake.

Figure 5. Forest plot of Cohen d effect size (standardized mean difference, SMD) for studies reporting change in vegetable intake separately. The diamond represents the overall effect size; the percentage weighting of each study toward the overall effect is indicated by the size of gray squares; and the 95% confidence limits are shown by horizontal lines. The overall intervention effect lies at the center of the larger clear diamond with right and left end points indicating the 95% confidence limits. Note: weights are from random effects analysis.



Validity of Dietary Assessment Tools

Of the reviewed studies, 5 used tools that had not been validated to assess changes in vegetable intake [68,69,71,73,75] (Table 4). While the majority of the tools were validated, only 1 was tested specifically in the young adult population [30]. Of the studies that used validated tools, short screeners were most popular, including the US National Cancer Institute's fruit and vegetable screener [53], as well as short questions adapted from the Australian and New Zealand national nutrition surveys [52,54,56]. Furthermore, only 2 studies defined what they classified as a serving [65,66], and the outcome measure for intake lacked consistency, with studies reporting change in terms of frequency, servings or cups of vegetables consumed, as well as the percentage meeting recommendations. No studies detailed which food composition databases they used for the analysis, or whether they checked records with respondents as per the requirements specified in the Nelson and colleagues' checklist [60]. All but 1 study [70] used a self-report measurement tool. The study by Gow and Colleagues [67] did not specify what the outcome measure was (servings vs score).

Table 4. Validity of tools used to measure fruit and vegetable intake and source of tools.

Author [citation]	Fruit and vegetable intake measurement tool and source [citation]	Tool validated for fruit and vegetables
Clifford et al [69]	Food frequency questionnaire adapted from US National Cancer Institute's health habits and history questionnaire [59]	No
Franko et al [73]	Single-item question measuring daily fruit and vegetable consumption [51]	No
Gow et al [67]	Block food screener [49]	Yes
Greene et al [62]	2-item screener and National Cancer Institute screener [53]	Yes
Hebden et al [65]	Web-based short survey using questions from Australian national survey [30,52,56]	Yes
Kattelmann et al [63]	National Cancer Institute's vegetable screener [43]	Yes
Kothe and Mullan [68]	Self-report measure of previous day's consumption	No
Kypri and McAnally [74]	2 questions from New Zealand National Survey questionnaire [54]	Yes
LaChausse [71]	US Centers for Disease Control and Prevention's youth risk behavior survey [58]	No
Nitzke et al [64]	5 A Day screener (7-item fruit and vegetable screener) from 5 A Day program [53]	Yes
Partridge et al [66]	Short questions adapted from the Australian National Nutrition Survey [30,52,56]	Yes
Richards et al [75]	1-item food frequency questionnaire [50]	No
Rompotis et al [72]	Short question on fruit and vegetable intake [57]	Yes
Shahril et al [70]	Diet history	NA[a]

[a] NA: Not applicable.

Discussion

This systematic review found preliminary evidence to suggest that eHealth and mHealth interventions may have a positive impact on fruit and vegetable intake among young adults. Meta-analyses revealed a small magnitude of effect on fruit and vegetable intake and a negligible effect on vegetable intake alone. Whether these effects have clinical or nutritional significance remains questionable. The quality of the body of evidence was rated low and therefore, findings should be interpreted with caution. Rather than making recommendations, we propose suggestions for improved research.

Among the studies that improved intake, only small changes were observed (<1 serving/day). This is consistent with conclusions from existing reviews, in which interventions appear to produce minor improvements in fruit and vegetable intake [26-28]. The effectiveness of the reviewed interventions in creating sustainable change in the long term remains unclear, as follow-up periods were short. The observed dose-response clinical outcomes associated with increasing vegetable intake [44-46] are likely to become evident only in the longer term. Additionally, the link between vegetable intake and weight maintenance during the transition to adulthood occurs over time [78]. Thus, investigators should integrate longer follow-up in intervention protocols. Future studies may also consider measuring secondary outcomes, such as weight and indicators of cardiovascular health, over time to understand the longer-term clinical implications of improved vegetable intake.

With the measurement and reporting of fruit and vegetable intakes as a summed value in most studies reviewed, the impact of the eHealth and mHealth strategies on vegetable consumption specifically remains unclear. Previous research has shown that knowledge of serving sizes is poorer for vegetables than for fruit [29], and for young adults, taste was a more important barrier to increasing vegetable consumption than it was for fruit [79]. Fruit and vegetables also have varying nutrient profiles and product attributes. Considering these factors, it is apparent that vegetables should be promoted and measured separately from fruit. Additionally, most of the reviewed studies targeted fruit and vegetable intake as part of a larger weight management program. Thus, the impact of an intervention focusing primarily on vegetables is an important question for future research.

Previous research established the importance of considering behavior change theory in intervention design [33,80]. The value of incorporating behavior change theory is reiterated by this review, where the majority of the successful studies incorporated behavior change constructs such as goal setting [62,66,73,75] and the provision of individually tailored advice and feedback was based on participants' stage of change [64,66,75]. While the transtheoretical model has been long established as an effective means of improving fruit and vegetable intake [81], these studies suggest its efficacy in eHealth and mHealth interventions where, for instance, motivational and confidence-enhancing text messages or phone calls can benefit individuals who are in the earlier contemplative stages of change. There was no clear pattern, however, to indicate that the incorporation of more behavior change techniques initiated larger improvements as previously suggested in the literature [34]. Researchers could consider investigating whether a combination of efficacious strategies and repeat exposure at a later date produces greater change to shed light on whether intensive short-duration or less-intensive, longer-duration interventions are more effective.

The mode of intervention delivery varied considerably between studies, making it difficult to determine which eHealth and mHealth strategies were most successful in supporting behavior change. However, 2 of the effective studies [66,75] used motivational phone counseling as part of their intervention. While details of the cost effectiveness of this design were not provided, generally, the individualized nature of this approach can be expensive, due to the necessity for trained staff and the monetary reimbursement required for their time. Consequently, the applicability of these studies to the whole population level may be limited. The use of other low-cost and convenient eHealth and mHealth techniques (texting and email) that can incorporate individually tailored information may be more feasible for interventions. Preliminary evidence suggests that these methods are successful [66,75]; however, further research is required to confidently determine their efficacy.

Our review was unable to identify social marketing campaigns targeted specifically at young adults. Addressing this gap is an opportunity for future public health promotion projects, with research indicating that young adults have poor awareness of population-wide campaigns and perceive considerable barriers to increasing their intake despite the promoted health benefits [82]. Additionally, we found no studies that incorporated social media platforms in their intervention. Using these high-reach and lower-cost information-sharing platforms can help to increase interactivity and collaborative content sharing. This may be the fastest and most wide-reaching way to engage young people, with approximately 89% of young adults using social media [19]. Effectiveness studies on the use of social media to improve health behaviors are limited, although preliminary reports are encouraging [83,84].

There is considerable uncertainty regarding the accuracy of the findings summarized by this review, due to the use of non-validated self-report measures of intake, which may not be sensitive enough to detect small changes and may be subject to reporter bias. Therefore, further effort is required to develop validated tools for the measurement of vegetable intake in young adults for consistent and accurate reporting of intervention outcomes. Researchers need to specify what is considered a serving of vegetables to allow easier comparison of outcomes and should use objective measures of intake for validation. Biomarkers such as vitamin C and beta-carotene are useful indicators of fruit and vegetable intake, respectively. While tests for these biomarkers are potentially costly for use in large interventions, they would be feasible and reliable in small validation studies [85].

The degree to which the interventions can be translated to the general young adult populations is questionable, as the majority of studies were conducted in the university or college setting in a sample of educated young adults. While the latest statistics indicate that an increasing proportion of young adults are enrolled in tertiary education [86], those of lower socioeconomic status remain underrepresented [87]. Future studies should limit the use of convenience sampling and aim to recruit a wider range of socioeconomic groups. Overall, the studies we reviewed did not consistently report on external validity, particularly program sustainability, costs, and long-term effects of the intervention. Process evaluations were also lacking. Consequently, the external validity of interventions for improving vegetable consumption in young adults is uncertain. There is a growing body of evidence in health research indicating that investigators are not reporting on external validity [88-90]. Improvements in this area are required to determine the potential for implementation of study designs in broader health promotion programs. Of particular importance is consideration and reporting of the costs involved in upscaling these interventions, which will have implications for health promotion officers and policy makers [91]. Furthermore, researchers should invest in conducting process evaluations to determine how to improve the efficacy of interventions and enhance their generalizability [92].

Strengths and Limitations

This is the first systematic review to report on the effect of eHealth and mHealth interventions on vegetable intake specifically and highlights relevant opportunities for future research. We conducted the review protocol in line with the PRISMA guidelines [35] and used a comprehensive search strategy. While we searched several electronic databases and made an effort to include gray literature, we may have missed some studies. The variability across interventions with differences in study designs and measures of vegetable intake, and the overall poor study quality, made it difficult to establish definitive conclusions. Consequently, we were reluctant to rule out any eHealth or mHealth approach as ineffective and rather discussed the outcomes as a means of highlighting gaps in the current literature and opportunities for future research to generate a stronger body of evidence on whether technology-based strategies are effective in this population. Finally, the lack of consistent reporting of external validity components prevented us from making conclusions about the potential for translating interventions to the wider young adult population.

Conclusions

Overall, this review revealed that young adults have been neglected in fruit and vegetable social marketing campaigns, and most interventions target fruit and vegetables concurrently. Very few good-quality eHealth and mHealth interventions using validated dietary assessment tools have been designed to support young adults in improving their vegetable intake. With preliminary evidence suggesting that eHealth and mHealth strategies may be an effective mode of delivering vegetable interventions, continued research using stronger and higher-quality study designs is required to better determine the efficacy of technology-based strategies for improving vegetable consumption in young adults. With previous research suggesting that multiple behavior change strategies should be used for greater improvements, researchers could consider combining promising strategies such as goal setting and tailored feedback in future interventions. The potential impact of using social media platforms to create awareness of the importance of eating enough vegetables also deserves attention. Finally, in light of the lack of reporting of external validity components in the reviewed papers, it is critical that future studies address key factors such as program costs, sustainability, and longer-term impact in order to determine the potential for upscaling interventions to the broader young adult population.

Acknowledgments

The authors thank school Librarian Rod Dyson for his assistance with setting up the database search strategy and Dr Kevin McGeechan for his advice on conducting meta-analyses in STATA. MN and JC are PhD students at The University of Sydney, funded by the Australian Postgraduate Award scholarship.

Authors' Contributions

MN, JC, and MAF developed the research question. MN drafted the review manuscript and JC assisted with screening, extraction, and data analysis. All authors have read and contributed to the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Additional supporting information Table s1. Search 1: e- and m-health interventions, databases searched, search terms, limits applied and results and Table s2. Search 2: social marketing and mass media interventions, databases searched, search terms, limits applied and results.

[PDF File (Adobe PDF File), 155KB - jmir_v18i4e58_app1.pdf]

Multimedia Appendix 2

List of references excluded by full-text with reasons (n=73).

[PDF File (Adobe PDF File), 47KB - jmir_v18i4e58_app2.pdf]

Multimedia Appendix 3

Table S3. Study descriptions of reach and representativeness of participants (n=14).

[PDF File (Adobe PDF File), 58KB - jmir_v18i4e58_app3.pdf]

Multimedia Appendix 4

Table S4.Study Description of intervention implementation and adaption (n=14).

[PDF File (Adobe PDF File), 284KB - jmir_v18i4e58_app4.pdf]

Multimedia Appendix 5

Table S5.Study maintenance and institutionalization (n=14).

[PDF File (Adobe PDF File), 165KB - jmir_v18i4e58_app5.pdf]

Multimedia Appendix 6

Table S6. Risk of bias as assessed by the cochrane collaboration tool for included studies.

[PDF File (Adobe PDF File), 54KB - jmir_v18i4e58_app6.pdf]

Multimedia Appendix 7

Table S7. Change in fruit and vegetable intake between baseline and follow-up for intervention and control arms with calculated effect size Cohen's d (95% CI) (n=14).

[PDF File (Adobe PDF File), 164KB - jmir_v18i4e58_app7.pdf]

References

1. Lock K, Pomerleau J, Causer L, Altmann DR, McKee M. The global burden of disease attributable to low consumption of fruit and vegetables: implications for the global strategy on diet. Bull World Health Organ 2005 Feb;83(2):100-108 [FREE Full text] [Medline: 15744402]

2. Ledoux TA, Hingle MD, Baranowski T. Relationship of fruit and vegetable intake with adiposity: a systematic review. Obes Rev 2011 May;12(5):e143-e150. [doi: 10.1111/j.1467-789X.2010.00786.x] [Medline: 20633234]
3. Rekhy R, McConchie R. Promoting consumption of fruit and vegetables for better health. Have campaigns delivered on the goals? Appetite 2014 Aug;79:113-123. [doi: 10.1016/j.appet.2014.04.012] [Medline: 24751917]
4. Casagrande SS, Wang Y, Anderson C, Gary TL. Have Americans increased their fruit and vegetable intake? The trends between 1988 and 2002. Am J Prev Med 2007 Apr;32(4):257-263. [doi: 10.1016/j.amepre.2006.12.002] [Medline: 17383556]
5. Cox DN, Anderson AS, Lean ME, Mela DJ. UK consumer attitudes, beliefs and barriers to increasing fruit and vegetable consumption. Public Health Nutr 1998 Mar;1(1):61-68. [Medline: 10555532]
6. Pomerleau J, Lock K, McKee M, Altmann DR. The challenge of measuring global fruit and vegetable intake. J Nutr 2004 May;134(5):1175-1180 [FREE Full text] [Medline: 15113966]
7. ABS. Australian Health Survey. First Results, 2011-12, cat. no. 4364.0.55.001 URL: http://www.abs.gov.au/ausstats/abs@.nsf/Lookup/4364.0.55.001main+features12011-12 [accessed 2015-08-26] [WebCite Cache ID 6b3kCMl2D]
8. Boffetta P, Couto E, Wichmann J, Ferrari P, Trichopoulos D, Bueno-de-Mesquita HB, van Duijnhoven Fränzel J B, Nielsen Michael R S, van Gils Carla H, et al. Fruit and vegetable intake and overall cancer risk in the European Prospective Investigation into Cancer and Nutrition (EPIC). J Natl Cancer Inst 2010 Apr 21;102(8):529-537 [FREE Full text] [doi: 10.1093/jnci/djq072] [Medline: 20371762]
9. Lunet N, Lacerda-Vieira A, Barros H. Fruit and vegetables consumption and gastric cancer: a systematic review and meta-analysis of cohort studies. Nutr Cancer 2005;53(1):1-10. [doi: 10.1207/s15327914nc5301_1] [Medline: 16351501]
10. Steinmetz KA, Potter JD. Vegetables, fruit, and cancer prevention: a review. J Am Diet Assoc 1996 Oct;96(10):1027-1039. [doi: 10.1016/S0002-8223(96)00273-8] [Medline: 8841165]
11. Dauchet L, Amouyel P, Hercberg S, Dallongeville J. Fruit and vegetable consumption and risk of coronary heart disease: a meta-analysis of cohort studies. J Nutr 2006 Oct;136(10):2588-2593 [FREE Full text] [Medline: 16988131]
12. Ness AR, Powles JW. Fruit and vegetables, and cardiovascular disease: a review. Int J Epidemiol 1997 Feb;26(1):1-13 [FREE Full text] [Medline: 9126498]
13. Appel LJ, Moore TJ, Obarzanek E, Vollmer WM, Svetkey LP, Sacks FM, et al. A clinical trial of the effects of dietary patterns on blood pressure. DASH Collaborative Research Group. N Engl J Med 1997 Apr 17;336(16):1117-1124. [doi: 10.1056/NEJM199704173361601] [Medline: 9099655]
14. Miura K, Greenland P, Stamler J, Liu K, Daviglus ML, Nakagawa H. Relation of vegetable, fruit, and meat intake to 7-year blood pressure change in middle-aged men: the Chicago Western Electric Study. Am J Epidemiol 2004 Mar 15;159(6):572-580 [FREE Full text] [Medline: 15003961]
15. Bazzano LA, He J, Ogden LG, Loria CM, Vupputuri S, Myers L, et al. Fruit and vegetable intake and risk of cardiovascular disease in US adults: the first National Health and Nutrition Examination Survey Epidemiologic Follow-up Study. Am J Clin Nutr 2002 Jul;76(1):93-99 [FREE Full text] [Medline: 12081821]
16. Mendis K, Forster T, Paxton K, Hyland K, Yelverton J, McLean R, et al. Large and forgotten in rural Australia: assessment, attitudes and possible approaches to losing weight in young adult males. BMC Public Health 2014;14:243 [FREE Full text] [doi: 10.1186/1471-2458-14-243] [Medline: 24612872]
17. Dumbrell S, Mathai D. Getting young men to eat more fruit and vegetables: a qualitative investigation. Health Promot J Austr 2008 Dec;19(3):216-221. [Medline: 19053939]
18. Harhay MO, King CH. Global burden of disease in young people aged 10-24 years. Lancet 2012 Jan 7;379(9810):27-8; author reply 28. [doi: 10.1016/S0140-6736(12)60019-1] [Medline: 22225664]
19. Fjeldsoe BS, Marshall AL, Miller YD. Behavior change interventions delivered by mobile telephone short-message service. Am J Prev Med 2009 Feb;36(2):165-173. [doi: 10.1016/j.amepre.2008.09.040] [Medline: 19135907]
20. 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]
21. Hurling R, Catt M, Boni MD, Fairley BW, Hurst T, Murray P, et al. Using internet and mobile phone technology to deliver an automated physical activity program: randomized controlled trial. J Med Internet Res 2007;9(2):e7 [FREE Full text] [doi: 10.2196/jmir.9.2.e7] [Medline: 17478409]
22. NIH. mHealth - Mobile Health Technologies URL: http://www.webcitation.org/6cYJhhyGD [accessed 2015-10-26] [WebCite Cache ID 6cYJhhyGD]
23. Eysenbach G. What is e-health? J Med Internet Res 2001;3(2):E20 [FREE Full text] [doi: 10.2196/jmir.3.2.e20] [Medline: 11720962]
24. Rainie L. Two-thirds of young adults and those with higher income are smartphone owners. Washington, DC: Pew Internet & American Life Project; 2012 Sep 11. URL: http://www.pewinternet.org/files/old-media/Files/Reports/2012/PIP_Smartphones_Sept12%209%2010%2012.pdf [accessed 2015-08-28] [WebCite Cache ID 6b6gtwgkH]
25. Duggan M, Smith A. Social media update 2013. Washington, DC: Pew Internet and American Life Project; 2013 Dec 30. URL: http://www.pewinternet.org/files/2013/12/PIP_Social-Networking-2013.pdf [accessed 2016-02-22] [WebCite Cache ID 6fUjv4sjP]

26. Pomerleau J, Lock K, Knai C, McKee M. Interventions designed to increase adult fruit and vegetable intake can be effective: a systematic review of the literature. J Nutr 2005 Oct;135(10):2486-2495 [FREE Full text] [Medline: 16177217]
27. Ammerman AS, Lindquist CH, Lohr KN, Hersey J. The efficacy of behavioral interventions to modify dietary fat and fruit and vegetable intake: a review of the evidence. Prev Med 2002 Jul;35(1):25-41. [Medline: 12079438]
28. Evans Charlotte E L, Christian MS, Cleghorn CL, Greenwood DC, Cade JE. Systematic review and meta-analysis of school-based interventions to improve daily fruit and vegetable intake in children aged 5 to 12 y. Am J Clin Nutr 2012 Oct;96(4):889-901 [FREE Full text] [doi: 10.3945/ajcn.111.030270] [Medline: 22952187]
29. Glasson C, Chapman K, James E. Fruit and vegetables should be targeted separately in health promotion programmes: differences in consumption levels, barriers, knowledge and stages of readiness for change. Public Health Nutr 2011 Apr;14(4):694-701. [doi: 10.1017/S1368980010001643] [Medline: 20576194]
30. Cook A, Roberts K, O'Leary F, Allman-Farinelli MA. Comparison of single questions and brief questionnaire with longer validated food frequency questionnaire to assess adequate fruit and vegetable intake. Nutrition 2015;31(7-8):941-947. [doi: 10.1016/j.nut.2015.01.006] [Medline: 26003391]
31. National Health Medical Research Council (NHMRC). Australian Dietary Guidelines: Summary. NHMRC Ref: N55. Canberra, Australia: Commonwealth Department of Health and Ageing; 2013.
32. NHS. Rough Guide: Fruit and Vegetable Portion Sizes.: NHS Choices URL: http://www.nhs.uk/livewell/5aday/documents/downloads/5aday_portion_guide.pdf [accessed 2015-09-25] [WebCite Cache ID 6bokXyRXx]
33. Abraham C, Michie S. A taxonomy of behavior change techniques used in interventions. Health Psychol 2008 May;27(3):379-387. [doi: 10.1037/0278-6133.27.3.379] [Medline: 18624603]
34. Bailey JV, Murray E, Rait G, Mercer CH, Morris RW, Peacock R, et al. Interactive computer-based interventions for sexual health promotion. Cochrane Database Syst Rev 2010(9):CD006483. [doi: 10.1002/14651858.CD006483.pub2] [Medline: 20824850]
35. 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]
36. Nour MM, Chen J, Allman-Farinelli M. Efficacy and External Validity of Electronic and Mobile Phone-Based Interventions Promoting Vegetable Intake in Young Adults: A Systematic Review Protocol. JMIR Res Protoc 2015;4(3):e92 [FREE Full text] [doi: 10.2196/resprot.4665] [Medline: 26220803]
37. Atkins D, Best D, Briss PA, Eccles M, Falck-Ytter Y, Flottorp S, et al. Grading quality of evidence and strength of recommendations. BMJ 2004 Jun 19;328(7454):1490 [FREE Full text] [doi: 10.1136/bmj.328.7454.1490] [Medline: 15205295]
38. Moschovitis C, Poole H, Senft T. History of the Internet: A Chronology, 1843 to the Present. USA: AB C-CLIO, Incorporated; 1999.
39. National Institutes of Health. Trials Use Technology to Help Young Adults Achieve Healthy Weights. 2010 Nov 29. URL: http://www.nih.gov/news/health/nov2010/nhlbi-29.htm [accessed 2015-09-27] [WebCite Cache ID 6bqMkhANC]
40. Higgins Julian P T, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, Sterne Jonathan A C, Cochrane Bias Methods Group, Cochrane Statistical Methods Group. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ 2011;343:d5928 [FREE Full text] [Medline: 22008217]
41. Wilson D. Practical Meta-Analysis. Effect Size Calculator URL: http://www.campbellcollaboration.org/escalc/html/EffectSizeCalculator-SMD1.php [accessed 2015-10-19] [WebCite Cache ID 6cPHsYSTq]
42. Lipsey M, Wilson D. Practical Meta-analysis. Thousand Oaks, CA: SAGE Publications; 2001.
43. Cohen. A power primer. Psychol Bull 1992 Jul;112(1):155-159. [Medline: 19565683]
44. Hu D, Huang J, Wang Y, Zhang D, Qu Y. Fruits and vegetables consumption and risk of stroke: a meta-analysis of prospective cohort studies. Stroke 2014 Jun;45(6):1613-1619 [FREE Full text] [doi: 10.1161/STROKEAHA.114.004836] [Medline: 24811336]
45. Wang X, Ouyang Y, Liu J, Zhu M, Zhao G, Bao W, et al. Fruit and vegetable consumption and mortality from all causes, cardiovascular disease, and cancer: systematic review and dose-response meta-analysis of prospective cohort studies. BMJ 2014;349:g4490 [FREE Full text] [Medline: 25073782]
46. Bertoia ML, Mukamal KJ, Cahill LE, Hou T, Ludwig DS, Mozaffarian D, et al. Changes in intake of fruits and vegetables and weight change in United States men and women followed for up to 24 years: analysis from three prospective cohort studies. PLoS Med 2015 Sep;12(9):e1001878 [FREE Full text] [doi: 10.1371/journal.pmed.1001878] [Medline: 26394033]
47. United States Department of Agriculture, United States Department of Health and Human Services. Nutrition and Your Health: Dietary Guidelines for Americans. 5th edition. Washington, DC: USDA; 2000:15.
48. Higgins JPT, Green S. Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0.: The Cochrane Collaboration; 2011 Mar. URL: http://handbook.cochrane.org/ [accessed 2016-02-19] [WebCite Cache ID 6fPQQWRRi]
49. Block G, Gillespie C, Rosenbaum EH, Jenson C. A rapid food screener to assess fat and fruit and vegetable intake. Am J Prev Med 2000 May;18(4):284-288. [Medline: 10788730]
50. Resnicow K, Odom E, Wang T, Dudley WN, Mitchell D, Vaughan R, et al. Validation of three food frequency questionnaires and 24-hour recalls with serum carotenoid levels in a sample of African-American adults. Am J Epidemiol 2000 Dec 1;152(11):1072-1080 [FREE Full text] [Medline: 11117617]

51. Thompson FE, Subar AF, Smith AF, Midthune D, Radimer KL, Kahle LL, et al. Fruit and vegetable assessment: performance of 2 new short instruments and a food frequency questionnaire. J Am Diet Assoc 2002 Dec;102(12):1764-1772. [Medline: 12487538]
52. Flood V, Webb K, Rangan A. Recommendations for Short Questions to Assess Food Consumption in Children for the NSW Health Surveys. Sydney, Australia: NSW Centre for Public Health Nutrition; 2005:1-109.
53. Kristal AR, Vizenor NC, Patterson RE, Neuhouser ML, Shattuck AL, McLerran D. Precision and bias of food frequency-based measures of fruit and vegetable intakes. Cancer Epidemiol Biomarkers Prev 2000 Sep;9(9):939-944 [FREE Full text] [Medline: 11008912]
54. Scott KM, Tobias MI, Sarfati D, Haslett SJ. SF-36 health survey reliability, validity and norms for New Zealand. Aust N Z J Public Health 1999 Aug;23(4):401-406. [Medline: 10462864]
55. Thompson FE, Kipnis V, Subar AF, Krebs-Smith SM, Kahle LL, Midthune D, et al. Evaluation of 2 brief instruments and a food-frequency questionnaire to estimate daily number of servings of fruit and vegetables. Am J Clin Nutr 2000 Jun;71(6):1503-1510 [FREE Full text] [Medline: 10837291]
56. Coles-Rutishauser I, Webb K, Abraham B, Allsopp R. Evaluation of Short Dietary Questions from the 1995 National Nutrition Survey. Brisbane, Australia: Australian Food and Nutrition Monitoring Unit, The University of Queensland; 2001:1-66.
57. Marks G. Monitoring Food Habits in the Australian Population Using Short Questions. Canberra, Australia: Ausinfo Canberra; 2001:1-112.
58. Brener ND, Kann L, Kinchen SA, Grunbaum JA, Whalen L, Eaton D, et al. Methodology of the youth risk behavior surveillance system. MMWR Recomm Rep 2004 Sep 24;53(RR-12):1-13 [FREE Full text] [Medline: 15385915]
59. Ma J, Betts NM, Horacek T, Georgiou C, White A, Nitzke S. The importance of decisional balance and self-efficacy in relation to stages of change for fruit and vegetable intakes by young adults. Am J Health Promot 2002;16(3):157-166. [Medline: 11802261]
60. Nelson M, Margetts BM, Black AE. Checklist for the methods section of dietary investigations. Metabolism 1993 Feb;42(2):258-259. [Medline: 8474324]
61. Green LW, Glasgow RE. Evaluating the relevance, generalization, and applicability of research: issues in external validation and translation methodology. Eval Health Prof 2006 Mar;29(1):126-153. [doi: 10.1177/0163278705284445] [Medline: 16510882]
62. Greene GW, White AA, Hoerr SL, Lohse B, Schembre SM, Riebe D, et al. Impact of an online healthful eating and physical activity program for college students. Am J Health Promot 2012;27(2):e47-e58. [doi: 10.4278/ajhp.110606-QUAN-239] [Medline: 23113786]
63. Kattelmann KK, Bredbenner CB, White AA, Greene GW, Hoerr SL, Kidd T, et al. The effects of Young Adults Eating and Active for Health (YEAH): a theory-based Web-delivered intervention. J Nutr Educ Behav 2014;46(6):S27-S41. [doi: 10.1016/j.jneb.2014.08.007] [Medline: 25457733]
64. Nitzke S, Kritsch K, Boeckner L, Greene G, Hoerr S, Horacek T, et al. A stage-tailored multi-modal intervention increases fruit and vegetable intakes of low-income young adults. Am J Health Promot 2007;22(1):6-14. [Medline: 17894257]
65. Hebden L, Cook A, van der Ploeg H P, King L, Bauman A, Allman-Farinelli M. A mobile health intervention for weight management among young adults: a pilot randomised controlled trial. J Hum Nutr Diet 2014 Aug;27(4):322-332. [doi: 10.1111/jhn.12155] [Medline: 23992038]
66. Partridge SR, McGeechan K, Hebden L, Balestracci K, Wong AT, Denney-Wilson E, et al. Effectiveness of a mHealth Lifestyle Program With Telephone Support (TXT2BFiT) to Prevent Unhealthy Weight Gain in Young Adults: Randomized Controlled Trial. JMIR Mhealth Uhealth 2015;3(2):e66 [FREE Full text] [doi: 10.2196/mhealth.4530] [Medline: 26076688]
67. Gow RW, Trace SE, Mazzeo SE. Preventing weight gain in first year college students: an online intervention to prevent the "freshman fifteen". Eat Behav 2010 Jan;11(1):33-39 [FREE Full text] [doi: 10.1016/j.eatbeh.2009.08.005] [Medline: 19962118]
68. Kothe EJ, Mullan BA. A randomised controlled trial of a theory of planned behaviour to increase fruit and vegetable consumption. Fresh Facts. Appetite 2014 Jul;78:68-75. [doi: 10.1016/j.appet.2014.03.006] [Medline: 24656949]
69. Clifford D, Anderson J, Auld G, Champ J. Good Grubbin': impact of a TV cooking show for college students living off campus. J Nutr Educ Behav 2009;41(3):194-200. [doi: 10.1016/j.jneb.2008.01.006] [Medline: 19411053]
70. Shahril MR, Wan Dali WP, Lua PL. A 10-week multimodal nutrition education intervention improves dietary intake among university students: cluster randomised controlled trial. J Nutr Metab 2013;2013:658642 [FREE Full text] [doi: 10.1155/2013/658642] [Medline: 24069535]
71. Lachausse RG. My student body: effects of an internet-based prevention program to decrease obesity among college students. J Am Coll Health 2012;60(4):324-330. [doi: 10.1080/07448481.2011.623333] [Medline: 22559092]
72. Rompotis CJ, Grove JR, Byrne SM. Benefits of habit-based informational interventions: a randomised controlled trial of fruit and vegetable consumption. Aust N Z J Public Health 2014 Jun;38(3):247-252. [doi: 10.1111/1753-6405.12232] [Medline: 24890483]

73. Franko DL, Cousineau TM, Trant M, Green TC, Rancourt D, Thompson D, et al. Motivation, self-efficacy, physical activity and nutrition in college students: randomized controlled trial of an internet-based education program. Prev Med 2008 Oct;47(4):369-377 [FREE Full text] [doi: 10.1016/j.ypmed.2008.06.013] [Medline: 18639581]
74. Kypri K, McAnally HM. Randomized controlled trial of a web-based primary care intervention for multiple health risk behaviors. Prev Med 2005;41(3-4):761-766. [doi: 10.1016/j.ypmed.2005.07.010] [Medline: 16120456]
75. Richards A, Kattelmann KK, Ren C. Motivating 18- to 24-year-olds to increase their fruit and vegetable consumption. J Am Diet Assoc 2006 Sep;106(9):1405-1411. [doi: 10.1016/j.jada.2006.06.005] [Medline: 16963345]
76. Carey L, Carey J, Dick W. The Systematic Design of Instruction. New York, NY: Longman; 2001.
77. Keller J. Motivational design of instruction. In: Reigeluth CM, editor. Instructional Design Theories and Models: An Overview of Their Current Status. Hillsdale, NJ: L Erlbaum Associates; 1983:383-434.
78. te Velde Saskia J, Twisk Jos W R, Brug J. Tracking of fruit and vegetable consumption from adolescence into adulthood and its longitudinal association with overweight. Br J Nutr 2007 Aug;98(2):431-438. [doi: 10.1017/S0007114507721451] [Medline: 17433126]
79. Strolla LO, Gans KM, Risica PM. Using qualitative and quantitative formative research to develop tailored nutrition intervention materials for a diverse low-income audience. Health Educ Res 2006 Aug;21(4):465-476 [FREE Full text] [doi: 10.1093/her/cyh072] [Medline: 16303783]
80. Baker PR, Francis DP, Soares J, Weightman AL, Foster C. Community wide interventions for increasing physical activity. Cochrane Database Syst Rev 2011(4):CD008366. [doi: 10.1002/14651858.CD008366.pub2] [Medline: 21491409]
81. Laforge RG, Greene GW, Prochaska JO. Psychosocial factors influencing low fruit and vegetable consumption. J Behav Med 1994 Aug;17(4):361-374. [Medline: 7966258]
82. Herbert G, Butler L, Kennedy O, Lobb A. Young UK adults and the 5 A DAY campaign: perceived benefits and barriers of eating more fruits and vegetables. Int J Consum Stud 2010;34(6):657-664. [doi: 10.1111/j.1470-6431.2010.00872.x]
83. Napolitano MA, Hayes S, Bennett GG, Ives AK, Foster GD. Using Facebook and text messaging to deliver a weight loss program to college students. Obesity (Silver Spring) 2013 Jan;21(1):25-31. [doi: 10.1002/oby.20232] [Medline: 23505165]
84. Hales SB, Davidson C, Turner-McGrievy GM. Varying social media post types differentially impacts engagement in a behavioral weight loss intervention. Transl Behav Med 2014 Dec;4(4):355-362 [FREE Full text] [doi: 10.1007/s13142-014-0274-z] [Medline: 25584084]
85. Drewnowski A, Rock CL, Henderson SA, Shore AB, Fischler C, Galan P, et al. Serum beta-carotene and vitamin C as biomarkers of vegetable and fruit intakes in a community-based sample of French adults. Am J Clin Nutr 1997 Jun;65(6):1796-1802 [FREE Full text] [Medline: 9174475]
86. OECD. Education at a Glance 2014: OECD Indicators.: OECD Publishing; 2014. URL: http://www.oecd.org/edu/Education-at-a-Glance-2014.pdf [accessed 2016-02-22] [WebCite Cache ID 6fUnrrpuB]
87. Australian Bureau of Statistics. Perspectives on Education and Training: Social Inclusion. Canberra, Australia: ABS; 2009.
88. Partridge SR, Juan SJ, McGeechan K, Bauman A, Allman-Farinelli M. Poor quality of external validity reporting limits generalizability of overweight and/or obesity lifestyle prevention interventions in young adults: a systematic review. Obes Rev 2015 Jan;16(1):13-31. [doi: 10.1111/obr.12233] [Medline: 25407633]
89. Laws RA, St George Alexis B, Rychetnik L, Bauman AE. Diabetes prevention research: a systematic review of external validity in lifestyle interventions. Am J Prev Med 2012 Aug;43(2):205-214. [doi: 10.1016/j.amepre.2012.04.017] [Medline: 22813687]
90. Klesges LM, Dzewaltowski DA, Glasgow RE. Review of external validity reporting in childhood obesity prevention research. Am J Prev Med 2008 Mar;34(3):216-223. [doi: 10.1016/j.amepre.2007.11.019] [Medline: 18312810]
91. Glasgow RE. eHealth evaluation and dissemination research. Am J Prev Med 2007 May;32(5 Suppl):S119-S126. [doi: 10.1016/j.amepre.2007.01.023] [Medline: 17466816]
92. Baranowski T, Cerin E, Baranowski J. Steps in the design, development and formative evaluation of obesity prevention-related behavior change trials. Int J Behav Nutr Phys Act 2009;6:6 [FREE Full text] [doi: 10.1186/1479-5868-6-6] [Medline: 19159476]

Abbreviations

GRADE: Grading of Recommendations Assessment, Development and Evaluation
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RCT: randomized controlled trial

Edited by S Kitsiou; submitted 27.08.15; peer-reviewed by S Hales, N Allen; comments to author 17.09.15; revised version received 07.11.15; accepted 22.01.16; published 08.04.16.

Please cite as:
Nour M, Chen J, Allman-Farinelli M
Efficacy and External Validity of Electronic and Mobile Phone-Based Interventions Promoting Vegetable Intake in Young Adults: Systematic Review and Meta-Analysis
J Med Internet Res 2016;18(4):e58
URL: http://www.jmir.org/2016/4/e58/
doi:10.2196/jmir.5082
PMID:27059765

Original Paper

Effectiveness of a Web-Based Computer-Tailored Multiple-Lifestyle Intervention for People Interested in Reducing their Cardiovascular Risk: A Randomized Controlled Trial

Vera Storm[1], MSc; Julia Dörenkämper[1], MSc; Dominique Alexandra Reinwand[1,2], MSc; Julian Wienert[3], PhD; Hein De Vries[2], PhD; Sonia Lippke[1,4], PhD

[1]Department of Psychology and Methods, Jacobs University Bremen, Bremen, Germany

[2]School for Public Health and Primary Care (CAPHRI), Department of Health Promotion, Maastricht University, Maastricht, Netherlands

[3]Institute for Social Medicine and Epidemiology, University of Lübeck, Lübeck, Germany

[4]Bremen International Graduate School for Social Sciences, Bremen, Germany

Corresponding Author:
Vera Storm, MSc
Department of Psychology and Methods
Jacobs University Bremen
Campus Ring 1
Bremen, 28759
Germany
Phone: 49 421200 ext 4721
Fax: 49 42120094721
Email: v.storm@jacobs-university.de

Abstract

Background: Web-based computer-tailored interventions for multiple health behaviors can improve the strength of behavior habits in people who want to reduce their cardiovascular risk. Nonetheless, few randomized controlled trials have tested this assumption to date.

Objective: The study aim was to test an 8-week Web-based computer-tailored intervention designed to improve habit strength for physical activity and fruit and vegetable consumption among people who want to reduce their cardiovascular risk. In a randomized controlled design, self-reported changes in perceived habit strength, self-efficacy, and planning across different domains of physical activity as well as fruit and vegetable consumption were evaluated.

Methods: This study was a randomized controlled trial involving an intervention group (n=403) and a waiting control group (n=387). Web-based data collection was performed in Germany and the Netherlands during 2013-2015. The intervention content was based on the Health Action Process Approach and involved personalized feedback on lifestyle behaviors, which indicated whether participants complied with behavioral guidelines for physical activity and fruit and vegetable consumption. There were three Web-based assessments: baseline (T0, N=790), a posttest 8 weeks after the baseline (T1, n=206), and a follow-up 3 months after the baseline (T2, n=121). Data analysis was conducted by analyzing variances and structural equation analysis.

Results: Significant group by time interactions revealed superior treatment effects for the intervention group, with substantially higher increases in self-reported habit strength for physical activity (F1,199=7.71, *P*=.006, Cohen’s d=0.37) and fruit and vegetable consumption (F1,199=7.71, *P*=.006, Cohen’s d=0.30) at posttest T1 for the intervention group. Mediation analyses yielded behavior-specific sequential mediator effects for T1 planning and T1 self-efficacy between the intervention and habit strength at follow-up T2 (fruit and vegetable consumption: beta=0.12, 95% CI 0.09-0.16, *P*<.001; physical activity: beta=0.04, 95% CI 0.02-0.06, *P*<.001).

Conclusions: Our findings indicate the general effectiveness and practicality of Web-based computer-tailored interventions in terms of increasing self-reported habit strength for physical activity and fruit and vegetable consumption. Self-efficacy and planning may play major roles in the mechanisms that facilitate the habit strength of these behaviors; therefore, they should be actively promoted in Web-based interventions. Although the results need to take into account the high dropout rates and medium effect sizes, a large number of people were reached and changes in habit strength were achieved after 3 months.

Trial Registration: Clinicaltrials.gov NCT01909349; https://clinicaltrials.gov/ct2/show/NCT01909349 (Archived by WebCite at http://www.webcitation.org/6g5F0qoft) and Nederlands Trial Register NTR3706 http://www.trialregister.nl/trialreg/admin/rctview.asp?TC=3706 (Archived by WebCite at http://www.webcitation.org/6g5F5HMLX)

(J Med Internet Res 2016;18(4):e78) doi:10.2196/jmir.5147

KEYWORDS

Web-based intervention; computer tailoring; cardiovascular disease; habit strength; self-efficacy; planning

Introduction

Cardiovascular diseases (CVD) are major causes of morbidity and mortality in men older than 45 years and women older than 65 years in most developed European countries, including Germany and the Netherlands [1-5]. Regular physical activity and a healthy diet play an important role in preventing CVD because of their wide range of beneficial effects on physical health. Results from different meta-analyses reveal that people who already suffer from CVD have a lower risk of reinfarction [6], cardiac mortality [6-8] and overall mortality [6,7,9] if they improve their physical activity levels. Sufficient fruit and vegetable consumption is also recommended because it reduces the risk of further cardiovascular issues, such as coronary heart disease [10,11] and ischemic heart disease [12], cardiac mortality [13], and overall mortality [11].

To reduce CVD risk, one of the main goals is to adopt a healthier lifestyle (ie, regular physical activity and sufficient fruit and vegetable consumption). Changes that make these behaviors more habitual are a desired goal in primary and secondary prevention because once a behavior has become habitual it requires less conscious effort and relapses become less likely [14,15]. Habituation of the healthy behavior may be the final phase in the health behavior change chain, whereby the behavior has stabilized and its strength has plateaued [14].

People who are aware of their risk for CVD are usually highly motivated to practice recommended health behaviors and break old, unhealthy habits [16]. However, the process of health behavior change involves motivational factors that extend beyond merely having knowledge about behavior change benefits [17] and good intentions [18]. Even when people successfully initiate recommended changes, the gains are often short term and, without intervention, behavior change adherence declines over time [19,20]. Thus, long-term studies investigating the underlying mechanisms of health behavior maintenance are needed.

There is ample empirical support that intentions for behavioral change may best provoke behavior initiation by increasing the use of self-regulation strategies (ie, self-efficacy [21,22] and planning [23-25]). According to Bandura [26], self-efficacy describes optimistic self-beliefs concerning the ability to cope with possible failure and recover from relapses. Perceived self-efficacy seems to be important at all points in the health behavior change process [21] and is not only important for behavior initiation, but also behavior maintenance, recovery, and habituation. Evidence for the relevance of techniques that increase self-efficacy can be found in intervention studies, which found that experimentally induced changes in selfefficacy were positively associated with behavior initiation at a later point [27-29]. However, research on the direct effect on habit strength is limited.

In addition to self-efficacy, it is likely that planning promotes habit strength because habits are assumed to result from frequent behavior enactment in stable settings [30-32]. For example, if a person plans to go swimming on Fridays after dinner, the behavior becomes closely tied to contextual cues such as the time and location for which he or she chose to perform the initial action plan and the behavior becomes automatized with minimal forethought [15]. Previous intervention studies using self-regulatory techniques have revealed effects on habit strength at short-term follow-up in the case of physical exercise [33] and nonsmoking [34], although no research is available addressing multiple behaviors.

Because habit strength is a relatively new concept in behavioral intervention research, it is not yet fully understood how planning and self-efficacy might interplay with habit strength. Interventions that make use of both self-efficacy and planning techniques may enhance social cognitions, thereby leading to increased habit strength. Thus, mediation analysis might unfold the underlying working mechanisms of such an intervention.

A growing area of research focuses on the incorporation of the Internet as a mode of delivery to allow for individualized behavior change interventions [35-38]. Because interventions cannot fit all populations and circumstances in the same way, tailoring intervention content and offering personal behavioral and action feedback might increase the effectiveness of such programs in comparison to generic interventions or so-called “one-size-fits-all” approaches [39,40]. Specific tailored feedback for individuals based on their perceptions about a given behavior may be similar to feedback provided in face-to-face interactions and thus hold a higher personal relevance for the participant [41-43]. In addition, compared to face-to-face interventions, tailored interventions are easily accessible when delivered via the Internet and provide a cost-effective means to reach a wide population [35,36]. Previous studies on computer-tailored Web-based health behavior change interventions provided positive results for a variety of health behaviors, including physical activity [44-48], fruit and vegetable consumption [49,50], and multiple health behaviors [51-53] among the general adult population as well as people with cardiovascular risk profiles [54,55]. Although these previous Web-based computer-tailored studies focused on behavioral achievement, this study extended this topic by directly assessing its ability to enhance habit strength.

In this study, the first objective is to investigate the effects of an 8-week Web-based computer-tailored intervention on improvements of self-reported physical activity habit strength and fruit and vegetable consumption habit strength among people who were interested in reducing their cardiovascular risk. Moreover, we also test the prediction that social cognitive variables targeted by the tailored intervention (ie, self-efficacy and planning) increase more from the baseline in the intervention group than in the control group. Finally, we investigate whether changes in self-efficacy and action planning mediate the effect of the intervention on improvements in habit strength after two follow-up measurements. Testing the mechanisms of how the intervention exhibits an effect on proximal indicators of habit strength is the added value of our research. It is only when we know whether interventions work in terms of supporting habit formation by successfully targeting self-efficacy and planning by model learning and concrete planning tasks that we can conclude what online interventions should address in the future.

Methods

A detailed description of the study protocol has been published previously [56]; therefore, only a summary of the study methodology and procedures is provided.

Study Design, Procedure, and Participants

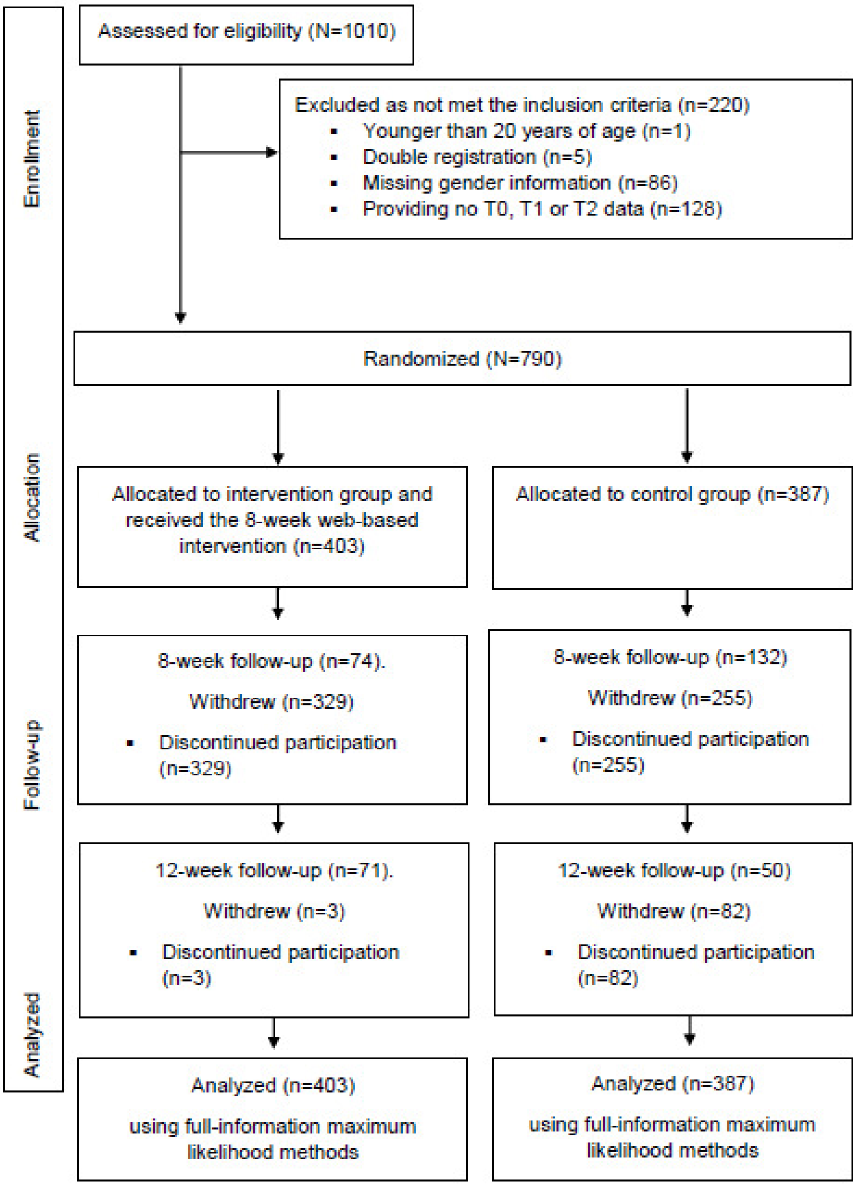
This study was a randomized controlled trial involving one intervention group and one waiting control group. There were three assessments: baseline (T0, N=790), a posttest at 8 weeks after the baseline (T1, n=206), and a follow-up 3 months after the baseline (T2, n=121). The waiting control group obtained access to the 8-week Web-based computer-tailored intervention at T2 after the intervention group had finished the intervention. The study received ethical approval by the Deutsche Gesellschaft für Psychologie in Germany (EK-A-SL022013) and the Medical Ethics Committee of Atrium Medical Centre Heerlen in the Netherlands (12-N-124).

Enrollment and follow-up took place from July 2014 to February 2015 in Germany and the Netherlands. We used different recruitment strategies: participants were recruited face-to-face by the authors of this study in 10 German and eight Dutch cardiac rehabilitation facilities and heart training groups. The authors of this study contacted the centers for acquisition and they were willing to participate. In addition, we called for participation via Internet platforms on diabetes and cardiovascular diseases as well as via an email invitation from two research agency online panels in Germany and the Netherlands. No data on how many participants were recruited through each strategy were available. The inclusion criteria were as follows: age between 20 and 85 years, no contraindications for physical activity and fruit and vegetable consumption, having an interest in improving physical activity and fruit and vegetable consumption, sufficient reading and writing skills in the relevant language (German or Dutch), and computer literacy and Internet access. Participation in the study was voluntary and data were anonymized.

Figure 1 shows the flow of participants from enrollment in the study to allocation to the two conditions (intervention group and waiting control group) and follow-up visits after 8 and 12 weeks. To obtain access to the Web-based questionnaires, the participants registered on the Rehabilitation-Aftercare for an optimal Transfer into Autonomous daily life (RENATA) website with a self-chosen nickname and password. The website was also open to the general public and provided broad information on the inclusion criteria and the procedure of the 8-week intervention as well as the duration of the questionnaires. Participants were made aware of the two-group design and the information provided was identical for all study participants, independent of the recruitment strategy.

After providing informed consent online, 1010 participants were then randomly assigned to either the intervention group or the waiting control group. Both groups took part in the identical baseline measurement (T0). Randomization into the intervention group and waiting control group was performed by the content management system, TailorBuilder, which was developed for Web-based tailored interventions. No block or cluster randomization was applied; rather, the randomization was conducted at the individual level. Participants and the authors of this study were blinded to their allocation for the duration of the study. The experiment was double blind. Overall, 220 datasets were excluded by the research team because of double registration (n=5), missing gender information (n=86), inadequate age (n=1 younger than 20 years), and no available T0, T1, or T2 data (n=128).

Figure 1. Flowchart of participants through the study.



Intervention Program

After registration, the 8-week intervention was delivered to the intervention group via the Internet, addressing physical activity in the first 4 weeks and fruit and vegetable consumption in the following 4 weeks. Once a week, the study participants were reminded to participate in the weekly intervention sessions and the follow-up measurements by automatically generated emails containing a link to the respective questionnaire.

The intervention was a Web-based computer-tailored intervention to increase physical activity and fruit and vegetable consumption among people who intended to change their physical activity and fruit and vegetable consumption. We used the Health Action Process Approach as a theoretical framework to develop the 8-week Web-based intervention [21,57]. The eight weekly sessions in the intervention period targeted the concepts of the different stages (nonintender, intender, and actor) via the use of behavior change techniques, such as providing information about behavioral risk, the benefits of behavior change, intention formation, barrier identification, prompting specific goal setting, and reviewing behavioral goals [58]. These techniques have proven effective in other computer tailoring programs [51,59]. Physical activity was discussed during the first four intervention sessions (sessions 1-4) and the last four sessions focused on fruit and vegetable consumption (sessions 5-8). In the following, we briefly describe the content of each session for both behaviors. In sessions 1 and 5, participants received tailored feedback about their risk perception, outcome expectancies, and their actual health behavior regarding physical activity and fruit and vegetable consumption based on their previous assessment. During the second and sixth sessions, participants were asked to determine personal goals and action plans for physical activity and fruit and vegetable consumption. During this session, participants received example plans and tailored feedback on how to structure a plan, and what elements a plan should contain (where, when, who, how long, with whom). Subsequently, participants had the opportunity to adapt their plans. Self-efficacy was also addressed during this time and the following sessions and participants were provided with motivating feedback on how to perform the desired behavior. During the third and seventh sessions, participants were asked whether they had succeeded in performing the action plan and if they would like to adjust the plan. People who indicated having problems formulating plans received role-model examples. Thereafter, coping planning was addressed in the third and seventh sessions, whereby participants were asked to identify personal barriers and generate coping plans. Again, these coping plans were evaluated during the next session and could be adjusted by the participants. The fourth and eighth sessions focused on social support. Participants developed an overview about people in their environment who could support them in achieving their plans. During the intervention, different types of feedback were provided. Ipsative feedback was used to provide participants with an overview of their development regarding physical activity and fruit and vegetable consumption during the intervention. This feedback was based on a short questionnaire that participants had to complete at the beginning of each session. Normative feedback was included to compare the participants' behavior with the norm of the population. Bar charts were included in each session to present the progress of the participant's behavior change. In this study, feedback and behavioral recommendations referred to the goals of five portions of fruit and vegetables a day and physical activity for at least 30 minutes five times a week because this constituted the best match for the target group [11,60]. Bar charts were included to monitor the behavior change progress and a personal tone was applied.

Measurement Instruments

All variables were self-assessed online at baseline (T0) and after the 8-week intervention period (T1) and during the 3-month follow-up (T2). Study participants indicated all social cognitive items on Likert scales ranging from "1=not true" to "7=exactly true."

Sociodemographic Variables

We assessed sociodemographic information such as gender (1=male, 2=female), year of birth, country (1=Netherlands, 2=Germany), employment status (1=working part-time, 2=working full-time, 3=in training, 4=unemployed, 5=retired, 6=housewife/-husband), marital status (1=single, 2=close relationship but not living together, 3=close relationship and living together, 4=marital partnership/common law marriage, 5=divorced, 6=widowed), and highest level of education (1=no school graduation, 2=primary school education, 3=secondary school education, 4=vocational school graduation, 5=university entrance diploma, 6=other) in the baseline questionnaire. The participants additionally reported body height and body weight to calculate their body mass index (BMI) at T0, T1, and T2.

Intentions

For physical activity, the three independent items used were "On 5 days a week for 30 minutes (or a minimum of 2.5 hours per week), I have the intention to perform..." (1) "strenuous physical activity," (2) "moderate physical activity," or (3) "walking activity" [59]. Intention about fruit and vegetable consumption was assessed using the item "I seriously intend to eat at least five portions of fruit and vegetable daily" [61].

Self-efficacy

Physical activity self-efficacy [62] was assessed with five items (Cronbach alpha=.88), such as "I am certain that I can be physically active permanently at a minimum of 5 days a week for 30 minutes." Self-efficacy for fruit and vegetable consumption [29] was assessed by five items (Cronbach alpha=.92), such as "I am certain that I can eat 5 portions of fruit and vegetable a day even if it is sometimes difficult."

Action Planning and Coping Planning

Action planning and coping planning were assessed using six items for physical activity (Cronbach alpha=.91) and six items for fruit and vegetable consumption (Cronbach alpha=.92). For both target behaviors, the question started with "For the next month, I have already planned in detail..." (1) "which physical activities I would like to do," (2) "when I have to be especially cautious not to stop being active," and (3) "what I can do in difficult situations to stick to my intentions" [63].

Planning for fruit and vegetable consumption started with the same phrase, followed by the three items: (1) "when I will eat 5 portions fruit and vegetables," (2) "which fruit and vegetables I will eat," and (3) "when I need to be especially cautious not to fall into my old eating habits" or "what I can do in difficult situations to stick to my intentions" [29].

Habit Strength

The strength of habit for physical activity (Cronbach alpha=.88) and fruit and vegetable consumption (Cronbach alpha=.93) was measured with an abbreviated version of the Self-Report Habit Index (SRHI) [64] and included the two items "Being physically active for at least 30 minutes on 5 days a week is something that..." and "Eating five portions of fruit and vegetable per day is something..." (1) "has become a confirmed habit" and (2) "I do without thinking about it."

Statistical Analyses

Data analysis was conducted with SPSS version 22. Dropout analysis was performed using ANOVAs for the quantitative variables age, baseline intentions, baseline habit strength, and BMI. Chi-square tests were performed for the categorical variables gender and country. Age, gender, country, employment status, marital status, highest level of education, and BMI were included as covariates in all analyses because we were not interested in subgroup differences.

To investigate the effectiveness of the intervention, we first conducted separate ANCOVAs with repeated measures analyses for the four outcome measures for habit strength for regular physical activity, habit strength for fruit and vegetable consumption, self-efficacy, and planning. In each analysis, time, group, and a group by time interaction were entered as independent variables and the group by time interactions were interpreted. Effect sizes for differences in means are presented as Cohen's *d*. Those effect sizes less than 0.30 were considered small, those between 0.30 and 0.80 were considered medium, and those larger than 0.80 were regarded as large [65].

To examine whether intervention effects on habit resulted from changes in social cognitive variables, we used mediation analysis to test indirect effects of intervention on change in habit strength through changes in those cognitions that the intervention aimed to modify. The mediation analyses were performed using SPSS AMOS mediation models according to Preacher and Hayes [66]. The bootstrapping approach (5000 bootstrap samples) was used to estimate 95% confidence intervals of the standardized effects of the intervention on habit strength through self-efficacy and planning.

Baseline habit strength and baseline levels of self-efficacy and planning were controlled for. The level of statistical significance was set at $P<.05$. All reported P values are two-tailed. We used no statistical measures to correct for multiple testing.

For the 16 variables used in the analyses, the missing data proportions were <21% at T0, <19% at T1, and <17% at T2. Therefore, missing study variables were estimated with the full-information maximum likelihood (FIML) method. We exported the estimated data to SPSS to perform further analyses. FIML is based on the maximum likelihood algorithm and, compared with other options (ie list-/pairwise deletion, regression imputation), maximum likelihood estimates exhibit the least bias [67]. For example, Demirtas et al [68] reported that parameters were estimated accurately with missing rates up to 25%.

Results

Participation and Sample Characteristics

The final sample consisted of 790 persons with a mean age of 50.9 years at baseline (SD 12.2, range 20-84). In all, 62.9% (497/790) of the participants were female, 71.8% (646/790) of the participants were married or in a relationship, and 569 participants (72.0%) were employed either full- or part-time. The mean BMI was 27.6 (SD 5.5, range 15.0-60.8), indicating that the participants as a group were considered overweight. Table 1 provides an overview of the main baseline variables in this study.

The t tests revealed small yet significant differences between German (n=371) and Dutch participants (n=419) regarding age (t_{789}=8.51, P=.004), BMI (t_{789}=5.38, P=.02), self-efficacy for physical activity (t_{789}=8.24, P=.004), and planning for fruit and vegetable consumption (t_{789}=18.91, P<.001). German participants were found to be slightly older and have a higher BMI and higher self-efficacy levels compared to the Dutch participants, whereas they reported less planning of their fruit and vegetable consumption. In addition, there were more women among the Dutch participants (χ^2_{790}=9.1, P=.004).

Dropout Analyses

Dropout analyses (1=dropout, 2=no dropout) showed no significant differences between the participants who completed all waves of data collection (T0, T1, and T2) and those who dropped out after T0 regarding age ($F_{1,789}$=1.11, P=.29), BMI ($F_{1,789}$=0.38, P=.54), baseline intentions (fruit and vegetable consumption: $F_{1,789}$=0.81, P=.78; physical activity: $F_{1,798}$=0.02, P=.90), and baseline habit strength (fruit and vegetable consumption: $F_{1,789}$=0.23, P=.63; physical activity: $F_{1,789}$=2.75, P=.10). In addition, the dropouts after T0 did not significantly differ from those who participated in the follow-up questionnaire in terms of gender (χ^2_{790}=2.1, P=.15) and country (χ^2_{790}=1.2, P=.27): Men and women, as well as German and Dutch participants dropped out after the baseline measurement T0 in equal numbers.

Those who dropped out after posttest T1 did not differ from those who participated in all measurement points in terms of age ($F_{1,789}$=3.36, P=.07), BMI ($F_{1,789}$=0.42, P=.52), baseline intentions (fruit and vegetable consumption: $F_{1,789}$=1.56, P=.21; physical activity: $F_{1,798}$=0.01, P=.98), and baseline habit strength for fruit and vegetable consumption ($F_{1,789}$=1.22, P=.27). In addition, the dropouts after T1 did not differ from those who participated in the follow-up questionnaire in terms of country (χ^2_{790} =1.2, P=.27); German and Dutch participants dropped out after the baseline measurement T0 in equal numbers. However, there were significant differences between those who dropped out after T1 and those who completed all measurement points in terms of baseline habit strength for physical activity ($F_{1,789}$=6.71, P=.01) and gender (χ^2_{790}=4.3, P=.04). Slightly more women than men dropped out after T1; those people who dropped out after T1 showed significantly lower baseline habit strength for physical activity than those who completed all measurement points. Overall, T0 to T1 attrition was 73.9% (584/790) and T0 to T2 attrition was 85.3% (705/790).

Table 2 presents the number of participants who participated in the single intervention sessions. Participation declined from 90.8% (314/790) in the first session to 19.9% (69/790) participation in the last session of the 8-week intervention. Participants completed a mean 2.0 (SD 2.4) intervention sessions of eight potential sessions. Most participants completed only one session (41.9%, 331/790), 15.3% (120/790) completed two sessions, 8.7% (68/790) completed three sessions, 4.6% (36/790) completed four sessions, 3.8% (30/790) completed five sessions, 5.8% (45/790) completed six sessions, 5.5% (43/790) completed seven sessions, and 8.1% (63/790) completed all eight sessions.

Intervention Effects on Baseline to Posttest Changes in Habit Strength

The assumption that the 8-week Web-based intervention would lead to an increase in habit strength for fruit and vegetable consumption and physical activity at posttest T1 was tested first. The results of the ANCOVA with repeated measurements (see Table 3) showed an interaction effect of group×time for habit strength for fruit and vegetable consumption ($F_{1,199}$=7.71, P=.006, Cohen’s d=0.30) as well as habit strength for physical activity ($F_{1,199}$=7.71, P=.006, Cohen’s d=0.37) with medium effect sizes. The intervention group showed a higher increase of dietary habit strength and physical activity habit strength from baseline (T0) to posttest (T1) than the waiting control group. This was then tested for the T2 follow-up, examining changes from T0 to T2. There was neither an intervention effect for habit strength for fruit and vegetable consumption ($F_{1,114}$=0.82, P=.14) nor for habit strength for physical activity ($F_{1,114}$=0.43, P=.24) at follow-up T2.

Table 1. Sample characteristics at baseline T0 (N=790).

Characteristics	Total (N=790)	Waiting control group (n=387)	Intervention group (n=403)
Age (years), mean (SD)	50.8 (12.2)	50.8 (12.3)	50.9 (12.0)
Gender, n (%)			
Male	293 (73.1)	151 (51.5)	142 (48.5)
Female	497 (62.9)	236 (47.5)	261 (52.5)
BMI (kg/m^2), mean (SD)	27.6 (5.4)	27.3 (5.2)	27.8 (5.6)
Physical activity, mean (SD)			
Intentions	3.9 (1.0)	3.9 (1.0)	4.0 (0.9)
Planning	4.3 (1.4)	4.3 (1.4)	4.3 (1.5)
Self-efficacy	4.6 (1.3)	4.5 (1.4)	4.7 (1.3)
Habit strength	3.5 (1.8)	3.4 (1.7)	3.6 (1.9)
Fruit and vegetable consumption, mean (SD)			
Intentions	4.5 (1.4)	4.5 (1.4)	4.6 (1.3)
Planning	3.8 (1.5)	3.7 (1.4)	3.8 (1.6)
Self-efficacy	4.7 (1.5)	4.6 (1.3)	4.7 (1.5)
Habit strength	3.7 (1.9)	3.6 (1.8)	3.8 (1.9)
Ethnicity, n (%)			
German	371 (47.0)	189 (50.9)	182 (49.1)
Dutch	419 (53.0)	198 (47.3)	221 (52.7)
Educational level, n (%)			
No school graduation	1 (0.1)	0 (0.0)	1 (100.0)
Primary school education	23 (2.9)	15 (65.2)	8 (34.8)
Secondary school education	86 (10.9)	40 (46.5)	46 (53.5)
Vocational school graduation	378 (47.8)	192 (50.8)	186 (49.2)
University entrance diploma	242 (30.6)	112 (46.3)	130 (53.7)
Other	60 (7.6)	28 (46.7)	32 (53.3)
Working situation, n (%)			
Working full-time	396 (50.1)	191 (48.2)	205 (51.8)
Working part-time	173 (21.9)	91 (52.6)	82 (47.4)
Schooling/vocational training	15 (1.9)	8 (53.3)	7 (46.7)
Unemployed	49 (6.2)	22 (44.9)	27 (55.1)
Retired	115 (14.6)	59 (51.3)	56 (48.7)
Housewife/-husband	42 (5.3)	16 (38.1)	26 (61.9)
Family status, n (%)			
Single	78 (9.9)	35 (44.9)	43 (55.1)
Close relationship but not living together	46 (5.8)	22 (47.8)	24 (52.2)
Close relationship and living together	76 (9.6)	41 (53.9)	35 (46.1)
Marital partnership/common law marriage	524 (66.3)	258 (49.2)	266 (50.8)
Divorced	54 (6.8)	27 (50.0)	27 (50.0)
Widowed	12 (1.5)	4 (33.3)	8 (66.7)

Table 2. Intervention use in terms of participation in the single sessions for physical activity and fruit and vegetable consumption.

Weekly intervention session	Participation in the specific session, n (%)
Physical activity	
Session 1	373 (47.2)
Session 2	240 (30.4)
Session 3	202 (25.6)
Session 4	148 (18.7)
Fruit and vegetable consumption	
Session 5	166 (21.0)
Session 6	141 (17.8)
Session 7	132 (16.7)
Session 8	123 (15.6)

Table 3. Changes of outcome measures from baseline (T0) to posttest (T1) (N=790).

Measure	Intervention group, mean (SD)	Waiting control group, mean (SD)	$F_{1,199}$	*P*	Cohen's *d*
Physical activity					
Self-efficacy	0.22 (1.47)	–0.06 (1.28)	2.11	.01	0.22
Planning	0.60 (1.76)	0.14 (1.25)	5.70	.02	0.35
Habit strength	1.00 (1.66)	0.34 (1.53)	7.71	.006	0.37
Fruit and vegetable consumption					
Self-efficacy	0.22 (1.47)	–0.06 (1.28)	1.40	.03	0.20
Planning	0.58 (1.72)	0.03 (1.60)	5.48	.02	0.36
Habit strength	0.83 (1.83)	0.26 (1.51)	7.71	.006	0.30

Intervention Effects on Baseline-Posttest Changes in Self-Efficacy and Planning

An ANCOVA with repeated measures revealed significant interaction effects of condition×time for self-efficacy for physical activity ($F_{1,199}$=2.11, *P*=.01, Cohen's *d*=0.22) and self-efficacy for fruit and vegetable consumption ($F_{1,199}$=1.40, *P*=.04, Cohen's *d*=0.20). The increase in self-efficacy from baseline (T0) to posttest (T1) was higher in the intervention group in comparison to the waiting control group. There was no intervention effect for self-efficacy for fruit and vegetable consumption ($F_{1,114}$=3.63, *P*=.06) nor for self-efficacy for physical activity ($F_{1,114}$=0.39, *P*=.54) at follow-up T2. For planning, we found a significant interaction effect of condition×time for both physical activity ($F_{1,199}$=5.70, *P*=.02, Cohen's *d*=.35) and fruit and vegetable consumption ($F_{1,199}$=5.48, *P*=.02, Cohen's *d*=0.36) with small to medium effect sizes. This indicates that the intervention led to a significantly higher increase in planning from baseline (T0) to posttest (T1) in the intervention group as compared to the waiting control group for both target behaviors.

Mediation Analyses

To address whether the intervention had an effect on habit strength through self-efficacy and planning, self-efficacy and planning at posttest T1 were considered to serve as sequential mediators between the intervention and habit strength at T2 follow-up. The entire hypothesized model is portrayed in Figure 2 and shows an acceptable fit to the data (χ^2_{190}=6.1, *P*<.001; comparative fit index=.91; Tucker-Lewis index=.81; root mean square error of approximation=.08). The intervention group condition significantly predicted T1 self-efficacy (physical activity: beta=0.32, *P*<.001; fruit and vegetable consumption: beta=0.39, *P*<.001), holding higher mean values in the intervention group. T1 self-efficacy was significantly interrelated with T1 planning (physical activity: beta=0.61, *P*<.001; fruit and vegetable consumption: beta=0.63, *P*<.001) for both target behaviors, whereas T1 planning predicted subsequent T2 habit strength (physical activity: beta=0.22, *P*<.001; fruit and vegetable consumption: beta=0.50, *P*<.001). Accordingly, people who planned more were also more likely to show strengthened habits later. Baseline habit strength also significantly predicted habit strength at T2 (physical activity: beta=0.79, *P*<.001; fruit and vegetable consumption: beta=0.43, *P*<.001). The standardized indirect effect of the intervention through T1 self-efficacy and T1 planning on T2 habit strength was beta=0.04 (95% CI 0.02-0.06) for physical activity and beta=0.12 (95% CI 0.09-0.16) for fruit and vegetable consumption. The multiple mediator model accounted for 68% of the variance in T2 physical activity habit strength and 44% of the variance in T2 fruit and vegetable consumption habit strength.

Figure 2. Conceptual model with standardized regression coefficients showing the effect of the intervention for fruit and vegetable consumption (FVC) and physical activity (PA) habit strength at follow-up controlling for age, gender, employment status, highest education, marital status, country, BMI, and baseline levels for self-efficacy and planning.

Discussion

Principal Results

The aim of this study was to test the effectiveness of a Web-based intervention in terms of improving habit strength for regular physical activity and fruit and vegetable consumption. The intervention led to significant increases in participants' self-reported physical activity habit strength as well as fruit and vegetable consumption habit strength 8 weeks after baseline. This is in line with the results from previous self-regulatory intervention studies, which yielded effects on habit strength with a short-term followup in the case of physical exercise [33] and smoking cessation [34]. However, our study extends these findings to the context of Web-based computer-tailored interventions for physical activity and also fruit and vegetable consumption, and shows the online practicality of a multiple behavior change intervention.

Previous research has mainly tested how habit strength is formed based on cues to action [14,15]. Such a cue can be medical treatment or advice received during an eHealth program. Instead, we investigated which self-regulatory mechanisms accounted for the effect of the intervention on habit strength. We were able to show that the intervention successfully addressed two key intervention variables, self-efficacy and planning, which subsequently mediated the intervention effect on habit strength 3 months after the baseline (ie, changes in self-efficacy and planning contributed additively to habit strength). This replicates the results of Fleig et al [31], who used a multiple health behavior model and theoretical assumptions [69]. Both self-efficacy and planning were predictive of habit strength at a later point; thus, they are not only important for behavior initiation [27-29], but also behavior maintenance. In future studies, this behavior maintenance should be researched in more depth with an Internet study design in a medical context (ie, with patients only).

The theoretical framework of habit strengthening used in this study may also be applied to processes when individuals try to break unhealthy habits, such as smoking and snacking. For example, Webb et al [34] showed that smokers with moderate or low smoking habits were successful in overriding their unhealthy habitual responses when accompanied with behavioral alternatives that they had specified in an action plan (eg, “If I am walking from the office to my car, then I will chew some gum instead of smoking a cigarette!”). These compensatory cognitions were not researched explicitly in this study, but should be addressed in the future [69].

Our results for the hypothesized mechanisms are important because they point toward the potential target constructs of Web-based interventions and how to make such interventions more efficient. It is imperative to address self-efficacy and planning to enable individuals to develop habits and translate intentions into behaviors. However, in future research this needs to be evaluated further. For instance, it should be tested whether the intervention effect only translates in the sequence via self-efficacy first and then planning or whether it could also be that self-efficacy moderates the mediation of intentions into behavior via planning. In addition, experimental designs should test whether the intervention addressing planning works only in intenders or people with high self-efficacy because it was found in previous studies [63].

Limitations

This study is subject to some limitations, including the measurement of our criterion variables fruit and vegetable consumption and physical activity habit strength. The SRHI offers a standardized and reliable measure to assess habituated action with evidence across behaviors and populations [70,71]. However, we relied on a short version of the habit strength measure referring to general physical activity and fruit and vegetable consumption. Future studies may include additional items of the SRHI to capture further facets of habitual automaticity (eg, lack of control). In addition, self-report for behavioral outcome measures of intervention studies can be criticized for their limitations, such as response and recall bias, underreporting, socially desirable answers, and measurement errors [72]. Thus, the inclusion of measures such as biomarkers or pedometers is advocated as an objective indicator of effectiveness [73,74].

Furthermore, the high dropout in our study needs to be addressed. Web-based interventions typically come with dropout rates [75] that can be very high (eg, up to 86% [76]). Although appropriate usage of the intervention differs among certain participant characteristics [76], we did not find any personal characteristics that could explain high dropout. Due to our widespread recruitment strategies, it can be assumed that a large number of the participants who signed up for the intervention did so out of curiosity rather than having a genuine interest in changing their health behavior, which could be one explanation for the high dropout rate [77]. Furthermore, dropout and not responding to questions could also be caused by intervention characteristics, such as the length of the questionnaire, layout, or navigation difficulties through the intervention [78]. Future studies should further investigate characteristics of dropout and nonresponse to eHealth interventions and consider how to tackle them to obtain larger sample sizes, including at follow-up. The results from a recent systematic review [79] show that the differences in technology and interaction predict user adherence in Web-based interventions.

One possible recommendation for future interventions is the inclusion of social media interaction, the integration of environmental components, and regular updates to promote adherence. Participants might show higher levels of engagement and complete program challenges in a Web-based program when they have social ties [80], the possibility to exchange experiences with others, and receive social support [81]. In addition, because environmental intervention components (eg, information on how to plan a cycling route for being physically active) might support people in finding possibilities to translate their goal intentions to behaviors [44], the integration of these environmental components could be useful in stimulating more active intervention participation.

Finally, our study participants form a rather heterogeneous group because we included participants via different recruitment channels. Unfortunately, no data on how many participants were recruited through each strategy are available. Although baseline intentions and sociodemographic data were controlled for in all analyses, a physician rating, medical diagnosis, or objective index of medical severity should be included as a control in future studies. In addition, there were small yet significant differences between German and Dutch participants regarding the measures used, although it is assumed that none of the differences in the results are due to country given that both countries have similar nutrition recommendations [82,83] and prevention campaigns (eg, “5 a day”) [84,85].

Conclusion

The results of this study are important for the future development of Web-based computer-tailored interventions to improve lifestyle behaviors that can reduce the risk of cardiovascular events. Web-based computer-tailored interventions can be a suitable delivery mode to successfully foster changes in self-efficacy and planning, which predict physical activity and fruit and vegetable consumption habit strength. This research adds to the growing literature on real-world habit strengthening, with our findings suggesting that planning and confidence in one’s action may aid the process of making behavior automatic. Our results add up to the current body of knowledge because they display mechanisms of how this intervention affects behavioral habit change. Future interventions should address habit formation by targeting self-efficacy and planning by model learning and concrete planning tasks.

Due to the high dropout rate, our results must be interpreted with caution, although these findings can guide further research. In particular, the investigated constructs and mechanisms should be further elaborated. Practical implications can be retrieved from the fact that habit-strengthening resources for plans and self-efficacy boosters can be delivered briefly via the Internet, are easy for people to implement, and theoretically have the potential for longer-term impact.

Future interventions may benefit from aids to support the creation and recall of plans, particularly when accompanied with self-efficacy prompting techniques, such as vicarious experience, personalized feedback, providing contingent rewards, self-monitoring (tracking one's own food- and exercise-related behavior), and becoming conscious of mastery experience [86,87].

Acknowledgments

This research was funded by the Wilhelm-Stiftung für Rehabilitationsforschung within the Donors' Association for the Promotion of Humanities and Sciences in Germany [Deutsches Stiftungszentrum im Stifterverband fÃ¼r die Deutsche Wissenschaft]. We would like to thank Amanda Whittal (Jacobs University Bremen) for proofreading this manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

CONSORT eHealth Checklist.

[PDF File (Adobe PDF File), 1MB - jmir_v18i4e78_app1.pdf]

References

1. Gößwald A, Schienkiewitz A, Nowossadeck E, Busch M. Prevalence of myocardial infarction and coronary heart disease in adults aged 40-79 years in Germany: results of the German Health Interview and Examination Survey for Adults (DEGS1). Bundesgesundheitsblatt 2013;56:650-655.
2. Leening MJ, Siregar S, Vaartjes I, Bots ML, Versteegh MI, van Geuns RJ, et al. Heart disease in the Netherlands: a quantitative update. Neth Heart J 2014 Jan;22(1):3-10 [FREE Full text] [doi: 10.1007/s12471-013-0504-x] [Medline: 24343132]
3. Mendis S, Puska P, Norvving B, editors. Global Atlas on Cardiovascular Disease Prevention and Control. Geneva: World Health Organization; 2011.
4. Nichols M, Townsend N, Scarborough P, Rayner M. Cardiovascular disease in Europe 2014: epidemiological update. Eur Heart J 2014 Nov 7;35(42):2950-2959. [doi: 10.1093/eurheartj/ehu299]
5. Organisation for Economic Co-operation and Development. Health at a Glance 2013: OECD Indicators. 2013. URL: http://www.oecd-ilibrary.org/social-issues-migration-health/health-at-a-glance-2013_health_glance-2013-en [accessed 2016-03-18] [WebCite Cache ID 6g62j3Ucn]
6. Lawler PR, Filion KB, Eisenberg MJ. Efficacy of exercise-based cardiac rehabilitation post-myocardial infarction: a systematic review and meta-analysis of randomized controlled trials. Am Heart J 2011 Oct;162(4):571-584. [doi: 10.1016/j.ahj.2011.07.017] [Medline: 21982647]
7. Heran BS, Chen JM, Ebrahim S, Moxham T, Oldridge N, Rees K, et al. Exercise-based cardiac rehabilitation for coronary heart disease. Cochrane Database Syst Rev 2011(7):CD001800 [FREE Full text] [doi: 10.1002/14651858.CD001800.pub2] [Medline: 21735386]
8. Li J, Siegrist J. Physical activity and risk of cardiovascular disease--a meta-analysis of prospective cohort studies. Int J Environ Res Public Health 2012 Feb;9(2):391-407 [FREE Full text] [doi: 10.3390/ijerph9020391] [Medline: 22470299]
9. Taylor RS, Brown A, Ebrahim S, Jolliffe J, Noorani H, Rees K, et al. Exercise-based rehabilitation for patients with coronary heart disease: systematic review and meta-analysis of randomized controlled trials. Am J Med 2004 May 15;116(10):682-692. [doi: 10.1016/j.amjmed.2004.01.009] [Medline: 15121495]
10. Mente A, de Koning L, Shannon HS, Anand SS. A systematic review of the evidence supporting a causal link between dietary factors and coronary heart disease. Arch Intern Med 2009 Apr 13;169(7):659-669. [doi: 10.1001/archinternmed.2009.38] [Medline: 19364995]
11. Wang X, Ouyang Y, Liu J, Zhu M, Zhao G, Bao W, et al. Fruit and vegetable consumption and mortality from all causes, cardiovascular disease, and cancer: systematic review and dose-response meta-analysis of prospective cohort studies. BMJ 2014;349:g4490 [FREE Full text] [Medline: 25073782]
12. Crowe FL, Roddam AW, Key TJ, Appleby PN, Overvad K, Jakobsen MU, European Prospective Investigation into CancerNutrition (EPIC)-Heart Study Collaborators. Fruit and vegetable intake and mortality from ischaemic heart disease:

results from the European Prospective Investigation into Cancer and Nutrition (EPIC)-Heart study. Eur Heart J 2011 May;32(10):1235-1243 [FREE Full text] [doi: 10.1093/eurheartj/ehq465] [Medline: 21245490]
13. Oyebode O, Gordon-Dseagu V, Walker A, Mindell JS. Fruit and vegetable consumption and all-cause, cancer and CVD mortality: analysis of Health Survey for England data. J Epidemiol Community Health 2014 Sep;68(9):856-862 [FREE Full text] [doi: 10.1136/jech-2013-203500] [Medline: 24687909]
14. 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]
15. Orbell S, Verplanken B. The automatic component of habit in health behavior: habit as cue-contingent automaticity. Health Psychol 2010 Jul;29(4):374-383. [doi: 10.1037/a0019596] [Medline: 20658824]
16. Johnston DW, Johnston M, Pollard B, Kinmonth A, Mant D. Motivation is not enough: prediction of risk behavior following diagnosis of coronary heart disease from the theory of planned behavior. Health Psychol 2004 Sep;23(5):533-538. [doi: 10.1037/0278-6133.23.5.533] [Medline: 15367073]
17. DiMatteo MR, Haskard KB, Williams SL. Health beliefs, disease severity, and patient adherence: a meta-analysis. Med Care 2007 Jun;45(6):521-528. [doi: 10.1097/MLR.0b013e318032937e] [Medline: 17515779]
18. Judah G, Gardner B, Aunger R. Forming a flossing habit: an exploratory study of the psychological determinants of habit formation. Br J Health Psychol 2013 May;18(2):338-353. [doi: 10.1111/j.2044-8287.2012.02086.x] [Medline: 22989272]
19. Froger-Bompas C, Laviolle B, Guillo P, Letellier C, Ligier K, Daubert J, et al. Sustained positive impact of a coronary rehabilitation programme on adherence to dietary recommendations. Arch Cardiovasc Dis 2009 Feb;102(2):97-104 [FREE Full text] [doi: 10.1016/j.acvd.2008.10.020] [Medline: 19303576]
20. Jackson L, Leclerc J, Erskine Y, Linden W. Getting the most out of cardiac rehabilitation: a review of referral and adherence predictors. Heart 2005 Jan;91(1):10-14 [FREE Full text] [doi: 10.1136/hrt.2004.045559] [Medline: 15604322]
21. 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]
22. Strecher VJ, McEvoy DeVellis B, Becker MH, Rosenstock IM. The role of self-efficacy in achieving health behavior change. Health Educ Behav 1986 Jan 01;13(1):73-92. [doi: 10.1177/109019818601300108]
23. Carraro N, Gaudreau P. Spontaneous and experimentally induced action planning and coping planning for physical activity: A meta-analysis. Psychol Sport Exerc 2013 Mar;14(2):228-248. [doi: 10.1016/j.psychsport.2012.10.004]
24. Hagger MS, Luszczynska A. Implementation intention and action planning interventions in health contexts: state of the research and proposals for the way forward. Appl Psychol Health Well Being 2014 Mar;6(1):1-47. [doi: 10.1111/aphw.12017] [Medline: 24591064]
25. de Vries H, Eggers SM, Bolman C. The role of action planning and plan enactment for smoking cessation. BMC Public Health 2013;13:393 [FREE Full text] [doi: 10.1186/1471-2458-13-393] [Medline: 23622256]
26. Bandura A. Health promotion by social cognitive means. Health Educ Behav 2004 Apr;31(2):143-164. [doi: 10.1177/1090198104263660] [Medline: 15090118]
27. Allison MJ, Keller C. Self-efficacy intervention effect on physical activity in older adults. West J Nurs Res 2004 Feb;26(1):31-46; discussion 47. [doi: 10.1177/0193945903259350] [Medline: 14984643]
28. Kreausukon P, Gellert P, Lippke S, Schwarzer R. Planning and self-efficacy can increase fruit and vegetable consumption: a randomized controlled trial. J Behav Med 2012 Aug;35(4):443-451. [doi: 10.1007/s10865-011-9373-1] [Medline: 21822980]
29. Luszczynska A, Tryburcy M, Schwarzer R. Improving fruit and vegetable consumption: a self-efficacy intervention compared with a combined self-efficacy and planning intervention. Health Educ Res 2007 Oct;22(5):630-638 [FREE Full text] [doi: 10.1093/her/cyl133] [Medline: 17060349]
30. Lally P, van Jaarsveld CH, Potts HW, Wardle J. How are habits formed: Modelling habit formation in the real world. Eur J Soc Psychol 2009 Jul 16;40(6):998-1009. [doi: 10.1002/ejsp.674]
31. Fleig L, Pomp S, Schwarzer R, Lippke S. Promoting exercise maintenance: how interventions with booster sessions improve long-term rehabilitation outcomes. Rehabil Psychol 2013 Nov;58(4):323-333. [doi: 10.1037/a0033885] [Medline: 24041250]
32. de Vries H, Eggers SM, Lechner L, van Osch L, van Stralen MM. Predicting fruit consumption: the role of habits, previous behavior and mediation effects. BMC Public Health 2014;14:730 [FREE Full text] [doi: 10.1186/1471-2458-14-730] [Medline: 25037859]
33. Fleig L, Lippke S, Pomp S, Schwarzer R. Intervention effects of exercise self-regulation on physical exercise and eating fruits and vegetables: a longitudinal study in orthopedic and cardiac rehabilitation. Prev Med 2011 Sep;53(3):182-187. [doi: 10.1016/j.ypmed.2011.06.019] [Medline: 21784096]
34. Webb TL, Sheeran P, Luszczynska A. Planning to break unwanted habits: habit strength moderates implementation intention effects on behaviour change. Br J Soc Psychol 2009 Sep;48(Pt 3):507-523. [doi: 10.1348/014466608X370591] [Medline: 18851764]
35. 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]

36. Bennett GG, Glasgow RE. The delivery of public health interventions via the Internet: actualizing their potential. Annu Rev Public Health 2009;30:273-292. [doi: 10.1146/annurev.publhealth.031308.100235] [Medline: 19296777]
37. Griffiths F, Lindenmeyer A, Powell J, Lowe P, Thorogood M. Why are health care interventions delivered over the internet? A systematic review of the published literature. J Med Internet Res 2006;8(2):e10 [FREE Full text] [doi: 10.2196/jmir.8.2.e10] [Medline: 16867965]
38. Wantland DJ, Portillo CJ, Holzemer WL, Slaughter R, McGhee EM. The effectiveness of Web-based vs. non-Web-based interventions: a meta-analysis of behavioral change outcomes. J Med Internet Res 2004 Nov 10;6(4):e40 [FREE Full text] [doi: 10.2196/jmir.6.4.e40] [Medline: 15631964]
39. Hawkins RP, Kreuter M, Resnicow K, Fishbein M, Dijkstra A. Understanding tailoring in communicating about health. Health Educ Res 2008 Jun;23(3):454-466 [FREE Full text] [doi: 10.1093/her/cyn004] [Medline: 18349033]
40. De Vries H, Brug J. Editorial. Patient Educ Couns 1999 Feb;36(2):99-105. [doi: 10.1016/S0738-3991(98)00127-X]
41. Krebs P, Prochaska JO, Rossi JS. A meta-analysis of computer-tailored interventions for health behavior change. Prev Med 2010;51(3-4):214-221 [FREE Full text] [doi: 10.1016/j.ypmed.2010.06.004] [Medline: 20558196]
42. Kreuter M, Wray R. Tailored and targeted health communication: strategies for enhancing information relevance. Am J Health Behav 2003;27 Suppl 3:S227-S232. [Medline: 14672383]
43. Smeets T, Kremers SP, Brug J, de Vries H. Effects of tailored feedback on multiple health behaviors. Ann Behav Med 2007 Apr;33(2):117-123. [doi: 10.1080/08836610701307801] [Medline: 17447863]
44. Peels DA, Bolman C, Golsteijn RH, de Vries H, Mudde AN, van Stralen MM, et al. Long-term efficacy of a printed or a Web-based tailored physical activity intervention among older adults. Int J Behav Nutr Phys Act 2013;10:104 [FREE Full text] [doi: 10.1186/1479-5868-10-104] [Medline: 24228862]
45. van Stralen MM, de Vries H, Mudde AN, Bolman C, Lechner L. The long-term efficacy of two computer-tailored physical activity interventions for older adults: main effects and mediators. Health Psychol 2011 Jul;30(4):442-452. [doi: 10.1037/a0023579] [Medline: 21639638]
46. Neville L, O'Hara B, Milat A. Computer-tailored physical activity behavior change interventions targeting adults: a systematic review. Int J Behav Nutr Phys Act 2009;6:30 [FREE Full text] [doi: 10.1186/1479-5868-6-30] [Medline: 19490649]
47. Bossen D, Veenhof C, Dekker J, de Bakker D. The effectiveness of self-guided web-based physical activity interventions among patients with a chronic disease: a systematic review. J Phys Act Health 2014 Mar;11(3):665-677. [doi: 10.1123/jpah.2012-0152] [Medline: 23493018]
48. 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]
49. Springvloet L, Lechner L, de VH, Candel Math J J M, Oenema A. Short- and medium-term efficacy of a Web-based computer-tailored nutrition education intervention for adults including cognitive and environmental feedback: randomized controlled trial. J Med Internet Res 2015;17(1):e23 [FREE Full text] [doi: 10.2196/jmir.3837] [Medline: 25599828]
50. Oenema A, Tan F, Brug J. Short-term efficacy of a web-based computer-tailored nutrition intervention: main effects and mediators. Ann Behav Med 2005 Feb;29(1):54-63. [doi: 10.1207/s15324796abm2901_8] [Medline: 15677301]
51. 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]
52. Broekhuizen K, Kroeze W, van Poppel M, 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 2012:259-286. [doi: 10.1007/s12160-012-9384-3]
53. Lustria ML, Noar SM, Cortese J, Van Stee SK, Glueckauf RL, Lee J. A meta-analysis of web-delivered tailored health behavior change interventions. J Health Commun 2013;18(9):1039-1069. [doi: 10.1080/10810730.2013.768727] [Medline: 23750972]
54. Widmer RJ, Allison TG, Keane B, Dallas A, Lerman LO, Lerman A. Using an online, personalized program reduces cardiovascular risk factor profiles in a motivated, adherent population of participants. Am Heart J 2014 Jan;167(1):93-100. [doi: 10.1016/j.ahj.2013.09.019] [Medline: 24332147]
55. Kuhl EA, Sears SF, Conti JB. Internet-based behavioral change and psychosocial care for patients with cardiovascular disease: a review of cardiac disease-specific applications. Heart Lung 2006;35(6):374-382. [doi: 10.1016/j.hrtlng.2006.02.004] [Medline: 17137938]
56. Reinwand D, Kuhlmann T, Wienert J, de Vries H, Lippke S. Designing a theory- and evidence-based tailored eHealth rehabilitation aftercare program in Germany and the Netherlands: study protocol. BMC Public Health 2013;13:1081 [FREE Full text] [doi: 10.1186/1471-2458-13-1081] [Medline: 24245493]
57. Schwarzer R. Modeling health behavior change: how to predict and modify the adoption and maintenance of health behaviors. Appl Psychol 2008 Jan;57(1):1-29. [doi: 10.1111/j.1464-0597.2007.00325.x]
58. Abraham C, Michie S. A taxonomy of behavior change techniques used in interventions. Health Psychol 2008 May;27(3):379-387. [doi: 10.1037/0278-6133.27.3.379] [Medline: 18624603]

59. Compernolle S, Vandelanotte C, Cardon G, De Bourdeaudhuij I, De Cocker K. Effectiveness of a web-based, computer-tailored, pedometer-based physical activity intervention for adults: a cluster randomized controlled trial. J Med Internet Res 2015;17(2):e38 [FREE Full text] [doi: 10.2196/jmir.3402] [Medline: 25665498]
60. Sattelmair J, Pertman J, Ding E, Kohl H, Haskell W, Lee I. Dose response between physical activity and risk of coronary heart disease: a meta-analysis. Circulation 2011;124(7):789-795.
61. Lippke S, Ziegelmann J, Schwarzer R, Velicer W. Validity of stage assessment in the adoption and maintenance of physical activity and fruit and vegetable consumption. Health Psychol 2009;28(2):183-193. [doi: 10.1037/a0012983]
62. Luszczynska A, Sutton S. Physical activity after cardiac rehabilitation: evidence that different types of self-efficacy are important in maintainers and relapsers. Rehabil Psychol 2006;51(4):314-321. [doi: 10.1037/0090-5550.51.4.314]
63. Lippke S, Schwarzer R, Ziegelmann JP, Scholz U, Schüz B. Testing stage-specific effects of a stage-matched intervention: a randomized controlled trial targeting physical exercise and its predictors. Health Educ Behav 2010 Aug;37(4):533-546. [doi: 10.1177/1090198109359386] [Medline: 20547760]
64. Verplanken B, Orbell S. Reflections on past behavior: A self-report index of habit strength. J Appl Soc Psychol 2003;33:1313-1330. [doi: 10.1111/j.1559-1816.2003.tb01951.x]
65. Cohen J. Statistical Power Analysis for the Behavioral Sciences. 2nd ed. New York: Lawrence Erlbaum Associates; 1988.
66. Preacher KJ, Hayes AF. SPSS and SAS procedures for estimating indirect effects in simple mediation models. Behav Res Methods Instrum Comput 2004 Nov;36(4):717-731. [Medline: 15641418]
67. Enders CK. Applied Missing Data Analysis (Methodology In The Social Sciences). New York: The Guilford Press; 2010.
68. Demirtas H, Freels SA, Yucel RM. Plausibility of multivariate normality assumption when multiply imputing non-Gaussian continuous outcomes: a simulation assessment. J Stat Comput Sim 2008 Feb;78(1):69-84. [doi: 10.1080/10629360600903866]
69. Lippke S. Modelling and supporting complex behavior change related to obesity and diabetes prevention and management with the compensatory carry-over action model. J Diabetes Obes 2014;1(2):1-5. [doi: 10.15436/2376-0494.14.009]
70. Thurn J, Finne E, Brandes M, Bucksch J. Validation of physical activity habit strength with subjective and objective criterion measures. Psychol Sport Exerc 2014 Jan;15(1):65-71. [doi: 10.1016/j.psychsport.2013.09.009]
71. 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]
72. Prince S, Adamo K, Hamel M, Hardt J, Connor GS, Tremblay M. A comparison of direct versus self-report measures for assessing physical activity in adults: a systematic review. Int J Behav Nutr Phys Act 2008;5:56 [FREE Full text] [doi: 10.1186/1479-5868-5-56] [Medline: 18990237]
73. Bravata DM, Smith-Spangler C, Sundaram V, Gienger AL, Lin N, Lewis R, et al. Using pedometers to increase physical activity and improve health: a systematic review. JAMA 2007 Nov 21;298(19):2296-2304. [doi: 10.1001/jama.298.19.2296] [Medline: 18029834]
74. Combs GJ, Trumbo P, McKinley M, Milner J, Studenski S, Kimura T, et al. Biomarkers in nutrition: new frontiers in research and application. Ann N Y Acad Sci 2013 Mar;1278:1-10 [FREE Full text] [doi: 10.1111/nyas.12069] [Medline: 23488557]
75. Eysenbach G. The law of attrition. J Med Internet Res 2005;7(1):e11 [FREE Full text] [doi: 10.2196/jmir.7.1.e11] [Medline: 15829473]
76. Reinwand D, Crutzen R, Elfeddali I, Schneider F, Schulz D, Stanczyk NE, et al. Impact of educational level on study attrition and evaluation of web-based computer-tailored interventions: results from seven randomized controlled trials. J Med Internet Res 2015;17(10):e228 [FREE Full text] [doi: 10.2196/jmir.4941] [Medline: 26446779]
77. Crutzen R, Ruiter R. Interest in behaviour change interventions: a conceptual model. Eur Health Psychol 2015;17(1):6-11.
78. Brouwer W, Kroeze W, Crutzen R, de Nooier J, de Vries NK, Brug J, et al. Which intervention characteristics are related to more exposure to internet-delivered healthy lifestyle promotion interventions? A systematic review. J Med Internet Res 2011;13(1):e2 [FREE Full text] [doi: 10.2196/jmir.1639] [Medline: 21212045]
79. Kelders SM, Kok RN, Ossebaard HC, Van Gemert-Pijnen JE. Persuasive system design does matter: a systematic review of adherence to web-based interventions. J Med Internet Res 2012;14(6):e152 [FREE Full text] [doi: 10.2196/jmir.2104] [Medline: 23151820]
80. Poirier J, Cobb NK. Social influence as a driver of engagement in a web-based health intervention. J Med Internet Res 2012;14(1):e36 [FREE Full text] [doi: 10.2196/jmir.1957] [Medline: 22356829]
81. Oh HJ, Lauckner C, Boehmer J, Fewins-Bliss R, Li K. Facebooking for health: An examination into the solicitation and effects of health-related social support on social networking sites. Comput Hum Behav 2013 Sep;29(5):2072-2080. [doi: 10.1016/j.chb.2013.04.017]
82. World Health Organization. Food Based Dietary Guidelines in the WHO European Region. Copenhagen: World Health Organization; 2003.
83. World Health Organization. Comparative Analysis of Nutrition Policies in the WHO European Region: A Comparative Analysis of Nutrition Policies and Plans of Action in WHO European. Geneva: World Health Organization; 2006. URL: http://www.euro.who.int/__data/assets/pdf_file/0004/149782/instanbul_conf_20ebd02.pdf [accessed 2016-03-18] [WebCite Cache ID 6g65EBrGo]

84. Deutsche Gesellschaft für Ernährung eV. Gemüse und Obst in der Prävention ausgewählter chronischer Krankheiten. Bonn: Deutsche Gesellschaft für Ernährung eV; 2012. URL: https://www.dge.de/fileadmin/public/doc/ws/stellungnahme/DGE-Stellungnahme-Gemuese-Obst-2012.pdf [accessed 2016-03-18] [WebCite Cache ID 6g65Bj6j9]
85. Voedingscentrum Den Haag. Gezonde basisvoeding met de schijf van vijf factsheet. Den Haag: Voedingscentrum Den Haag; 2011. URL: http://www.voedingscentrum.nl/Assets/Uploads/Documents/Factsheet%20Gezonde%20basisvoeding%20met%20de%20Schijf%20van%20Vijf.pdf [accessed 2016-03-18] [WebCite Cache ID 6g659XZOp]
86. Ashford S, Edmunds J, French DP. What is the best way to change self-efficacy to promote lifestyle and recreational physical activity? A systematic review with meta-analysis. Br J Health Psychol 2010 May;15(Pt 2):265-288. [doi: 10.1348/135910709X461752] [Medline: 19586583]
87. Prestwich A, Kellar I, Parker R, MacRae S, Learmonth M, Sykes B, et al. How can self-efficacy be increased? Meta-analysis of dietary interventions. Health Psychol Rev 2013 Jul 15;8(3):270-285. [doi: 10.1080/17437199.2013.813729#.VVEoB6NwamE]

Abbreviations

BMI: body mass index
CVD: cardiovascular disease
FIML: full-information maximum likelihood
RENATA: Rehabilitation-Aftercare for an optimal Transfer into Autonomous daily life
SRHI: Self-Report Habit Index

Edited by G Eysenbach; submitted 20.09.15; peer-reviewed by R Soellner, A Jander, F Schneider, K Broekhuizen; comments to author 15.10.15; revised version received 17.11.15; accepted 17.01.16; published 11.04.16.

Please cite as:
Storm V, Dörenkämper J, Reinwand DA, Wienert J, De Vries H, Lippke S
Effectiveness of a Web-Based Computer-Tailored Multiple-Lifestyle Intervention for People Interested in Reducing their Cardiovascular Risk: A Randomized Controlled Trial
J Med Internet Res 2016;18(4):e78
URL: http://www.jmir.org/2016/4/e78/
doi:10.2196/jmir.5147
PMID:27068880

Original Paper

Enhancement of Self-Monitoring in a Web-Based Weight Loss Program by Extra Individualized Feedback and Reminders: Randomized Trial

Melinda Jane Hutchesson[1], B Nutr & Diet (hons), PhD; Chor Yin Tan[1], B Nutr & Diet (hons); Philip Morgan[1,2], PhD; Robin Callister[1,3], PhD; Clare Collins[1,4], PhD

[1]Priority Research Centre in Physical Activity and Nutrition, School of Health Sciences, Faculty of Health and Medicine, University of Newcastle, Callaghan, Australia

[2]School of Education, Faculty of Education and Arts, University of Newcastle, Callaghan, Australia

[3]School of Biomedical Sciences and Pharmacy, Faculty of Health and Medicine, University of Newcastle, Callaghan, Australia

[4]School of Health Sciences, Faculty of Health and Medicine, University of Newcastle, Callaghan, Australia

Corresponding Author:
Clare Collins, PhD
School of Health Sciences
Faculty of Health and Medicine
University of Newcastle
Advanced Technology Centre, Level 3
University Drive
Callaghan, 2308
Australia
Phone: 61 2 49215646
Fax: 61 2 4921 7053
Email: clare.collins@newcastle.edu.au

Abstract

Background: Self-monitoring is an essential behavioral strategy for effective weight loss programs. Traditionally, self-monitoring has been achieved using paper-based records. However, technology is now more frequently used to deliver treatment programs to overweight and obese adults. Information technologies, such as the Internet and mobile phones, allow innovative intervention features to be incorporated into treatment that may facilitate greater adherence to self-monitoring processes, provide motivation for behavior change, and ultimately lead to greater weight loss success.

Objective: The objective of our study was to determine whether the consistency of self-monitoring differed between participants randomly assigned to a basic or an enhanced 12-week commercial Web-based weight loss program.

Methods: We randomly assigned a sample of 301 adults (mean age 42.3 years; body mass index 31.3 kg/m2; female 176/301, 58.5%) to the basic or enhanced group. The basic program included tools for self-monitoring (online food and exercise diary, and a weekly weigh-in log) with some feedback and reminders to weigh in (by text or email). The enhanced program included the basic components, as well as extra individualized feedback on self-monitoring entries and reminders (by text, email, or telephone) to engage with self-monitoring tools. We evaluated the level of self-monitoring by examining the consistency of self-monitoring of food, exercise, and weight during the 12 weeks. Consistency was defined as the number of weeks during which participants completed a criterion number of entries (ie, ≥3 days of online food or exercise diary records per week and ≥1 weigh-in per week).

Results: The enhanced group’s consistency of use of self-monitoring tools was significantly greater than that of the basic group throughout the 12 weeks (median consistency for food 8 vs 3 weeks, respectively, $P<.001$; for exercise 2.5 vs 1 weeks, respectively, $P=.003$).

Conclusions: Enhanced features, including additional individualized feedback and reminders, are effective in enhancing self-monitoring behaviors in a Web-based weight loss program.

ClinicalTrial: Australian New Zealand Clinical Trials Registry (ANZCTR): ACTRN12610000197033; https://www.anzctr.org.au/Trial/Registration/TrialReview.aspx?id=335159 (Archived by WebCite at http://www.webcitation.org/6gCQdj21G)

(J Med Internet Res 2016;18(4):e82) doi:10.2196/jmir.4100

KEYWORDS

weight loss; Web-based intervention; randomized controlled trial; commercial weight loss program; self-monitoring

Introduction

Self-monitoring is a behavioral strategy considered essential for effective weight loss programs [1]. Self-monitoring typically involves systematic observation, measurement, and recording of dietary intake, exercise, and weight [1]. This monitoring may raise individuals' awareness of their own actions, how and when these actions occur, and the initial and lasting impact on their actions [1]. It allows individuals to evaluate their progress toward goal attainment, reinforces behavior changes made, and highlights behaviors that may require further attention [2,3].

A systematic review evaluated the effects of self-monitoring of diet, exercise, and self-weighing on weight loss as part of a behavioral intervention program [4]. All 22 included studies supported the hypothesis that greater frequency of self-monitoring is associated with greater weight loss [4]. Due to the diversity of measurements of self-monitoring of dietary intake and exercise in the included studies, the reviewers could not determine an optimal frequency of self-monitoring of dietary intake and physical activity necessary for weight loss. However, the review concluded that individuals who weighed themselves at least once per week lost significantly more weight, which is consistent with an earlier systematic review [5]. Notably, the later review identified that very few studies have examined participants' adherence to self-monitoring over time (ie, how consistently they self-monitored over time) and any association with weight loss [4].

Behavioral weight loss programs are traditionally delivered in a face-to-face format and self-monitoring is completed using paper-based diaries [6]. However, new treatment modalities using technologies, such as the Internet or mobile phone apps, have been developed. The multimedia capabilities of such technologies have the potential to minimize the obstacles associated with paper-based self-monitoring, such as reducing the participant's burden by simplifying the recording process [7]. Furthermore, technology provides an opportunity for inclusion of features that may facilitate greater adherence to the self-monitoring process, enhance motivation for behavior change, and ultimately lead to greater weight loss success [4]. Such features include the provision of automated or tailored feedback on weight, dietary intake, or exercise levels or reminders (eg, text messages, emails) to complete program tasks such as self-monitoring [4,8]. Using meta-analysis, a recent systematic review demonstrated that eHealth weight loss programs with additional features achieved 1.46 kg greater weight loss postintervention than those providing a standard eHealth program alone [8]. However, few studies have investigated whether the provision of these additional self-monitoring-related features improves adherence to self-monitoring and facilitates greater weight loss. Burke et al [9] randomly assigned participants to self-monitor their dietary intake using 3 approaches: a personal digital assistant with no feedback; a personal digital assistant with daily tailored automated feedback; or a paper-based diary with no feedback. They found that after 24 months there was no significant difference in weight loss between the 3 groups, although the 2 personal digital assistant groups self-monitored on a significantly greater proportion of days than the paper-based monitoring group over the 24-month period [9]. This study highlighted the ability of technology to engage individuals in the self-monitoring process and, as the authors conclude, was an important "early step" in understanding how technology can be used for self-monitoring in weight loss programs. Further examination of whether specific components of technology-based weight loss programs can improve participants' adherence to self-monitoring of diet, exercise, and weight is required.

We previously conducted a randomized controlled trial (RCT) comparing the efficacy of 2 versions (basic vs enhanced) of a commercial Web-based weight loss program for 12 weeks [10]. Both versions of the program included tools for self-monitoring (online food and exercise diary, and a weekly weigh-in log) with automated feedback on self-monitoring records and once-weekly reminders via text message or email to weigh in. The enhanced group also received additional weekly automated individualized feedback reports on current diet and exercise based on their previous week's self-monitoring records, as well as extra reminders via text message, email, or phone to complete all self-monitoring records. Both groups lost weight; however, we found no significant difference in mean weight change between groups (basic -2.7, SD 4.0, enhanced -3.3, SD 4.5, *P*=.21) or the proportion of participants who achieved a clinically significant weight loss of 5% (basic 24.5%, enhanced 32.9%, *P*=.11) at 12 weeks based on intention-to-treat analysis [11].

Therefore, the aims of this investigation were to determine whether consistency (ie, number of weeks during which participants completed a criterion number of online entries) of self-monitoring of food intake, exercise, and weight differed between participants randomly assigned to the basic version and those assigned to the enhanced version of the commercial Web-based weight loss program, and whether the consistency of self-monitoring was related to weight loss after 12 weeks [10]. We hypothesized that the enhanced group would achieve significantly greater frequency and consistency of self-monitoring of dietary intake, exercise, and weight than the basic group.

Methods

Study Design

We collected data for this analysis as part of a commercial Web-based weight loss program RCT. The methods of the RCT have been published in detail elsewhere [10]. We investigated online self-monitoring behaviors in adults allocated to 1 of 2 versions of a commercial weight loss program with basic or enhanced features for 12 weeks.

Participants and Recruitment

We recruited overweight and obese (body mass index, BMI, of 25 to 40 kg/m^2) adults (18 to 60 years old) in the Hunter region of New South Wales, Australia, through media advertising (radio, TV, newspaper, flyers in general practitioner clinics, university website) from October to December 2009. To be included in the study, participants had to agree not to take part in other weight loss programs for the study duration; pass a health screening questionnaire; have access to a computer with Internet and an email account; and be able to attend assessment sessions at the University of Newcastle campus (Callaghan, Australia). Participants were ineligible for the study if they were pregnant or trying to conceive; had major medical illnesses; had physical inabilities such as orthopedic or joint problems; had lost 4.5 kg or more in the preceding 6 months; or were taking medications that affected or were affected by weight loss. We obtained written informed consent from all participants before their enrollment.

Random Allocation to Groups

We initially randomly assigned participants to 1 of 3 groups (basic or enhanced treatment group, or a waiting list control) using a stratified randomized block design. Blocks of variable length (either 3 or 6) were used to stratify participants according to their sex and baseline BMI category (25 to <30; ≥30 to <35 or ≥35 to 40 kg/m^2). After 12 weeks, we randomly reallocated participants in the control group to either the basic or enhanced group, using the same procedures. We analyzed the self-monitoring behaviors of all participants during their participation in the basic or enhanced group. Participants were informed of their group allocation in sealed envelopes, which contained their online program login details. Participants in the basic and enhanced groups and researchers assessing outcomes were blinded to participants’ assignment to treatment groups.

Weight Loss Interventions

Participants were given free access to a basic or enhanced version of the commercial Web-based weight loss program The Biggest Loser Club [12] provided by SP Health Co (Sydney, NSW, Australia). The features of the Web-based program were designed based on social cognitive theory [13]. The program targeted the major factors of behavioral change, including self-efficacy, goal setting, self-monitoring, and social support. Both the basic and enhanced programs were conducted through the Web-based program for 12 weeks. Table 1 describes the key features of the basic and enhanced programs, specifically highlighting the self-monitoring tools, as well as features designed to encourage participants to self-monitor.

Table 1. Comparison of features of the basic and enhanced commercial Web-based weight loss programs.

	Basic and enhanced	Enhanced only
Self-monitoring tools	*Online food and exercise diary* to monitor energy intake and energy expenditure: participants were encouraged to self-monitor their dietary intake and exercise using an online diary at least 4 days per week. Participants recorded the type and amount of food or exercise by searching a database for the most appropriate item, selecting the appropriate measurement unit, and entering the amount. Participants *recorded weight* (weigh-in) as well as other body measurements (waist and hip girths) via website or text message, and were encouraged to record at least once per week.	No additional features were available.
Tools to enhance self-monitoring: feedback	*Online food and exercise diary:* Automated calculations of energy intake, energy expenditure, and energy balance were provided on the online diary page. Automated nutrition summaries were available via link on online diary page. Reported intake was compared with recommended nutrient targets for key nutrients: energy, total fats, saturated fat, protein, carbohydrate, sugars, fiber, sodium, calcium, iron, zinc, magnesium, iodine, selenium, vitamins B1, B2, B3, B6, B12, A, C, and folate (if entries made in online diary). *Weigh-ins:* entered weight data (and other measurements) were tracked and displayed graphically and in a body (body mass index) silhouette to demonstrate change over time.	A weekly automated individualized feedback report based on *online food and exercise diary* entries was provided via the website for the previous week. Feedback for key elements of diet and exercise (ie, weekly summary of energy intake and expenditure, saturated fat, fruit and vegetable intakes, frequency and intensity of physical activity, and time spent being active compared with national recommendations), usage patterns of the website (ie, cumulative average website visits, diary entries, and forum posts), and level of success with weight loss (ie, weight loss to date) was provided. The feedback used a color-coded traffic light system (green, amber, red) to indicate whether a participant was meeting recommendations (green), moving in the right direction (amber), or not meeting recommendations (red).
Tools to enhance self-monitoring: reminders	Participants were encouraged to *weigh-in* via once-weekly email or short message service text messaging reminders to enter weight on the website on the due date.	Weekly reminders to further motivate participants to log in to the website, *weigh in*, and use the *online food and exercise diary* were sent. The reminders escalated with urgency, starting with an initial reminder email, then a text message, and lastly a phone call if participants did not engage with the program. Reminders commenced when weigh-in was 2 days overdue or if no site visits were made in 3 days or site visits but no diary entries were made in 4 days.
Other tools	Participants set a weight loss goal and were assigned individualized daily calorie targets to facilitate a 0.5–1 kg weight loss per week (~2600 kJ less than their estimated energy requirements). Online education in the form of weekly tutorials, fact sheets, meal and exercise plans, and weekly challenges were provided. Access to weekly low-fat menu plan and grocery lists designed to meet nutrient reference values and assigned calorie target was available. Social support was available via online discussion forums.	An individualized weekly automated enrollment report based on responses to the enrollment survey was sent. It included an assessment of current weight and suggestions for appropriate weight loss goals; an energy balance assessment and recommended calorie target; an assessment of eating habits and behaviors, including saturated fat and fiber intake, daily servings of fruits and vegetables, high-risk eating behaviors (eg, skipping meals, not eating breakfast, drinking soft drinks), and nonhungry eating triggers; and weight loss motivation assessment.

Measures

All self-monitoring data were collected by SP Health Co and provided to the researchers. Data stored by SP Health Co included the date a participant submitted a food exercise or weigh-in entry. To make a food or exercise entry, participants searched for and selected a food or exercise item from the database, selected a unit of measurement (eg, grams or cups for food, or minutes or distance in kilometers for exercise), and then recorded the amount. A weigh-in entry required a weight to be entered by participants either online or via text message. For the purposes of this study, we required a participant to make 1 entry per day (ie, enter 1 food item, 1 exercise, or 1 weight) for the day to be counted as self-monitoring and therefore be included in the calculation of consistency of self-monitoring.

Consistency of self-monitoring for this study refers to the number of weeks during which a criterion number of entries was made. For food and exercise the criterion number was ≥3 days of entries per week, as per the previous definition of Peterson et al [14]. For weigh-ins the criterion number was ≥1 weigh-in per week, as per previous systematic reviews indicating that individuals who weighed themselves at least once weekly lost significantly more weight [4,5].

To determine the relationship between categories of self-monitoring consistency and weight loss, we grouped participants into 3 levels of self-monitoring consistency over the 12 weeks, based on the number of weeks when they met the criterion number of self-monitoring entries (ie, ≥3 days of online food or exercise diary records per week and ≥1 weigh-in per week). The levels were defined as low consistency if participants met the criterion number of self-monitoring entries in ≤4 weeks; as moderate if they met the criterion between 5 and 8 weeks; or as high if they met the criterion for ≥9 weeks. Low, medium, and high were defined a priori, by dividing the number of intervention weeks (ie, 12 weeks) into 3 groups covering an equal number of weeks.

All other measurements were taken at the Human Performance Laboratory at the University of Newcastle, Callaghan Campus, with assessments at baseline and 12 weeks of the study. Height was measured to 0.1 cm using the stretch stature method on a Harpenden portable stadiometer (Holtain Limited, Dyfed, UK). Weight was measured with the participant wearing light clothing, without shoes, on a digital scale to 0.01 kg (model CH-150kp; A&D Mercury Pty Ltd, Adelaide, Australia). BMI was calculated as weight (kg)/height $(m)^2$. Participants completed a survey at baseline that captured sociodemographic characteristics (age, sex, education level, ethnicity and income).

Statistical Analysis

We analyzed the data using Stata 11.0 (StataCorp LP). Descriptive statistics are described as mean and SD for normally distributed continuous variables, median and interquartile range for nonnormal continuous data, and number (n) and percentage for categorical variables. We used chi square tests to compare the self-monitoring consistency each week (weeks 1 to 12) by treatment group. Due to the multiple comparisons, we applied the Bonferroni correction, with *P*<.004 (*P*=.05/12) considered statistically significant.

We also used chi-square tests to compare the total number of consistent weeks and consistency of self-monitoring groups (low, moderate, high) by treatment group. Analysis of variance tested for differences in percentage weight loss at 12 weeks between consistency of self-monitoring groups (low, moderate, high). We performed post hoc comparisons using the Tukey-Kramer method. An intention-to-treat approach was used for calculating percentage weight loss, with baseline observation carried forward for those lost to follow-up at 12 weeks. We considered *P*<.05 to be statistically significant for these single comparisons.

Results

Baseline Characteristics

Of the 591 adults who expressed interest in participating in the study, 309 were randomly assigned to 1 of 3 groups (basic n=99, enhanced n=106, waiting list control n=104). We randomly reassigned waiting list control group participants to a treatment group after 12 weeks (basic n=44, enhanced n=52, lost to follow-up n=8); therefore, this analysis included 301 participants (basic n=143, enhanced n=158). We previously reported participants' characteristics at baseline [11]. In summary, 58.5% of participants were female (176/301), with a mean (SD) age of 42 (10.2) years, and most were born in Australia (273/301, 90.7%), were classified as obese (195/301, 64.8%), had an educational level higher than high school (210/301, 69.8%), and had a weekly household income of more than A$1500 (194/301, 64.5%). At baseline (treatment group entry), characteristics of the basic and enhanced group participants did not differ significantly [11], nor were there any differences between participants initially randomly assigned to the waiting list control group and their respective intervention groups.

Attrition Rates

A total of 62 participants did not have their weight assessed at 12 weeks, resulting in 20.6% (62/301) attrition. There was no significant difference (*P*=.7) in attrition rates between the basic (31/143, 21.7%) and enhanced groups (31/158, 19.6%) at 12 weeks.

Consistency of Self-monitoring

Table 2 describes the consistency of self-monitoring by treatment group. The median number of weeks during which participants made food and exercise entries on ≥3 days per week and weighed in once per week was significantly greater in the enhanced than in the basic group (food: 8 vs 3 weeks, respectively, *P*<.001; exercise: 2.5 vs 1 week, respectively, *P*=.003; weigh-ins: 11 vs 8 weeks, respectively, *P*<.001). The enhanced and basic groups differed significantly in the proportion of participants classified as having low, moderate, and high self-monitoring consistency for food (χ^2_2=18.9, *P*<.001), exercise (χ^2_2=10.0, *P*=.007) and weight (χ^2_2=16.5, *P*<.001) entries (Table 2).

Table 2. Frequency and level of consistency of self-monitoring[a] by basic and enhanced groups of a commercial Web-based weight loss program over 12 weeks.

Data entered by participant	Basic (n=143)	Enhanced (n=158)	*P* value[b,c]
Food			
No. of weeks ≥3 days of entries, median (IQR[d])	3 (0–12)	8 (0–2)	<.001
Low (n=133), n (%)	55.94 (80)	33.54 (53)	<.001
Moderate (n=54), n (%)	18.18 (26)	17.72 (28)	
High (n=114), n (%)	25.87 (37)	48.73 (77)	
Exercise			
No. of weeks ≥3 days of entries, median (IQR)	1 (0–12)	2.5 (0–12)	.003
Low (n=207), n (%)	77.62 (111)	60.76 (96)	.007
Moderate (n=36), n (%)	8.39 (12)	15.19 (24)	
High (n=58), n (%)	13.99 (20)	24.05 (38)	
Weigh-ins			
No. weeks with 1 weigh-in, median (IQR)	8 (2–12)	11 (7–12)	<.001
Low (n=79), n (%)	37.06 (53)	16.46 (26)	<.001
Moderate (n=50), n (%)	13.99 (20)	18.99 (30)	
High (n=172), n (%)	48.95 (70)	64.56 (102)	

[a] Levels were defined as low consistency if participants met the criterion number of self-monitoring entries in ≤4 weeks; as moderate if they met the criterion between 5 and 8 weeks; or as high if they met the criterion for ≥9 weeks.

[b] Wilcoxon test populations to compare between groups for median number of days, entries, and weeks with ≥3 days of entries.

[c] Chi square to compare between groups for number who used the self-monitoring feature, and number having low, moderate, or high consistency.

[d] IQR: interquartile range.

The figures illustrate the proportion of participants from the basic and enhanced groups who consistently made an online diary food entry (Figure 1) or exercise entry (Figure 2) or weighed in (Figure 3) during each week of the program (weeks 1 to 12). A significantly higher proportion of the enhanced group than of the basic group made food entries ≥3 days per week from weeks 4 to 12 of the program. A significantly higher proportion of enhanced group participants than of basic group participants made exercise entries ≥3 days per week during weeks 8 and 11, and a significantly higher proportion of enhanced group participants than of basic group participants weighed in during weeks 2, 5, 7, 9, 10, and 11 (Multimedia Appendix 1). Notably, both the basic and enhanced groups' consistency of self-monitoring declined from weeks 1 to 12 (Figure 1, Figure 2, Figure 3, and Multimedia Appendix 1). For example, in week 1 of the program, 79.6% (126/158) of enhanced group participants made food entries to the online diary ≥3 days compared with 46.8% (74/158) in week 12. In comparison, 69.2% (99/143) of basic group participants made food entries to the online diary in week 1 compared with 28.7% (41/143) in week 12.

Figure 1. Proportion of participants from the basic (full line) and enhanced (dashed line) groups who consistently (3 or more days/week) made food entries to the online diary from weeks 1 to 12.

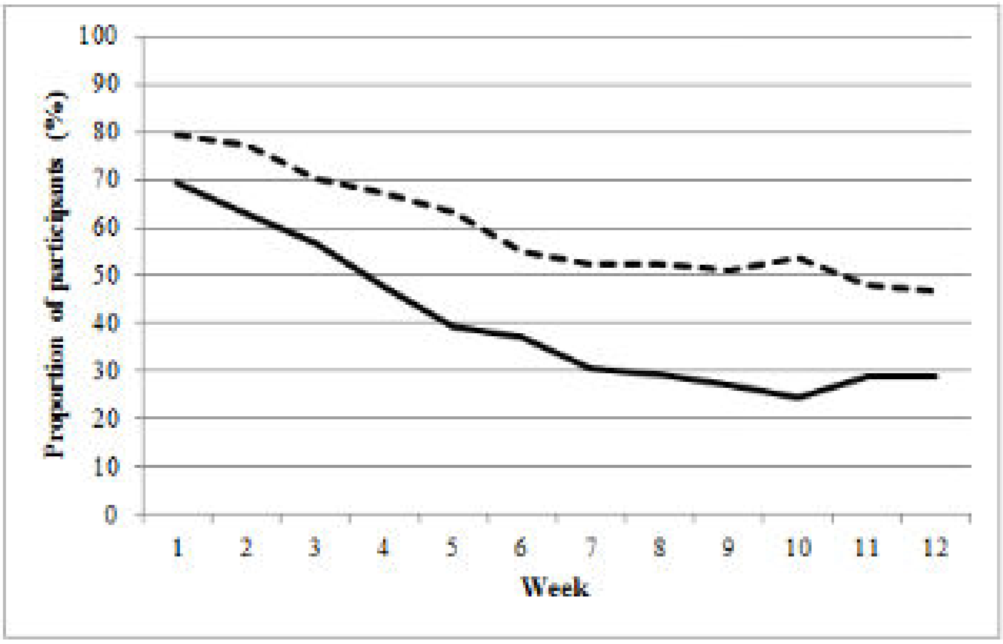


Figure 2. Proportion of participants from the basic (full line) and enhanced (dashed line) groups who consistently (3 or more days/week) made exercise entries to the online diary from weeks 1 to 12.

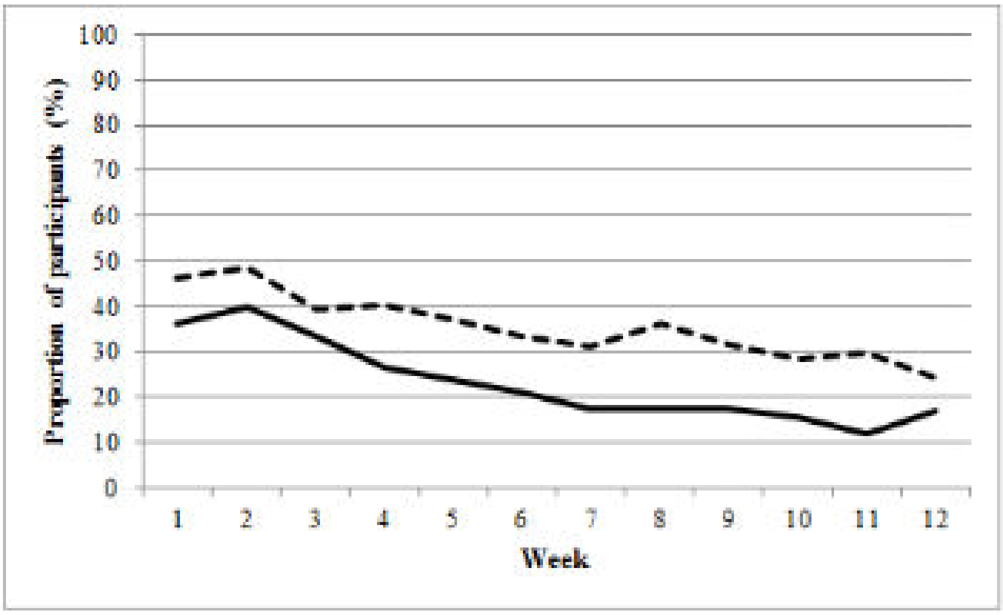
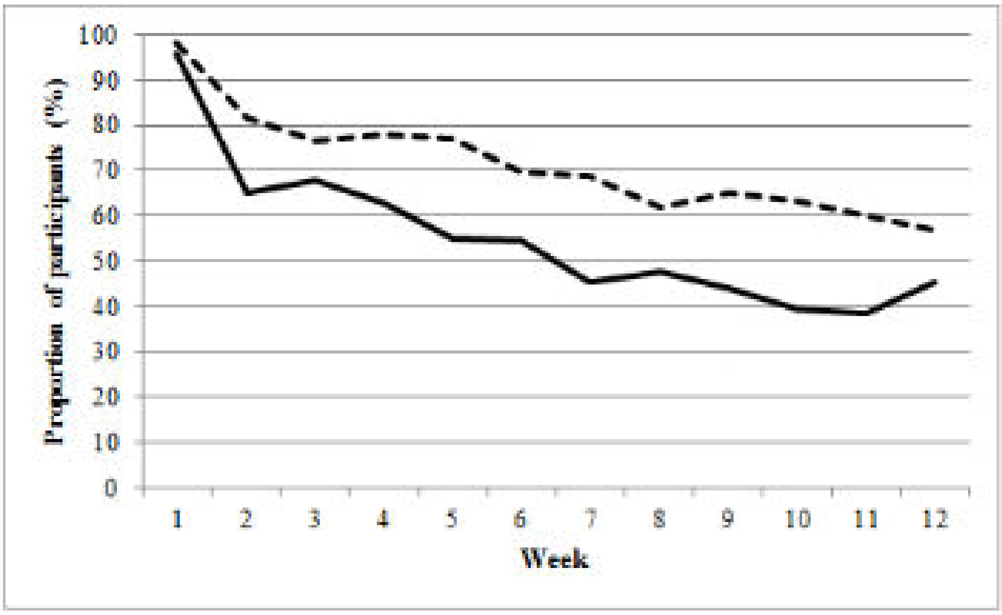


Figure 3. Proportion of participants from the basic (full line) and enhanced (dashed line) groups who consistently (1 or more days/week) weighed in from weeks 1 to 12.



Consistency of Self-monitoring and Weight Loss

Consistency of self-monitoring strongly predicted weight loss. There was a significant difference in percentage weight loss by consistency of self-monitoring for food entries ($F_{2,298}$=57.39, P<.001), exercise entries ($F_{2,298}$=66.20, P<.001), and weigh-ins ($F_{2,298}$=33.49, P<.001). Post hoc comparisons revealed that, for food and exercise entries to the online diary, participants whose consistency was high lost significantly more weight than those whose consistency was low or moderate, and those whose consistency was classified as moderate lost significantly more weight than those classified as low (food entries online diary for high: mean –6.2, SD 4.5%; moderate: –2.6, SD 3.7%; low –1.1, SD 3.0%; exercise entries online diary for high: mean –7.9, SD 4.7%; moderate –4.5, SD 3.8%; low: –1.7, SD 3.3%). For weigh-ins, participants whose consistency was high or moderate lost more weight (mean –4.9, SD 4.7% and –1.7, SD 3.0%, respectively) than those whose consistency was low (–0.7, SD 2.7%).

Discussion

Our study found greater consistency of self-monitoring among participants randomly assigned to an enhanced version of a 12-week commercial Web-based weight loss program that included additional individualized feedback on self-monitoring records and reminders to engage with the self-monitoring tools. The 2 groups differed in weigh-ins, and in food and exercise entries to the online diary, although the enhanced features had the least impact on exercise entries. Greater consistency of self-monitoring of all 3 monitoring behaviors was associated with increased weight loss. However, there was no significant difference in weight loss between the enhanced and basic intervention groups.

The greater use of self-monitoring tools by the enhanced group may be associated with the push factors, such as the reminders, or the positive feedback or encouragement provided by the weekly feedback reports, which is supported by the research of Eysenbach [15].

Therefore, more in-depth evaluation of the multiple intervention components used in the enhanced program is warranted, to examine reasons for greater engagement with the self-monitoring tools by the enhanced group participants. Furthermore, the consistency of self-monitoring varied with the behavior (weight, food, or exercise) being monitored and decreased over the 12 weeks for all 3 behaviors in both groups. Self-monitoring entries of weight were much more consistent in both groups throughout the 12 weeks than either food or exercise entries. Reasons for this greater consistency of weight entries are not clear but possible explanations are the relative ease of providing weight entries and positive reinforcement of effort if weight continues to decline. The consistency of food entries was also reasonably high initially, particularly in the enhanced group, but decreased substantially over time, especially in the basic group. Consistency of exercise entries was relatively poor from week 1. Although a greater proportion of enhanced group participants consistently self-monitored exercise compared with the basic group over the 12 weeks, it was only during 2 weeks (weeks 8 and 11) that a higher proportion of enhanced group participants made exercise entries ≥3 days per week. The additional features (eg, reminders, feedback) provided in the enhanced intervention may not have prompted self-monitoring of exercise and suggests that a greater understanding of the barriers to self-monitoring of exercise is required. However, it is also possible that the poor consistency of exercise entries may be due to a lack of exercise being performed by participants.

There is evidence from systematic reviews that greater self-monitoring within weight loss interventions is associated with greater weight loss [4,5]. The provision of enhanced features within the Web-based weight loss program appears to have facilitated greater consistency of self-monitoring; however, this did not result in significantly greater weight loss in this study [11]. This is consistent with several recent studies that have been unable to demonstrate a significant difference in weight loss between Web-based interventions providing additional features, such as individualized feedback, online support groups, or behavioral lessons to facilitate greater weight loss success [16-19]. The similar rates of weight loss between the 2 groups in this study and in other studies [16-19] may have been due to both the standard and basic versions of these programs having included key features known to contribute to weight loss success (eg, goal setting, self-monitoring with some feedback, and social support). In this study, although the enhanced features motivated more participants to use the self-monitoring tools, the additional feedback provided may not have assisted all participants to adequately self-regulate their behaviors in order to lose weight. For example, the extra feedback provided was individualized to behaviors reported by the participants as part of their self-monitoring records. Therefore, if they reported the same behavior each week, they would continue to receive an identical feedback message in subsequent weekly feedback reports. Furthermore, although the enhanced group's use of the self-monitoring tools was significantly higher than that of the basic group, the use of the tools varied among study participants in both groups, as evidenced by the large interquartile ranges reported for all self-monitoring metrics. Clearly not all study participants were motivated to self-monitor by the enhanced program features, nor did all participants require the enhanced features to facilitate adequate levels of self-monitoring. Notably, both groups demonstrated declining rates of self-monitoring over time, which is consistent with previous reports of the use of online public health interventions [20] and Web-based weight loss programs (eg, [16,21]), including the program in this study [22]. This may have further contributed to the nonsignificant difference in weight loss between the 2 groups.

In a cohort study, we previously identified that a higher median number of days when participants used the self-monitoring features of the basic version of the program was associated with significantly greater weight loss [23]. A limitation of that study was that it relied on self-reported weight data. The results of the RCT reported here, with objective assessment of weight and use of self-monitoring tools, confirm our previous findings [23]. Participants who were highly consistent (≥9 weeks out of 12) in self-monitoring weight (≥1 day/week) and in reporting food or exercise in the online diary (≥3 days/week) lost significantly more weight (5% to 8%). Our findings are supported by Krukowski et al [24], who demonstrated that participants who consistently self-monitored during a 6-month online behavioral weight control program were significantly more likely to achieve clinically important weight loss. They found that participants who self-monitored on ≥6 days during the initial stages of the program (weeks 1 to 4) were more likely to achieve clinically important weight loss after 6 months, as well as those who self-monitored during the later weeks of the intervention (weeks 9 to 24) [24]. Further experimental research is required to determine whether the association between self-monitoring consistency and weight loss observed in this study is indicative of a cause-effect relationship between self-monitoring and weight loss success.

Study Strengths and Limitations

A strength of this RCT was the large sample including substantial proportions of both male and female participants, and use of objective measures of self-monitoring. Potential limitations include that we evaluated self-monitoring behaviors during the 12-week weight loss intervention. Therefore, we do not know whether self-monitoring was continued or maintained beyond the 12-week period, and whether this was associated with further weight loss. This is an area where future research is warranted, as it was recently suggested that frequency and consistency of self-monitoring of dietary intake after a 6-month weight loss program improved weight loss success at 12 months [14]. Our study focused on 1 behavioral strategy (self-monitoring) within the commercial weight loss program, and did not consider the potential influence of other key behavioral strategies (eg, social support via the discussion forum, or provision of tailored feedback via automated feedback reports). Due to errors with tracking of participant's usage of other program components during the trial, this analysis is not possible. The definition of consistency of self-monitoring was based on previous literature where possible; however, as no universally accepted metrics for self-monitoring exist, the results may vary with the use of different cut points.

Conclusion

Enhanced program features, such as reminders and tailored feedback, facilitated greater consistency of self-monitoring of food, exercise, and weight during a 12-week commercial Web-based weight loss program. However, there were no significant differences in weight loss between the enhanced and basic intervention groups. Given the strong association between self-monitoring of these behaviors and successful weight change outcomes, further evaluation of individuals' experiences with self-monitoring and intervention components designed to promote self-monitoring (ie, reminders, tailored feedback) is warranted. This will provide greater insight into factors contributing to group and individual variations in engagement with self-monitoring tools and facilitate the design of Web-based weight loss interventions that are adaptive and provide individually tailored features to optimize self-monitoring (eg, frequency and mode of reminders, language used in feedback).

Acknowledgments

CC, PM, and RC designed the study and obtained funding. MH and CT conceptualized the secondary analysis, conducted the analysis, and drafted the paper. All authors contributed to and approved the final manuscript. We would like to thank the study participants for participating in the trial: research assistants (Julia Martin, Kate Fletcher, Elroy Aguiar, Ashlee Lucas, Rebecca Collins, Trevor Cripps, James Dower, Sharenjit Gill, Jenna Hannan, Skye Huxley, Hannah Mackay, Bryana Melnick, Justin Nicol, Hannah Lucas, Tom Mitchell, Huiru Teoh, Janine Wright, and Mei Yap) who assisted with data collection; and Scott Penn, Anna Crook, Julian Barton, Penelope Jones, Sandra Mitchell, and Laura Welsford from SP Health Co. This trial was funded by an Australian Research Council Linkage Project grant (2009-2012) (LP0990414, G0189752), with SP Health Co as the Industry Partner Organization (G0189753). Dr Melinda Hutchesson is supported by a Postdoctoral Research Fellowship (#100177) from the National Heart Foundation of Australia.

Conflicts of Interest

CC has been a nutrition consultant to SP Health Co and MH was funded by a Penn Health postdoctoral fellowship. CT, PM, and RC declare that they have no competing interests.

Multimedia Appendix 1

Consistency of use of the commercial web-based weight loss program's self-monitoring tools (online diary to monitor food and exercise and weigh-ins) by treatment group (basic vs enhanced) from weeks 1 to 12.

[PDF File (Adobe PDF File), 40KB - jmir_v18i4e82_app1.pdf]

References

1. Foster GD, Makris AP, Bailer BA. Behavioral treatment of obesity. Am J Clin Nutr 2005 Jul;82(1 Suppl):230S-235S [FREE Full text] [Medline: 16002827]
2. Kirschenbaum DS. Self-regulatory failure: a review with clinical implications. Clin Psychol Rev 1987 Jan;7(1):77-104. [doi: 10.1016/0272-7358(87)90005-5]
3. Baker RC, Kirschenbaum DS. Self-monitoring may be necessary for successful weight control. Behav Ther 1993 Jun;24(3):377-394. [doi: 10.1016/S0005-7894(05)80212-6]
4. 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]
5. Vanwormer JJ, French SA, Pereira MA, Welsh EM. The impact of regular self-weighing on weight management: a systematic literature review. Int J Behav Nutr Phys Act 2008;5:54 [FREE Full text] [doi: 10.1186/1479-5868-5-54] [Medline: 18983667]
6. Wadden TA, Webb VL, Moran CH, Bailer BA. Lifestyle modification for obesity: new developments in diet, physical activity, and behavior therapy. Circulation 2012 Mar 6;125(9):1157-1170 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.111.039453] [Medline: 22392863]
7. Illner A, 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]
8. Hutchesson M, 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]
9. Burke LE, Styn MA, Sereika SM, Conroy MB, Ye L, Glanz K, et al. Using mHealth technology to enhance self-monitoring for weight loss: a randomized trial. Am J Prev Med 2012 Jul;43(1):20-26 [FREE Full text] [doi: 10.1016/j.amepre.2012.03.016] [Medline: 22704741]
10. Collins CE, Morgan PJ, Jones P, Fletcher K, Martin J, Aguiar EJ, et al. Evaluation of a commercial web-based weight loss and weight loss maintenance program in overweight and obese adults: a randomized controlled trial. BMC Public Health 2010;10:669 [FREE Full text] [doi: 10.1186/1471-2458-10-669] [Medline: 21047432]
11. 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]
12. SP Health. The Biggest Loser Club. North Sydney, NSW: SP Health Co URL: http://www.biggestloserclub.com.au [accessed 2016-03-22] [WebCite Cache ID 6gCTUzHoz]
13. Bandura A. Social Foundations of Thought and Action: A Social Cognitive Theory. Englewood Cliffs, NJ: Prentice-Hall; 1986.
14. Peterson ND, Middleton KR, Nackers LM, Medina KE, Milsom VA, Perri MG. Dietary self-monitoring and long-term success with weight management. Obesity (Silver Spring) 2014 Sep;22(9):1962-1967. [doi: 10.1002/oby.20807] [Medline: 24931055]
15. Eysenbach G. The law of attrition. J Med Internet Res 2005;7(1):e11 [FREE Full text] [doi: 10.2196/jmir.7.1.e11] [Medline: 15829473]
16. Brindal E, Freyne J, Saunders I, Berkovsky S, Smith G, Noakes M. Features predicting weight loss in overweight or obese participants in a web-based intervention: randomized trial. J Med Internet Res 2012;14(6):e173 [FREE Full text] [doi: 10.2196/jmir.2156] [Medline: 23234759]
17. Hersey JC, Khavjou O, Strange LB, Atkinson RL, Blair SN, Campbell S, et al. The efficacy and cost-effectiveness of a community weight management intervention: a randomized controlled trial of the health weight management demonstration. Prev Med 2012 Jan;54(1):42-49. [doi: 10.1016/j.ypmed.2011.09.018] [Medline: 22001689]

18. Pullen CH, Hageman PA, Boeckner L, Walker SN, Oberdorfer MK. Feasibility of Internet-delivered weight loss interventions among rural women ages 50-69. J Geriatr Phys Ther 2008;31(3):105-112. [Medline: 19856616]
19. Wing RR, Crane MM, Thomas JG, Kumar R, Weinberg B. Improving weight loss outcomes of community interventions by incorporating behavioral strategies. Am J Public Health 2010 Dec;100(12):2513-2519. [doi: 10.2105/AJPH.2009.183616] [Medline: 20966375]
20. Glasgow RE. eHealth evaluation and dissemination research. Am J Prev Med 2007 May;32(5 Suppl):S119-S126. [doi: 10.1016/j.amepre.2007.01.023] [Medline: 17466816]
21. Ware LJ, Hurling R, Bataveljic O, Fairley BW, Hurst TL, Murray P, et al. Rates and determinants of uptake and use of an internet physical activity and weight management program in office and manufacturing work sites in England: cohort study. J Med Internet Res 2008;10(4):e56 [FREE Full text] [doi: 10.2196/jmir.1108] [Medline: 19117828]
22. Neve MJ, Collins CE, Morgan PJ. Dropout, nonusage attrition, and pretreatment predictors of nonusage attrition in a commercial Web-based weight loss program. J Med Internet Res 2010;12(4):e69 [FREE Full text] [doi: 10.2196/jmir.1640] [Medline: 21156470]
23. Neve M, Morgan PJ, Collins CE. Weight change in a commercial web-based weight loss program and its association with website use: cohort study. J Med Internet Res 2011;13(4):e83 [FREE Full text] [doi: 10.2196/jmir.1756] [Medline: 21993231]
24. Krukowski RA, Harvey-Berino J, Bursac Z, Ashikaga T, West DS. Patterns of success: online self-monitoring in a web-based behavioral weight control program. Health Psychol 2013 Feb;32(2):164-170. [doi: 10.1037/a0028135] [Medline: 22545978]

Abbreviations

BMI: body mass index
RCT: randomized controlled trial

Edited by G Eysenbach; submitted 04.12.14; peer-reviewed by R Krukowski, E Brindal; comments to author 22.02.15; revised version received 31.08.15; accepted 23.01.16; published 12.04.16.

Please cite as:
Hutchesson MJ, Tan CY, Morgan P, Callister R, Collins C
Enhancement of Self-Monitoring in a Web-Based Weight Loss Program by Extra Individualized Feedback and Reminders: Randomized Trial
J Med Internet Res 2016;18(4):e82
URL: http://www.jmir.org/2016/4/e82/
doi:10.2196/jmir.4100
PMID:27072817

Original Paper

The Impact of Automated Brief Messages Promoting Lifestyle Changes Delivered Via Mobile Devices to People with Type 2 Diabetes: A Systematic Literature Review and Meta-Analysis of Controlled Trials

Carukshi Arambepola[1], MD; Ignacio Ricci-Cabello[2], PhD; Pavithra Manikavasagam[1], MBBS; Nia Roberts[3], MSc; David P French[4], PhD; Andrew Farmer[2], DM

[1]University of Colombo, Faculty of Medicine, Colombo, Sri Lanka
[2]University of Oxford, Nuffield Department of Primary Care Health Sciences, Oxford, United Kingdom
[3]University of Oxford, Bodleian Libraries, Oxford, United Kingdom
[4]University of Manchester, School of Psychological Sciences, Manchester, United Kingdom

Corresponding Author:
Ignacio Ricci-Cabello, PhD
University of Oxford
Nuffield Department of Primary Care Health Sciences
Radcliffe Observatory Quarter
Woodstock Road, Oxford.
Oxford, OX2 6GG
United Kingdom
Phone: 44 01865 617190
Fax: 44 01865 289412
Email: ignacio.riccicabello@phc.ox.ac.uk

Abstract

Background: Brief automated messages have the potential to support self-management in people with type 2 diabetes, but their effect compared with usual care is unclear.

Objective: To examine the effectiveness of interventions to change lifestyle behavior delivered via automated brief messaging in patients with type 2 diabetes.

Methods: A systematic literature review of controlled trials examined the impact of interventions, delivered by brief messaging, and intended to promote lifestyle change in people with type 2 diabetes, on behavioral and clinical outcomes. Bibliographic databases searched included Medline, Embase, CINAHL, PsycINFO, and ISI WoK. Two reviewers independently screened citations. We extracted information on study risk of bias, setting (high versus low- and middle-income countries) and intervention characteristics (including use of theory and behavior-change techniques). Outcome measures included acceptability of the interventions and their impact on 1) determinants of lifestyle behavior (knowledge about diabetes, self-efficacy, attitudes towards self-management), 2) lifestyle behavior (diet, physical activity), and 3) clinical and patient-reported outcomes. Where possible, we pooled data using random-effects meta-analyses to obtain estimates of effect size of intervention compared to usual care.

Results: We identified 15 trials (15 interventions) meeting our inclusion criteria. Most interventions were delivered via short message service text messaging (n=12) and simultaneously targeted diet and physical activity (n=11). Nine interventions consisted of unidirectional messages, whereas six consisted of bidirectional messages, with patients receiving automated tailored feedback based on self-reported data. The acceptability of the interventions, and their impact on lifestyle behavior and its determinants, were examined in a low proportion of trials, with heterogeneous results being observed. In 13 trials (1155 patients) where data were available, there was a difference in glycated hemoglobin of -0.53% (95% CI -0.59% to -0.47%) between intervention groups compared to usual care. In five trials (406 patients) there was a non-significant difference in body mass index of -0.25 kg/m2 (95% CI -1.02 to 0.52). Interventions based on unidirectional messages produced similar effects in the outcomes examined, compared to those based on bidirectional messages. Interventions conducted in low- and middle-income countries showed a greater impact than those conducted in high-income countries. In general, trials were not free of bias and did not use explicit theory.

Conclusions: Automated brief messages strategies can improve health outcomes in people with type 2 diabetes. Larger, methodologically robust trials are needed to confirm these positive results.

(J Med Internet Res 2016;18(4):e86) doi:10.2196/jmir.5425

KEYWORDS

Diabetes mellitus, type 2; mobile health; text messaging; systematic review; diet; physical activity; self-care

Introduction

The number of people with type 2 diabetes worldwide is currently estimated to be 387 million, and is expected to increase to 592 million by 2035 [1]. This prevalence imposes a substantial burden of disease, mainly due to life-long multi-organ complications [2], leading to increased disability and premature deaths in low- and middle-income countries (LMIC) as well as high-income countries (HIC) [3].

Available evidence suggests that better control of blood glucose, blood pressure, and cholesterol levels would delay the onset of complications, and thereby prevent premature deaths among those already diagnosed with diabetes [4]. Lifestyle modification focusing on healthy diet is an accepted component of management [5], alongside promotion of physical activity [6]. However, patients with diabetes do not always follow advice about recommended changes in diet and physical activity, and therefore do not achieve optimal control of risk factors. Reasons for this are multifactorial, including psychological, social, and health care related factors [7,8].

Recommendations for supporting diabetes self-management are now widely incorporated in clinical practice guidelines [9,10]. However, strategies for providing effective continuing support and motivation are not well developed, and facilitating sustained behavior-changes remains an important challenge. Emerging evidence suggests that mobile health (mHealth) interventions may improve cardiovascular-related lifestyle behaviors and disease management [11]. Interventions based on brief messages delivered via mobile device technologies, such as short message service (SMS) technology, are one of the most studied types of mHealth interventions, and available evidence suggests that they may contribute to behavior-change [11,12]. Messages can be readily delivered at a wide-scale and at a low cost, and can be used to provide information through one-way (unidirectional) systems, but also to facilitate two-way communication (interactive or bidirectional). In contrast to more resource-intensive one-to-one clinician-patient contacts in clinics, the use of brief messages can be attractive for patients in terms of convenience, acceptability, and user-friendliness [13,14]. This type of intervention could address non-adherence to lifestyle recommendations by providing frequent reminders, motivational support and prompts to action, as well as timely access and feedback to relevant health information, while making patient-provider communication much easier [15]. Although their impact in different resource settings is still unclear, automated messaging technologies could be especially relevant in low-resource settings, given their ubiquity, low cost, and potential to underpin a developing health care infrastructure (eg electronic medical records systems). All of these features are leading to an increasing interest in the use of brief messages as a part of public health interventions.

A number of systematic reviews have provided evidence for the effectiveness of mobile-phone based interventions on self-management of long term conditions [11,16-19]. However, so far no study has specifically reviewed the potential impact of automated brief messages on promoting lifestyle modifications in patients with type 2 diabetes. In addition, the theoretical basis for this type of intervention is not well established, and the usefulness of basing messages on established behavior-change techniques is not known. These gaps in knowledge may partly be due to inadequate characterization of the techniques being used, which prevents the identification of those that might be helpful.

The primary objective of this systematic review was to examine the effectiveness of brief messages in improving glycemic control through promoting healthy eating and increasing physical activity, compared to usual care. We focused specifically on interventions delivered via mobile devices to people with type 2 diabetes. Secondary objectives included 1) examining the extent to which interventions have used explicit theory, 2) examining the behavior-change techniques used, 3) examining the acceptability of the interventions, 4) examining their impact on lifestyle-change behavior and its determinants, 5) examining their impact on other clinical (blood pressure, lipids, and weight) and patient-reported outcomes, 6) comparing the impact of unidirectional vs bidirectional messages, and 7) exploring the specific impact of the interventions in HIC and LMIC.

Methods

The study was planned, conducted, and reported according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [20]. The review protocol was registered in the PROSPERO International prospective register of systematic reviews (registration number CRD42015024302).

Data Sources and Searches

Specific search strategies were designed for the following databases (Multimedia Appendix 1): Medline, Embase, Cumulative Index of Nursing and Allied Health (CINAHL), PsycINFO, Cochrane Central Register of Controlled Trials (CENTRAL), and Science Citation Index & Conference proceedings Citation Index (ISI Web of Knowledge). To ensure the identification of relevant studies carried out in LMIC, we also searched the following databases: African Index Medicus, Index Medicus for the Eastern Mediterranean Region, Index Medicus for South-East Asia Region, Inter-Science Latin American and Caribbean Health Sciences Literature, Western

Pacific Region Index Medicus, and World Health Organization Library Database (all accessed via the Global Health Library).

The search strategy combined Medical Subject Headings terms and free-text keywords (Multimedia Appendix 2). Databases were searched from inception to April 2015 and no language restriction was applied. In addition, potentially relevant studies were identified using a snowball technique initiated by the examination of 52 previous systematic reviews on the broader area of telehealth and diabetes. A bibliographical database was created using EndNote X7, which was used to store and manage the references.

Study Selection

We included controlled trials examining the impact of interventions intended to promote lifestyle changes on diet and physical activity among people with type 2 diabetes. All interventions were delivered by brief messaging using mobile devices, and were compared on behavioral and clinical outcomes, versus usual care.

The main component of eligible interventions was the provision of information via brief messaging systems, characterized by automated messages (including computer-generated messages following an algorithm), which were tailored/custom-made personalized messages, or bulk messages. Messages had to be received via devices using mobile technology, such as mobile phones, smart-phones or hand-held computers. Messaging systems were those using the following technologies: SMS, automated email, or software apps. These systems could be used to provide information through one-way (unidirectional) messages or to facilitate two-way communication (interactive or bidirectional). Unidirectional messages were conceptualized as messages sent from the providers or researchers to the participants. Bidirectional messages were conceptualized as those involving a two-way communication regarding self-monitoring data that was sent by the participants, who in return received real-time automated brief messages providing tailored feedback. Eligible interventions targeted healthy eating, physical activity, or both.

Inclusion criteria for study design specifically focused on controlled clinical trials, including randomized controlled trials (RCTs), cluster RCTs, non-randomized controlled trials, and crossover studies. Only studies with control or comparator arms that consisted of patients receiving usual (standard) clinical care, or a minimal intervention (ie an intervention that is unlikely to produce any effect, such as sending non-health related messages, but allows blinding participants to condition allocation) were eligible for inclusion. Eligible studies included adult (at least 18 years old) patients with type 2 diabetes mellitus (with or without comorbid conditions) and had to be set in the community or in any primary, secondary, or tertiary care setting.

Studies reporting at least one of the following outcome measures were included: acceptability by recipients, determinants of change in lifestyle (namely knowledge, attitudes and self-efficacy on lifestyle modifications), lifestyle behaviors that impact on diabetic control (diet and physical activity), and clinical outcomes (glycated hemoglobin [HbA1c], body mass index [BMI] or body weight, lipids, blood pressure, and waist circumference).

Trials were excluded if they examined the use of messages created by a clinician based on a clinical judgement of a patient’s disease status (ie, not automated), had a proportion of patients with type 2 diabetes lower than 90%, evaluated a multifaceted intervention in which brief messages were not the main component of the intervention, or were published only in the form of conference abstracts. No language restrictions were applied.

A preliminary screen for eligibility was followed by retrieval and assessment of full texts of the selected articles. Studies that met the inclusion criteria were included for data extraction. All citations were independently screened by two reviewers. Any disagreements were solved by consensus with a third reviewer.

Data Extraction and Quality Assessment

Structured forms were used to extract data about the trial design, trial setting (HIC vs LMIC according to the World Bank classification [21]), number of participants in each group, length of follow-up, key elements of the intervention, type of comparison group, and acceptability of the intervention. In addition, the impact on determinants of change in lifestyle, diet and physical activity behavior, and clinical outcomes were recorded.

The extent to which the trials used theory explicitly in the development and evaluation of the interventions was assessed using an established coding-scheme which contains 19 items. This scheme assessed whether a theory was mentioned, how theories were used in intervention design and in the selection of intervention techniques, how intervention evaluations tested theory, and the implications of the results for future theory development [22]. The behavior-change techniques used in each intervention were classified using an established taxonomy [23].

We used the Cochrane Collaboration’s tool for risk of bias assessment [24]. Two reviewers independently extracted all information and assessed the risk of bias and the use of theory. Disagreements were discussed with a third reviewer until consensus was reached.

Data Synthesis and Analysis

We examined and synthesized the acceptability of the interventions, and their impact on 1) determinants of change in lifestyle, 2) behavior (diet and/or physical activity), and 3) clinical outcomes. Outcomes in all studies were examined and classified as measuring one of these three domains. Variables that measured other domains were not included in the analysis. For all pooled outcomes we used subgroup analyses to examine potential differences between types of messages (unidirectional vs bidirectional) and settings (HIC vs LMIC).

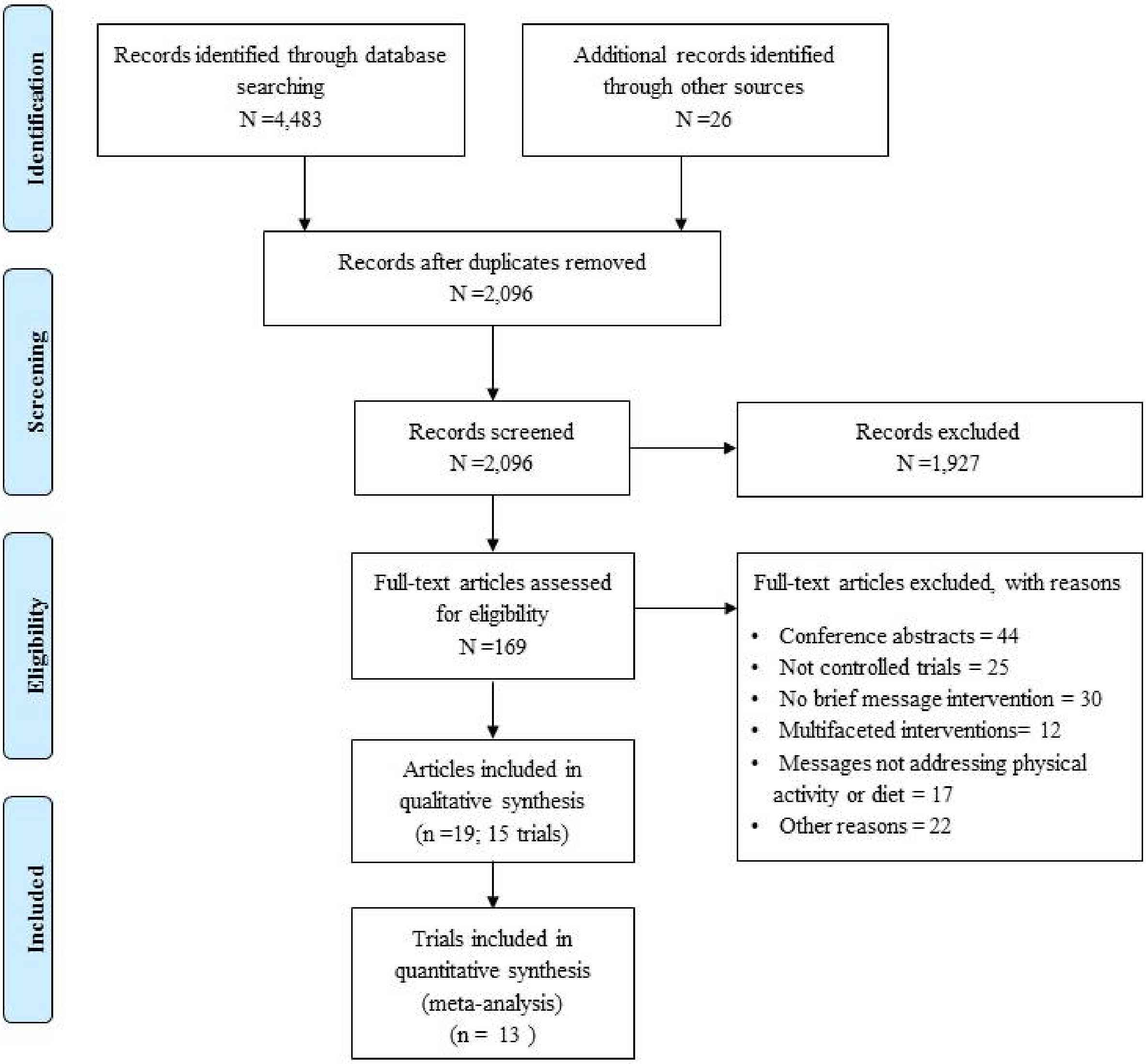
From each study we extracted the mean and standard deviation (SD) of HbA1c levels and BMI, contacting study authors when the information was not available. We transformed this information into weighted mean difference (95% CI), and pooled the data using random-effects models. Where SD of the change between intervention and control group for an outcome was not provided, we derived them from baseline and final SDs, assuming a correlation of 0.5 [25]. A sensitivity analysis was undertaken using different values of correlation to determine whether the overall result of the analysis was robust to the use of imputed correlation coefficients. Heterogeneity was quantified by the I^2statistic, where I^2>50% was considered evidence of substantial heterogeneity [26]. Publication bias was examined with funnel plots and presence of asymmetry tested with Begg [27] and Egger tests [28]. Meta-analyses were conducted with Stata, version 12.0. We set a threshold of P=0.05 to accept statistical significance.

Results

Trial Identification

Search results are summarized in the PRISMA flow diagram (Figure 1). The initial search identified a total of 2096 unique citations. Title and abstract screening of these citations resulted in the inclusion of 169 citations for further review. Following full text screening, 19 articles [29-47] reporting on 15 separate trials (evaluating 15 separate interventions) were finally included.

Figure 1. Flowchart of articles included at each stage of the screening process.



Characteristics of Trials and Interventions

Each trial assessed only one intervention. Thirteen trials used an RCT design, whereas the remaining two used a cluster-RCT [39,40] and a crossover design [44]. The average number of participants per trial was 92 (SD=52), ranging from 19 to 215. Nine of the trials were conducted in HIC, whereas six were conducted in LMIC (Multimedia Appendix 3).

On average, interventions lasted 7 months (SD=4). The majority of trials (n=12) used SMS technology to deliver the messages, whereas the rest were based on graphical information presented to the patients [44], texts available in a website [37], or brief video-messages [31]. Approximately three quarters of the interventions (n=11) addressed both diet and physical activity. Nine interventions consisted of unidirectional messages and six of bidirectional messages. Bidirectional messages were usually initiated by the study participants, who were instructed to use specific devices to conduct the glucose (and in some interventions blood pressure) measurements and transmit the results to the study database using their mobile devices. Other systems used by patients to transmit self-management information included a hand-held electronic diary that allowed participants to describe their meals by selecting the food ingredients [44], or an SMS-based system that allowed them to record their physical activity [47]. In all interventions for each reported measurement, the patient received automated, real-time educational, behavioral, and motivational messaging specific to the entered data, on the basis of a decision-support algorithm.

Twelve behavior-change techniques were used in total. Most interventions used multiple techniques. Most frequently used techniques consisted of 1) providing information about the consequences of inaction, 2) providing instructions about how to perform a behavior, and 3) providing feedback on performance – each of which were used in eight interventions. Whereas all studies examined the impact of the interventions on clinical outcomes, behavior-change was only measured in six of them, with a wide range of instruments being used.

Risk of Bias

Only a minority of the trials presented low risk of bias, and none was completely free of bias (Multimedia Appendices 4 and 5). Most frequent biases were related to blinding of participants and personnel to the interventions (eight trials with high risk of bias), and other sources of bias (seven trials), which were mostly related to small sample sizes that did not allow the detection of clinically meaningful differences.

Acceptability of the Interventions

Five trials examined whether unidirectional messaging interventions were acceptable to participants. Three [29,30,32,33,42] reported high acceptability and satisfaction. However, one [29,33] reported moderate usability, with 40% of the participants requesting to stop receiving the messages before the end of the intervention. Another trial, in which messages were available through a website, reported low acceptability due to the lack of a user-friendly interface and inexperience with mobile web use [37]. In a trial evaluating the use of video-messages [31] it was observed that 47% of the participants in the intervention group did not view videos at all, or did so only briefly at the beginning of their participation and then stopped in the first two months.

Four trials examined participants' acceptability of bidirectional messaging interventions, consistently observing high acceptability in terms of easiness to use systems, usefulness, and general satisfaction [38,44,45,47].

Effectiveness of the Interventions

Impact on Determinants of Behavior-Change

The impact of unidirectional messages on behavior-change determinants was only examined by two trials. One study observed significant improvement in diabetes knowledge and self-efficacy, but not in self-management attitudes [34], whereas the other observed no effect on knowledge or self-efficacy [30,32]. None of the trials evaluating bidirectional messages examined the potential impact on determinants of behavior-change (Multimedia Appendix 6).

Impact on Behavior-Change

Four trials examined the impact of unidirectional messages on diet and physical activity. Two studies reported no effects [30,32,42], whereas the remaining two reported statistically significant improvements in both diet and physical activity [34,43]. Only one trial examined the impact of bidirectional messages on behavior-change [45], reporting no effects.

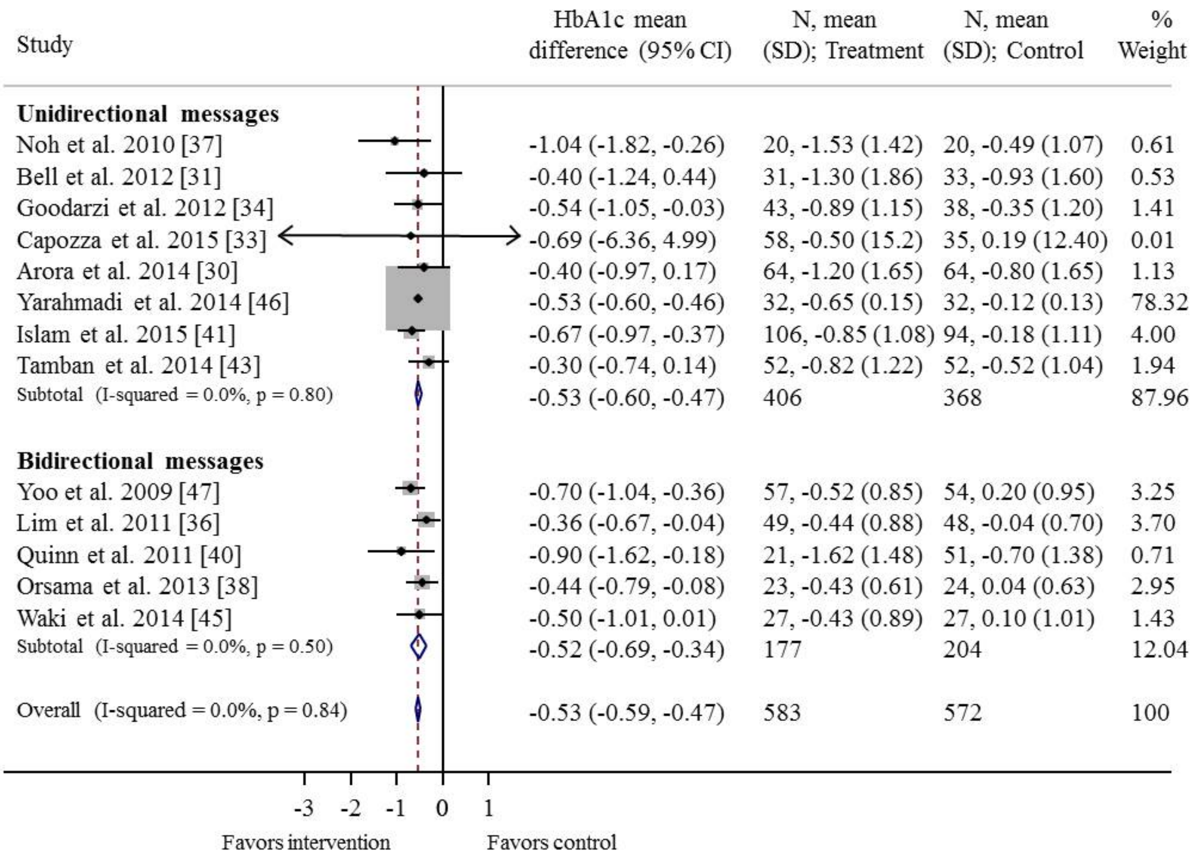
Impact on Clinical Outcomes

Data from thirteen trials reporting the impact of the interventions on HbA1c [29-41,43,45-47] were pooled in a meta-analysis (Figure 2). The trials included thirteen comparisons assessing the impact of unidirectional and bidirectional messages. The weighted HbA1c mean difference between intervention (n=583) and control group (n=572) was -0.53% (95% CI -0.59% to -0.47%). There was no observed heterogeneity in HbA1c among the trials (I^2=0%). Very similar effects were produced by unidirectional (-0.53%, 95% CI -0.60% to -0.47%) compared with bidirectional messages (-0.52%, 95% CI -0.69% to -0.34%).

A second meta-analysis examined the impact of the interventions on BMI. Five trials [36,37,43,45,47] with a total of 406 participants were included. The BMI mean difference between the intervention and control group was -0.25 kg/m^2(-1.02 to 0.52) and not statistically significant (Multimedia Appendix 7). There was no observed heterogeneity among the trials (I^2=0%). Unidirectional messages produced a smaller effect than bidirectional messages (0.08 [-1.76 to 1.93] vs -0.32 [-1.16 to 0.53], respectively), but the difference was not statistically significant. For both meta-analyses, sensitivity analyses confirmed that the overall results were robust to the use of imputed correlation coefficients, and Egger and Begg tests indicated an absence of publication bias.

Other clinical outcomes were too heterogeneous to pool. Unidirectional messaging interventions led to significant reductions of blood pressure in one of the two trials examining this outcome [31,37], and in one [47] of the four bidirectional message-based trials [38-40,45,47]. Improvement in lipid levels was reported for two [29,37] out of three unidirectional messaging interventions, and for three [36,39,40,47] out of four bidirectional messaging interventions. Patient-reported outcomes (diabetes-related distress, diabetes symptoms, and depression) did not significantly improve in either of the two trials that examined these parameters [30,32,39,40].

Figure 2. Weighted mean difference in size of effect of intervention compared with “no treatment” for glycated hemoglobin. HbA1c, glycated hemoglobin; CI, confidence interval; N, number of participants; SD, standard deviation



Differences in Impact Between High-Income and Low- and Middle-Income Countries

The proportion of studies reporting positive effects was consistently higher for trials carried out in LMIC compared to HIC in all domains examined, including acceptability of the interventions (100% in LMIC vs 57% in HIC), impact on determinants of behavior-change (100% vs 0%), on physical activity and diet (67% vs 0%), and on clinical outcomes (100% vs 55%, see Multimedia Appendix 8). Subgroup meta-analysis showed a similar reduction in HbA1c in HIC (-0.53 [-0.60 to -0.47]) compared to LMIC (-0.53 [-0.69 to -0.37]).

Use of Theory in Included Studies

The extent to which the trials explicitly used theory in relation to a number of criteria is reported in Multimedia Appendix 9. In general, theory was not used extensively. Only three trials [35,38-41] explicitly reported that the interventions were based on theory. Two interventions [35,39-41] were based on the transtheoretical model of behavioral change [48] (one study [35,41] in conjunction with the behavioral learning theory [49]), whereas the remaining intervention [38] was based on the information-motivation-behavioral skills model [50,51].

Where theory was explicitly mentioned, two trials [35,38,41] used theory to develop the intervention techniques. Only one study [38] mentioned the targeted construct that the intervention was hypothesized to change, and linked the theoretical constructs to at least one intervention technique. None of the trials measured theory-relevant constructs, used adequate measures of behavior-change, carried out a mediational analysis of constructs, or used their results to refine theory.

Discussion

This systematic review identified 15 controlled trials examining the effectiveness of interventions to promote healthy eating and physical activity in people with type 2 diabetes, delivered via automated brief messaging sent to mobile devices. The interventions predominantly used SMS technology, addressed both diet and physical activity, and were not based on theoretical models of behavior-change. Our meta-analyses showed that automated brief messaging produced a clinically important and statistically significant effect on glycemic control (pooled effect on HbA1C= -0.53%, $P<0.001$), but not on weight loss (BMI= -0.25 kg/m^2, $P=0.53$). In general, interventions based on the use of unidirectional messages produced similar effects compared to those based on bidirectional messages. Interventions conducted in LMIC generally showed a more positive impact than those conducted in HIC.

Strengths and Limitations of the Review

This is the first systematic review specifically examining the impact of automated brief messages on self-management behavior in people with type 2 diabetes. Additional novel aspects of this review include an assessment of the extent to which interventions were based on theory, an assessment of the behavioral change techniques used, and an examination of the relative impact of this type of intervention in countries with different levels of economic development. Relevant trials were identified using a comprehensive search strategy and a large number of bibliographic sources.

In terms of limitations, our meta-analyses were restricted to glycemic control and BMI. Although we intended to conduct meta-analyses on other relevant outcomes (namely diabetes knowledge, attitudes toward diabetes self-management, and change in diet and physical activity-related behavior), these parameters were seldom measured, which represents a gap in evaluations to date. Finally, although formal tests on publication bias seemed to exclude its presence, we cannot completely rule out its existence.

Comparison with Previous Reviews and Implications

The positive findings observed in our diabetes-specific review are consistent with findings from reviews examining a wide range of conditions. For example, a recent meta-analysis observed that SMS messages produced a small, positive, significant effect (g=0.29) on a broad range of healthy behaviors in patients with different types of long-term conditions [12]. A recent study reviewed 15 systematic reviews and meta-analyses, observing that the majority of published text-messaging interventions were effective when addressing weight loss, physical activity, smoking cessation, and medication adherence for antiretroviral therapy [52].

The estimated 0.53% reduction in HbA1c observed in our meta-analysis is clinically important, as evidence suggests that every percentage point decrease in HbA1c over 10 years is associated with a risk reduction of 21% for deaths related to diabetes, 14% for myocardial infarctions, and 37% for microvascular complications [53]. Our result is consistent with findings from a previous systematic review of computer-based interventions to improve diabetes self-management, which showed that interventions based on the use of mobile phones (although not specifically text messages) produced the largest HbA1c reductions (-0.5%) [19]. We deliberately focused this review on interventions to improve physical activity and healthy diet. Medication adherence is also a key aspect of diabetes self-management, and adherence behavior can also be targeted by messaging interventions. We examined the impact of brief messages to improve adherence to diabetes medication in a separate systematic review, which provided evidence that messages produced a moderate positive effect on medication adherence and clinical outcomes [17].

Interventions based on the use of one-way messages produced a very similar effect to those based on two-way messages, which can be more tailored and usually require more complex technology, and are more resource intensive. This similarity is also consistent with findings from a recent trial evaluating text messages to improve treatment adherence in people with hypertension. This result is also consistent with findings from our systematic review on messages to improve adherence to diabetes medication, which observed that interventions exclusively based on brief messages produced a similar effect compared to more complex interventions combining messaging with monitoring strategies [17].

More than half of the trials did not include measures of behavior-change, and those that did reported mixed results. Previous systematic reviews also reported mixed results. For example, Cassimatis et al [54] observed that only five out of eight trials examining the effects of type 2 diabetes behavioral telehealth interventions showed significant improvements in dietary adherence and physical activity. Cotter et al [55] observed that only two out of nine studies based on Internet interventions to support lifestyle modification for type 2 diabetes management demonstrated improvements in diet or physical activity.

Although we did not observe a statistically significant reduction in BMI, we cannot exclude a small reduction. Since the text messages specifically targeted diet and physical activity behavior, we expected a greater effect. However, there were only a low number of trials reporting BMI as an outcome. Evidence from previous systematic reviews examining the impact of text messages on weight reduction is mixed, with some suggesting lack of consistent effects [56,57] and others reporting significant weight loss [58].

The interventions seem to have been acceptable to the recipients. There are many features related to mobile technology that may engage patients with the intervention. Some of these features include ease of use, convenience (eg messages need not be retrieved immediately), mobility (eg read at home or away), and frequent reinforcement (eg can read more than once). However, measures of acceptance and usability of the interventions assessed in the review were not obtained using a validated tool.

Our results suggested that the interventions were more effective in LMIC than in HIC. As far as we know, this is the first study comparing the impact of automated brief messages on long-term condition self-management between countries with different levels of economic development. Two recent reviews evaluated the impact of mHealth interventions in patients with long-term conditions living in LMICs, concluding that they are cost-effective and can produce a positive impact on clinical outcomes, health-related quality of life [59], and treatment adherence [60]. Increasing evidence suggests that mHealth interventions are a useful tool to address health care system constraints in developing countries, namely limited health care workforce, limited financial resources, high burden of disease, and difficulties in providing health care to hard-to-reach populations [61]. These factors may partially explain the more positive impact in LMIC observed in our review.

Limitations of Available Evidence and Future Research Needs

The studies in our review consistently supported the use of brief messages to promote healthier lifestyle behavior in patients with type 2 diabetes. However, available evidence is limited by several factors. First, most of the trials presented moderate or high risk of bias, mainly due to small sample sizes and inadequate blinding. To confirm the positive findings observed in our review, methodologically robust trials of greater size are very much needed. Second, although all the interventions specifically aimed to improve lifestyle behavior, behavior-change was measured in less than half of the studies. Where measured, a wide range of instruments were used, most having been designed *ad hoc* and not meeting adequate standards for validity or reliability. Third, only a small fraction of trials reported use of explicit behavior-change theory. Where it was mentioned, theory was used to design the intervention, but not to examine process measures that might indicate effect, or to subsequently refine theory. There is a current debate about whether or not the evidence base for behavior-change interventions can be enhanced by applying relevant theory. It has been suggested that doing so may focus attention on the mechanisms by which interventions are effective [62]. Finally, interventions used a relatively narrow range of behavior-change techniques, focusing on provision of information. Techniques such as those involving goal-setting and planning how to enact behavior or elicit social support were seldom considered, despite evidence that such techniques are generally effective at increasing physical activity in people with diabetes or obese people [63,64].

Additional research needs include an estimation of the cost-effectiveness of the interventions, an examination of their long-term impact, an understanding of what circumstances are effective (which features of the underlying health system and target population are helpful, and which features mitigate against them working), assessment of intervention safety, and an examination of their potential contribution to more comprehensive, multifaceted interventions [52].

Conclusions

Interventions based on the use of automated brief messages sent to mobile devices to promote lifestyle behavior can improve glycemic control in patients with type 2 diabetes, both in developed and developing countries. Larger and methodologically robust trials are needed to confirm these positive findings.

Acknowledgments

AJF, CA, IR, DPF and PM designed the study. NR undertook the literature search. IR, PM, and CA were involved in the short-listing of identified studies, with input from AJF. Data extraction was undertaken by IR and PM, with input from AJF. Statistical analyses were undertaken by IR. All authors contributed to the final manuscript.

This work was funded through an NIHR Senior Investigator Award to AJF, who also received funding from the NIHR Oxford Biomedical Research Centre.

This review was carried out in partial fulfilment of CA’s fellowship at the University of Oxford under the Commonwealth Academic Fellowship scheme.

The funders had no role in study design, data collection or analysis, decision to publish, or preparation of the manuscript. The views expressed are those of the authors and not necessarily those of the funders.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Bibliographic searches - registry of searches.

[PDF File (Adobe PDF File), 30KB - jmir_v18i4e86_app1.pdf]

Multimedia Appendix 2

Bibliographic searches - search strategy (Medline).

[PDF File (Adobe PDF File), 30KB - jmir_v18i4e86_app2.pdf]

Multimedia Appendix 3

Characteristics of the identified trials and interventions.

[PDF File (Adobe PDF File), 31KB - jmir_v18i4e86_app3.pdf]

Multimedia Appendix 4

Cochrane summary risk of bias for the included trials (n=15).

[PDF File (Adobe PDF File), 18KB - jmir_v18i4e86_app4.pdf]

Multimedia Appendix 5

Cochrane individual risk of bias for the included trials (n=15).

[PDF File (Adobe PDF File), 145KB - jmir_v18i4e86_app5.pdf]

Multimedia Appendix 6

Main results of the studies identified.

[PDF File (Adobe PDF File), 36KB - jmir_v18i4e86_app6.pdf]

Multimedia Appendix 7

Weighted mean difference in size of effect of intervention compared with “no treatment” for body mass index.

[PDF File (Adobe PDF File), 57KB - jmir_v18i4e86_app7.pdf]

Multimedia Appendix 8

Differences between low- and middle- income countries and high-income countries in the impact of the interventions.

[PDF File (Adobe PDF File), 22KB - jmir_v18i4e86_app8.pdf]

Multimedia Appendix 9

Degree of use of theory in the development of the interventions.

[PDF File (Adobe PDF File), 42KB - jmir_v18i4e86_app9.pdf]

References

1. International Diabetes Federation. 2014. Diabetes atlas sixth edition URL: https://www.idf.org/sites/default/files/EN_6E_Atlas_Full_0.pdf [accessed 2015-12-09] [WebCite Cache ID 6de5S62dj]
2. Fowler MJ. Microvascular and macrovascular complications of diabetes. Clin Diabetes 2008 Apr 01;26(2):77-82. [doi: 10.2337/diaclin.26.2.77]
3. van Dieren S, Beulens JW, van der Schouw YT, Grobbee DE, Neal B. The global burden of diabetes and its complications: an emerging pandemic. Eur J Cardiovasc Prev Rehabil 2010 May;17 Suppl 1:S3-S8. [doi: 10.1097/01.hjr.0000368191.86614.5a] [Medline: 20489418]
4. Ray KK, Seshasai SR, Wijesuriya S, Sivakumaran R, Nethercott S, Preiss D, et al. Effect of intensive control of glucose on cardiovascular outcomes and death in patients with diabetes mellitus: a meta-analysis of randomised controlled trials. Lancet 2009 May 23;373(9677):1765-1772. [doi: 10.1016/S0140-6736(09)60697-8] [Medline: 19465231]
5. Andrews RC, Cooper AR, Montgomery AA, Norcross AJ, Peters TJ, Sharp DJ, et al. Diet or diet plus physical activity versus usual care in patients with newly diagnosed type 2 diabetes: the Early ACTID randomised controlled trial. Lancet 2011 Jul 9;378(9786):129-139 [FREE Full text] [doi: 10.1016/S0140-6736(11)60442-X] [Medline: 21705068]
6. Thomas DE, Elliott EJ, Naughton GA. Exercise for type 2 diabetes mellitus. Cochrane Database Syst Rev 2006(3):CD002968. [doi: 10.1002/14651858.CD002968.pub2] [Medline: 16855995]
7. Campbell R, Pound P, Pope C, Britten N, Pill R, Morgan M, et al. Evaluating meta-ethnography: a synthesis of qualitative research on lay experiences of diabetes and diabetes care. Soc Sci Med 2003 Feb;56(4):671-684. [Medline: 12560003]
8. Frost J, Garside R, Cooper C, Britten N. A qualitative synthesis of diabetes self-management strategies for long term medical outcomes and quality of life in the UK. BMC Health Serv Res 2014;14:348 [FREE Full text] [doi: 10.1186/1472-6963-14-348] [Medline: 25127714]
9. American Diabetes Association. Standards of medical care in diabetes - 2010. Diabetes Care 2010 Jan;33 Suppl 1:S11-S61 [FREE Full text] [doi: 10.2337/dc10-S011] [Medline: 20042772]
10. National Institute for Health and Care Excellence. 2015 Aug. Type 1 diabetes in adults: diagnosis and management URL: http://www.nice.org.uk/guidance/ng17 [accessed 2015-12-09] [WebCite Cache ID 6de5vct4w]
11. Piette JD, List J, Rana GK, Townsend W, Striplin D, Heisler M. Mobile health devices as tools for worldwide cardiovascular risk reduction and disease management. Circulation 2015 Nov 24;132(21):2012-2027. [doi: 10.1161/CIRCULATIONAHA.114.008723] [Medline: 26596977]
12. Orr JA, King RJ. Mobile phone SMS messages can enhance healthy behaviour: a meta-analysis of randomised controlled trials. Health Psychol Rev 2015 Dec;9(4):397-416. [doi: 10.1080/17437199.2015.1022847] [Medline: 25739668]

13. International Telecommunication Union. Geneva, Switzerland; 2014. Measuring the information society report 2014: executive summary URL: https://www.itu.int/dms_pub/itu-d/opb/ind/D-IND-ICTOI-2014-SUM-PDF-E.pdf [accessed 2015-12-09] [WebCite Cache ID 6de64WcdD]
14. Naughton F, Sutton S. The European Health Psychologist. 2011. Going mobile: delivering behavioural support via SMS text messages URL: http://openhealthpsychology.net/ehp/issues/2011/v13iss2_June2011/13_2_NaughtonSutton.pdf [accessed 2015-12-09] [WebCite Cache ID 6de66qKaY]
15. Heron KE, Smyth JM. Ecological momentary interventions: incorporating mobile technology into psychosocial and health behaviour treatments. Br J Health Psychol 2010 Feb;15(Pt 1):1-39 [FREE Full text] [doi: 10.1348/135910709X466063] [Medline: 19646331]
16. Burke LE, Ma J, Azar KM, Bennett GG, Peterson ED, Zheng Y, American Heart Association Publications Committee of the Council on Epidemiology and Prevention, Behavior Change Committee of the Council on Cardiometabolic Health, Council on Cardiovascular and Stroke Nursing, Council on Functional Genomics and Translational Biology, Council on Quality of Care and Outcomes Research, and Stroke Council. Current science on consumer use of mobile health for cardiovascular disease prevention: a scientific statement from the American Heart Association. Circulation 2015 Sep 22;132(12):1157-1213. [doi: 10.1161/CIR.0000000000000232] [Medline: 26271892]
17. Farmer AJ, McSharry J, Rowbotham S, McGowan L, Ricci-Cabello I, French DP. Effects of interventions promoting monitoring of medication use and brief messaging on medication adherence for people with type 2 diabetes: a systematic review of randomized trials. Diabet Med 2015 Oct 15;132(12):1157-1213. [doi: 10.1111/dme.12987] [Medline: 26470750]
18. Free C, Phillips G, Galli L, Watson L, Felix L, Edwards P, et al. The effectiveness of mobile-health technology-based health behaviour change or disease management interventions for health care consumers: a systematic review. PLoS Med 2013;10(1):e1001362 [FREE Full text] [doi: 10.1371/journal.pmed.1001362] [Medline: 23349621]
19. Pal K, Eastwood SV, Michie S, Farmer A, Barnard ML, Peacock R, et al. Computer-based interventions to improve self-management in adults with type 2 diabetes: a systematic review and meta-analysis. Diabetes Care 2014 Jun;37(6):1759-1766. [doi: 10.2337/dc13-1386] [Medline: 24855158]
20. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Ann Intern Med 2009 Aug 18;151(4):264-9, W64. [Medline: 19622511]
21. The World Bank. 2015. Country and lending groups 2015 URL: http://data.worldbank.org/about/country-and-lending-groups [accessed 2015-12-09] [WebCite Cache ID 6de6AmuEP]
22. Michie S, Prestwich A. Are interventions theory-based? Development of a theory coding scheme. Health Psychol 2010 Jan;29(1):1-8. [doi: 10.1037/a0016939] [Medline: 20063930]
23. Abraham C, Michie S. A taxonomy of behavior change techniques used in interventions. Health Psychol 2008 May;27(3):379-387. [doi: 10.1037/0278-6133.27.3.379] [Medline: 18624603]
24. Higgins JPT, Green S, editors. Assessing risk of bias in included studies. In: Cochrane handbook for systematic reviews of interventions, version 5.1.0. Chichester, UK: John Wiley & Sons, Ltd; 2008:187-242.
25. Rice J. Mathematical statistics and data analysis, 3rd edition. Belmont, CA: Brooks/Cole; 2007.
26. Higgins JPT , Green S , editors. Analysing data and undertaking meta-analyses. In: Cochrane handbook for systematic reviews of interventions, Version 5.1.0. Chichester, UK: John Wiley & Sons, Ltd; 2008:243-298.
27. Begg CB, Mazumdar M. Operating characteristics of a rank correlation test for publication bias. Biometrics 1994 Dec;50(4):1088-1101. [Medline: 7786990]
28. Egger M, Davey SG, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ 1997 Sep 13;315(7109):629-634 [FREE Full text] [Medline: 9310563]
29. Abebe NA, Capozza KL, Des Jardins TR, Kulick DA, Rein AL, Schachter AA, et al. Considerations for community-based mHealth initiatives: insights from three Beacon Communities. J Med Internet Res 2013;15(10):e221 [FREE Full text] [doi: 10.2196/jmir.2803] [Medline: 24128406]
30. Arora S, Peters AL, Burner E, Lam CN, Menchine M. Trial to examine text message-based mHealth in emergency department patients with diabetes (TExT-MED): a randomized controlled trial. Ann Emerg Med 2014 Jun;63(6):745-54.e6. [doi: 10.1016/j.annemergmed.2013.10.012] [Medline: 24225332]
31. Bell AM, Fonda SJ, Walker MS, Schmidt V, Vigersky RA. Mobile phone-based video messages for diabetes self-care support. J Diabetes Sci Technol 2012 Mar;6(2):310-319 [FREE Full text] [Medline: 22538140]
32. Burner ER, Menchine MD, Kubicek K, Robles M, Arora S. Perceptions of successful cues to action and opportunities to augment behavioral triggers in diabetes self-management: qualitative analysis of a mobile intervention for low-income Latinos with diabetes. J Med Internet Res 2014;16(1):e25 [FREE Full text] [doi: 10.2196/jmir.2881] [Medline: 24476784]
33. Capozza K, Woolsey S, Georgsson M, Black J, Bello N, Lence C, et al. Going mobile with diabetes support: a randomized study of a text message-based personalized behavioral intervention for type 2 diabetes self-care. Diabetes Spectr 2015 May;28(2):83-91. [doi: 10.2337/diaspect.28.2.83] [Medline: 25987806]
34. Goodarzi M, Ebrahimzadeh I, Rabi A, Saedipoor B, Jafarabadi MA. Impact of distance education via mobile phone text messaging on knowledge, attitude, practice and self efficacy of patients with type 2 diabetes mellitus in Iran. J Diabetes Metab Disord 2012;11(1):10 [FREE Full text] [doi: 10.1186/2251-6581-11-10] [Medline: 23497632]

35. Islam SM, Lechner A, Ferrari U, Froeschl G, Alam DS, Holle R, et al. Mobile phone intervention for increasing adherence to treatment for type 2 diabetes in an urban area of Bangladesh: protocol for a randomized controlled trial. BMC Health Serv Res 2014;14:586 [FREE Full text] [doi: 10.1186/s12913-014-0586-1] [Medline: 25424425]
36. Lim S, Kang SM, Shin H, Lee HJ, Won YJ, Yu SH, et al. Improved glycemic control without hypoglycemia in elderly diabetic patients using the ubiquitous healthcare service, a new medical information system. Diabetes Care 2011 Feb;34(2):308-313 [FREE Full text] [doi: 10.2337/dc10-1447] [Medline: 21270188]
37. Noh JH, Cho Y, Nam H, Kim J, Kim D, Yoo H, et al. Web-based comprehensive information system for self-management of diabetes mellitus. Diabetes Technol Ther 2010 May;12(5):333-337. [doi: 10.1089/dia.2009.0122] [Medline: 20388042]
38. Orsama AL, Lähteenmäki J, Harno K, Kulju M, Wintergerst E, Schachner H, et al. Active assistance technology reduces glycosylated hemoglobin and weight in individuals with type 2 diabetes: results of a theory-based randomized trial. Diabetes Technol Ther 2013 Aug;15(8):662-669. [doi: 10.1089/dia.2013.0056] [Medline: 23844570]
39. 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]
40. 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]
41. Shariful Islam SM, Niessen LW, Ferrari U, Ali L, Seissler J, Lechner A. Effects of mobile phone SMS to improve glycemic control among patients with type 2 diabetes in Bangladesh: a prospective, parallel-group, randomized controlled trial. Diabetes Care 2015 Aug;38(8):e112-e113. [doi: 10.2337/dc15-0505] [Medline: 26207059]
42. Shetty AS, Chamukuttan S, Nanditha A, Raj RK, Ramachandran A. Reinforcement of adherence to prescription recommendations in Asian Indian diabetes patients using short message service (SMS) - a pilot study. J Assoc Physicians India 2011 Nov;59:711-714. [Medline: 22616337]
43. Tamban C, Isip-Tan I, Jimeno C. Journal of the ASEAN Federation of Endocrine Societies. 2013 Dec 28. Use of short message services (SMS) for the management of type 2 diabetes mellitus: a randomized controlled trial URL: http://www.asean-endocrinejournal.org/index.php/JAFES/article/view/68 [accessed 2015-12-09] [WebCite Cache ID 6de6PnZgo]
44. Tsang MW, Mok M, Kam G, Jung M, Tang A, Chan U, et al. Improvement in diabetes control with a monitoring system based on a hand-held, touch-screen electronic diary. J Telemed Telecare 2001;7(1):47-50. [Medline: 11265938]
45. 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]
46. Yarahmadi A, Zare-Farashbandi F, Kachuei A, Nouri R, Hassanzadeh A. The effects of non-attendance information therapy on the control of glycosylated hemoglobin (HbA1C) in type 2 diabetic patients. J Educ Health Promot 2014;3:90 [FREE Full text] [doi: 10.4103/2277-9531.139650] [Medline: 25250356]
47. Yoo HJ, Park MS, Kim TN, Yang SJ, Cho GJ, Hwang TG, et al. A ubiquitous chronic disease care system using cellular phones and the internet. Diabet Med 2009 Jun;26(6):628-635. [doi: 10.1111/j.1464-5491.2009.02732.x] [Medline: 19538239]
48. Prochaska JO, DiClemente CC. Transtheoretical therapy: toward a more integrative model of change. Psychother Theor Res 1982;19(3):276-288. [doi: 10.1037/h0088437]
49. Leventhal H, Cameron L. Behavioral theories and the problem of compliance. Patient Educ Couns 1987 Oct;10(2):117-138. [doi: 10.1016/0738-3991(87)90093-0]
50. Fisher JD, Fisher WA. Changing AIDS-risk behavior. Psychol Bull 1992 May;111(3):455-474. [Medline: 1594721]
51. Osborn CY, Rivet AK, Fisher WA, Egede LE, Fisher JD. An information-motivation-behavioral skills analysis of diet and exercise behavior in Puerto Ricans with diabetes. J Health Psychol 2010 Nov;15(8):1201-1213 [FREE Full text] [doi: 10.1177/1359105310364173] [Medline: 20453056]
52. Hall AK, Cole-Lewis H, Bernhardt JM. Mobile text messaging for health: a systematic review of reviews. Annu Rev Public Health 2015 Mar 18;36:393-415 [FREE Full text] [doi: 10.1146/annurev-publhealth-031914-122855] [Medline: 25785892]
53. UK Prospective Diabetes Study Group. Intensive blood-glucose control with sulphonylureas or insulin compared with conventional treatment and risk of complications in patients with type 2 diabetes (UKPDS 33). UK Prospective Diabetes Study (UKPDS) Group. Lancet 1998 Sep 12;352(9131):837-853. [Medline: 9742976]
54. Cassimatis M, Kavanagh DJ. Effects of type 2 diabetes behavioural telehealth interventions on glycaemic control and adherence: a systematic review. J Telemed Telecare 2012 Dec;18(8):447-450. [doi: 10.1258/jtt.2012.GTH105] [Medline: 23209266]
55. Cotter AP, Durant N, Agne AA, Cherrington AL. Internet interventions to support lifestyle modification for diabetes management: a systematic review of the evidence. J Diabetes Complications 2014;28(2):243-251 [FREE Full text] [doi: 10.1016/j.jdiacomp.2013.07.003] [Medline: 24332469]
56. Keating SR, McCurry MK. Systematic review of text messaging as an intervention for adolescent obesity. J Am Assoc Nurse Pract 2015 Apr 25;27(12):714-720. [doi: 10.1002/2327-6924.12264] [Medline: 25914399]

57. Shaw R, Bosworth H. Short message service (SMS) text messaging as an intervention medium for weight loss: a literature review. Health Informatics J 2012 Dec;18(4):235-250 [FREE Full text] [doi: 10.1177/1460458212442422] [Medline: 23257055]
58. Siopis G, Chey T, Allman-Farinelli M. A systematic review and meta-analysis of interventions for weight management using text messaging. J Hum Nutr Diet 2015 Feb;28 Suppl 2:1-15. [doi: 10.1111/jhn.12207] [Medline: 24480032]
59. 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]
60. 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? Glob Health Action 2014;7:25606 [FREE Full text] [Medline: 25361730]
61. Mechael P, Batavia H, Kaonga N, Searly S, Kwan A, Goldberger A. mHealth Alliance. 2010 Mar 10. Barriers and gaps affecting mhealth in low and middle income countries: policy white paper URL: https://www.ghdonline.org/uploads/Barriers_Gaps_to_mHealth_in_LMICs_-_White_Paper_-_May_2010.pdf [accessed 2016-04-03] [WebCite Cache ID 6gVVIELrs]
62. Prestwich A, Webb TL, Conner M. Using theory to develop and test interventions to promote changes in health behaviour: evidence, issues, and recommendations. Curr Opin Psychol 2015 Oct;5:1-5. [doi: 10.1016/j.copsyc.2015.02.011]
63. Avery L, Flynn D, van Wersch A, Sniehotta FF, Trenell MI. Changing physical activity behavior in type 2 diabetes: a systematic review and meta-analysis of behavioral interventions. Diabetes Care 2012 Dec;35(12):2681-2689 [FREE Full text] [doi: 10.2337/dc11-2452] [Medline: 23173137]
64. Olander EK, Fletcher H, Williams S, Atkinson L, Turner A, French DP. What are the most effective techniques in changing obese individuals' physical activity self-efficacy and behaviour: a systematic review and meta-analysis. Int J Behav Nutr Phys Act 2013;10:29 [FREE Full text] [doi: 10.1186/1479-5868-10-29] [Medline: 23452345]
65. Hall AK, Cole-Lewis H, Bernhardt JM. Mobile text messaging for health: a systematic review of reviews. Annu Rev Public Health 2015 Mar 18;36:393-415 [FREE Full text] [doi: 10.1146/annurev-publhealth-031914-122855] [Medline: 25785892]

Abbreviations

BMI: body mass index
HbA1c: glycated hemoglobin
HIC: high-income country
LMIC: low- and middle-income countries
mHealth: mobile health
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RCT: randomized controlled trial
SD: standard deviation
SMS: short message service

Edited by G Eysenbach; submitted 09.12.15; peer-reviewed by P Aslani, J Lähteenmäki, E Aramaki; comments to author 07.01.16; revised version received 29.01.16; accepted 30.01.16; published 19.04.16.

Please cite as:
Arambepola C, Ricci-Cabello I, Manikavasagam P, Roberts N, French DP, Farmer A
The Impact of Automated Brief Messages Promoting Lifestyle Changes Delivered Via Mobile Devices to People with Type 2 Diabetes: A Systematic Literature Review and Meta-Analysis of Controlled Trials
J Med Internet Res 2016;18(4):e86
URL: http://www.jmir.org/2016/4/e86/
doi:10.2196/jmir.5425
PMID:27095386

Original Paper

Evaluation of QuitNow Men: An Online, Men-Centered Smoking Cessation Intervention

Joan L Bottorff[1,2*], RN, PhD, FCAHS, FAAN; John L Oliffe[3*], RN, PhD; Gayl Sarbit[1*], PhD; Paul Sharp[1], MSc; Cristina M Caperchione[4*], PhD; Leanne M Currie[3*], RN, PhD; Jonathan Schmid[5*], MA; Martha H Mackay[3], RN, PhD; Sean Stolp[1], MA

[1]Institute for Healthy Living and Chronic Disease Prevention, University of British Columbia, Kelowna, BC, Canada
[2]Australian Catholic University, Faculty of Health Sciences, Melbourne, Australia
[3]School of Nursing, Faculty of Applied Science, University of British Columbia, Vancouver, BC, Canada
[4]School of Health and Exercise Science, Faculty of Health and Social Development, University of British Columbia, Kelowna, BC, Canada
[5]Context Ltd, Vancouver, BC, Canada
[*]these authors contributed equally

Corresponding Author:
Joan L Bottorff, RN, PhD, FCAHS, FAAN
Institute for Healthy Living and Chronic Disease Prevention
University of British Columbia
1147 Research Road, Art 223
Kelowna, BC, V1V 1V7
Canada
Phone: 1 250 807 8627
Fax: 1 250 807 8629
Email: joan.bottorff@ubc.ca

Abstract

Background: Men continue to smoke cigarettes in greater numbers than women. There is growing evidence for the value of developing targeted, men-centered health promotion programs. However, few smoking cessation interventions have been designed for men. A gender-specific website, QuitNow Men, was developed based on focus group interview findings, stakeholder feedback, and evidence-based cessation strategies. The website was designed to incorporate a masculine look and feel through the use of images, direct language, and interactive content. Usability experts and end-users provided feedback on navigation and functionality of the website prior to pilot testing.

Objectives: The objectives of the pilot study were to describe (1) men's use and evaluations of the interactive resources and information on the QuitNow Men website, and (2) the potential of QuitNow Men to engage men in reducing and quitting smoking.

Methods: A one-group, pretest-posttest study design was used. Men who were interested in quitting were recruited and invited to use the website over a 6-month period. Data were collected via online questionnaires at baseline, 3-month, and 6-month follow-up. A total of 117 men completed the baseline survey. Over half of those (67/117, 57.3%) completed both follow-up surveys.

Results: At baseline, participants (N=117) had been smoking for an average of 24 years (SD 12.1) and smoked on average 15 cigarettes a day (SD 7.4). The majority had not previously used a quit smoking website (103/117, 88.0%) or websites focused on men's health (105/117, 89.7%). At the 6-month follow-up, the majority of men used the QuitNow Men website at least once (64/67, 96%). Among the 64 users, 29 (43%) reported using the website more than 6 times. The men using QuitNow Men agreed or strongly agreed that the website was easy to use (51/64, 80%), the design and images were appealing (42/64, 66%), they intended to continue to use the website (42/64, 66%), and that they would recommend QuitNow Men to others who wanted to quit (46/64, 72%). Participants reported using an average of 8.76 (SD 4.08) of the 15 resources available on the website. At 6-month follow-up, 16 of the 67 participants (24%) had quit, 27 (40%) had reduced their smoking and 24 (36%) had not changed their smoking habits. Repeated measures general linear model showed a significant decrease in the number of cigarettes smoked between the 3-month and 6-month follow-up ($F_{1,63}$=6.41, P=.01, eta squared=0.09). Number of resources used on the website, quit confidence, nicotine dependence and age significantly predicted number of quit attempts by those still smoking at 6 months ($F_{4,45}$=2.73, P=.04), with number of resources used being the strongest predictor (P=.02).

Conclusions: The results of this research support efforts to integrate gender-sensitive approaches in smoking cessation interventions and indicate that this novel Web-based resource has potential in supporting men's smoking cessation efforts.

(J Med Internet Res 2016;18(4):e83) doi:10.2196/jmir.5076

KEYWORDS

smoking cessation; web-based health promotion; internet; masculinity; men's health

Introduction

Tobacco use remains a serious and persistent health risk and is a leading preventable cause of death related to cancer, heart disease, and other chronic diseases [1,2]. Since most tobacco is consumed in the form of manufactured cigarettes, this risk is highest among men because they smoke in greater numbers [3]. By estimates that include 21 countries, there are over 879 million current smokers, of whom 721 million are men [4].

Despite decreases in smoking rates in Canada over the past decade, men (18%) continue to smoke in greater numbers than women (14%) [5]. These trends point to a critical need to rethink how cessation services are delivered to men. Although there is increasing evidence for the effectiveness of gender-specific health promotion programs, few men-centered tobacco reduction and cessation interventions have been developed and/or formally evaluated [6,7].

Experts suggest that dynamic and interactive Web-based technologies have the potential to encourage and assist men in accessing and using health promotion information [8,9]. Although findings from a systematic review of Web-based smoking cessation programs indicate their potential usefulness [10], the majority of programs have not targeted specific subgroups of individuals or tailored interventions to their preferences and personal characteristics. Evidence suggests that online "one-size fits-all" cessation programs do not garner better outcomes than usual care, printed self-help materials, or email notification interventions [11,12].

Current research reveals men turn to the Internet for health information [8,9,13]. Websites and social networks influence men's lifestyles, enable information-gathering, and provide opportunities to engage in mutual help and content co-creation [14,15]. Specific to smoking cessation, Web-based and mobile apps have emerged as commonplace, making possible on-demand and just-in-time information retrieval [16-18]. This, in turn, fosters "collective intelligence," drawing on men's preferences for autonomy in decision making related to mapping and monitoring their quit strategies. There is also evidence signaling the effectiveness of tailored, interactive online cessation programs. A three-arm randomized trial (N=1236) conducted in the United States compared the psychosocial variables mediating 30-day abstinence among adult participants assigned to either a basic cessation website, an interactive cessation website enhanced with online social networking, or an interactive and enhanced website with the addition of telephone counseling [19]. Increased levels of website utilization predicted higher abstinence rates. However, the authors caution that more work is needed to enhance the effectiveness of online smoking cessation resources [19].

To assist with reducing the number of men who smoke, we designed a men-centered smoking cessation website in collaboration with experts from QuitNow, the British Columbia Lung Association's QuitNow.ca site. The purpose of the current paper is to (1) describe the development and usability testing of QuitNow Men, a novel, tailored, evidence-based smoking cessation website optimized for use on desktop and mobile devices, and (2) report on a pilot study conducted to garner insights about the acceptability and potential effectiveness of the QuitNow Men website. The pilot study objectives were (1) to describe men's use and evaluation of the interactive resources and information included on the QuitNow Men website, and (2) to evaluate the potential of the QuitNow Men website to engage men in reducing and quitting smoking. The study protocol for all phases of this project was approved by the University of British Columbia Behavioural Research Ethics Board.

Methods

Initial Development of the QuitNow Men Website

Four focus groups were conducted with men who smoke or had recently quit (N=56) to gather information on men's motivations for quitting and their suggestions for developing men-centered resources [20]. The men were recruited from three locations in British Columbia, Canada, using posters and online strategies, and represented diverse sociodemographic and ethnocultural backgrounds. Their suggestions for the kinds of support they would find helpful to quit smoking, along with evidence-based cessation strategies, guided the development of gender-specific online resources for the website. QuitNow Men was optimized for mobile use and designed to incorporate a look and feel that would appeal to men 18-45 years of age by using masculine images, direct language, and content that included interactive video dramas [21] (see Table 1). In addition, to tailor content to men's preferences, we used strong, positive messages to promote change (eg, "Put these tactics and tools to good use to get the job done"); connected positive identities, such as being healthy and strong, with being smoke free; and included men's stories about quitting to show common challenges and create a community of mutual help. An interactive discussion forum, a section for sharing stories about reducing and quitting, and a formatted email to "get a buddy on board" were developed for men to engage each other, provide support, and share personal experiences. To reflect preferences for autonomous decision making, the website was structured to offer men choices by providing an array of resources to map, monitor, and maintain their quit. A section with tactics and tools to support cessation efforts was provided along with quizzes, videos, expert chat, text support, and a smoking calculator. Table 1 provides an overview of resources available on the QuitNow Men website.

Table 1. QuitNow Men website resources.

Resource	Description
Man-friendly informational resources	
Tactics for getting started	Man-tailored information about different quitting methods, strategies for dealing with withdrawal, cravings, and smoking triggers.
Tactics for surviving quit day	Man-tailored information about lining up support, gathering supplies, avoiding risky situations, staying busy, and changing routines.
Tactics for staying on track	Man-tailored information with suggestions for staying smoke free. Topics include slips vs relapses, keeping one’s guard up, and strategies for dealing with stress.
Patches, gum, and more	Information about quit aids and where to obtain them, including; nicotine patches, gum, inhalers, lozenges, mouth sprays, and prescription drugs.
Videos on addiction	Inspiring, informational, and motivating videos from the Web on the science behind addiction.
Man-friendly interactive resources	
Pick a quit date tool	Participants select a quit date that will appear in the top corner of the webpage and a digital timer begins a countdown to the chosen date. Emails are sent to the participants: two before the quit date, one on the quit date, and four follow-up emails after the quit date. Once a quit date is reached, the timer begins to count the number of day, hours, and minutes of being smoke free.
Smoking dependence mini-quiz	A 6-item self-report questionnaire that evaluates participant’s addiction level [19] and provides suggestions for next steps.
Smoking cost calculator	A smoking calculator that calculates the cost spent on cigarettes and provides suggestions of things that money could be used to buy (eg, a new truck). Cost savings are also calculated based on a 50% reduction in smoking behavior.
If I were Nick (interactive video drama)	Seven brief scenarios about a character, Nick, on his first day of a quit. The audience is encouraged to put themselves in Nick’s shoes and answer reflective questions on how they would think and feel in certain situations [18].
Social support	
Tune into the forum	A community forum where participants can connect and discuss topics surrounding reducing and quitting smoking.
Read shared stories	A space where men can read and share quit stories with other users on the website.
Text messages with quit tips	A 3-month automated text messaging system that is designed to provide advice and tips to help participants prepare for a quit and stay smoke free. Information is tailored to the specific quit date provided.
Challenge a friend to quit	Participants are encouraged to challenge a friend to quit with them. Participants are provided with a sample email that can be sent to a buddy through the Web app.
Expert chat (online quit coach)	Trained quit coaches, available 7 days a week to answer questions or provide advice through a live chat room.
Quitline (talk to a quit coach)	A toll-free one-on-one consultation with a trained quit coach designed to help participants prepare for a quit and stay smoke free. Following the initial contact, the quit coach schedules five follow-up calls.

Usability Testing

Prior to launching the pilot study, usability testing was conducted by usability experts (n=4) and with end-users (n=9) to obtain feedback on the navigation and functionality of the QuitNow Men website. We focused on uncovering usability issues related to well-known design principles, as well as those that could impact the end-user’s overall ability to navigate the website, understand the feedback received, and progress successfully through a range of resources. Four usability experts conducted an expert review in pairs and then conducted a “cognitive walkthrough” using two target population personas and scenarios: (1) a man working in an outdoor job who would be using a smartphone, and (2) a man working in a desk job who would be using a desktop computer. A cognitive walkthrough is a specific method of expert review, in which the reviewer completes a series of tasks that the interface is meant to support from the perspective of a target user [22]. Potential issues related to expected end-user actions and associated cues/feedback were identified and categorized as being issues related to either navigation, information architecture, or functionality. Issues were rated on a scale of 1-4 ranging from suggestions/opportunities for improvement to urgent usability problems. A total of 27 issues were identified, which developers addressed through revisions to the website resource.

End-user usability testing was then conducted with 9 men using iOS and Android smartphones, a laptop, and a desktop computer. The men were recruited using e-postings on social media. A think-aloud protocol was followed, using Morae to capture audio and video of device screens as men completed tasks related to the expected use of the system (eg, enter quit date, calculate how much smoking has cost you so far). Recordings were analyzed and violations of usability principles were again classified based on usability best practices [23]. The main issues were related to not recognizing that the QuitNow Men logo was the method to return to the home page (recognition not recall), not having a back button on some pages (user control and freedom), and mixed use of first and third person in the text (match between system and real world). After the tasks were completed, the majority of the men (8/9, 89%) also completed a short questionnaire evaluating the usability of the website. Every man agreed or strongly agreed that the website was easy to use, and the majority (7/8, 88%) felt very confident using the website. Half of the men agreed or strongly agreed with the statement, “I think that I would like to use this website frequently.” Participants also indicated a preference for content on the site to be phrased in the first person (eg, How much cash am I blowing on smoking? How dependent am I?), and the website was modified accordingly. The QuitNow Men website was then readied for pilot testing.

Pilot Test

Study Design and Sample

A one-group, pretest-posttest study design was used. Men were recruited using social media (Facebook and Twitter) and online classified advertisements (Kijiji, Craigslist, and Castanet) between June 26, 2014, and September 8, 2014. Eligible participants were male, lived in Canada, read and understood English, had a valid email address and Internet access, and currently smoked but were interested in quitting. Participants provided online informed consent prior to participation.

Study Procedures

All baseline and follow-up data were collected using online questionnaires administered through a secure, password-protected link. Upon completion of the baseline questionnaire, participants were provided with a unique ID and password to access the QuitNow Men website and asked to interact with the site over a 6-month period. Study participants were invited to complete follow-up questionnaires at 3 months and 6 months. Two subsequent emails were sent to non-responders at 1-week intervals reminding them to complete the questionnaire. Two intervention emails were sent out between the 3-month and 6-month follow-ups to remind participants to continue to use the website. Intervention emails included a humorous anecdote to encourage participants to return to the site as well as a link to the homepage. Participants received a maximum of CAN $50 for participating in the study ($25 gift card for completing the 3-month follow-up and $25 gift card for completing the 6-month follow-up).

Measures

Baseline Measures

Participants were asked about their current use of technology and preferences for connecting to the Internet, as well as previous use and interest in using online smoking cessation resources. Questions were included to gather data on smoking patterns, and dependence was measured using the Fagerstom Test for Nicotine Dependence [24], a standardized instrument for assessing intensity of physical addition to nicotine. Self-efficacy to avoid smoking temptation across various situations (eg, at a party, when under stress) was assessed using Velicer et al’s [25] instrument in its 9-item version [26]. Cronbach alpha for this measure was acceptable for this sample (alpha=.810). Demographic data were also collected including age, education level, and marital status.

Follow-Up Measures

Usability and Acceptability

Participants reported on the number of visits to the website between baseline and 3 months, and between 3-month and 6-month follow-up, with the following options: not at all, once, 2-3 times, 4-6 times, more than 6 times. For each resource they reported using, men were asked to rate its helpfulness using a 3-point scale (not helpful, somewhat helpful, very helpful). Acceptability of the website was measured by assessing level of agreement with 4 statements using a 5-point Likert response format (strongly disagree to strongly agree): “QuitNow Men is easy to use,” “The design and images used on QuitNow Men are appealing,” “I intend to continue to use QuitNow Men resource,” and “I will recommend QuitNow Men to other men who want to quit smoking.”

Smoking Behavior

Participants were asked if they currently smoked daily, occasionally, or not at all. Those who smoked daily or occasionally were asked about number of quit attempts lasting at least 24 hours in the past 3 months, number of cigarettes per day on the days that they smoked, readiness to quit, and confidence in ability to quit during the next month. Readiness to quit was measured using the Contemplation Ladder [27], which assesses a smoker’s position on a continuum ranging from no thought of quitting (0) to being engaged in taking action to quit (10). Finally, confidence in smoking cessation was measured using a single item that asked: “If you decided to quit smoking during the next month, how confident are you that you could do it?” Responses to this item were coded “not at all confident”=1, “not very confident”=2, “somewhat confident”=3, and “very confident”= 4.

Data Analysis

All data were analyzed using SPSS version 22. Descriptive statistics were used to describe website usage and participant evaluations of QuitNow Men. One-way analysis of variance (ANOVA) was conducted to determine if those who dropped out of the study significantly differed in their demographic profile from those who completed both 3-month and 6-month follow-up. Repeated-measures general linear model, assuming compound symmetry, was used to investigate within-subject differences between 3-month and 6-month follow-ups on number of cigarettes smoked, quit confidence, and readiness to quit score. One-way ANOVA was conducted to assess whether or not those who quit smoking, reduced smoking, or had no change in smoking status at 6 months differed on number of website resources used at 6-month follow up. Finally, multiple logistic regression was use to investigate predictors of the number of quit attempts at 6 months. Standardized predictor variables included number of resources used on the website (at 6-month follow-up), number of visits to the website, quit confidence (at 6-month follow-up), nicotine dependence, numbers of years smoking, and age.

Results

Sample

A total of 117 men completed the baseline survey of which 67 (57.3%) completed both the 3-month and 6-month follow-up survey. The 117 men had an average age of 39.82 (SD 11.08) with a range of 21-68 years. The majority of the men (78/117, 66.7%) were 45 years of age or younger. Sample demographic characteristics at baseline are provided in Table 2. At baseline, participants reported smoking for an average of 24 years (range 2-55, SD 12.12) and smoked on average 15 cigarettes per day (range 3-40, SD 7.4). Based on Fagerstrom scores, 48% were classified as minimally dependent, followed by 27% moderately dependent, and 25% highly dependent. The average smoking self-efficacy/temptation score was 2.7 out of a possible 5, where higher scores demonstrate lower levels of self-efficacy (higher temptation). Participants reported spending an average of 2.8 hours per day on the Internet, the majority connected to the Internet via home computer (105/117, 89.7%), followed by a smartphone (89/117, 76.1%). The majority of participants (103/117, 88.0%) had not previously used a quit smoking website or websites that focus on men's health (105/117, 89.7%). A statistically significant age difference was found between those who completed the follow-up surveys and those who dropped out ($F_{1,113}$=6.75, P=.01, eta squared=0.056). Those who dropped out were significantly older (mean 43.12, SD 11.53, range 23-68) than those who completed follow-up measures (mean 37.83, SD 10.18, range 21-58). As well, those who dropped out had been smoking for significantly longer than those who completed follow-up ($F_{1,113}$=8.16, P=.005, eta squared=0.067). Yet there were no significant differences on Fagerstrom scores (P=.84), average number of hours per day on the Internet (P=.28), number of quit attempts in the past 2 years (P=.67), and self-efficacy/temptation score (P=.70).

Table 2. Baseline demographics characteristics (N=117).

Characteristic	n	%
Age		
20-45 years	78	66.7
>45 years	39	33.3
Cultural identity		
Canadian/Caucasian	85	73
First Nations/Metis/Inuit	5	4
Other	17	14
Missing/Invalid	10	9
Highest level of education		
Incomplete high school	14	12
Complete high school	42	36
Complete non-university (vocational, technical, trade)	23	20
Complete university degree / diploma / certificate	34	29
Other	4	3
Marital status		
Married	40	35
Single	39	33
Common law/live-in partner	24	20
Divorced or separated	14	12
Main activity		
Working for pay or profit	71	61
Caring for family and working for pay or profit	19	16
Recovering from illness or disability	12	10
Looking for work	6	5
Going to school	5	4
Unemployed and not looking for work	2	2
Caring for family	1	1
Retired	1	1

Website Usage and Evaluation of QuitNow Men

Of the 67 men who completed both follow-up surveys, 6-month data showed that 64 men (96%) reported using the QuitNow Men website at least once during the pilot study. Nearly half of the users (29/64, 45%) reported using the website more than 6 times, while 10 (16%) used the website 4-6 times, 24 (38%) used it 2-3 times, and 1 man (2%) used it once. At 6-month follow-up, the majority of users (45/64, 70%) reported that they were either satisfied or very satisfied with the website. In addition, they agreed or strongly agreed that the website was easy to use (51/64, 80%), the design and images were appealing (42/64, 66%), they intended to continue to use the resource (42/64, 66%), and that they would recommend QuitNow Men to others who wanted to quit (46/64, 72%). Participants reported using an average of 8.76 (SD 4.08) of the 15 resources available on the website. Figure 1 details the percentage of participants who reported using each of the QuitNow Men resources during the 6 months. Figure 2 shows men’s ratings of the helpfulness of the resources that were used.

Figure 1. Percentage of participants reporting use of resources on QuitNow Men at 6 months (n=67).

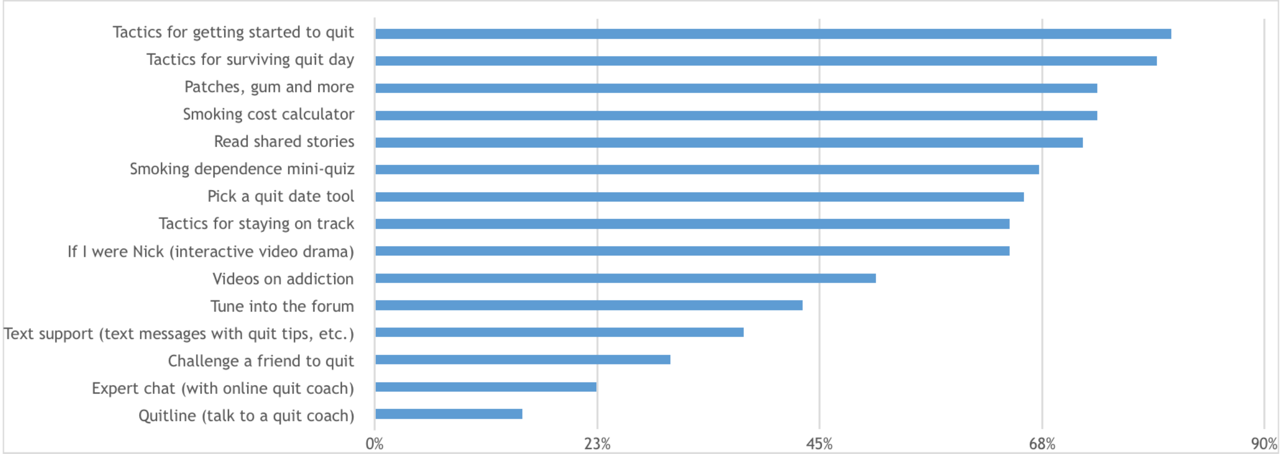
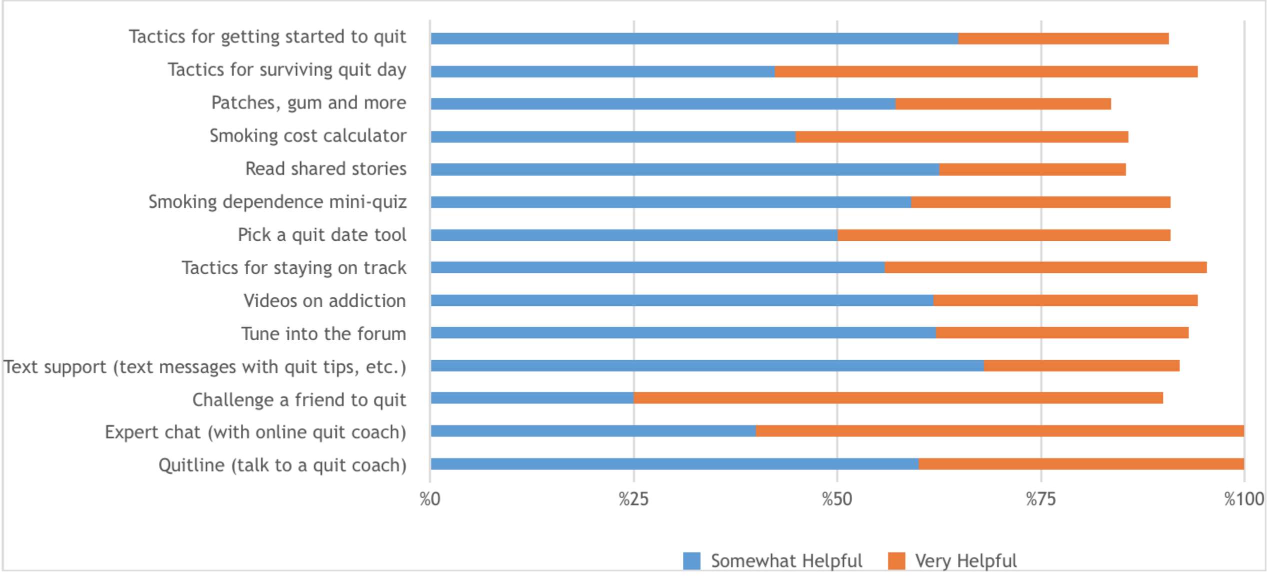


Figure 2. Percentage of participants rating resources they used on QuitNow Men as "somewhat helpful" or "very helpful" (n=64).



Use of QuitNow Men and Smoking Behavior

Self-reported smoking status at 6-month follow-up indicated that 16 of the 67 participants (24%) had quit, 27 (40%) had reduced their smoking, and 24 (36%) had not changed their smoking habits. Among those 51 men who were still smoking at 6-month follow-up, 44 (86%) reported that they had quit for 24 hours or longer since using the QuitNow Men website, with an average number of quit attempts of 3.38 (SD 2.33). The average readiness to quit score among those still smoking at 6-month follow-up was 7.61 out of a possible 10. The majority (39/51, 77%) were somewhat confident or very confident that they could quit smoking during the next month, with no significant differences between 3-month and 6-month follow-up. Repeated measures general linear model showed a significant reduction in the average number of cigarettes smoked between 3-month and 6-month follow-up ($F_{1,63}$=6.41, P=.01, eta squared=0.09). At 3-month follow-up, the average number of cigarettes smoked was 6.84 (95% CI 5.24-8.43) and at 6-month was 5.25 (95% CI 3.82-6.68).

Those who quit smoking, reduced smoking, or had no change in smoking status at 6 months differed on the total number of website resources used ($F_{2,64}$=3.34, P=.042, eta squared=0.094).

Post-hoc least significant difference analysis showed that men who reduced the number of cigarettes smoked used significantly more website resources than those that had no change between time points (*P*=.024). Means and confidence intervals can be seen in Table 3.

Table 3. Number of website resources used by smoking status.

Smoking status	Number of website resources used	
	Mean	95% CI
Quit smoking	7.25	5.52-8.98
Reduced smoking	10.0	8.37-11.63[a]
No change in smoking	7.31	5.20-9.43[a]

[a] *P*=.03.

Multiple logistic regression was conducted to identify if number of resources used on the website, number of website visits, quit confidence, nicotine dependence, and participant age predicted the number of quit attempts among men who were smoking at 6-month follow-up. Potential predictors were converted to Z scores. Multicollinearity diagnostics showed that the number of resources used on the website and website visits were highly correlated (r_{s51}=65.9, *P*<.001) resulting in unacceptably low tolerance levels (<.865; 1-R^2). It was determined that examining the number of resources used on the website would be more representative of examining the number of times the user visited the website and as such, website visits was removed from the regression (see Table 4). The combination of remaining variables significantly predicted number of quit attempts ($F_{4,45}$=2.73, *P*=.04). The adjusted R^2 value was 0.124. Number of website resources used was the strongest predictor of quit attempts (*P*=.02).

Table 4. Logistic regression predicting quit attempts (6-month follow-up).

Variable	B (CI)	SE	β	*P*
Number of resources used on website at 6-month follow-up	.81 (0.12-1.50)	0.34	.35	.02
Quit confidence at 6-month follow-up	.50 (1.2-0.20)	0.35	.21	.16
Nicotine dependence	.33 (-0.33 to 1.0)	0.33	.15	.32
Age	.58 (-0.14 to 1.31)	0.36	.23	.11
Constant	3.34 (2.71-3.96)	0.31		<.001

Discussion

Principal Findings

This work adds to the small but growing field of online health promotion innovations directed specifically to promote men's health. To our knowledge, QuitNow Men is the first men-centered smoking cessation website in the world, and the pilot study findings reveal that the site is appealing to men who want to quit and demonstrates potential as a self-guided smoking cessation resource. The majority of men reported they had quit smoking for 24 hours or longer since using the QuitNow Men website, and the number of resources used on the website and quit confidence predicted the number of quit attempts. At the 6-month follow-up, 24% reported having quit smoking. Although QuitNow Men was designed for adult men 45 years and younger, men over 45 years of age were interested in using this online resource suggesting that this approach to supporting smoking cessation may have broad appeal among men. Since older men were more likely to drop out of the study, it is possible QuitNow Men did not have the same appeal for this group given that images, text, etc, were designed for a younger audience. Nevertheless, given older men's interest in trying this online resource, it is worth considering how an online smoking cessation resource could be tailored to reflect the life stage and needs of this group and what additional supports they may require. It should also be noted that among those men that remained in the study, the oldest was 58 years of age and multiple regression results show that age was not a significant predictor of number of quit attempts made. Overall, in the broader context of men's utilization of eHealth apps, these results afford important empirical insights that could guide future online smoking cessation programs and provide direction to other eHealth resources aimed at reducing men's health risks (eg, abstaining from high-fat foods, addressing alcohol overuse).

In this study, the 64 men who completed both the 3-month and 6-month follow-up surveys and used the resources on QuitNow Men rated them as helpful. This suggests that the specific design elements used in the QuitNow Men website hold value for increasing and improving men's engagement in the core components of this self-guided intervention. For example, setting up the website to allow for quick exploration of a number of resources and strategies (as distinct from in-depth exploration) was a design element used to entice men to learn about a variety of approaches. In turn, the finding that exploring more resources increased men's likelihood of having made a quit attempt suggests that the variety of resources included on QuitNow Men may be a key factor in the overall positive response to the site. It is interesting to note two of the least-used resources (ie, expert chat and the quitline) were the highest rated in terms of helpfulness among men who used them. Although few men desired personal contact with experts either through chat sessions or the quitline, these supplementary delivery modes provide an important component of Internet smoking cessation programs for men given that they offer a degree of anonymity and convenience (in comparison to face-to-face counseling) and have been found to be effective in providing proactive support to smokers [28,29]. In addition, there is empirical support for the value of including multiple modes of delivery in Internet programs to promote health behavior change [30,31]. However, determining the optimal balance between self-management resources and interactive components, as well as the relationship of activity complexity to cessation rates, warrants further investigation with a larger sample.

More broadly, the findings drawn from this pilot study remind us that the Internet provides a suitable medium for men's health promotion and, more specifically, for delivering smoking cessation interventions to men. Tobacco reduction and cessation research has revealed strong evidence supporting therapist-led interventions, and intensive group and individual counseling interventions to assist with cessation [32-34]. These intensive interventions, however, are often dependent on primary healthcare professionals delivering or facilitating the sessions at a specific time and place, which is often inconvenient to individuals. As a result, the reach of this type of intervention is limited [35]. Although Internet-based smoking cessation interventions offer a low cost, accessible option, and have been found to be acceptable to users and effective in aiding cessation [36], we have found through experiences in our region that the majority of users of an online cessation resource for the general population are women. The QuitNow Men smoking cessation website is a promising and potentially powerful resource toward balancing this gender inequity and engaging men in taking actions to become smoke free. Determining the best means to promote the use of this novel program to male smokers will be important.

Limitations

The findings of this study should be considered in light of several limitations. The sample of men that participated in this study may not be representative of all male smokers. In addition, self-report measures (eg, with respect to smoking patterns, quit attempts) may have introduced recall and reporting bias. Smoking cessation outcomes were not biochemically verified. Nevertheless, the findings provide important estimates of outcomes that reflect the potential value of the QuitNow Men smoking cessation website and a basis for conducting a full-scale evaluation.

Conclusion

The results of this research support efforts to integrate gender-sensitive approaches in health promotion interventions. The QuitNow Men smoking cessation website is highly acceptable and engaging to men interested in reducing and quitting smoking. Results indicate that this novel resource, tailored to men, has potential to support and perhaps catalyze men's smoking cessation efforts. Furthermore, given the tailored nature of the QuitNow Men smoking cessation website and that it caters to the specific values of a particular population (ie, men) [37-39], it has the potential to attract, engage, and retain men interested in quitting smoking.

Acknowledgments

This research was funded by the Canadian Cancer Society (Grant #701615) and conducted in collaboration with the BC Lung Association's and QuitNow BC. The authors would also like to recognize other members of the QuitNow Men research team for their collective contributions: Jack Boomer and Sharon Hammond from the BC Lung Association's QuitNow Services and Context Research, and Blye Frank from the University of British Columbia.

Conflicts of Interest

None declared.

References

1. Lee YA, Hashibe M. Tobacco, alcohol, and cancer in low and high income countries. Ann Glob Health 2014;80(5):378-383 [FREE Full text] [doi: 10.1016/j.aogh.2014.09.010] [Medline: 25512153]
2. Schroeder SA. New evidence that cigarette smoking remains the most important health hazard. N Engl J Med 2013 Jan 24;368(4):389-390. [doi: 10.1056/NEJMe1213751] [Medline: 23343069]
3. Giovino GA, Mirza SA, Samet JM, Gupta PC, Jarvis MJ, Bhala N, et al. Tobacco use in 3 billion individuals from 16 countries: an analysis of nationally representative cross-sectional household surveys. Lancet 2012 Aug 18;380(9842):668-679. [doi: 10.1016/S0140-6736(12)61085-X] [Medline: 22901888]
4. Asma S, Mackay J, Song S, Zhao L, Morto J, Palipudi K. The GATS Atlas. 2015. URL: http://gatsatlas.org/ [accessed 2015-09-02] [WebCite Cache ID 6bEOMiwKz]
5. Health C. Canadian Tobacco Use Monitoring Survey (CTUMS). 2012. URL: http://www.hc-sc.gc.ca/hc-ps/tobac-tabac/research-recherche/stat/_ctums-esutc_2012/ann_summary-sommaire-eng.php [accessed 2015-09-02] [WebCite Cache ID 6bEOlUE7r]
6. Okoli CT, Torchalla I, Oliffe JL, Bottorff JL. Men's smoking cessation interventions: a brief review. Journal of Men's Health 2011 May;8(2):100-108. [doi: 10.1016/j.jomh.2011.03.003]

7. Bottorff JL, Haines-Saah R, Kelly MT, Oliffe JL, Torchalla I, Poole N, et al. Gender, smoking and tobacco reduction and cessation: a scoping review. Int J Equity Health 2014;13:114 [FREE Full text] [doi: 10.1186/s12939-014-0114-2] [Medline: 25495141]
8. Pollard J. Working with men via the internet. In: White A, Pettifer M, editors. Hazardous waist: tackling male weight problems. Oxford, UK: Radcliffe; 2007:186-197.
9. Robinson M, Robertson S. Young men's health promotion and new information communication technologies: illuminating the issues and research agendas. Health Promot Int 2010 Sep;25(3):363-370 [FREE Full text] [doi: 10.1093/heapro/daq022] [Medline: 20427372]
10. Myung S, McDonnell DD, Kazinets G, Seo HG, Moskowitz JM. Effects of Web- and computer-based smoking cessation programs: meta-analysis of randomized controlled trials. Arch Intern Med 2009 May 25;169(10):929-937. [doi: 10.1001/archinternmed.2009.109] [Medline: 19468084]
11. Civljak M, Stead LF, Hartmann-Boyce J, Sheikh A, Car J. Internet-based interventions for smoking cessation. Cochrane Database Syst Rev 2013;7:CD007078. [doi: 10.1002/14651858.CD007078.pub4] [Medline: 23839868]
12. Shahab L, McEwen A. Online support for smoking cessation: a systematic review of the literature. Addiction 2009 Nov;104(11):1792-1804. [doi: 10.1111/j.1360-0443.2009.02710.x] [Medline: 19832783]
13. Lohan M, Aventin A, Oliffe JL, Han CS, Bottorff JL. Knowledge translation in men's health research: development and delivery of content for use online. J Med Internet Res 2015;17(1):e31 [FREE Full text] [doi: 10.2196/jmir.3881] [Medline: 25642787]
14. Tong SF, Khoo EM, Low WY, Ng CJ, Wong CH, Yusoff HM, et al. Health innovation project: A concept paper on a virtual health promotion program for men. Journal of Men's Health 2014 Mar;11(1):4-9. [doi: 10.1089/jomh.2013.0020]
15. Tyler RE, Williams S. Masculinity in young men's health: exploring health, help-seeking and health service use in an online environment. J Health Psychol 2014 Apr;19(4):457-470. [doi: 10.1177/1359105312473784] [Medline: 23493865]
16. Thackeray R, Neiger BL, Hanson CL, McKenzie JF. Enhancing promotional strategies within social marketing programs: use of Web 2.0 social media. Health Promot Pract 2008 Oct;9(4):338-343. [doi: 10.1177/1524839908325335] [Medline: 18936268]
17. Kerns J. Fred Hutch going into tech business with 'SmartQuit' smoking cessation app. 2014. URL: http://mynorthwest.com/11/2509848/Fred-Hutch-going-into-tech-business-with-SmartQuit-smoking-cessation-app [accessed 2015-06-01] [WebCite Cache ID 6YxrzmWkq]
18. BinDhim NF, McGeechan K, Trevena L. Who uses smoking cessation apps? A feasibility study across three countries via smartphones. JMIR Mhealth Uhealth 2014;2(1):e4 [FREE Full text] [doi: 10.2196/mhealth.2841] [Medline: 25098439]
19. Graham AL, Papandonatos GD, Cobb CO, Cobb NK, Niaura RS, Abrams DB, et al. Internet and telephone treatment for smoking cessation: mediators and moderators of short-term abstinence. Nicotine Tob Res 2015 Mar;17(3):299-308 [FREE Full text] [doi: 10.1093/ntr/ntu144] [Medline: 25156528]
20. Bottorff JL, Oliffe JL, Sarbit G, Sharp P, Kelly M. Smoke free men: completing and connecting to quit. Am J Health Promot 2016:1-21 (forthcoming).
21. Bottorff JL, Oliffe JL, Sarbit G, Kelly MT, Cloherty A. Men's responses to online Smoking cessation resources for new fathers: The influence of masculinities. JMIR Res Protoc 2015;4(2):e54 [FREE Full text] [doi: 10.2196/resprot.4079] [Medline: 25971202]
22. Wharton C, Rieman J, Lewis C, Polson P. The cognitive walkthrough method: a practitioner's guide. In: Nielson J, Mack RL, editors. Usability inspection methods. New York: Wiley; 1994:105-140.
23. Johnson JG. Designing with the mind in mind: a simple guide to understanding user interface design guidelines. Amsterdam: Elsevier; 2014.
24. Heatherton TF, Kozlowski LT, Frecker RC, Fagerström KO. The Fagerström Test for Nicotine Dependence: a revision of the Fagerström Tolerance Questionnaire. Br J Addict 1991 Sep;86(9):1119-1127. [Medline: 1932883]
25. Velicer WF, Diclemente CC, Rossi JS, Prochaska JO. Relapse situations and self-efficacy: an integrative model. Addict Behav 1990;15(3):271-283. [Medline: 2378287]
26. Cancer Prevention Research Centre. Smoking: self-efficacy/temptation (short form). University of Rhode Island; 2001. URL: http://web.uri.edu/cprc/smoking-self-efficacy-temptation-short-form/ [accessed 2016-04-04] [WebCite Cache ID 6giQbIEzc]
27. Biener L, Abrams DB. The Contemplation Ladder: validation of a measure of readiness to consider smoking cessation. Health Psychol 1991;10(5):360-365. [Medline: 1935872]
28. Newman MG, Szkodny LE, Llera SJ, Przeworski A. A review of technology-assisted self-help and minimal contact therapies for drug and alcohol abuse and smoking addiction: is human contact necessary for therapeutic efficacy? Clin Psychol Rev 2011 Feb;31(1):178-186. [doi: 10.1016/j.cpr.2010.10.002] [Medline: 21095051]
29. Stead L, Hartmann-Boyce J, Perera R, Lancaster T. Telephone counselling for smoking cessation. Cochrane Database of Systematic Reviews 2013 Aug 12(8):1-134 [FREE Full text] [doi: 10.1002/14651858.CD002850.pub3]
30. 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]

31. Graham AL, Papandonatos GD, Cobb CO, Cobb NK, Niaura RS, Abrams DB, et al. Internet and telephone treatment for smoking cessation: mediators and moderators of short-term abstinence. Nicotine Tob Res 2015 Mar;17(3):299-308 [FREE Full text] [doi: 10.1093/ntr/ntu144] [Medline: 25156528]
32. Stead LF, Lancaster T. Group behaviour therapy programmes for smoking cessation. Cochrane Database Syst Rev 2005(2):CD001007. [doi: 10.1002/14651858.CD001007.pub2] [Medline: 15846610]
33. Lancaster T, Stead LF. Individual behavioural counselling for smoking cessation. Cochrane Database Syst Rev 2005(2):CD001292. [doi: 10.1002/14651858.CD001292.pub2] [Medline: 15846616]
34. Stead LF, Bergson G, Lancaster T. Physician advice for smoking cessation. Cochrane Database Syst Rev 2008(2):CD000165. [doi: 10.1002/14651858.CD000165.pub3] [Medline: 18425860]
35. Civljak M, Sheikh A, Stead LF, Car J. Internet-based interventions for smoking cessation. Cochrane Database Syst Rev 2010(9):CD007078. [doi: 10.1002/14651858.CD007078.pub3] [Medline: 20824856]
36. Shahab L, McEwen A. Online support for smoking cessation: a systematic review of the literature. Addiction 2009 Nov;104(11):1792-1804. [doi: 10.1111/j.1360-0443.2009.02710.x] [Medline: 19832783]
37. Dijkstra A. Working mechanisms of computer-tailored health education: evidence from smoking cessation. Health Educ Res 2005 Oct;20(5):527-539 [FREE Full text] [doi: 10.1093/her/cyh014] [Medline: 15701665]
38. Kreuter MW, Wray RJ. Tailored and targeted health communication: strategies for enhancing information relevance. Am J Health Behav 2003;27 Suppl 3:S227-S232. [Medline: 14672383]
39. Stanczyk N, Bolman C, van AM, Candel M, Muris J, de VH. Comparison of text and video computer-tailored interventions for smoking cessation: randomized controlled trial. J Med Internet Res 2014;16(3):e69 [FREE Full text] [doi: 10.2196/jmir.3016] [Medline: 24589938]

Edited by G Eysenbach; submitted 18.11.15; peer-reviewed by V Ramirez, B Ansa, W Yoo; comments to author 17.12.15; revised version received 16.02.16; accepted 20.02.16; published 20.04.16.

Please cite as:
Bottorff JL, Oliffe JL, Sarbit G, Sharp P, Caperchione CM, Currie LM, Schmid J, Mackay MH, Stolp S
Evaluation of QuitNow Men: An Online, Men-Centered Smoking Cessation Intervention
J Med Internet Res 2016;18(4):e83
URL: http://www.jmir.org/2016/4/e83/
doi:10.2196/jmir.5076
PMID:27097991

Original Paper

Effectiveness of Social Media Interventions for People With Schizophrenia: A Systematic Review and Meta-Analysis

Maritta Välimäki[1,2], Professor; Christina Athanasopoulou[1], PhD(c); Mari Lahti[1,3], PhD; Clive E Adams[4], Professor

[1]Department of Nursing Science, University of Turku, Turku, Finland
[2]Turku University Hospital, Turku, Finland
[3]Faculty of Health and Well-being, Turku University of Applied Sciences, Turku, Finland
[4]Institute of Mental Health, Division of Psychiatry, University of Nottingham, Nottingham, United Kingdom

Corresponding Author:
Maritta Välimäki, Professor
Department of Nursing Science
University of Turku
Lemminkäisenkatu 1
Turku, 20014
Finland
Phone: 358 02 333 8495
Fax: 358 02 333 6363
Email: mava@utu.fi

Abstract

Background: Recent studies have shown that people with serious mental disorders spend time online for the purposes of disclosure, information gathering, or gaming. However, coherent information on the effects of social media on treatment for people with schizophrenia is still lacking.

Objective: Our aim was to determine the effects of social media interventions for supporting mental health and well-being among people with schizophrenia.

Methods: A systematic review and meta-analysis were undertaken to determine the effects of social media interventions for supporting mental health and well-being among people with schizophrenia. Ten databases were searched, while search parameters included English-only manuscripts published prior to June 25, 2015. Study appraisals were made independently by 2 reviewers, and qualitative and quantitative syntheses of data were conducted.

Results: Out of 1043 identified records, only two randomized studies of moderate quality (three records, total N=331, duration 12 months) met the inclusion criteria. Participants were people with schizophrenia spectrum or an affective disorder. Social media was used as part of Web-based psychoeducation, or as online peer support (listserv and bulletin board). Outcome measures included perceived stress, social support, and disease-related distress. At 3 months, participants with schizophrenia in the intervention group reported lower perceived stress levels (P=.04) and showed a trend for a higher perceived level of social support (P=.06). However, those who reported more positive experiences with the peer support group also reported higher levels of psychological distress (P=.01).

Conclusions: Despite using comprehensive searches from 10 databases, we found only two studies, whereas numerous reports have been published citing the benefits of social media in mental health. Findings suggest the effects of social media interventions are largely unknown. More research is needed to understand the effects of social media, for users with and without mental illness, in order to determine the impact on mental well-being ofsocial media use as well as its risks.

(J Med Internet Res 2016;18(4):e92) doi:10.2196/jmir.5385

KEYWORDS

social media; effectiveness; technology; Internet; Web 2.0; schizophrenia; mental health

Introduction

People with severe mental illness, including schizophrenia, have symptoms that may seriously affect them for life. Evidence-based recommendations of current schizophrenia treatment guidelines include approaches such as medication use, psychotherapy, and family psychoeducational interventions [1]. For those with serious mental illness, online interventions have been shown to have the potential to disseminate care, support the patient's participation in group interactions [2], and serve as an alternative for individuals [3]. Combining online psychoeducation with various other online tools, including social media, is becoming more popular, and online psychotherapy and videogames have been evaluated for validity and efficacy [4].

Recently, the online world has been experiencing a huge growth of user-generated content, which has increased the acceptance of social media [5]. Although doubts in professional discourse about Internet use [6] or the use of social media in treatment and prevention exist—mostly concerning challenges regarding controlling user behavior, accurately assessing risks, and privacy and confidentiality issues [7]—the use of online communities is increasing constantly. More than 64% of Internet users access social media services online, while social networking is one of the most popular ways to contact friends and family [8]. Currently, social media users total approximately 2.22 billion [9], and 31% of all Internet users spend their time specifically on social networking sites [10]. Facebook is the most popular social networking site, with close to 1.6 billion active users as of the fourth quarter of 2015 [11]. It has been estimated that in 2017 there will be around 2.39 billion online social network users [9]. Given the high penetration rates of social media, it is important to acquire a deeper understanding of the applications of social media in health care [12].

Investigating the trend of using social networking interventions is important, as peer-support services are viewed as best practice based solid theory and are supported by policy makers [13-15]. Lal and Adair's rapid review [16] indicates that many find an enormous potential for e-mental health to help address the gap between the identified need for services and the limited capacity of resources to provide conventional treatment. Numerous social networking sites have already been developed to change behavior and improve health outcomes [17]. Systematic reviews of online interventions have also been conducted, focusing on, for example, online peer support [18], Internet support groups for depression [19], online communication, social media and adolescent well-being [20], social media to address Asian immigrants' mental health needs [21], online and social networking for the treatment of depression in youth [22], and social media and suicide prevention [7]. Although the results seem promising, the reviews share a concern about poor designs of the original studies with underpowered samples [18,19,22], mixed findings [23], or a lack of intervention studies [7]. A Cochrane review by Välimäki et al [24] also stresses the importance of high-quality studies in assessing the effects of novel interventions, in particular, the case of virtual reality for people with serious mental illness.

There are few studies that have produced coherent knowledge of the effects of social media use in treatment [7]. Research on the effects of online support for people with mental illnesses is even more scarce [25], especially when the illness is severe. Although online social media and mobile technology have shown some promise in improving positive psychotic symptoms, hospital admissions, socialization, social connectedness, depression, and medication adherence of people with psychosis, the issues of heterogeneity, poor study quality, and the early state of current research preclude any definite conclusions [26]. As far as we are aware, there have been no systematic reviews and meta-analyses to date of social media interventions for people with schizophrenia or similar disorders. Since people with psychosis spend more time in chat rooms or playing online games than other people [27], and online social networking can be used for establishing or maintaining relationships or reconnecting with people and online peer support [27], this review is important, as it presents a quantitative analysis describing relevant interventions and seeks to determine the effects of social media for supporting mental health and well-being in this group.

Methods

This systematic review was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-analysis (PRISMA) [28]. The data extraction was in line with the CONSORT-EHEALTH checklist [29]. Social media interventions were extracted by using the Template for Intervention Description and Replication (TIDieR) checklist and guide [30].

Search Methodology

Ten databases (PubMed, MEDLINE, Cochrane Database of Systematic Reviews, PsycInfo, CINAHL, JBI, Scopus, ISI Web of Science, SOCIndex, Sociological abstracts) were searched for potentially relevant abstracts. These databases cover a wide range of published research from the field of health and social care. Search parameters included English-only manuscripts published (or in press) prior to June 25, 2015.

The search terms (or equivalent index terms and free-text words) for each of the databases were used to ensure a broad coverage of studies in our review. For example, search terms for schizophrenia included "schizophreni" OR "schizoaffective disorder" OR "schizophrenia" OR "schizophrenia-like illness" OR "schizo-affective disorder" OR "severe mental illness" OR "severe mental illnesses" OR "serious mental illness" OR "schizophrenia and disorders with psychotic features". Search terms for social media included "social media" OR "Internet" OR "world wide web applications" OR "blogging" OR "blog" OR "wiki" OR "facebook" OR "twitter" OR "youtube" OR "Instagram" OR "web 2.0" OR "chat" OR "chats" OR "chatting." The detailed terms for each database were searched by the information specialist at the University of Turku. Search terms for each database are presented in Table 1.

Table 1. Databases, search terms, and references found on June 25, 2015 (N=1043).

Database	Search terms	N
PubMed	(schizophreni* OR schizoaffective disorder* OR “severe mental illness” OR “severe mental illnesses” OR “serious mental illness” OR “serious mental illnesses” OR “schizophrenia-like illnesses” OR “schizophrenia-like illness” OR “Schizo-affective disorder” OR “Schizo-affective disorders” OR “Schizophrenia and Disorders with Psychotic Features”[Mesh]) AND (“Blogging”[Mesh] OR “Social Media”[Mesh] OR “Internet”[Mesh] OR Social media* OR Wiki* OR Facebook* OR Twitter * OR Youtube* OR Instagram* OR web 2.0 OR blogging OR blog* OR chat OR chats OR chatting*)	324
Ovid Medline	((schizophreni* or schizoaffective disorder* or severe mental illness* or serious mental illness* or schizophrenia-like illness* or Schizo-affective disorder*).mp. or exp Schizophrenia/) and (exp Blogging/ or exp social media/ or exp Internet/ or (Social media* or Wiki* or Facebook* or Twitter* or Youtube* or Instagram* or web 2* or blogging or blog* or chat*).mp.)	166
JBI	(schizophreni* or schizoaffective disorder* or severe mental illness* or serious mental illness* or schizophrenia-like illness* or Schizo-affective disorder*).mp. and (Social media* or Wiki* or Facebook* or Twitter* or Youtube* or Instagram* or web 2* or blogging or blog* or chat*).mp.	8
CINAHL	((MH “Schizophrenia+”) OR schizophreni* OR “schizoaffective disorder*” OR “severe mental illness*” OR “serious mental illness*” OR “schizophrenia-like illness*” OR “Schizo-affective disorder*”) AND ((MH “Social Media”) OR (MH “World Wide Web Applications+”) OR (MH “Instant Messaging”) OR (MH “Blogs”) OR “Social media*” OR Wiki* OR Facebook* OR Twitter* OR Youtube* OR Instagram* OR “web 2.0” OR blogging OR blog* OR chat*)	8
Cochrane	(schizophreni* or schizoaffective NEXT disorder* or severe NEXT mental NEXT illness* or serious NEXT mental NEXT illness* or schizophrenia NEXT like NEXT illness* or Schizo NEXT affective NEXT disorder*) and (Social NEXT media* or Wiki* or Facebook* or Twitter* or Youtube* or Instagram* or web NEXT 2* or blogging or blog* or chat*)	24
PsycInfo	(SU.EXACT.EXPLODE(“Schizophrenia”) OR schizophreni* OR “schizoaffective disorder*” OR “severe mental illness*” OR “serious mental illness*” OR “schizophrenia-like illness*” OR “Schizo-affective disorder*”) AND (SU.EXACT.EXPLODE(“Social Media”) OR “Social media*” OR Wiki* OR Facebook* OR Twitter* OR Youtube* OR Instagram* OR “web 2.0” OR blogging OR blog* OR chat*)	276
Web of Science	(schizophreni* OR “schizoaffective disorder*” OR “severe mental illness*” OR “serious mental illness*” OR “schizophrenia-like illness*” OR “Schizo-affective disorder*”) AND (“Social media*” OR Wiki* OR Facebook* OR Twitter* OR Youtube* OR Instagram* OR “web 2*” OR blog* OR chat*)	121
Scopus	(schizophreni* OR “schizoaffective disorder*” OR “severe mental illness*” OR “serious mental illness*” OR “schizophrenia like illness*” OR “Schizo affective disorder*”) AND (“Social media*” OR Wiki* OR Facebook* OR Twitter* OR Youtube* OR Instagram* OR “web 2.0” OR blog* OR chat*)	84
SOCIndex	(DE “SCHIZOPHRENIA”OR schizophreni* OR “schizoaffective disorder*” OR “severe mental illness*” OR “serious mental illness*” OR “schizophrenia-like illness*” OR “Schizo-affective disorder*”) AND (DE “SOCIAL media” OR DE “BACKCHANNELS (Social media)” OR DE “BLOGS” OR DE “COMPUTER bulletin boards” OR DE “ONLINE chat” OR DE “SOCIAL bookmarks” OR DE “WEB 2.0” OR OR DE “ONLINE comments” OR DE “ELECTRONIC discussion groups” OR “Social media*” OR Wiki* OR Facebook* OR Twitter* OR Youtube* OR Instagram* OR “web 2.0” OR blogging OR blog* OR chat*)	12
Sociological abstract	(SU.EXACT.EXPLODE(“Schizophrenia”) OR SU.EXACT.EXPLODE(“Paranoia” OR “Psychosis” OR “Schizophrenia”) OR schizophreni* OR “schizoaffective disorder*” OR “severe mental illness*” OR “serious mental illness*” OR “schizophrenia-like illness*” OR “Schizo-affective disorder*”) AND (SU.EXACT(“Computer Mediated Communication”) OR SU.EXACT.EXPLODE(“Internet”) OR “Social media*” OR Wiki* OR Facebook* OR Twitter* OR Youtube* OR Instagram* OR “web 2.0” OR blogging OR blog* OR chat*)	20

Some differences between databases and search words used are due to available thesaurus terms in the specific databases, that is, the descriptor/thesaurus term “electronic discussion groups” is used only in SOCIndex, in an attempt to translate the Medical Subject Headings (MeSH) terms used.

Inclusion and Exclusion Criteria

The review was limited to texts published in English with an abstract available (published on or before June 25, 2015). Participants were people with schizophrenia spectrum disorders. If the study included other participants, data from those people were included only if reported separately. The review topic was limited to studies concerning interactivity and social media. Interactivity refers to user-to-user contact [31], such as patients with peers, staff members, or their nearest or public social media. Social media was understood as a broader term including collaborative projects (eg, user-generated content, content communities, content sharing, and social online networking sites) [5,32-34], social networking (eg, the broader concepts of Health 2.0 and Medicine 2.0) [35], or interventions involving Facebook, Twitter, YouTube, Instagram, blogs, Wiki, chats, the Internet, or Web 2.0. The concept was used in the health care domain and targeted adult persons with various schizophrenia spectrum diagnoses. If the study included adolescents, it was included only if the mean age of the participants was over 30 years. Only peer-reviewed, published papers with randomized clinical trials were included.

Studies were excluded if the information and communication technology was used only in interventions for data collection purposes (eg, online surveys, electronic medical records), for patient education without any online social networking or to print the paper material for participants. Further, papers describing the design or the development process of the social media intervention, theoretical or methodological papers, books or book chapters, letters, dissertations, editorials, or study protocols were excluded.

In the case of multiple publications from the same study, we combined data to avoid double counting. When required, we contacted the corresponding author to acquire more detailed data if the data we were interested in were not available in the publication.

Study Identification

Out of 1043 hits, duplicates across all databases were removed, leaving us with 727 abstracts. First, 2 authors (MV, ML) screened titles and abstracts independently for eligibility; ineligible hits were excluded (n=720). Second, we found 4 additional papers when an additional hand search of lists of references was conducted. Third, 11 full papers were obtained and screened by the 2 authors for the inclusion and exclusion criteria. Finally, eight papers had to be excluded (see Table 2). The systematic review and meta-analysis were conducted on the three retrieved papers (two studies). In cases of discrepancy concerning the decisions made between the reviewers, the papers were discussed until a consensus was reached with the support of CA. Figure 1 outlines the search process of the literature, according to PRISMA [28].

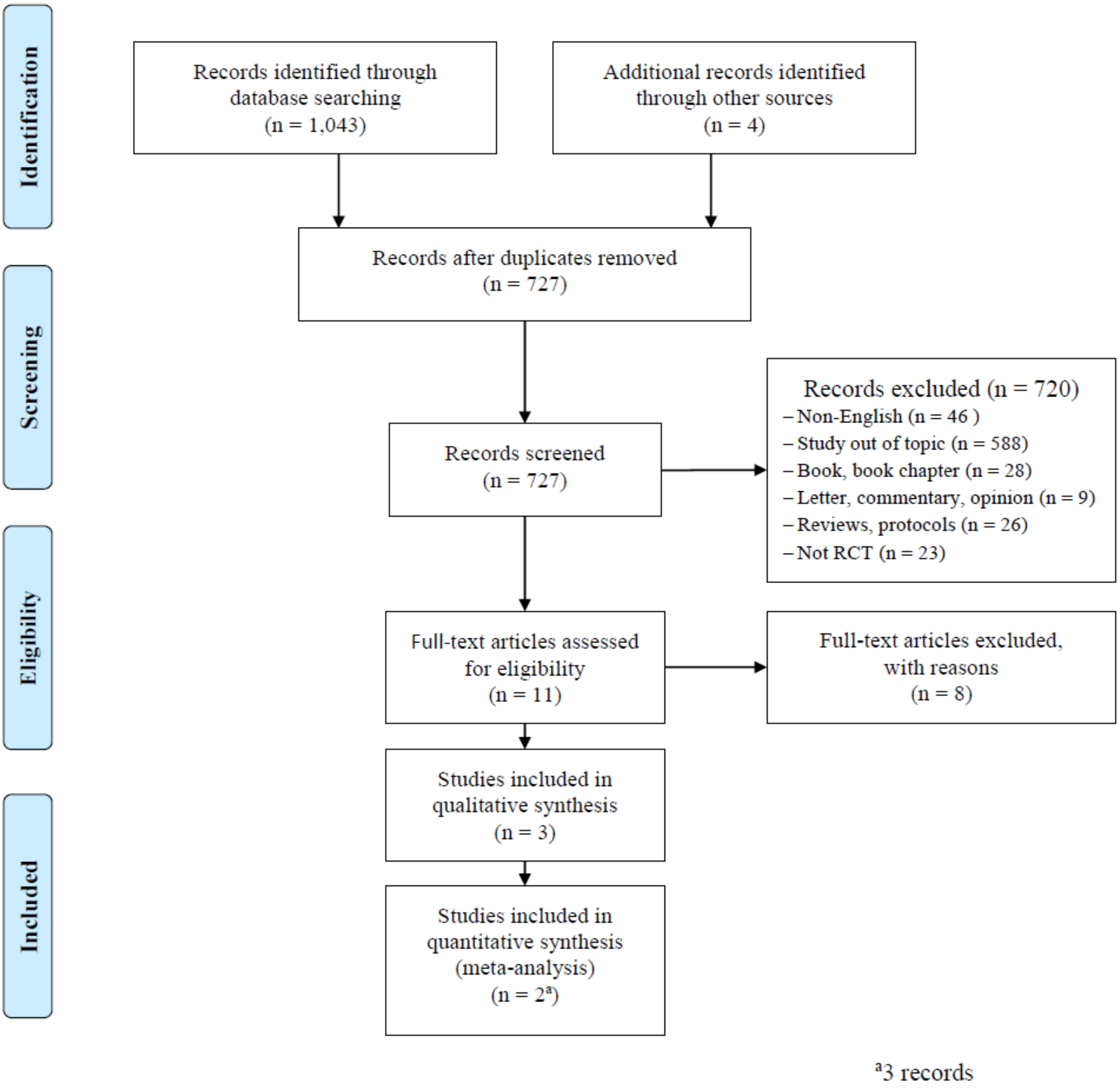
Excluded studies (n=8) are categorized based on Higgins [36]. The specific reasons for exclusion are described in Table 2 [37-44].

Table 2. Description of excluded studies.

Study	Description	Reason for being excluded[a]
Kilbourne et al 2013 [37]	Cluster randomized adaptive implementation trial comparing a standard versus enhanced implementation intervention to improve uptake of an effective re-engagement program for people with serious mental illness	Allocation: cluster randomization Participants: people with serious mental illness Intervention: no social media
Kim & Stout 2010 [38]	The effects of interactivity on information processing and attitude change: implications for mental health stigma	Allocation: non-randomized Participants: undergraduate students Intervention: no social media
McFarlane et al 1995 [39]	Multiple-Family Groups and Psychoeducation in the Treatment of Schizophrenia	Allocation: randomized Participants: people with schizophrenia, schizoaffective disorder, or schizophrenia form disorder (DSM-IH-R) Intervention: no social media
Sicilia et al 2005 [40]	Effects of interactivity in a website—The moderating effect of need for cognition	Allocation: non-randomized Participants: consumers Intervention: no social media
Spinzy et al 2012 [41]	Does the Internet offer social opportunities for individuals with schizophrenia? A cross sectional pilot study	Allocation: non-randomized Participants: people with psychotic disorders and affective disorder, anxiety disorders F20-F48 (ICD-10) Intervention: no social media
Steinwachs et al 2011 [42]	A Web-based program to empower people with schizophrenia to discuss quality of care with mental health providers	Allocation: non-randomized Participants: people with schizophrenia (no classification code specified, ICD-10 or DSM-V) Intervention: no social media
van der Krieke et al 2012 [43]	Usability evaluation of a Web-based support system for people with schizophrenia diagnosis	Allocation: non-randomized Participants: people with schizophrenia or a related psychotic disorder (eg, schizo-affective disorder, schizophreniform disorder, schizotypal disorder) (no classification code specified, ICD-10 or DSM-V) Intervention: no social media
van der Krieke et al 2013 [44]	A Web-based tool to support shared decision making for people with a psychotic disorder: Randomized controlled trial and process evaluation	Allocation: randomized Participants: people with non-affective psychosis (DSM-IV) Intervention: no social media

[a] ICD-10=International Classification of Diseases, 10th Revision; DSM-IV=Diagnostic and Statistical Manual of Mental Disorders, 4th Edition.

Figure 1. PRISMA flow diagram outlining the review process.



Data Extraction and Analyses

We created a specific data extraction matrix to collect information. Where possible, data extraction was based on the CONSORT-EHEALTH checklist [29]. First, characteristics of the studies were described (country of origin, purpose of the study, setting of patient recruitment, patient group, inclusion criteria, number of patients enrolled, follow-up period, and number of drop-outs). Second, social media interventions were extracted by using the TIDieR checklist [30] and guide. The matrix used in our extraction was based on a 12-item checklist and was constructed with the following information: brief name, why (rationale or theory), what (materials, procedures), who provided intervention, how, where, when and how much, tailoring, modifications, and how well (planned, actual) [30]. Third, outcomes and instruments used in each study were identified. Fourth, reasons for exclusion of studies were described.

For narrative analysis, data on each included study were entered into the data extraction matrix; each study was treated as a separate case. Descriptive characteristics of the studies were categorized manually. In addition, social media usage and stress after social media use were described. Quantitative analysis was undertaken using the Review Manager 5.3 [45], which is the software used for preparing and maintaining Cochrane Reviews. The data analysis was divided into two phases. First, the analysis explored the descriptive characteristics of the individual studies included. Second, for continuous outcomes we estimated the mean difference between groups. When similar scales, such as symptom checklists, were used, we presumed there was a small difference in measurement tools and combined the measurements. This decision was made to answer the overall question of whether there is evidence that social media can be an effective intervention among people with schizophrenia [36]. In this approach, standard deviations were used together with the sample sizes to compute the weight given to each study. Random effect was used instead of fixed effect because random effect allows the outcomes of studies to vary more than fixed effects. In other words, random effects can be seen to be a more natural way of explaining outcomes [46]. Heterogeneity was checked by calculating I^2statistic. Where the I^2estimated was greater than, or equal to 50%, it was interpreted as indicating the presence of high levels of heterogeneity [36].

Quality Assessment

Quality of the included studies was assessed based on the Cochrane Collaboration's tool for assessing risk of bias [36]. This tool assesses the possible bias in randomized controlled trials including random sequence generation, allocation concealment, blinding of participants and blinding of outcome assessment, incomplete outcome data, selective reporting, and other possible biases. The assessment was undertaken independently by 2 reviewers (ML, MV). Again, in cases of disagreement, differing views were discussed, decisions were documented decisions, and if necessary, authors of studies were contacted for clarification.

The data were extracted from all included studies. Data presented only in graphs and figures were extracted. In case of any discrepancy, the solution was based on discussion with the third author (CA).

Results

Characteristics of the Studies

Both studies were conducted in the United States. Participants were recruited from out-patient facilities [47,48] or using websites and e-newslists [25]. Rotondi's study [47,48] included people with schizophrenia or schizoaffective disorder (Diagnostic and Statistical Manual of Mental Disorders IV criteria), with a mean age of 38 years (SD 11). Ten participants were male (32%), and 15 were white (48%). In the study by Kaplan et al [25], the participants were diagnosed with schizophrenia spectrum or affective disorder, their mean age was 47 years, one third were male (n=102), and 87% were white (n=260). Information about the studies' characteristics and methodology are presented in Table 3 [25,47,48].

Intervention Characteristics

Interventions included interactivity and social media. The participants in the intervention group used peer-support forums [25,48], while the participants in the control group [48] underwent treatment as usual or were under waiting list control conditions [25]. In Rotondi at al's study [48], "the usual care" was not further described. In Kaplan's study [25], those who were assigned to the waitlist control condition were asked to refrain from using all forms of Internet peer support for the duration of their participation in the study. Interventions are described in detail in Table 4 [25,48].

Outcome Measures

The studies included specific outcomes: symptoms, knowledge level, recovery, quality of life, empowerment, social support, or stress. A variety of instruments were used in each study. The outcomes and instruments used in the studies are listed in Table 5 [47-58].

In addition, in Kaplan et al's study [25], participants' experiences with the online group were assessed with the 7-item measure (Online Group Questionnaire [OGQ], adapted from Chang et al [59]). The OGQ contained questions about how often they thought the discussion topic was relevant, whether they felt supported by the group members, or they were satisfied with the group overall.

Table 3. Study characteristics.

Categories	Rotondi et al 2005, 2010 [47,48]	Kaplan et al 2011 [25]
Country of origin	United States	United States
Purpose of the study	To examine use of websites and home computers to deliver online multifamily psychoeducational therapy to people with schizophrenia (and their informal supports).	To test the effects of unmoderated, unstructured Internet peer support for people with psychiatric disabilities.
Setting of patient recruitment	Community mental health centers and inpatient units.	Websites, e-newslists, study advertisements via mental health provider agencies and programs.
Inclusion criteria	People diagnosis of schizophrenia or schizoaffective disorder, 14 years of age or older, one or more psychiatric hospitalizations or emergency department visits within the previous 2 years, ability to speak and read English, living in the community at the time of study entry, and absence of physical limitations that would preclude using a computer.	People diagnosed with a schizophrenia spectrum or an affective disorder, access to both a computer and the Internet, no use of Internet peer support in the past year, US resident, fluent in English.
Randomization	Randomly assigned	Randomly assigned, block randomization
Number enrolled	31	300
Follow-up period	3, 6, and 12 months	4 and 12 months
Number leaving early	1	41

Table 4. Description of the interventions (modified based on TIDieR checklist and guide).

Categories	Rotondi et al 2010 [48]	Kaplan et al 2011 [25]
Brief name	Schizophrenia Online Access to Resources (SOAR) intervention, specifically made for the study (the telehealth group).	Peer support Listserv or Peer support bulletin board. A Listserv and bulletin board were specifically made for the study.
Rationale/theory	Designed to provide key elements of family psychoeducation: empathetic engagement of participants, education about the illness and treatments, a supportive safety net, and coping strategies. Previous theoretical and empirical work: meeting and individual's needs can reduce stress, promote better adaption to illness-related difficulties, and improve outcomes; promotion of self-efficacy, self-management, and problem solving.	Participation to the online support group would result in improved well-being and decreased distress.
Materials	Computer access to the Internet via a dial-up modem and local Internet service provider, Schizophrenia Guide website/software.	Computer and Internet access. Participants received a message describing how their participation was to occur, how it worked, security information, warnings, advice, resources, and contact information.
Procedures	Participants received dial-up Internet access and a computer (if not having already). They were granted access to the following information and services via the "Schizophrenia Guide Web Site": 1. 3 online therapy groups, 2. Ask Our Experts Your Questions, 3. Questions and Answers Library, 4. Educational and Reading Materials; and 5. What's New. Participants were interviewed and provided their subjective evaluations of the website, several aspects of social support were also accessed.	1. Participants assigned to the experimental Listserv group communicated anonymously with each other using a group distribution email list; 2. Participants in an experimental peer support bulletin board group were given instructions on how to use the bulletin board. The content of Listserv and the bulletin board were entirely peer directed.
Providers	Project team members answered the questions of "Ask Our Experts Your Questions" module; therapy forums were facilitated by experienced mental health professionals (master of social work and PhD clinicians) trained in the monitoring and management of Web-based interventions; trained interviewers collected self-report data from participants.	Interventions were directed for study participants only and not facilitated by staff. Research staff was available for technical help.
How	The website provided 3 therapy forums (one for patients, one for support persons, one for both groups), a capability for asking questions of and receiving answers from a project team within 24-48 hours. The therapist emphasized discussions that focused on problem solving, and interacting with peers to develop a supportive forum where members could work together to address problems.	With both the Listserv and bulletin board, individuals communicated anonymously with one another using a group distribution email list specifically created for the group. The participants were encouraged to read and respond to email messages.
Where	The participants worked at home (Pittsburgh area, Pennsylvania, United States) and had access to the SOAR intervention via a desktop icon.	The participants worked at home and had access to Listserv or the bulletin board.
When and how much	Telehealth participants attended a joint, 4-hour workshop. Participants were in the study for up to 1 year.	Participants were in the study for 12 months.
Tailoring and modifications	N/A	N/A

Table 5. Outcomes and instruments used in the studies.

Outcomes	Rotondi et al 2005, 2010 [47,48]	Kaplan et al 2011 [25]
Symptoms	Scale for the Assessment of Positive Symptoms [49][a]	The Hopkins Symptom Checklist [50]
Knowledge level	Knowledge About Schizophrenia Instrument [51]	-
Recovery	-	The Recovery Assessment Scale [52]
Quality of life	-	The Quality of Life Interview, QoL [53]
Empowerment	-	The Empowerment Scale [54]
Social support	Perceived social support [55-58]	The Medical Outcome Study [58]
Stress	Self-rated stress [55-58]	-

[a] Information not available.

Narrative Analysis

An analysis of social media usage was reported in both studies [25,48]. The time that people with schizophrenia spent online on the SOAR website was, in total, 43,789 minutes (730 hours); the time involved 47,630 page views. The average time spent on the SOAR website was 46 hours, with an average of 2977+/-4.5 page views [48]. The users of the SOAR website asked on average 113 questions and read 69 articles. They used an average of 124 minutes asking questions and 1643 minutes for reading articles [48].

Analysis of engagement in social media forums used in the studies [25,48] showed an engagement of people with schizophrenia. They were active in therapy forums during the 3300 sessions [48]. They also sent 11,105 messages in different bulletin board forums [25].

Kaplan et al [25] categorized the participants into “high” and “low” dose participants. The participants categorized in the “high participation” group (57/185) reported having read the messages at least weekly and sent at least 5 messages at the 12-month post baseline point. People in this “high participation” group showed significantly higher distress levels at 4 months and at 12 months, while participants in the “low participation” group reported less distress at 12 months than at baseline. Further, the participants were grouped into those who had a “positive online experience” (OGQ scores ≥3) and those with a “less positive experience” (OGQ scores <3). People who reported positive experiences using social media forums (n=90) were significantly more distressed than participants in the less positive experience group [25].

Effectiveness of the Social Media Interventions

Meta-analysis was performed on both the Kaplan et al [25] and Rotondi et al [48] studies. In Kaplan et al’s study [25], comparisons for symptoms less than 6 months from baseline showed some improvement in the social media intervention group (*P*<.001, median -0.14, 95% CI -0.15 to -0.13) (see Figure 2). The Rotondi et al study [48] did not provide the information required regarding symptoms, thus meta-analysis was not performed.

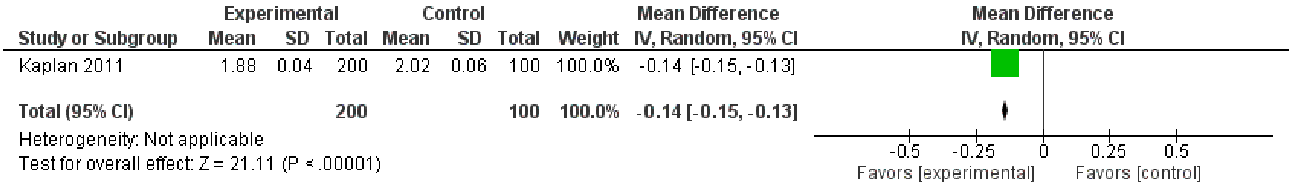
Comparisons for social support after 6 months from baseline [25,48] showed some improvement in the treatment as usual group (*P*=.03, median 0.22, 95% CI 0.02-0.42) (see Figure 3). When self-rated stress was compared 6 months from baseline, Rotondi et al [48] reported some effects in the social media intervention group (*P*=.01, median -0.51, 95% Cl -0.90 to -0.12) (see Figure 4).

Regarding self-management, Kaplan et al [25] compared self-management between groups after 4 and 12 months from baseline. They found that the treatment as usual group was slightly more effective than the social media group (*P*<.001, median 0.07, 95% CI 0.07-0.089) (see Figure 5). Moreover, Kaplan et al [25] compared quality of life after 4 and 12 months from baseline and reported that participants in the social media group had significantly higher QoL scores than participants in the control group (*P*<.001, median 0.15, 95% CI 0.14-0.17) (see Figure 6).

Assessment of Methodological Quality

The methodological quality of the two studies varied. Incomplete details in reporting the sequence generation and allocation concealment, decreases the methodological quality of both studies. Neither study was blinded, nor was an attempt made at blinding because of the nature of the intervention. Selective reporting may be possible as study protocols were not available. Intention-to-treat analyses were used in both the Kaplan [25] and Rotondi [47,48] studies. More details about possible risk of bias can be found in Table 6 and Figure 7.

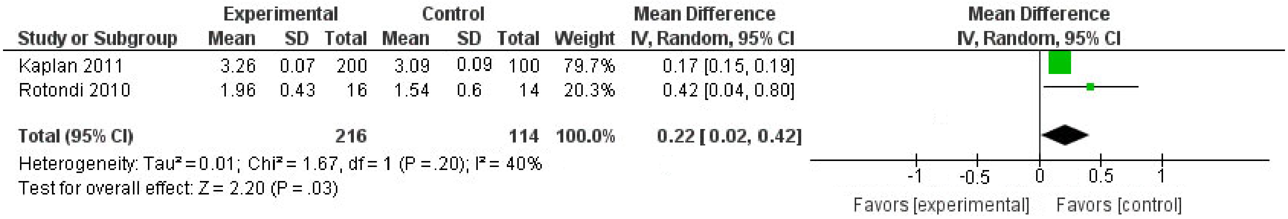
Figure 2. Positive symptoms for experimental and control groups by 6 months or less.



	Experimental			Control				Mean Difference
Study or Subgroup	**Mean**	**SD**	**Total**	**Mean**	**SD**	**Total**	**Weight**	**IV, Random, 95% CI**
Kaplan 2011	1.88	0.04	200	2.02	0.06	100	100.0%	-0.14 [-0.15, -0.13]
Total (95% CI)			**200**			**100**	**100.0%**	**-0.14 [-0.15, -0.13]**

Heterogeneity: Not applicable
Test for overall effect: Z = 21.11 (P < .00001)

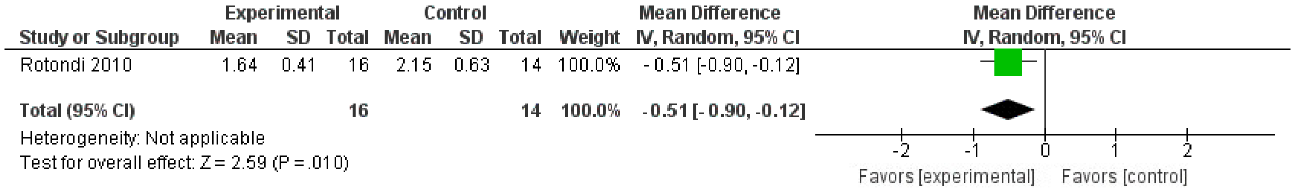
Figure 3. Social support for experimental and control groups by 6 months or less.



	Experimental			Control				Mean Difference
Study or Subgroup	**Mean**	**SD**	**Total**	**Mean**	**SD**	**Total**	**Weight**	**IV, Random, 95% CI**
Kaplan 2011	3.26	0.07	200	3.09	0.09	100	79.7%	0.17 [0.15, 0.19]
Rotondi 2010	1.96	0.43	16	1.54	0.6	14	20.3%	0.42 [0.04, 0.80]
Total (95% CI)			**216**			**114**	**100.0%**	**0.22 [0.02, 0.42]**

Heterogeneity: Tau² = 0.01; Chi² = 1.67, df = 1 (P = .20); I² = 40%
Test for overall effect: Z = 2.20 (P = .03)

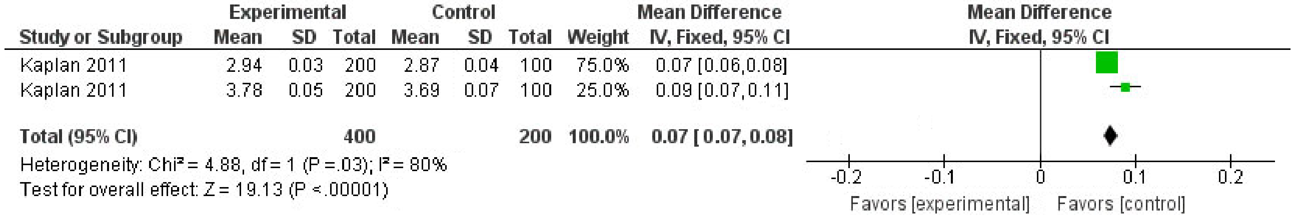
Figure 4. Stress for experimental and control groups by 6 months or less.



	Experimental			Control				Mean Difference
Study or Subgroup	**Mean**	**SD**	**Total**	**Mean**	**SD**	**Total**	**Weight**	**IV, Random, 95% CI**
Rotondi 2010	1.64	0.41	16	2.15	0.63	14	100.0%	-0.51 [-0.90, -0.12]
Total (95% CI)			**16**			**14**	**100.0%**	**-0.51 [-0.90, -0.12]**

Heterogeneity: Not applicable
Test for overall effect: Z = 2.59 (P = .010)

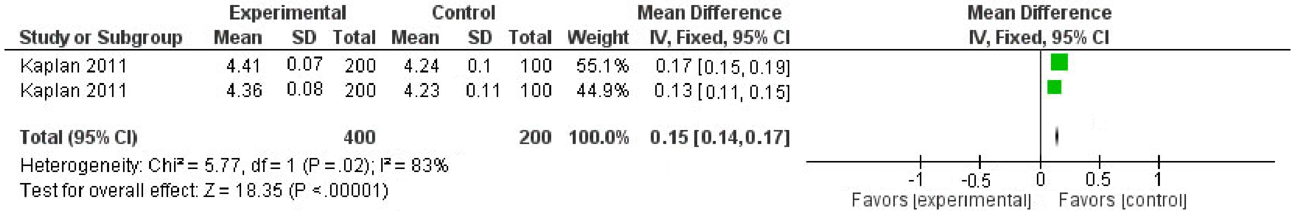
Figure 5. Self-management for experimental and control groups by 4 and 12 months (total score).



	Experimental			Control				Mean Difference
Study or Subgroup	**Mean**	**SD**	**Total**	**Mean**	**SD**	**Total**	**Weight**	**IV, Fixed, 95% CI**
Kaplan 2011	2.94	0.03	200	2.87	0.04	100	75.0%	0.07 [0.06,0.08]
Kaplan 2011	3.78	0.05	200	3.69	0.07	100	25.0%	0.09 [0.07,0.11]
Total (95% CI)			**400**			**200**	**100.0%**	**0.07 [0.07,0.08]**

Heterogeneity: Chi² = 4.88, df = 1 (P = .03); I² = 80%
Test for overall effect: Z = 19.13 (P < .00001)

Figure 6. Quality of life for experimental and control groups by 4 and 12 months (total score).



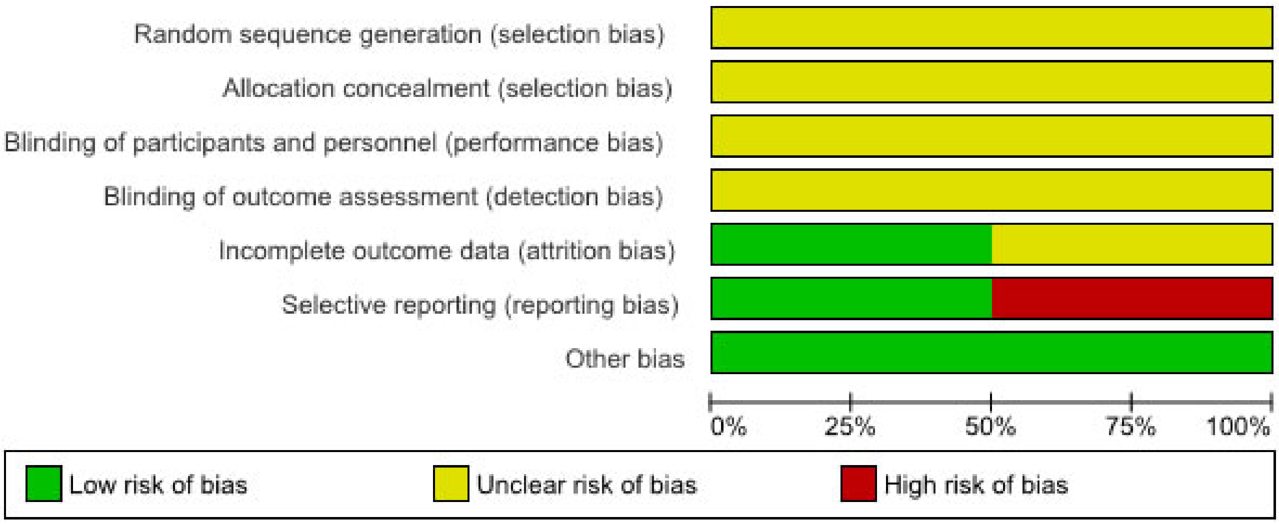
	Experimental			Control				Mean Difference
Study or Subgroup	**Mean**	**SD**	**Total**	**Mean**	**SD**	**Total**	**Weight**	**IV, Fixed, 95% CI**
Kaplan 2011	4.41	0.07	200	4.24	0.1	100	55.1%	0.17 [0.15, 0.19]
Kaplan 2011	4.36	0.08	200	4.23	0.11	100	44.9%	0.13 [0.11, 0.15]
Total (95% CI)			**400**			**200**	**100.0%**	**0.15 [0.14,0.17]**

Heterogeneity: Chi² = 5.77, df = 1 (P = .02); I² = 83%
Test for overall effect: Z = 18.35 (P < .00001)

Table 6. Outcomes and instruments used in the studies.

Bias	Rotondi et al 2005, 2010 [47,48]	Kaplan et al 2011 [25]
Random sequence generation (selection bias)	Randomly assigned. No further details.	Block randomization (nine in each block). No further details.
Allocation concealment (selection bias)	No further details.	No further details.
Blinding of participants and personnel (performance bias)	Non-blinded interviews. No further details.	No further details.
Blinding of outcome assessment (detection bias)	No further details.	Survey conducted using self-report measures. No further details about blinding of outcome assessment.
Incomplete outcome data (attrition bias)	Protocol published in 2003 NCT00051233, no outcomes provided.	No available study protocol.
Selective reporting (reporting bias)	Perceived social support [55-58].	Missing outcome data balanced in numbers across 3 groups. Missing data have been inputted. Retention rate varies slightly (11-18%).

Figure 7. Risk of bias assessment.



Discussion

Principal Findings

To our knowledge, this is the first systematic review and meta-analysis on the effects of social media interventions for people with schizophrenia. Our search strategy allowed us to capture and screen a large number of studies and explore their characteristics, interventions, outcome measures, and quality. Only two studies (three records) fulfilled the criteria and were included, while both showed a strong engagement in social media forums. However, social media use was found to be generally less effective than treatment as usual. Nonetheless, there is not enough evidence to arrive at a definite conclusion.

Existing trials suggest that people with schizophrenia are active in therapy forums and on bulletin boards. They have high participation in social media forums, similar to the general population [8]. These findings are in line with previous research, which reports use of forums and online chats by people with psychosis and suggests that such media could play a role in reducing the risk of isolation [27]. However, people who were in the "high participation" group showed significantly higher distress levels at 4 and 12 months compared to those in the "low participation" group. Also, those who reported positive experiences using social media forums were more distressed than participants in the negative experience group. As Kaplan et al [25] discussed, some clinicians fear that patient participation in online peer support without professional moderation may cause harm by fostering anxiety. Whether these findings really demonstrate this requires more studies.

After comparing the group allocated to social media with treatment as usual, social support, self-management, and quality of life ratings were better for those in treatment as usual. However, another recent systematic review indicated that positive psychotic symptoms, hospital admissions, socialization, social connectedness, and medication adherence have the potential to improve via online and mobile-based interventions for people with psychosis [26]. Thus, although based on our review, it appears that treatment as usual might be more effective in many aspects than treatment through social media use, other systematic reviews show that online and mobile-based interventions have a positive impact [26]. In our review, social media was part of wider intervention; it was not a pure intervention used for this patient group. Therefore, it would be useful to further investigate the engagement and effects of new technologies, the Internet, and their elements, especially social media, on people with mental disorders, in order to clarify their potential effectiveness for this population.

Limitations

There are limitations to our review. Only two studies fulfilled the inclusion criteria. Using English language studies might have resulted in our results being biased toward Western countries. It is possible that analysis of studies produced in languages other than English could yield different findings, although we think this unlikely. With both studies originating from the United States, it is unclear if the same findings would be reflected in other countries or cultures. The methodological quality of the included studies was assessed, and we identified a high risk of reporting bias because of missing outcomes or non-availability of study protocols.

Conclusion

Our findings suggest the effects of social media interventions are largely unknown. Use of social media forums is ubiquitous and increasing, but the relation between social media and mental health is complex [60], not well understood, and potentially detrimental. Thus, we suggest that this is reason enough to support further investigation. Emerging evidence suggests that online social networking can be related to major mental health problems such as depression [60], but at the same time online and mobile-based interventions for people with psychosis seem to improve depression [26]. Given the constant increase of social networking sites, it is understandable that recent studies have identified the need for exploring such sites and mental health [12,34]. Future research should comprehensively assess social media use for people with mental illness to determine the impact of mental well-being for social media use, as well as its risks. In these studies, interventions should be simple and usable to ensure patient engagement. Before suggesting specific social media interventions, however, factors and elements that may foster anxiety and distress by the use of social media should be identified. This knowledge would be an important resource for those who develop and evaluate mental health interventions involving social media, by knowing what elements to avoid in order to make patients' engagement more pleasant and less distressful. It is also obvious that there is an absence of reliable data coming from high-quality studies to help draw clear conclusions. Thus, robust reporting of the outcome assessment and the study protocol are essential for increasing the quality of studies in this emerging field.

Acknowledgments

The study has been funded by the University of Turku, Turku University Hospital, State Research Funding, University of Turku Graduate School – UTUGS, and the Academy of Finland. Special thanks to Elise Johansson, Information Services Librarian (University of Turku, Finland), for her invaluable contribution to the development and implementation of the systematic search and Leigh Ann who checked the language of the manuscript.

Conflicts of Interest

None declared.

References

1. NICE guidelines.: National Institute for Health and Care Excellence; 2014. Psychosis and schizophrenia in adults: treatment and management URL: http://www.nice.org.uk/guidance/cg178 [accessed 2015-11-30] [WebCite Cache ID 6dQIMb7Pb]
2. Rotondi A, Haas G, Anderson C, Mueser K. Effectiveness of web-based multi-family treatment delivered to the homes of persons with schizophrenia and their supporters. Schizophr Bull 2013;39:S350.
3. Ritterband L, Gonder-Frederick L, Cox D, Clifton A, West R, Borowitz S. Internet interventions: In review, in use, and into the future. Psychol Res Pr 2003 Oct;34(5):527-534.
4. Parikh S, Huniewicz P. E-health: an overview of the uses of the Internet, social media, apps, and websites for mood disorders. Curr Opin Psychiatry 2015 Jan;28(1):13-17. [doi: 10.1097/YCO.0000000000000123] [Medline: 25420193]
5. Kaplan AM, Haenlein M. Users of the world, unite! The challenges and opportunities of social media. Business Horizons 2010 Jan;53(1):59-68. [doi: 10.1016/j.bushor.2009.09.003]
6. Daker-White G, Rogers A. What is the potential for social networks and support to enhance future telehealth interventions for people with a diagnosis of schizophrenia: a critical interpretive synthesis. BMC Psychiatry 2013;13:279 [FREE Full text] [doi: 10.1186/1471-244X-13-279] [Medline: 24180273]
7. Robinson J, Rodrigues M, Fisher S, Bailey E, Herrman H. Social media and suicide prevention: findings from a stakeholder survey. Shanghai Arch Psychiatry 2015 Feb 25;27(1):27-35 [FREE Full text] [doi: 10.11919/j.issn.1002-0829.214133]
8. Pew Research Center. 2015. Internet seen as positive influence on education but negative influence on morality in emerging and developing nations URL: http://www.pewglobal.org/files/2015/03/Pew-Research-Center-Technology-Report-FINAL-March-19-20151.pdf [accessed 2015-11-30] [WebCite Cache ID 6dQKM1Fw7]
9. Statista. 2016. Number of social network users worldwide from 2010 to 2019 (in billions) URL: http://www.statista.com/statistics/278414/number-of-worldwide-social-network-users/ [accessed 2016-04-19] [WebCite Cache ID 6gsxekNBq]
10. Statista. 2016. Active social network penetration in selected countries as of January 2016 URL: http://www.statista.com/statistics/282846/regular-social-networking-usage-penetration-worldwide-by-country/ [accessed 2016-04-19] [WebCite Cache ID 6gsxmrkKj]
11. Statista. 2016. Number of monthly active Facebook users worldwide as of 4th quarter 2015 (in millions) URL: http://www.statista.com/statistics/264810/number-of-monthly-active-facebook-users-worldwide/ [accessed 2016-04-19] [WebCite Cache ID 6gsxvvirC]
12. Grajales FJ, Sheps S, Ho K, Novak-Lauscher H, Eysenbach G. Social media: a review and tutorial of applications in medicine and health care. J Med Internet Res 2014;16(2):e13 [FREE Full text] [doi: 10.2196/jmir.2912] [Medline: 24518354]
13. Fast I, Sørensen K, Brand H, Suggs LS. Social media for public health: an exploratory policy analysis. Eur J Public Health 2015 Feb;25(1):162-166 [FREE Full text] [doi: 10.1093/eurpub/cku080] [Medline: 24942532]
14. Davidson L, Chinman M, Sells D, Rowe M. Peer support among adults with serious mental illness: a report from the field. Schizophr Bull 2006 Jul;32(3):443-450 [FREE Full text] [doi: 10.1093/schbul/sbj043] [Medline: 16461576]
15. Salzer MS, The Mental Health Association of Southeastern Pennsylvania Best Practices Team. Consumer-delivered services as a best practice in mental health care and the development of practice guidelines. Psychiatr Rehabil Skills 2002;6:355-382.
16. Lal S, Adair CE. E-mental health: a rapid review of the literature. Psychiatr Serv 2014 Jan 1;65(1):24-32. [doi: 10.1176/appi.ps.201300009] [Medline: 24081188]
17. Maher CA, Lewis LK, Ferrar K, Marshall S, De BI, Vandelanotte C. Are health behavior change interventions that use online social networks effective? A systematic review. J Med Internet Res 2014;16(2):e40 [FREE Full text] [doi: 10.2196/jmir.2952] [Medline: 24550083]

18. Eysenbach G, Powell J, Englesakis M, Rizo C, Stern A. Health related virtual communities and electronic support groups: systematic review of the effects of online peer to peer interactions. BMJ 2004 May 15;328(7449):1166 [FREE Full text] [doi: 10.1136/bmj.328.7449.1166] [Medline: 15142921]
19. Griffiths KM, Calear AL, Banfield M. Systematic review on Internet Support Groups (ISGs) and depression (1): Do ISGs reduce depressive symptoms? J Med Internet Res 2009;11(3):e40 [FREE Full text] [doi: 10.2196/jmir.1270] [Medline: 19793719]
20. Best P, Manktelow R, Taylor B. Online communication, social media and adolescent wellbeing: A systematic narrative review. Child Youth Serv Rev 2014 Jun;41:27-36. [doi: 10.1016/j.childyouth.2014.03.001]
21. Leung R, Li J. Using social media to address Asian immigrants' mental health needs: A systematic literature review. J Nat Sci 2015 Apr;1(4):e66.
22. Rice SM, Goodall J, Hetrick SE, Parker AG, Gilbertson T, Amminger GP, et al. Online and social networking interventions for the treatment of depression in young people: a systematic review. J Med Internet Res 2014;16(9):e206 [FREE Full text] [doi: 10.2196/jmir.3304] [Medline: 25226790]
23. Bessell TL, McDonald S, Silagy CA, Anderson JN, Hiller JE, Sansom LN. Do Internet interventions for consumers cause more harm than good? A systematic review. Health Expect 2002 Mar;5(1):28-37. [Medline: 11906539]
24. Välimäki M, Hätönen HM, Lahti ME, Kurki M, Hottinen A, Metsäranta K, et al. Virtual reality for treatment compliance for people with serious mental illness. Cochrane Database Syst Rev 2014;10:CD009928. [doi: 10.1002/14651858.CD009928.pub2] [Medline: 25300174]
25. Kaplan K, Salzer MS, Solomon P, Brusilovskiy E, Cousounis P. Internet peer support for individuals with psychiatric disabilities: A randomized controlled trial. Soc Sci Med 2011 Jan;72(1):54-62. [doi: 10.1016/j.socscimed.2010.09.037] [Medline: 21112682]
26. Alvarez-Jimenez M, Alcazar-Corcoles MA, González-Blanch C, Bendall S, McGorry PD, Gleeson JF. Online, social media and mobile technologies for psychosis treatment: a systematic review on novel user-led interventions. Schizophr Res 2014 Jun;156(1):96-106. [doi: 10.1016/j.schres.2014.03.021] [Medline: 24746468]
27. Highton-Williamson E, Priebe S, Giacco D. Online social networking in people with psychosis: A systematic review. Int J Soc Psychiatry 2015 Feb;61(1):92-101. [doi: 10.1177/0020764014556392] [Medline: 25381145]
28. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Ann Intern Med 2009 Aug 18;151(4):264-269, W64. [Medline: 19622511]
29. Eysenbach G, CONSORT- E. 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]
30. Hoffmann TC, Glasziou PP, Boutron I, Milne R, Perera R, Moher D, et al. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. BMJ 2014;348:g1687 [FREE Full text] [Medline: 24609605]
31. McMillan S. Exploring models of interactivity from multiple research traditions: Users, documents, and systems. In: Handbook of new media: Social shaping and consequences of ICTS. London: Sage; 2002:163-181.
32. Aichner T, Jacob F. Measuring the Degree of Corporate Social Media Use. IJMR 2015;57(2):257-275.
33. Chou WS, Hunt YM, Beckjord EB, Moser RP, Hesse BW. Social media use in the United States: implications for health communication. J Med Internet Res 2009;11(4):e48 [FREE Full text] [doi: 10.2196/jmir.1249] [Medline: 19945947]
34. Toseeb U, Inkster B. Online social networking sites and mental health research. Front Psychiatry 2015;6:36 [FREE Full text] [doi: 10.3389/fpsyt.2015.00036] [Medline: 25814958]
35. Balatsoukas P, Kennedy CM, Buchan I, Powell J, Ainsworth J. The Role of Social Network Technologies in Online Health Promotion: A Narrative Review of Theoretical and Empirical Factors Influencing Intervention Effectiveness. J Med Internet Res 2015;17(6):e141 [FREE Full text] [doi: 10.2196/jmir.3662] [Medline: 26068087]
36. Higgins JPT, Green S. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0.: The Cochrane Collaboration; 2011 Mar. URL: http://handbook.cochrane.org/ [accessed 2015-11-30] [WebCite Cache ID 6dQUeS6ft]
37. Kilbourne AM, Abraham KM, Goodrich DE, Bowersox NW, Almirall D, Lai Z, et al. Cluster randomized adaptive implementation trial comparing a standard versus enhanced implementation intervention to improve uptake of an effective re-engagement program for patients with serious mental illness. Implement Sci 2013;8:136 [FREE Full text] [doi: 10.1186/1748-5908-8-136] [Medline: 24252648]
38. Kim H, Stout PA. The effects of interactivity on information processing and attitude change: implications for mental health stigma. Health Commun 2010 Mar;25(2):142-154. [doi: 10.1080/10410230903544936] [Medline: 20390680]
39. McFarlane WR, Lukens E, Link B, Dushay R, Deakins SA, Newmark M, et al. Multiple-family groups and psychoeducation in the treatment of schizophrenia. Arch Gen Psychiatry 1995 Aug;52(8):679-687. [Medline: 7632121]
40. Sicilia M, Ruiz S, Munuera J. Effects of Interactivity in a Web Site: The Moderating Effect of Need for Cognition. J Advertising 2005;34(3):31-45.
41. Spinzy Y, Nitzan U, Becker G, Bloch Y, Fennig S. Does the Internet offer social opportunities for individuals with schizophrenia? A cross-sectional pilot study. Psychiatry Res 2012 Jul 30;198(2):319-320. [doi: 10.1016/j.psychres.2012.02.022] [Medline: 22440545]

42. Steinwachs DM, Roter DL, Skinner EA, Lehman AF, Fahey M, Cullen B, et al. A web-based program to empower patients who have schizophrenia to discuss quality of care with mental health providers. Psychiatr Serv 2011 Nov;62(11):1296-1302 [FREE Full text] [doi: 10.1176/ps.62.11.pss6211_1296] [Medline: 22211208]
43. van der Krieke L, Emerencia AC, Aiello M, Sytema S. Usability evaluation of a web-based support system for people with a schizophrenia diagnosis. J Med Internet Res 2012;14(1):e24 [FREE Full text] [doi: 10.2196/jmir.1921] [Medline: 22311883]
44. van der Krieke L, Emerencia AC, Boonstra N, Wunderink L, de Jong JP, Sytema S. A web-based tool to support shared decision making for people with a psychotic disorder: randomized controlled trial and process evaluation. J Med Internet Res 2013;15(10):e216 [FREE Full text] [doi: 10.2196/jmir.2851] [Medline: 24100091]
45. Cochrane Informatics & Knowledge Management Department. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration; 2014. Review Manager (RevMan) Computer program Version 5 URL: http://tech.cochrane.org/revman/download [accessed 2016-04-08] [WebCite Cache ID 6gc8ERIoj]
46. Ades AE, Lu G, Higgins JPT. The interpretation of random-effects meta-analysis in decision models. Med Decis Making 2005;25(6):646-654. [doi: 10.1177/0272989X05282643] [Medline: 16282215]
47. Rotondi AJ, Haas GL, Anderson CM, Newhill CE, Spring MB, Ganguli R, et al. A Clinical Trial to Test the Feasibility of a Telehealth Psychoeducational Intervention for Persons With Schizophrenia and Their Families: Intervention and 3-Month Findings. Rehabil Psychol 2005 Nov;50(4):325-336 [FREE Full text] [doi: 10.1037/0090-5550.50.4.325] [Medline: 26321774]
48. Rotondi AJ, Anderson CM, Haas GL, Eack SM, Spring MB, Ganguli R, et al. Web-based psychoeducational intervention for persons with schizophrenia and their supporters: one-year outcomes. Psychiatr Serv 2010 Nov;61(11):1099-1105. [doi: 10.1176/ps.2010.61.11.1099] [Medline: 21041348]
49. Andreasen N. Scale for the Assessment of Negative Symptoms (SANS). Iowa City: University of Iowa; 1984.
50. Derogatis LR, Lipman RS, Rickels K, Uhlenhuth EH, Covi L. The Hopkins Symptom Checklist (HSCL): a self-report symptom inventory. Behav Sci 1974 Jan;19(1):1-15. [Medline: 4808738]
51. Barrowclough C, Tarrier N. Families of schizophrenic patients, cognitive behavioural intervention. In: London: Chapman & Hall. London: Chapman & Hall; 1992.
52. Corrigan PW, Gifford D, Rashid F, Leary M, Okeke I. Recovery as a psychological construct. Community Ment Health J 1999 Jun;35(3):231-239. [Medline: 10401893]
53. Lehman AF. A quality of life interview for the chronically mentally ill. Evaluation and Program Planning 1988 Jan;11(1):51-62. [doi: 10.1016/0149-7189(88)90033-X]
54. Rogers ES, Chamberlin J, Ellison ML, Crean T. A consumer-constructed scale to measure empowerment among users of mental health services. Psychiatr Serv 1997 Aug;48(8):1042-1047. [doi: 10.1176/ps.48.8.1042] [Medline: 9255837]
55. Finch JF, Okun MA, Barrera M, Zautra AJ, Reich JW. Positive and negative social ties among older adults: measurement models and the prediction of psychological distress and well-being. Am J Community Psychol 1989 Oct;17(5):585-605. [Medline: 2534023]
56. Krause N, Markides K. Measuring social support among older adults. Int J Aging Hum Dev 1990;30(1):37-53. [Medline: 2407665]
57. Lubben J. Assessing social networks among elderly populations. Fam Community Health 1988;11:42-52.
58. Sherbourne CD, Stewart AL. The MOS social support survey. Soc Sci Med 1991;32(6):705-714. [Medline: 2035047]
59. Chang T, Yeh CJ, Krumboltz JD. Process and outcome evaluation of an on-line support group for Asian American male college students. J Counc Psychol 2001;48(3):319-329.
60. Pantic I, Damjanovic A, Todorovic J, Topalovic D, Bojovic-Jovic D, Ristic S, et al. Association between online social networking and depression in high school students: behavioral physiology viewpoint. Psychiatr Danub 2012 Mar;24(1):90-93 [FREE Full text] [Medline: 22447092]

Abbreviations

CONSORT-EHEALTH: Consolidated Standards of Reporting Trials of Electronic and Mobile Health Applications and onLine TeleHealth
DSM: Diagnostic and Statistical Manual of Mental Disorders
ICD-10: International Statistical Classification of Diseases, 10th Revision
MeSH: Medical Subject Headings
OGQ: Online Group Questionnaire
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-analysis
QoL: Quality of Life
SOAR: Schizophrenia Online Access to Resources
TIDieR: Template for Intervention Description and Replication

Edited by G Eysenbach; submitted 01.12.15; peer-reviewed by Y Khazaal, A Rotondi; comments to author 30.12.15; accepted 01.02.16; published 22.04.16.

Please cite as:
Välimäki M, Athanasopoulou C, Lahti M, Adams CE
Effectiveness of Social Media Interventions for People With Schizophrenia: A Systematic Review and Meta-Analysis
J Med Internet Res 2016;18(4):e92
URL: http://www.jmir.org/2016/4/e92/
doi:10.2196/jmir.5385
PMID:27105939

Original Paper

An Internet-Based Physical Activity Intervention to Improve Quality of Life of Inactive Older Adults: A Randomized Controlled Trial

Karen Broekhuizen[1,2], PhD; Jelle de Gelder[1], MD, MSc; Carolien A Wijsman[1], MD, PhD; Liselotte W Wijsman[1], MD, MSc; Rudi GJ Westendorp[3], MD, PhD; Evert Verhagen[4], PhD; Pieternella E Slagboom[5], PhD; Anton J de Craen[1†], PhD; Willem van Mechelen[4], MD, PhD; Diana van Heemst[1], PhD; Frans van der Ouderaa[3], PhD; Simon P Mooijaart[1,2], MD, PhD

[1]Leiden University Medical Center, Department of Gerontology and Geriatrics, Leiden, Netherlands

[2]Institute for Evidence-Based Medicine in Old Age, Leiden, Netherlands

[3]Leyden Academy on Vitality and Ageing, Leiden, Netherlands

[4]ExtraMuraal Gezondheids Onderzoek (EMGO+) Institute for Health and Care Research, Department of Public and Occupational Health, Vrije Universiteit (VU) University Medical Center, Amsterdam, Netherlands

[5]Department of Medical Statistics, Section of Molecular Epidemiology, Leiden University Medical Center, Leiden, Netherlands

[†]deceased

Corresponding Author:
Simon P Mooijaart, MD, PhD
Leiden University Medical Center
Department of Gerontology and Geriatrics
Albinusdreef 2
2333 ZA
Leiden,
Netherlands
Phone: 31 715266640
Fax: 31 715266640
Email: s.p.mooijaart@lumc.nl

Abstract

Background: Increasing physical activity is a viable strategy for improving both the health and quality of life of older adults.

Objective: The aim of this study was to assess if an Internet-based intervention aimed to increase physical activity was effective in improving quality of life of inactive older adults. In addition, we analyzed the effect of the intervention on quality of life among those participants who successfully reached their individually targeted increase in daily physical activity as indicated by the intervention program, as well as the dose-response effect of increasing physical activity on quality of life.

Methods: The intervention was tested in a randomized controlled trial and was comprised of an Internet program—DirectLife (Philips)—aimed at increasing physical activity using monitoring and feedback by accelerometry and feedback by digital coaching (n=119). The control group received no intervention (n=116). Participants were inactive 60-70-year-olds and were recruited from the general population. Quality of life and physical activity were measured at baseline and after 3 months using the Research ANd Development 36-item health survey (RAND-36) and wrist-worn triaxial accelerometer, respectively.

Results: After 3 months, a significant improvement in quality of life was seen in the intervention group compared to the control group for RAND-36 subscales on emotional and mental health (2.52 vs -0.72, respectively; *P*=.03) and health change (8.99 vs 2.03, respectively; *P*=.01). A total of 50 of the 119 participants (42.0%) in the intervention group successfully reached their physical activity target and showed a significant improvement in quality of life compared to the control group for subscales on emotional and mental health (4.31 vs -0.72, respectively; *P*=.009) and health change (11.06 vs 2.03, respectively; *P*=.004). The dose-response analysis showed that there was a significant association between increase in minutes spent in moderate-to-vigorous physical activity (MVPA) and increase in quality of life.

Conclusions: Our study shows that an Internet-based physical activity program was effective in improving quality of life in 60-70-year-olds after 3 months, particularly in participants that reached their individually targeted increase in daily physical activity.

Trial Registration: Nederlands Trial Register: NTR 3045; http://www.trialregister.nl/trialreg/admin/rctview.asp?TC=3045 (Archived by WebCite at http://www.webcitation.org/6fobg2sjJ)

(J Med Internet Res 2016;18(4):e74) doi:10.2196/jmir.4335

KEYWORDS

Internet; physical activity; quality of life

Introduction

Increasing physical activity is a viable strategy for improving both health and quality of life in inactive older adults, who are a growing public health concern [1]. It is estimated that the proportion of adults aged 65 years and over will account for about 11% (939 million) of the total global population by 2030 [2]. Increased life expectancy is associated with an increase in multiple chronic conditions, translating into functional disability, need for assistance, reduced mobility, depression, isolation, and loneliness [3]. These outcomes are related to functioning and well-being and fall under the umbrella term *quality of life* [4]. Earlier efforts of health promotion have primarily focused on lower mortality rates or reduced disease risk. In the past decade, there is increasing concern that quality of life deserves attention as well [5].

Previously, we investigated the effects of a 3-month Internet-assisted intervention directed at increasing daily physical activity on objectively measured physical activity and metabolic health in 60-70-year-old inactive individuals. The intervention was tested in a randomized controlled trial and comprised of an Internet program—DirectLife (Philips)—aimed at increasing physical activity using monitoring and feedback by accelerometry and feedback by digital coaching. Results showed that the intervention was effective in increasing physical activity and in improving metabolic health [6,7].

In this study, our aim was threefold. First, we aimed to assess if the intervention was also effective in improving quality of life. Second, we analyzed the effect on quality of life among those participants who successfully reached their individually targeted increase in daily physical activity as indicated by the DirectLife program. Finally, we performed a dose-response analysis of increasing physical activity on quality of life among all participants.

Methods

Overview

Analyses were performed with data obtained from a previously reported randomized controlled trial—Nederlands Trial Register: NTR 3045—on the effects of a 3-month Internet-assisted intervention directed at increasing daily physical activity on objectively measured physical activity and metabolic health in 60-70-year-old inactive individuals. The CONSORT-EHEALTH Checklist for this trial is included as Multimedia Appendix 1 [8]. Details on study design and intervention content have been published elsewhere [6]. In short, the study recruited inactive participants aged 60-70 years from the region of Leiden, the Netherlands. The presence of an inactive lifestyle was assessed before randomization by a self-reported physical activity questionnaire: the General Practice Physical Activity Questionnaire (GPPAQ) [9]. This yielded four categories of physical activity: inactive, moderately inactive, moderately active, and active. We defined inactive as having less than 3 hours per week of exercise and cycling combined, corresponding to the inactive, moderately inactive, or moderately active category. Participants in the active category of the GPPAQ did not meet inclusion criteria for our definition of an inactive lifestyle. Participants were considered eligible if they (1) had no history of diabetes or did not use glucose-lowering medication, (2) had no disability impeding increase in physical activity, and (3) possessed and used a personal computer with Internet connection. At baseline, participants were randomly assigned to the intervention group or to a waiting list control group by the study physician or research nurse. Participants were randomized, via computerized program, into intervention versus waiting list control groups at a ratio of 1:1, with a block size of 12; stratification was performed by sex. Concealment of treatment allocation was ensured by randomizing at the end of the first study visit, after all baseline measurements and instructions at the study center were completed. Written informed consent was obtained from all participants. The study was approved by the Medical Ethical Committee of Leiden University Medical Center, the Netherlands.

Intervention

Participants in the intervention group received a commercially available Internet-based physical activity program—DirectLife (Philips, Consumer Lifestyle, Amsterdam, the Netherlands)—directed at increasing daily physical activity. The DirectLife program comprised components that are based on the stages of change and the I-Change model. Briefly, this model assumes that behavioral change is the result of individual awareness of one's behavior, motivation to change behavior and action, and taking into account the individual’s current activity level; it subsequently provides a personal goal [6,10,11]. DirectLife consists of three elements: (1) an accelerometer-based activity monitor, (2) a personal website, and (3) a personal e-coach, who provides regular updates of the individual’s physical activity status by email and gives advice on how to increase physical activity. By means of these elements, the program aims to increase awareness about one’s own physical activity behavior, to give feedback on recent actual physical activity, and to provide support to make sustainable changes in physical activity behavior. The activity monitor of DirectLife is based on the Tracmor triaxial accelerometer, and has been validated against double-labeled water for the estimation of total daily life energy expenditure [12]. The DirectLife monitor is the consumer version of the Tracmor accelerometer. Participants of the program were instructed to wear the activity monitor continuously throughout the day to measure daily physical activity. Data were uploaded through a secure Internet connection to the database of the commercial provider. After an initial 8-day *assessment period* starting 1 week after the study visit, in which the current level of daily activity was measured, a target was set to increase the level of daily activity during a 12-week Internet-based interactive coaching program. Personalized targets were set by the DirectLife program and were defined as the absolute increase in physical activity compared to the individual's baseline assessment data. For the whole group, this corresponded to a mean increase of approximately 10% in daily physical activity at week 12, increasing at a linear rate per week. All participants were given the option to decrease the personalized goal, within limits (ie, minimum of 5% increase in physical activity versus 10%), or to increase their personalized end goal, dependent on physical activity level of the last week.

Participants were given a target for daily activity, which increased weekly, and data from the accelerometer were used for regular feedback. Coaching included general recommendations on physical activities and coaches were available for further questions and advice by email correspondence. The control group was placed on a 3-month waiting list after which they received access to the intervention program at the end of the study. During the trial, no specific instructions regarding daily physical activity were given to the control group.

Measurements

Enrollment and follow-up took place from November 2011 to August 2012.

Baseline Questionnaire

In preparation of the first visit to the study center, all participants completed an Internet-delivered questionnaire on education, smoking status, and medical history, including medication use. Education was categorized as low (primary education and lower vocational education), intermediate (secondary education and intermediate vocational education), or high (high vocational education and university).

Quality of Life

Health-related quality of life was assessed at baseline and at 3-month follow-up, with the use of the self-administered standard Dutch paper version of the Research ANd Development (RAND) 36-item health survey (RAND-36) [13]. The RAND-36 questionnaire entails eight domains of health-related quality of life pertaining to both physical and mental health. The domains of physical functioning (10 items), limitations on usual role-related activities due to physical health problems (four items), pain (two items), and general health perception (five items) comprise the physical component; the domains of vitality (four items), social functioning (two items), limitations on usual role-related activities due to emotional or mental problems (three items), and emotional or mental problems (three items) comprise the mental component. In addition to the eight subscales, participants were asked to compare their current general health with their general health one year earlier. Scores on the subscales range from 0 to 100, with higher scores indicating better health or functioning. The total RAND-36 score is the sum of the scores on the subscales and ranges from 0 to 800, where 0 is the poorest quality of life and 800 the best imaginable. The RAND-36 is a reliable and valid measure for determining health-related quality of life in the elderly [13].

Physical Activity

The primary outcome was the individual's relative change in activity counts after the intervention compared to baseline, measured at the right wrist by the activity monitor. At baseline and 3-month follow-up, daily physical activity was measured during 7 days following the visit at the study center, using a wrist-worn triaxial accelerometer—GeneActiv (Kimbolton, Cambridgeshire, UK). To assess the primary outcome, we used accelerometers other than the one included in the intervention program to avoid interpretation of the intervention as an outcome. As a derivative outcome, we calculated from the wrist accelerometer the minutes per day spent in moderate-to-vigorous physical activity (MVPA), which has been validated against indirect calorimetry [14]. We chose not to report physical activity in counts, but in the amount of minutes spent in MVPA. For the elderly, general recommendations entail at least half an hour of MVPA at least five days a week. While the recommendations are formulated in terms of minutes spent in MVPA, the outcome of our study is better interpretable in terms of guideline adherence. In addition, minutes spent in MVPA as a derivative outcome for activity counts has been validated against indirect calorimetry. A detailed description of processing the collected accelerometer data into activity counts and average number of minutes, daily, spent in MVPA is described elsewhere [6]. Outcome assessment was done by an independent researcher who was blind to study arm allocation.

Successful Use of DirectLife Program

From the intervention group, a subgroup was created including participants who successfully reached their individually targeted increase in daily physical activity as indicated by the intervention program. An average level of physical activity per week was calculated from the last 3 weeks of the program and was compared with the personalized target of the corresponding week. Because a substantial number of participants reached the targeted personalized goals at the end of the 12-week program, but with some variation in the last 3 weeks, we labeled participants as being successful if they reached their target in at least 2 of the 3 last weeks of the program [7].

Statistical Analyses

Normally distributed data are shown as means with standard deviation, skewed data as medians with interquartile range (IQR). Between-group differences in quality of life after 3 months were analyzed by the intention-to-treat principle with an independent samples *t* test. For relative change in MVPA, the Mann-Whitney nonparametric test was performed. In a secondary analysis, we included only those participants in the intervention group who successfully reached the individual, personalized end goal to increase average physical activity that was set as part of the intervention program. To investigate if an increase in physical activity was associated with an improvement in quality of life, linear regression models were used, adjusted for age, sex, and body mass index (BMI) [15]. For this purpose, physical activity was divided into tertiles based on the change in minutes spent in MVPA. All analyses were performed with SPSS version 20.0 (IBM Corp, Armonk, NY, USA). Statistical significance was accepted at $P<.05$.

Results

Figure 1 shows the flowchart of included participants. A total of 235 participants were randomized—119/235 (50.6%) intervention, 116/235 (49.4%) control—and 96.2% (226/235) completed the trial—114/226 (50.4%) intervention, 111/226 (49.1%) control. Wijsman and colleagues already reported the main intervention effects [6]. In short, significant changes in favor of the intervention group were found for minutes of MVPA, weight loss, fat percentage, and glycated hemoglobin (HbA1c) [6]. In addition, in a dose-response analysis, Vroege and colleagues showed that there was a significant association between an increase in minutes spent in MVPA and body weight loss, reduction of BMI, waist circumference reduction, increase of high-density lipoprotein (HDL) cholesterol, and lowering of low-density lipoprotein (LDL)/HDL ratio [7].

Table 1 shows the baseline characteristics of the 235 study participants. The intervention and control groups were similar for all characteristics. In both groups, most participants were male, with a mean age of approximately 65 years old. The study population was overweight, with a mean BMI of 28.9 kg/m^2 (SD 4.7) and 29.1 kg/m^2 (SD 4.7) in the intervention and control groups, respectively.

Table 2 describes the change in quality of life for both the intervention and control groups. After 3 months, a significant improvement in quality of life was seen in the intervention group compared to the control group for subscales on emotional and mental health (2.52 vs -0.72, respectively; P=.03, 95% CI 0.39-6.09) and health change (8.99 vs 2.03, respectively; P=.01, 95% CI 1.60-12.32). No significant between-group differences were found for all other subscales, nor for the total RAND-36 score. From the results of the main analyses of Wijsman and colleagues, we know that accelerometer data were available for 107 intervention and 109 control participants, and that after 3 months, there was a mean increase of 11.1 minutes (SE 2.1) of MVPA per day in the intervention group, compared to a mean decrease of 0.1 minutes (SE 1.5) in the control group (P=.001) [6].

Further analysis included only those participants in the intervention group who were successful in reaching their personalized target after finishing the 3-month intervention program. This was the case for 50 (42.0%) of the 119 participants in the intervention group. Similar to the results from our primary analysis, a significant improvement in quality of life was seen in the successful intervention group compared to the control group for subscales on emotional and mental health (4.31 vs -0.72, respectively; P=.009, 95% CI 1.26-8.79) and health change (11.06 vs 2.03, respectively; P=.004, 95% CI 2.93-15.13). Overall, improvements in quality of life were larger in the successful intervention group compared to the overall intervention group for all subscales, as well as for the total RAND-36 score (see Table 2).

Table 3 describes the relationship between increase in physical activity and improvement in quality of life. Increase in physical activity in the entire sample was divided into tertiles based on the change in minutes spent in MVPA (see Table 3). Because of technical errors, data on activity counts were available for only 211 of the 235 participants (89.8%). With an increase of MVPA, the total RAND-36 score improved significantly (P_{trend}=.001, 95% CI 0.02-0.09), as well as quality of life regarding the subscales' usual role-related activities due to emotional health problems (P_{trend} =.03, 95% CI 0.01-0.20), emotional or mental health (P_{trend}=.005, 95% CI 0.10-0.58), pain (P_{trend}=.008, 95% CI 0.06-0.38), vitality (P_{trend}=.004, 95% CI 0.11-0.54), and general health perception (P_{trend}=.04, 95% CI 0.01-0.41). Other subscales were not associated with an increase in MVPA.

Table 1. Baseline characteristics of 235 study participants.

Characteristics	Intervention group (n=119)	Control group (n=116)
Demographics		
Sex (female), n (%)	47 (39.5)	49 (42.2)
Age (years), mean (SD)	64.7 (3.0)	64.9 (2.8)
Degree of self-reported activity, n (%)		
Moderately active	41 (34.5)	48 (41.4)
Moderately inactive	36 (30.3)	34 (29.3)
Inactive	42 (35.3)	34 (29.3)
Level of education, n (%)		
Low	7 (5.9)	2 (1.7)
Intermediate	45 (38.1)	46 (40.0)
High	66 (55.9)	67 (58.3)
Clinical parameters, mean (SD)		
Height (cm)	173.6 (9.9)	172.1 (9.3)
Weight (kg)	87.4 (15.8)	86.3 (15.8)
BMI[a] (kg/m^2)	28.9 (4.7)	29.1 (4.7)
Physical activity per day, mean (SD)		
Minutes spent in moderate-to-vigorous activity, median (IQR[b])	16.8 (18.6)	14.4 (23.8)
Quality of life, [c] **mean (SD)**		
Physical functioning	83.40 (14.98)	84.61 (15.05)
Social functioning	88.03 (17.63)	86.42 (18.76)
Role limitations (physical problem)	81.30 (32.09)	82.97 (29.41)
Role limitations (emotional problem)	85.99 (28.95)	86.21 (28.51)
Emotional or mental health	77.24 (15.52)	77.31 (15.58)
Vitality	67.48 (17.07)	67.03 (17.69)
Pain	80.04 (20.20)	84.29 (17.03)
General health perception	68.15 (16.34)	67.72 (14.94)
Health change	53.57 (20.40)	51.51 (13.31)
Total RAND-36[d,e] score	630.86 (120.12)	639.68 (118.64)

[a] BMI: body mass index.

[b] IQR: interquartile range.

[c] Scores on the subscales of quality of life range from 0 to 100, with higher scores indicating better health or functioning.

[d] RAND-36: Research ANd Development 36-item health survey.

[e] The total RAND-36 score is the sum of the scores on the subscales and ranges from 0 to 800.

Table 2. Changes in quality of life in the control, intervention, and successful intervention group.

Quality-of-life subscales[a]	All (n=225)	Control group (n=111)	Intervention group (n=114)		Successful intervention group (n=50)	
	Mean (SE)	Mean (SE)	Mean (SE)	*P* (95% CI) (between groups[b,c])	Mean (SE)	*P* (95% CI) (between groups[b,d])
Δ[e] Physical functioning	1.40 (0.77)	0.95 (1.01)	1.84 (1.16)	.56 (-2.14 to 3.93)	3.37 (1.60)	.19 (-1.21 to 6.06)
Δ Social functioning	-0.44 (1.06)	-1.13 (1.62)	0.22 (1.37)	.53 (-2.82 to 5.51)	1.20 (1.95)	.39 (-3.04 to 7.69)
Δ Role limitations (physical problem)	0.11 (1.94)	2.03 (2.93)	-1.75 (2.56)	.33 (-11.44 to 3.88)	-1.92 (4.28)	.45 (-14.20 to 6.30)
Δ Role limitations (emotional problem)	-1.04 (1.97)	-2.40 (2.62)	0.29 (2.95)	.50 (-5.09 to 10.48)	0 (3.88)	.61 (-6.80 to 11.60)
Δ Emotional or mental health	0.92 (0.73)	-0.72 (1.19)	2.52 (0.83)	*.03*[f] (0.39 to 6.09)	4.31 (1.14)	*.009*(1.26 to 8.79)
Δ Vitality	1.84 (0.83)	0.90 (1.09)	2.76 (1.24)	.26 (-1.40 to 5.13)	4.62 (1.89)	.07 (-0.34 to 7.77)
Δ Pain	-0.12 (1.09)	-0.74 (1.50)	0.48 (1.58)	.58 (-3.07 to 5.51)	-0.47 (2.04)	.92 (-4.87 to 5.40)
Δ General health perception	0.42 (0.77)	0.22 (1.22)	-0.09 (2.29)	.86 (-3.83 to 3.21)	0.38 (3.32)	.94 (-4.22 to 4.54)
Δ Health change	5.56 (1.38)	2.03 (1.44)	8.99 (2.29)	*.01*(1.60 to 12.32)	11.06 (2.28)	*.004*(2.93 to 15.13)
Δ Total RAND-36[g,h] score	2.75 (5.55)	-0.89 (7.57)	6.28 (8.13)	.52 (-14.75 to 29.09)	11.48 (11.96)	.37 (-14.80 to 35.54)

[a] Scores on the subscales range from 0 to 100, with higher scores indicating better health or functioning.

[b] Independent samples *t* test.

[c] Difference between control and intervention group.

[d] Difference between control and successful intervention group.

[e] Represents change in subscale score from baseline to follow-up.

[f] *P* values in italics represent significant values.

[g] RAND-36: Research ANd Development 36-item health survey.

[h] The total RAND-36 score is the sum of the scores on the subscales and ranges from 0 to 800.

Table 3. Dose-response relationship of the change in minutes spent in moderate-to-vigorous physical activity with quality of life.

Characteristics	Tertiles of Δ[a] moderate-to-vigorous physical activity (minutes per day)			
	Low (n=69)	Middle (n=72)	High (n=70)	*P*[b] (95% CI)
Δ Moderate-to-vigorous physical activity (minutes per day), median (IQR[c])	-7.80 (10.65)	2.20 (3.20)	19.40 (20.00)	N/A[d]
Quality-of-life subscales, [e] mean (SD)				
Δ Physical functioning	1.64 (1.39)	0.73 (1.36)	2.07 (1.38)	.26 (-0.10 to 0.37)
Δ Social functioning	-1.22 (1.88)	-0.84 (1.84)	-1.68 (1.86)	.46 (-0.11 to 0.24)
Δ Role limitations (physical problem)	-4.01 (3.60)	1.15 (3.51)	3.48 (3.55)	.09 (-0.01 to 0.17)
Δ Role limitations (emotional problem)	-3.62 (3.49)	-2.73 (3.41)	0.66 (3.45)	*.03*[f] (0.01 to 0.20)
Δ Emotional or mental health	-0.73 (1.35)	0.78 (1.32)	2.26 (1.33)	*.005* (0.10 to 0.58)
Δ Vitality	1.98 (1.49)	0.62 (1.45)	3.55 (1.47)	*.004* (0.11 to 0.54)
Δ Pain	-2.53 (1.97)	0.47 (1.93)	3.00 (1.95)	*.008* (0.06 to 0.38)
Δ General health perception	-2.29 (1.62)	1.07 (1.59)	1.44 (1.60)	*.04* (0.01 to 0.41)
Δ Health change	5.14 (2.40)	2.37 (2.35)	9.29 (2.37)	.07 (-0.01 to 0.26)
Δ Total RAND-36[g,h] score	-10.77 (9.93)	1.24 (9.71)	14.79 (9.81)	*.001* (0.02 to 0.09)

[a] Represents change in values from baseline to follow-up.

[b] Linear regression, adjusted for sex, age, and body mass index (BMI).

[c] IQR: interquartile range.

[d] N/A: not applicable.

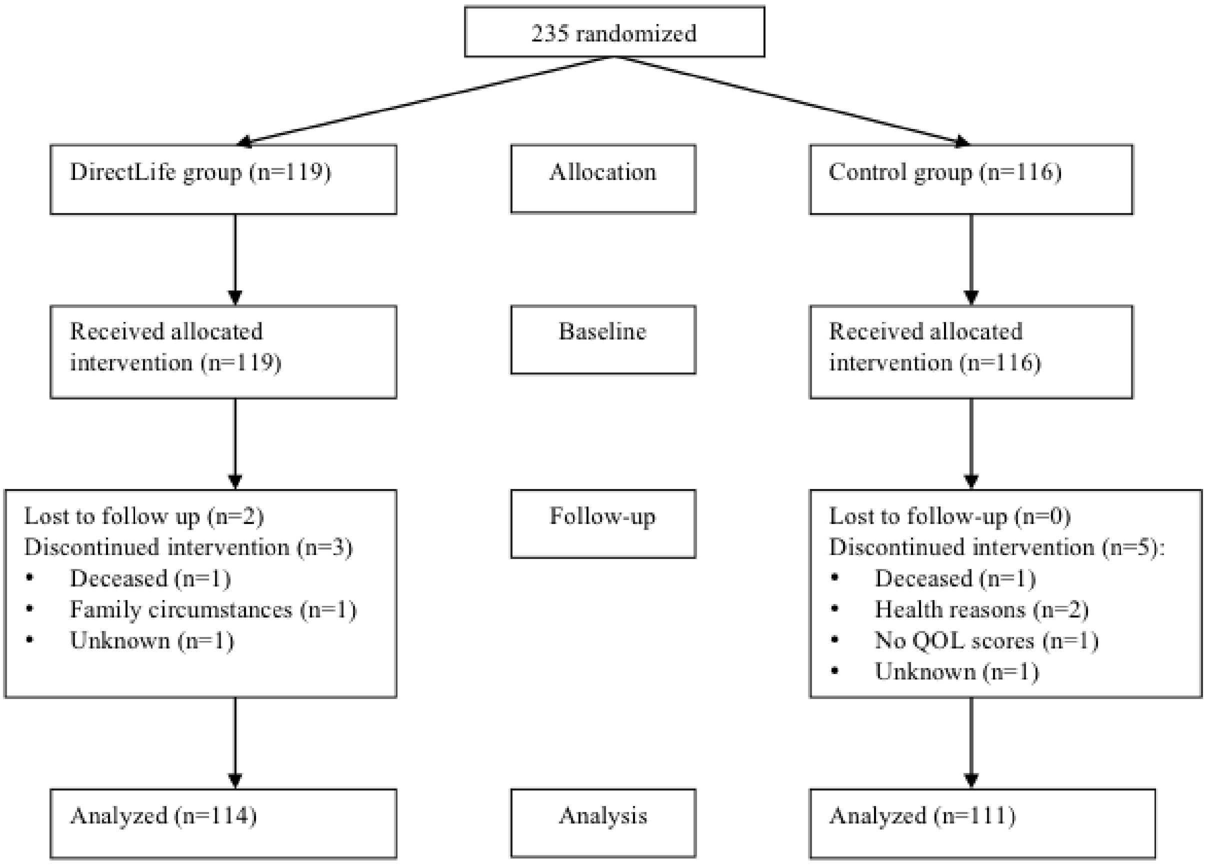
[e] Scores on the subscales range from 0 to 100, with higher scores indicating better health or functioning.

[f] *P* values in italics represent significant values.

[g] RAND-36: Research ANd Development 36-item health survey.

[h] The total RAND-36 score is the sum of the scores on the subscales and ranges from 0 to 800.

Figure 1. Flowchart of participants. QOL: quality of life.



Discussion

Principal Findings

In this randomized controlled trial assessing the effect of an Internet-based intervention on improving quality of life in inactive older adults, we found a significant improvement in quality of life in the intervention group compared to the control group for subscales on *emotional or mental health* and *health change* of the RAND-36 health survey. Improvements in quality of life were particularly observed in those participants in the intervention group who had successfully reached their personalized target with the intervention. Furthermore, our results show that more MVPA is associated with higher quality-of-life levels.

Earlier, we found that the DirectLife intervention led to significant improvements of metabolic health and showed a significant association between an increase in minutes spent in MVPA and metabolic health outcomes [6,7]. Given the observed dose-response relationship we have shown between physical activity and the majority of subscales of quality of life, we are confident that the DirectLife program has induced an improvement in quality of life through an increase in physical activity. Our results are in accordance with results from previous experimental studies that showed that an increase in physical activity had a positive effect on quality of life in older adults [16,17]. These studies performed no subgroup analyses that included successful intervention adopters only. Evidence from a series of randomized controlled trials conducted with frail older adults to test the effects of physical activity on quality of life has indicated that there is a positive effect on emotional and social functioning in particular, and that physical activity did not exacerbate perceptions of pain [18]. Our results also confirm these findings, as the intervention particularly induced an improvement of mentally related subscales of the RAND-36 and did not lead to increased perceptions of pain. It is actually noteworthy that the effects of DirectLife induced the largest increases in subscales regarding the emotional component of quality of life. One of the explanations for this might be the type of intervention. DirectLife primarily focused on personal goal setting, aiming for an increase of perceived control, self-efficacy, and mastery, which probably will induce an improvement of mental functioning in particular [1,19,20]. Secondly, our sample contained inactive older people, who will probably more directly experience improvement of mental functioning, rather than physical functioning. As the empirical evidence regarding the mechanisms underlying the association between increased physical activity and quality of life is limited, these hypotheses remain speculations and further research on this topic is warranted.

Assuming that an increase of ≥5 points in any of the RAND-36 subscales is clinically relevant [21], the clinical relevance of the significant increases in the *emotional or mental health* (2.52 points) and *health change* (8.99 points) subscales in the intervention group is moderate to high. Baseline scores on the RAND-36 were higher than scores observed in a healthy Dutch population in a similar age category, particularly for physical functioning and limitations on usual role-related activities due to physical health problems [13]. Apparently, inactive 60-70-year-olds do not feel hampered in daily functioning because of physical limitations. For that reason, it is less likely that scores on these subscales could have increased to a higher level. This so-called ceiling effect could have contributed to the lack of intervention effect on RAND-36 subscales regarding physical functioning.

A limitation of this study arises from the representativeness of our study sample. Participation was voluntary, which unintentionally could have led to an overrepresentation of participants who are highly motivated to increase physical activity. Also, selecting participants who were able to use the Internet led to a sample with a relatively high education level. As a consequence, generalizability of the results toward the general elderly population is limited. On the other hand, our study sample contained overweight, inactive older adults with comorbidities—with exclusion of diabetes—which is representative of the general population and consequently leads to increased generalizability. Adults in our study sample are categorized as being inactive, moderately inactive, and moderately active based on the GPPAQ categories and corresponds with less than 3 hours of physical exercise and/or cycling per week. According to the official GPPAQ guidance document, the *active* category is taken as consistent with achieving goals set by the physical activity (PA) guidelines relating to time spent in MVPA or vigorous physical activity (VPA) [9]. It is important to mention that all other categories require a PA intervention, which was the rationale behind our choice for this GPPAQ-guided cutoff point for inclusion in the DirectLife study. The relative short duration of the DirectLife program (ie, 3 months) should be considered while interpreting the results of this study. Although the effects found on quality of life seem promising, these need to be reinforced by results from a longer-term study. In addition, we have not specifically emphasized the unraveling of determinants of quality of life in this study, such as socioeconomic status, family support, computer literacy, type of environment, and housing status. An in-depth and adequately designed determinant study would be useful to unravel the influence of these determinants on physical activity, biomedical outcomes, and quality of life.

An important strength is that we used validated instruments to measure physical activity and quality of life. However, although accelerometry is one of the state-of-the-art objective measurements of physical activity, two comments are noteworthy. First, our choices for wearing the accelerometers on the right side, as well for considering an assessment day as valid if >10 hours were registered, were arbitrary. A study by Masse and colleagues critically used the criteria of different algorithms to reduce accelerometer data on physical activity and showed that the algorithm we used—minimal amount needed: 5 days; minimal daily wear: 10 hours—will not lead to different MVPA scores in comparison to more or less stringent algorithms [22]. Second, the conversion of accelerometer data into an accurate and reliable PA measurement is an evolving topic in PA research. Whether accelerometer data should be used in the evaluation of PA guideline adherence remains under doubt [23]. In our study, still-limited information is available on the type of physical activity that was performed. Knowing whether reduced sitting time or increased cycling or walking time are responsible for improvements of metabolic outcomes and quality of life could contribute to more individually tailored advice on how to improve health and quality of life in the elderly. For example, ankle-worn accelerometers might yield data that is more predictive for other types of physical activity, such as cycling. Activity counts assessed with use of ankle-worn accelerometers are available from a subset of our study population. This topic has our attention and is one of the planned further investigations with data from the study.

Our findings show the feasibility and potential of Internet-assisted physical activity interventions in an older adult population. Based on our findings, we encourage further implementation of the DirectLife program. The DirectLife program 2.0 could function as an open-access, Web-based lifestyle advice tool that can be referred to by general practitioners and medical specialists, or can be consulted by inactive older adults. Further development of the program, however, requires reinforcement of the found effects at a longer term and optimization of the personally tailored advice by including more determinants of physical (in)activity, such as sitting time.

Conclusions

Our study shows that an Internet-based physical activity program was effective in improving health-related quality of life among inactive 60-70-year-olds that successfully reached their individually targeted increase in daily physical activity after 3 months, as indicated by the intervention program. Our results indicate that significant improvements in physical activity can lead to an improvement of mental functioning in particular.

Acknowledgments

The Actief en Gezond Oud study was financially supported by Philips Consumer Lifestyle, and the Netherlands Genomics Initiative/Netherlands Organisation for Scientific Research (NGI/NWO) (grant numbers: 05040202 and 050-060-810). The funders had no role in the design and performance of the study, nor in the analyses or interpretation of the data or in the drafting of the manuscript.

Conflicts of Interest

Willem van Mechelen is director-shareholder of VU University Medical Center Amsterdam spin-off company Evalua Nederland B.V. (www.evalua.nl) and non-executive board member of Arbo Unie B.V. (www.arbounie.nl). Both companies operate on the Dutch occupational health care market.

Multimedia Appendix 1

CONSORT-EHEALTH Checklist v.1.6.1.

[PDF File (Adobe PDF File), 985KB - jmir_v18i4e74_app1.pdf]

References

1. Rejeski WJ, Mihalko SL. Physical activity and quality of life in older adults. J Gerontol A Biol Sci Med Sci 2001 Oct;56 Spec No 2:23-35. [Medline: 11730235]
2. Nelson ME, Rejeski WJ, Blair SN, Duncan PW, Judge JO, King AC, et al. Physical activity and public health in older adults: Recommendation from the American College of Sports Medicine and the American Heart Association. Med Sci Sports Exerc 2007 Aug;39(8):1435-1445. [doi: 10.1249/mss.0b013e3180616aa2] [Medline: 17762378]
3. Ory MD, Cox DM. Forging ahead: Linking health and behavior to improve quality of life in older people. In: Romney DM, Brown RI, Fry PS, editors. Improving the Quality of Life. Dordrecht, the Netherlands: Springer Publishing Company; 1994:89-120.
4. Stewart AL, King AC. Conceptualizing and measuring quality of life in older populations. In: Abeles RP, Gift HC, Ory MG, editors. Aging and Quality of Life. New York, NY: Springer Publishing Company; 1994:27-56.
5. Drewnowski A, Evans WJ. Nutrition, physical activity, and quality of life in older adults: Summary. J Gerontol A Biol Sci Med Sci 2001 Oct;56 Spec No 2:89-94. [Medline: 11730242]
6. Wijsman CA, Westendorp RG, Verhagen EA, Catt M, Slagboom PE, de Craen AJ, et al. Effects of a web-based intervention on physical activity and metabolism in older adults: Randomized controlled trial. J Med Internet Res 2013;15(11):e233 [FREE Full text] [doi: 10.2196/jmir.2843] [Medline: 24195965]
7. Vroege DP, Wijsman CA, Broekhuizen K, de Craen AJ, van Heemst D, van der Ouderaa FJ, et al. Dose-response effects of a Web-based physical activity program on body composition and metabolic health in inactive older adults: Additional analyses of a randomized controlled trial. J Med Internet Res 2014;16(12):e265 [FREE Full text] [doi: 10.2196/jmir.3643] [Medline: 25486673]
8. 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]
9. Department of Health. General Practise Physical Activity Questionnaire. London, UK: Physical Activity Policy Team; 2006. URL: http://webarchive.nationalarchives.gov.uk/20130107105354/http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/documents/digitalasset/dh_063813.pdf [accessed 2016-03-13] [WebCite Cache ID 6fz5uUy1I]
10. Prochaska JO, Velicer WF, Rossi JS, Goldstein MG, Marcus BH, Rakowski W, et al. Stages of change and decisional balance for 12 problem behaviors. Health Psychol 1994 Jan;13(1):39-46. [Medline: 8168470]
11. De Vries H, Mudde AN, Dijkstra A, Willemsen MC. Differential beliefs, perceived social influences, and self-efficacy expectations among smokers in various motivational phases. Prev Med 1998;27(5 Pt 1):681-689. [doi: 10.1006/pmed.1998.0344] [Medline: 9808799]
12. Bonomi AG, Plasqui G, Goris AH, Westerterp KR. Estimation of free-living energy expenditure using a novel activity monitor designed to minimize obtrusiveness. Obesity (Silver Spring) 2010 Sep;18(9):1845-1851. [doi: 10.1038/oby.2010.34] [Medline: 20186133]
13. VanderZee KI, Sanderman R, Heyink JW, de Haes H. Psychometric qualities of the RAND 36-Item Health Survey 1.0: A multidimensional measure of general health status. Int J Behav Med 1996;3(2):104-122. [doi: 10.1207/s15327558ijbm0302_2] [Medline: 16250758]
14. Esliger DW, Rowlands AV, Hurst TL, Catt M, Murray P, Eston RG. Validation of the GENEA accelerometer. Med Sci Sports Exerc 2011 Jun;43(6):1085-1093. [doi: 10.1249/MSS.0b013e31820513be] [Medline: 21088628]
15. Brown DW, Balluz LS, Heath GW, Moriarty DG, Ford ES, Giles WH, et al. Associations between recommended levels of physical activity and health-related quality of life. Findings from the 2001 Behavioral Risk Factor Surveillance System (BRFSS) survey. Prev Med 2003 Nov;37(5):520-528. [Medline: 14572437]
16. Acree LS, Longfors J, Fjeldstad AS, Fjeldstad C, Schank B, Nickel KJ, et al. Physical activity is related to quality of life in older adults. Health Qual Life Outcomes 2006;4:37 [FREE Full text] [doi: 10.1186/1477-7525-4-37] [Medline: 16813655]
17. Spirduso WW, Cronin DL. Exercise dose-response effects on quality of life and independent living in older adults. Med Sci Sports Exerc 2001 Jun;33(6 Suppl):S598-S608; discussion S609. [Medline: 11427784]
18. Schechtman KB, Ory MG, Frailty and Injuries: Cooperative Studies of Intervention Techniques. The effects of exercise on the quality of life of frail older adults: A preplanned meta-analysis of the FICSIT trials. Ann Behav Med 2001;23(3):186-197. [Medline: 11495219]
19. McAuley E, Doerksen SE, Morris KS, Motl RW, Hu L, Wójcicki TR, et al. Pathways from physical activity to quality of life in older women. Ann Behav Med 2008 Aug;36(1):13-20 [FREE Full text] [doi: 10.1007/s12160-008-9036-9] [Medline: 18677453]

20. Dechamps A, Diolez P, Thiaudière E, Tulon A, Onifade C, Vuong T, et al. Effects of exercise programs to prevent decline in health-related quality of life in highly deconditioned institutionalized elderly persons: A randomized controlled trial. Arch Intern Med 2010 Jan 25;170(2):162-169. [doi: 10.1001/archinternmed.2009.489] [Medline: 20101011]
21. Ware 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]
22. Mâsse LC, Fuemmeler BF, Anderson CB, Matthews CE, Trost SG, Catellier DJ, et al. Accelerometer data reduction: A comparison of four reduction algorithms on select outcome variables. Med Sci Sports Exerc 2005 Nov;37(11 Suppl):S544-S554. [Medline: 16294117]
23. 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]

Abbreviations

BMI: body mass index
EMGO+: ExtraMuraal Gezondheids Onderzoek
GPPAQ: General Practice Physical Activity Questionnaire
HbA1C: glycated hemoglobin
HDL: high-density lipoprotein
IQR: interquartile range
LDL: low-density lipoprotein
MVPA: moderate-to-vigorous physical activity
N/A: not applicable
NGI/NWO: Netherlands Genomics Initiative/Netherlands Organisation for Scientific Research
PA: physical activity
QOL: quality of life
RAND: Research ANd Development
RAND-36: Research ANd Development 36-item health survey
VPA: vigorous physical activity
VU: Vrije Universiteit

Edited by G Eysenbach; submitted 11.02.15; peer-reviewed by T Halliday, C Marsaux, S Bonn, S Wangberg; comments to author 05.08.15; revised version received 29.09.15; accepted 21.01.16; published 27.04.16.

Please cite as:
Broekhuizen K, de Gelder J, Wijsman CA, Wijsman LW, Westendorp RGJ, Verhagen E, Slagboom PE, de Craen AJ, van Mechelen W, van Heemst D, van der Ouderaa F, Mooijaart SP
An Internet-Based Physical Activity Intervention to Improve Quality of Life of Inactive Older Adults: A Randomized Controlled Trial
J Med Internet Res 2016;18(4):e74
URL: http://www.jmir.org/2016/4/e74/
doi:10.2196/jmir.4335
PMID:27122359

Original Paper

Interprofessional Communication of Clinicians Using a Mobile Phone App: A Randomized Crossover Trial Using Simulated Patients

Bhavesh Patel[1], BSc, MB BS, MRCS; Maximilian Johnston[2], MB BCh, MRCS, PhD; Natalie Cookson[3], MSc, MBBS; Dominic King[2,4], MBChB, MRCS, PhD; Sonal Arora[2], MBBS, MRCS, PhD; Ara Darzi[1,2], FACS, FRCS, MD

[1]Department of Surgery and Cancer, St Mary's Campus, Imperial College London, London, United Kingdom
[2]Imperial Patient Safety Translational Research Centre, Department of Surgery and Cancer, Imperial College London, London, United Kingdom
[3]Academic Surgical Unit, St Marys Hospital, Imperial College Healthcare NHS Trust, London, United Kingdom
[4]Google DeepMind, London, United Kingdom

Corresponding Author:
Bhavesh Patel, BSc, MB BS, MRCS
Department of Surgery and Cancer
St Mary's Campus
Imperial College London
QEQM Building
London, W2 1NY
United Kingdom
Phone: 44 0 207 594 792
Fax: 44 0 0 7594 3137
Email: bhavesh.patel07@imperial.ac.uk

Abstract

Background: Most hospitals use paging systems as the principal communication system, despite general dissatisfaction by end users. To this end, we developed an app-based communication system (called Hark) to facilitate and improve the quality of interpersonal communication.

Objective: The objectives of our study were (1) to assess the quality of information transfer using pager- and app-based (Hark) communication systems, (2) to determine whether using mobile phone apps for escalation of care results in additional delays in communication, and (3) to determine how end users perceive mobile phone apps as an alternative to pagers.

Methods: We recruited junior (postgraduate year 1 and 2) doctors and nurses from a range of specialties and randomly assigned them to 2 groups who used either a pager device or the mobile phone-based Hark app. We asked nurses to hand off simulated patients while doctors were asked to receive handoff information using these devices. The quality of information transfer, time taken to respond to messages, and users' satisfaction with each device was recorded. Each participant used both devices with a 2-week washout period in between uses.

Results: We recruited 22 participants (13 nurses, 9 doctors). The quality of the referrals made by nurses was significantly better when using Hark (Hark median 118, range 100–121 versus pager median 77, range 39–104; *P*=.001). Doctors responded to messages using Hark more quickly than when responding to pagers, although this difference was not statistically significant (Hark mean 86.6 seconds, SD 96.2 versus pager mean 136.5 seconds, SD 201.0; *P*=.12). Users rated Hark as significantly better on 11 of the 18 criteria of an information transfer device (*P*<.05) These included "enhances interprofessional efficiency," "results in less disturbance," "performed desired functions reliably," and "allows me to clearly transfer information."

Conclusions: Hark improved the quality of transfer of information about simulated patients and was rated by users as more effective and efficient, and less distracting than pagers. Using this device did not result in delay in patient care.

(J Med Internet Res 2016;18(4):e79) doi:10.2196/jmir.4854

KEYWORDS

communication; mobile phone; pager; applications; apps; escalation of care; simulation

Introduction

Failures in communication can have serious and damaging implications for patient safety [1,2]. Data published by The Joint Commission in the United States has consistently highlighted errors in communication as the most common root cause of sentinel events, with handoff errors in particular being implicated in as many as 80% of all adverse events [3-6]. As such, there is a need for health systems to prioritize the improvement of communication skills and modalities within acute health care settings.

In hospitals worldwide, the most widely used communication system continues to be the 1-way pager device, first introduced over 50 years ago [7]. Despite its widespread adoption, users of pagers report the devices to be highly disruptive and inefficient [8-10]. In most paging systems, a member of staff sends an alert to a colleague requesting to be called back. The alert is received as a loud sound with an extension number to be called shown on the pager's display. As the only information displayed by the pager is a number to be called, pagers do not effectively allow the urgency of a message to conveyed. Furthermore, information cannot be transferred until the receiver returns the call. Delays in response to a page can cause frustration, prompting additional pages [11]. The frequent interruptions caused by repeated pages alongside ineffective information transfer lead directly to adverse events and safety events [12,13]. More generally, frontline clinicians report that pagers have a negative impact on communication, quality of work, and efficiency [11].

The negative sentiment toward pagers in parallel with advances in communication technologies has encouraged many hospitals to explore alternative solutions to pager systems. Such solutions include 2-way alphanumeric pagers, secure short message service (SMS) text messaging platforms, and task management systems [14-17]. While an ideal communication tool would enable bidirectional, rapid, secure, and nondisruptive transmission of content-rich messages, existing systems have different limitations that fall short of this ideal [18].

As a result, the uptake of the new technologies in the health care setting is not yet widespread. Rogers proposed that new innovations diffuse through five distinct sections of consumer in the process of widespread adoption: innovators, early adopters, early majority, late majority, and laggards. Health care organizations have invested heavily in older communication systems that are still functional and have been slow to transition to emergent technologies. New technologies have therefore not moved beyond the innovators and early adopters [19,20]. In contrast, in hospitals where pagers remain the principal method of communication, health professionals are increasingly taking things into their own hands and using personal mobile phones to communicate with colleagues. This can be through a combination of voice calls, SMS messages, and increasingly messaging platforms such as WhatsApp (WhatsApp Inc, Mountain View, CA, USA) [21,22].

Previous systematic reviews of communication systems have demonstrated improvements in clinicians' perception of communication when using communication systems other than pagers. [14] A more recent randomized controlled trial highlighted the greater employee satisfaction when using mobile phones for intrateam communication [15]. Mobile phones, therefore, present a significant opportunity for advances in interprofessional communication. Clinicians recognize the advantages that mobile phones offer over paging devices, specifically increased accessibility, clearer communication, and the ability to triage messages based on urgency [23,24].

To address the shortcomings of existing communication systems and take advantage of widespread access to mobile phone devices, an app-based communication system (ABCS) called Hark has been developed by Imperial College London, London, United Kingdom. Hark runs over mobile phones and tablet devices to support clinical task management. In doing so it builds on the specifications we defined to reduce adverse events associated with clinical handoff, prioritization of clinical tasks, failure to rescue, and escalation of care [3,11,25,26].

Hark was designed and developed by a multidisciplinary team of researchers, clinicians, technologists, and software developers according to the Imperial clarify, design, and evaluate approach to the development of digital health (mHealth) solutions [27]. It was developed building on extensive feedback from focus group discussions with multiple stakeholders (including nursing staff, doctors, and hospital managers) to ensure that all features of the app are tailored toward end users [11].

Through its design, Hark aims to address the shortcomings of the pager and the negative perceptions of users specifically around the time taken to respond to communication episodes and the quality of information transferred through the device [10,11,25]. This study aimed to assess the feasibility of using Hark in the clinical environment by determining whether the identified shortcomings had been addressed and whether using an alternative communication system would result in additional delays to patient care.

The aims of this study were (1) to assess the quality of information transfer using pager-based and app-based communication systems, (2) to determine whether using a mobile phone app for escalation of care would result in additional delays in communication, and (3) to determine how end users perceive mobile phone apps as an alternative means of communication to pagers.

Methods

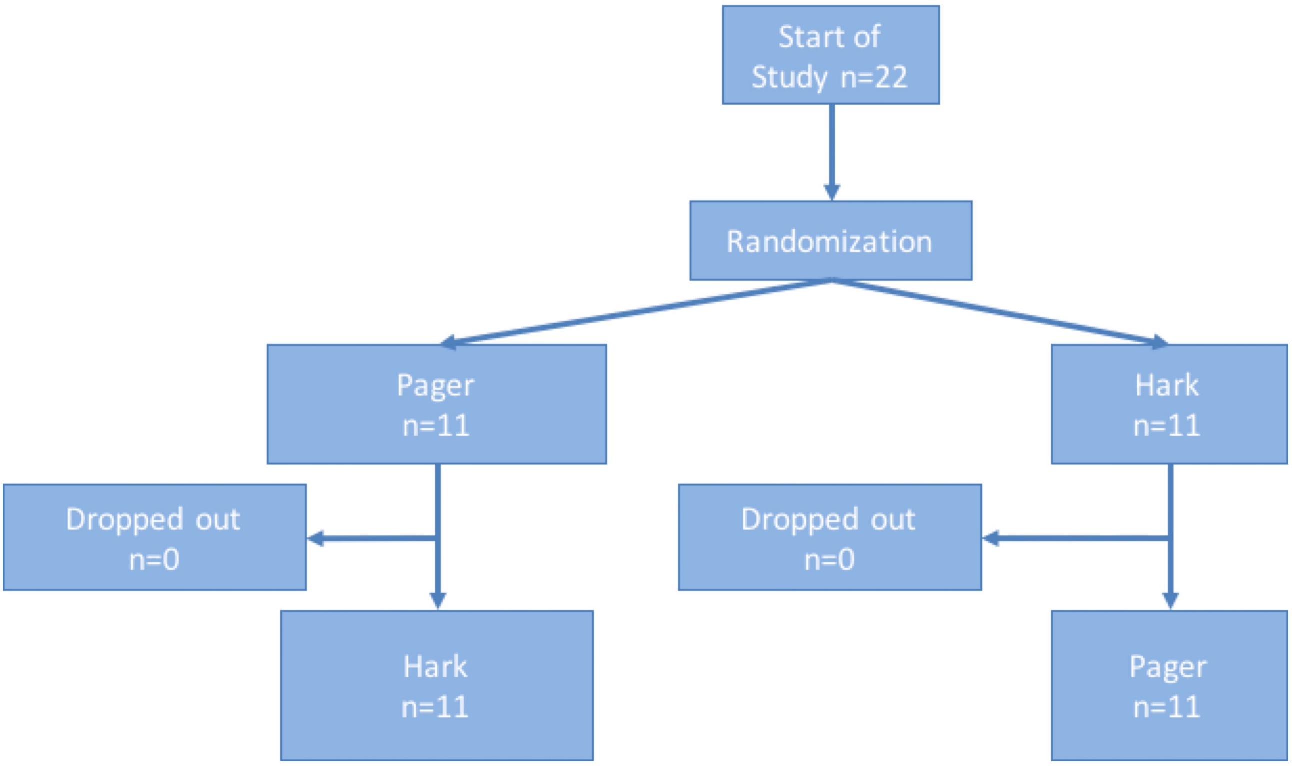
Participants

We approached doctors and nurses working at a 500-bed tertiary hospital for participation in the study. Participants were recruited against the following criteria. Inclusion criteria were (1) being a doctor or nurse working in the hospital and (2) working in a specialty that serves inpatients. Exclusion criteria were (1) being a medical or nursing student, (2) being a nurse not directly involved with the routine care of inpatients (eg, operating room staff, scrub nurses, (3) being agency (temporary) staff, and (4) being computer illiterate. Before recruitment, we deemed a sample size of a minimum of 40 data points to be sufficient on the basis of previously published literature involving simulation in the methodology [28-32].

Study Design

This study used a randomized crossover design. Participants were randomly assigned into 1 of 2 groups according to a random number generator. We asked one group to use a conventional pager device to transfer information and the other to use Hark. Once participants completed the task, they underwent a 2-week washout period before being asked to perform similar tasks using the other device. By using such a study design, we obtained a greater number of data points using a smaller sample size. Furthermore, as all participants used both devices, they acted as their own internal control to minimize sampling bias between the groups. Figure 1 illustrates the study design.

Figure 1. Design of the study.



Data Collection

Participants were requested to use either a pager or Hark to give or receive information about simulated patients. We produced 6 scenarios, each describing a patient who required input from a doctor. In order to allow both urgent escalation of care and routine task management functions to be investigated, the scenarios varied in terms of the how quickly they required a response. In the 6 scenarios, 2 patients required immediate attention, 2 required urgent attention, and 2 required routine attention. All participates sent or received information about all 6 scenarios. Figure 2 details the scenarios.

In clinical practice referrals are usually made from nurses to doctors [8]; therefore, we presented nurse participants with the above scenarios and asked them to refer the patient to a member of the research team, who acted as a doctor. Referrals were made using either a paging device or Hark.

We asked doctor participants to carry a device, either a pager or a mobile phone with Hark installed, during a typical day at work. A member of the research team then randomly contacted them through the device to refer the patients from the scenarios. Doctors were able to differentiate messages sent for simulated patients from those about actual patients, as all pages were sent from a single telephone number reserved for the purposes of the study. All messages sent through Hark were about simulated patients.

These tasks were designed to model the day-to-day use of a pager device for doctors and nurses. Each participant was asked to perform the tasks on 2 separate occasions: once using a pager device and another time using Hark.

We collected data from observation and monitoring of participants during the task and from feedback questionnaires given to participants after the task was completed.

Figure 2. Outline of clinical scenarios. SpO_2: oxygen saturation.

Immediate

Scenario 1
2 days post operative Hartmann's procedure for perforated diverticulitis

Previous Medical History: Chronic Obstructive Pulmonary Disease, Diabetes Mellitus, Hypertension
Patient appears breathless, distressed and anxious. Complaining of chest pain in right side of chest.

Latest Observations
Blood Pressure 125/78; Heart Rate 120; Respiratory Rate 35; SpO2 88% on 2L via nasal cannula; Temperature 37.2

Current medications- Acetaminophen, *Oramorph, Ceftrioxone, Metronidazole, Metformin, Ramipril, Hartmann's Solution*

Immediate

Scenario 2
Admitted 2 days ago with Pulmonary Edema
Previous Medical History: High Cholesterol, Hypertension, Heart Failure, Angina

Complaining of chest pain in left side of chest similar to previous angina. Patient appears to be in pain and sweaty. Developed sudden onset chest pain 10 minutes ago.

Latest Observations
Blood Pressure 110/75; Heart Rate 120; Respiratory Rate 28; SpO2 96% on 2L via nasal cannula; Temperature 36.8

Current Medication: Simvastatin, Bisoprolol, Ramipril, Aspirin, Furosemide, Glyceryl Trinitrate, Acetaminophen

Urgent

Scenario 3
Patient was admitted on todays surgical take with appendicitis. She has been kept nil by mouth to be operated tomorrow morning and has been put on a Variable Rate Insulin infusion (Sliding scale Actrapid).
Previous Medical History: Diabetes Mellitus

Her latest Capillary Glucose was 7.6 and her current bag of 5% dextrose will be finished in 45 minutes. She has no further bags of 5% dextrose written up to continue after this bag.

Latest Observations
Blood Pressure 120/80; Heart Rate 84; Respiratory Rate 18; SpO2 96% on air; Temperature 36.0

Current medications- Acetaminophen, *Oramorph, Variable Rate Insulin Infusion*

Urgent

Scenario 4
Admitted from Nursing Home today with fevers and dysuria. Working diagnosis is a Urinary Tract Infection
Previous Medical History: High Cholesterol, previous fracture of neck of left femur

Patient was walking back from toilet when slipped while trying to get back into bed. Did not hit her head, able to stand back up herself and get back into bed. Moving all 4 limbs without pain. Oriented to time and place.

Latest Observations
Blood Pressure 124/82; Heart Rate 83; Respiratory Rate 15; SpO2 98% on air; Temperature 36.4

Current Medication: Trimethoprim, Simvastatin, 0.9% Saline

Non-Urgent

Scenario 5
4 days post Femoro-Popliteal bypass surgery
Previous Medical History: Atrial Fibrillation; Hypertension

Day team forgot to prescribe warfarin. The patient requires warfarin as he has Atrial Fibrillation. The latest INR was 1.8 taken yesterday. Patient is otherwise well.

Latest Observations
Blood Pressure 135/89; Heart Rate 89; Respiratory Rate 18; SpO2 96% on air; Temperature 36.5

Current medications- Acetaminophen, *Oramorph, Warfarin, Ramipril, Bisoprolol, 0.9% Saline*

Non-Urgent

Scenario 6
Post Inguinal Hernia repair (operated today)
Previous Medical History: Nil

Patient had to stay in overnight as he developed post-operative urinary retention. The patient and his family are very concerned as to why the patient is being kept overnight and would like to speak to a doctor. The patient has been catheterized by the day team and is otherwise well

Latest Observations
Blood Pressure 117/75; Heart Rate 67; Respiratory Rate 18; SpO2 98% on air; Temperature 36.4

Current Medication: Acetaminophen, *Oramorph*

Features of Hark

A full overview of the Hark platform and product overview can be seen at the Hark website [33]. Individuals register with Hark, creating a user profile from which they are able to send or receive messages. Their profiles contain information about their clinical role, experience level, and contact details.

To assign a task to colleagues, senders first select a patient and then enter details about the task required into a “new task” form in the app. As Hark integrates with the hospital's electronic medical records, the form is prepopulated with the selected patient's demographic information, recent vital signs, and test results. Users complete the remaining fields by selecting the type of task required and filling in free-text spaces with details about the task. At the end of the form, the sender selects a colleague to whom to send the task, indicating the urgency of the task by selecting a time frame for completion.

Once the form is sent, the colleague to whom the task was assigned receives a notification that a task has been sent to him or her. As soon as the recipient opens the form in the app, a notification is generated on the requesting device that the task has been read. The receiver can choose to accept the task, delegate it to another colleague, or send comments to clarify any further questions about the task. Once the task is completed it is marked as complete on the devices of both the sender and the receiver. A full audit trail of all communication episodes is made providing a contemporaneous record of all events associated with a clinical task.

Measures

We created specific measures to allow the aims of the project to be achieved. To determine whether the quality of referrals differed between the devices, we audio-recorded the referrals made by nurses using the pager while the data from referrals made by using Hark were automatically recorded on the app itself. We then assessed the data against a previously validated assessment tool called the QUality of Information Transfer Tool (QUIT) [26]. The QUIT is a 25-item, 7-category assessment tool comprising core components to measure the quality and content of information transfer. To ensure consistency when using the QUIT, 2 independent researchers rated each referral.

The time taken for doctors to respond to referrals was recorded to determine whether using a new technology for escalation of care results in additional delays in communication. For the pager, we defined this as the time between when a page was sent and when a call was received in response to the page. For Hark, we defined this as the time between when a task was sent and when the device received a "message read" notification. The message read notification was deemed appropriate as a response, as this was the point (when using Hark) where information has been transferred from the sender to the receiver. By reading the message, the receiver can decide how quickly he or she needs to reply on the basis of that information. For nonurgent task requests, it would be appropriate for them to prioritize completing an urgent task in front of them above immediately replying on Hark. Furthermore, the message read notification provides assurance to the sender that the information has been viewed and appropriate action will be taken.

As referrals are usually made from nurses to doctors, we asked only nurses to make referrals, and therefore we used the QUIT only for nurse participants. Similarly, as we asked only doctors to respond to referrals, we collected data about the time taken to respond only for doctor participants.

Finally, to determine how end users perceive mobile phone apps as an alternative means of communication to pagers, we asked all participants to complete a questionnaire. The questionnaire asked participants to express their degree of agreement with 18 statements on a 5-point Likert scale. The statements correlated with the key requirements of a communication system that emerged from previously published research [11]. We sought trust ethical approval from the Imperial Joint Research Compliance Office, but because it was a simulated study that didn't involve real patients, project approval alone was advised to be sufficient.

Data Analysis

We analyzed data using SPSS statistics version 22 (IBM Corporation). We conducted descriptive analysis of the sociodemographic information.

Statistical analysis was performed to assess whether any significant differences existed in the performance of either device according to the QUIT scores, response times, and participants' feedback. Aside from the time taken to respond to messages, the data were not normally distributed; therefore, we report median values and used the Wilcoxon test for within-group analysis. For the time taken to respond, we report mean values and used the paired-samples *t* test for within-group analysis. Spearman correlation coefficient was used to assess whether there was any relationship between the participants' age and their evaluation of the device.

Statistical significance was accepted at a level of $P<.05$.

Results

Participants and Demographics

We included a total of 22 participants in the study, of whom 9 were doctors and 13 were nurses. All participants used both Hark and the pager device, which gave a total of 44 data points.

All doctors who participated were junior doctors (postgraduate year 1 and 2) with a mean age of 25 (range 24–28) years. Of the 9 doctor participants, 7 were female and 2 were male.

The seniority of nurses ranged from licensed practical nurses (pay scale band 5 in the United Kingdom) to advanced practice registered nurses (band 8) with a mean age of 38 (range 25–55) years. Of the 13 nurse participants, 7 were female and 6 were male.

The sample included representatives from medical and surgical specialties, along with psychiatry and pediatrics.

Statistical Analyses

Quality of Information Transfer

The quality of the referrals made by nurses was significantly better when using Hark, with higher overall scores when using Hark than when using a pager device (Hark median 118, range 100–121 versus pager median 77, range 39–104; $P=.001$) (Figure 3). For all 7 categories, nurses were found to perform significantly better when using Hark ($P<.05$). Furthermore, 22 of the 25 QUIT items were conveyed significantly more frequently when using Hark. There was no significant difference in the frequency of communicating the "patient name," "patient location," "responsible consultant [clinician]," and "current treatment to date." These results provide good evidence that Hark improves the quality of referrals (see Table 1).

Time Taken to Respond to Device

Doctors responded to messages using Hark more quickly than when responding to pagers, although this difference was not statistically significant (see Table 2). There was also no statistically significant difference when comparing response times to messages requiring immediate, urgent, and nonurgent response.

Evaluation of the Device

Hark was found to perform significantly better on 11 of the 18 criteria of an information transfer device (*P*<.05; Table 3). These included "enhances interprofessional efficiency," "results in less disturbance," "performed desired functions reliably," and "allows me to clearly transfer information." On the remaining 7 items, users rated Hark as better than or equal to the pager, but the difference was not statistically significant. Users also reported a greater overall satisfaction when using Hark, although this difference was also not significant (Hark median 4, range 1–5 versus pager median 3.5, range 1–5; *P*=.24).

Older users were more likely to rate the pager highly on enabling them "to both send and receive communications" (correlation coefficient 0.538 *P*=.01). Age did not correlate with ratings of either device on any other criteria (Table 4).

Figure 3. Box plots of overall quality of information transfer (QUIT) scores between devices. Asterisks and dot represent outliers.

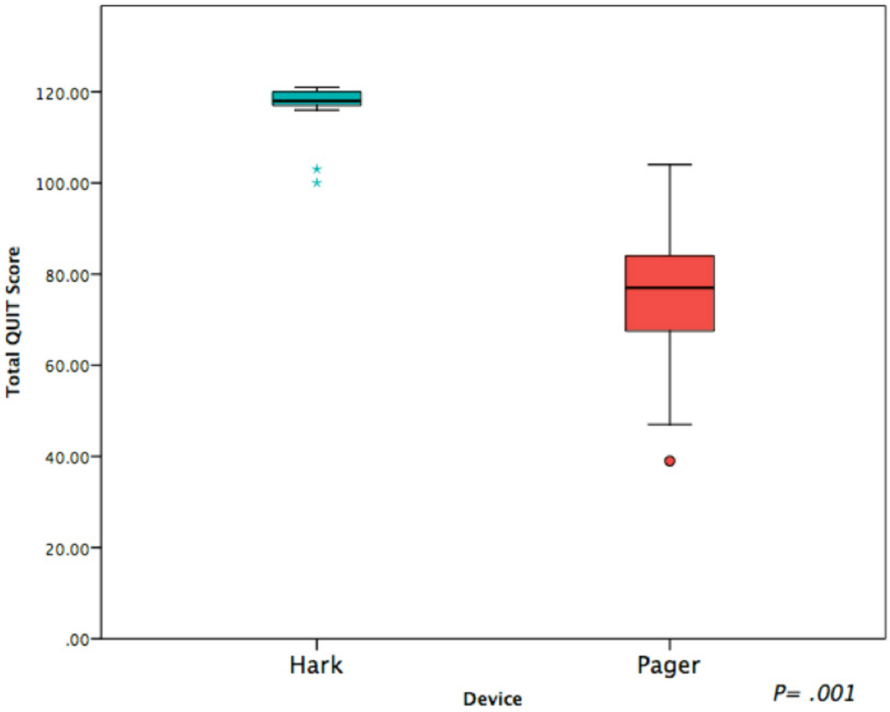


Table 1. Quality of information transfer scores for use of Hark versus a pager.

Categories and items	Hark		Pager		*P* value
	Median	Range	Median	Range	
1) Communicator identities	15	11–15	7	3–13	.001
Clearly communicates initiator identity (include grade and specialty)	5	5–5	4	1–5	.004
Clearly confirms appropriate receiver identity (include grade and specialty)	5	5–5	2	1–4	.001
Establishes rapport and mutual respect	5	1–5	3	1–5	.004
2) Patient identity	16	11–18	11	7–17	.003
Clearly communicates patient name	5	4–5	4	1–5	.03
Clearly communicates patient location	5	1–5	5	2–5	.33
Clearly communicates responsible consultant	1	1–3	1	1–5	.91
Clearly communicates age/date of birth	5	4–5	1	1–5	.002
3) Clinical details	14	9–15	11	3–14	.003
Clearly articulates working diagnosis	5	3–5	4	1–5	.009
Clearly communicates relevant history (including recent operation and date)	5	1–5	4	1–5	.004
Outlines current treatment to date	5	2–5	3	1–5	.03
4) Problem	30	25–30	18	11–28	.001
Clearly describes current problem with patient	5	5–5	4	3–5	.01
Communicates relevant vital signs and fluid balance	5	3–5	4	2–5	.004
Describes patient assessment and examination findings	5	2–5	3	1–4	.002
Outlines relevant investigation results to date	5	3–5	3	1–4	.001
Effectively prioritizes clinical issues	5	5–5	2	1–5	.002
Clearly communicates degree of urgency	5	5–5	2	1–5	.003
5) Plan	15	12–15	11	5–14	.002
Clearly defines the reason for the call (eg, advice, patient review, transfer)	5	3–5	4	2–5	.002
Definitively resolves questions and ambiguities about patient care	5	4–5	3	1–4	.002
Agrees plan for ongoing care for patient	5	3–5	4	1–5	.02
6) Overall quality of information presentation	30	17–30	18	8–29	.001
Uses clear, understandable language throughout	5	5–5	4	2–5	.002
Presents information in a structured and logical order	5	5–5	3	1–5	.002
Used available documentation to structure handoff	5	3–5	5	1–5	.003
Selected and communicated all relevant information	5	5–5	3	1–5	.002
Completed information transfer without digressing	5	5–5	3	1–5	.002
Overall quality of information transfer	5	4–5	3	1–4	.002
Total score	118	100–121	77	39–104	.001

Table 2. Time taken to respond to various types of messages transmitted by Hark versus a pager.

Type of message	Hark			Pager			*P* value
	Mean	SD	Range	Mean	SD	Range	
All messages, time in seconds	86.6	96.2	2–416	136.5	201.0	4–900	.11
Immediate messages, time in seconds	126.0	128.4	2–416	80.6	81.2	4–300	.22
Urgent messages, time in seconds	68.1	84.3	4–286	130.2	168.0	9–585	.12
Nonurgent messages, time in seconds	65.8	52.9	8–195	198.8	289.4	15–900	.07

Table 3. Performance of Hark versus a pager as evaluated by agreement on a 5-point Likert scale with criteria for an information transfer device.

Statement	Hark		Pager		*P* value
	Median	Range	Median	Range	
I do not need to be in a specific location within the hospital to initiate or receive communication through this system	5	1–5	5	1–5	.15
I am able to send or receive sufficient levels of detail through this system	5	1–5	4	1–5	.31
It enhances interprofessional collaboration and efficiency	4	1–5	3	2–5	.01
It results in fewer interruptions	5	2–5	2	1–5	.001
It results in less disturbance from interruptions	4	2–5	2	1–5	.003
It minimizes the time between sending a message and receiving the desired response	4	1–5	3	1–5	.22
It makes it easy to contact colleagues in times of need	4	1–5	3.5	2–5	.37
It discourages transfer of unnecessary information	4	1–5	3	1–5	.07
It is simple to operate	4	2–5	4	2–5	.75
It allows me to both send and receive communication	5	1–5	4	1–5	.003
It allows me to clearly transfer information about tasks and patients	5	1–5	4	1–5	.01
It allows me to easily delegate tasks or patients to colleagues	5	3–5	3	1–5	.001
It allows me to access patient information	4	2–5	2	1–5	.001
It allows me to prioritize messages according to urgency	5	1–5	2.5	1–5	.002
It performs the desired functions reliably, with minimal occurrence of malfunctions	4	2–5	3.5	1–5	.01
It can be stored as evidence that communication occurred	5	1–5	2	1–5	.001
It can allow a third person to differentiate between different senders and receivers	5	1–5	2	1–5	.001
I would be satisfied if this was the primary system used for communication between wards staff and doctors	4	1–5	3.5	1–5	.24

Table 4. Correlation between age of participants and their rating of a device by agreement with evaluation criteria.

Statement	Hark		Pager	
	Spearman correlation coefficient	*P* value	Spearman correlation coefficient	*P* value
I do not need to be in a specific location within the hospital to initiate or receive communication through this system	–.160	.48	–.125	.58
I am able to send or receive sufficient levels of detail through this system	–.162	.47	.155	.49
It enhances interprofessional collaboration and efficiency	.118	.60	.196	.38
It results in fewer interruptions	–.175	.44	.381	.08
It results in less disturbance from interruptions	–.203	.37	.417	.05
It minimizes the time between sending a message and receiving the desired response	.032	.89	.216	.34
It makes it easy contact colleagues in times of need	.100	.66	–.260	.24
It discourages transfer of unnecessary information	.096	.67	.400	.07
It is simple to operate	–.051	.82	.172	.45
It allows me to both send and receive communication	–.301	.17	.538	.01
It allows me to clearly transfer information about tasks and patients	–.076	.74	.041	.86
It allows me to easily delegate tasks or patients to colleagues	–.013	.96	.417	.05
It allows me to access patient information	–.087	.70	.287	.20
It allows me to prioritize messages according to urgency	.066	.77	.075	.74
It performs the desired functions reliably, with minimal occurrence of malfunctions	–.147	.51	.254	.26
It can be stored as evidence that communication occurred	–.376	.09	–.394	.07
It can allow a third person to differentiate between different senders and receivers	–.405	.06	.262	.24
I would be satisfied if this was the primary system used for communication between wards staff and doctors	.035	.88	.271	.22

Discussion

Our study aimed to investigate the efficiency of a user-centered ABCS compared with a pager device for interprofessional communication in hospitals. It compared the quality of information transfer during a patient referral, the time taken for users to respond to referrals, and user perceptions of interprofessional communication using mobile phone apps as an alternative means of communication to pagers.

We found the quality of information transfer to be higher when using the Hark app than when using the pager. Further analysis of the results revealed that. aside from 3 parameters (patient name, location, and responsible clinician), all relevant pieces of information were transferred significantly more frequently when using Hark than when using a pager. To transfer information using Hark, users must complete several fields, each of which corresponds to a piece of information that needs to be transferred. As such, the task entry form is a template. with each field serving as a prompt to ensure more complete information transfer [34]. This is in contrast to the pager, where the only information transferred is a number to be called back (see Figure 4).

There was no field for entering the name of the responsible clinician, which may explain the equal frequency of information transfer of this parameter when using Hark or the pager. The other 2 parameters that we found to be nonsignificant (patient name and location) were so essential and intuitive for a referral that prompting should not be necessary to ensure transfer of the information.

The data demonstrated no significant difference in the time taken to respond to messages sent through either device. Is it noteworthy that, despite using a newer device with little training and limited prior experience of it, users did not take longer to respond to referrals sent using Hark. It should be noted that the time taken to respond to messages was defined differently according to the device used. When using Hark, response time was defined as the time taken to receive a message read notification. When using the pager, this was defined as the time taken for the receiver to return the call. This may introduce some bias in favor of Hark; however, this is a reflection of how reading a message through Hark results in earlier transfer of information. This suggests that in a clinical setting using Hark would not result in any additional delay over using a pager when contacting colleagues for escalation of care purposes. Indeed, Joseph et al found that at a hospital that used mobile phones in the clinical environment, users reported a reduction in physician response time to both routine and critical patients compared with when using pagers [23].

Users of Hark reported that it performed significantly better than the pager on most of the parameters we assessed.

Specifically, they perceived Hark to be significantly better at enhancing interprofessional collaboration, reducing disruption, enabling task prioritization, and improving reliability. There were no instances in which users reported better performance when using the pager. This suggests that the participants of this study would support switching the devices used on wards from pagers to app-based systems [35]. These findings support the conclusions of other studies in the literature that explore altnernatives to the pager. Interventions such as 2-way alphanumeric pagers, task management systems, and secure text messaging platforms all result in users reporting improved efficiency in communication and reduced disruption [14,15,36].

Using an ABCS offers several benefits for patients, health care professionals, and managers alike. As we have shown, using Hark results in a higher quality of information transfer, which results in clearer communication and may lead to improved patient care. Using Hark allows health care professionals to communicate more effectively, resulting in greater interprofessional collaboration and less disruption to their work. As all communication is stored and indexed according to the type of tasks requested, Hark offers managers fresh insights about the most frequent types of tasks performed in their hospitals and thus supports decision making about resource workforce management. The audit trail provided by Hark would also facilitate incident investigation and keeping staff portfolios and logbooks. It is important to be aware of possible unintended consequences when using an ABCS. As more information is transferred over the technology, the number of verbal and face-to-face communications may decrease, negatively affecting workplace relationships. Furthermore, complex situations may be oversimplified when communicated over an ABCS, resulting in greater back-and-forth messaging for clarification [37].

While there are many demonstrable benefits of using an ABCS, some essential pieces of infrastructure need to be in place before it can be widely adopted in the health care setting. First, there must be secure Internet or mobile phone coverage throughout the hospital to allow users to be able to send and receive information regardless of their location. This is particularly important if time-critical information regarding deteriorating patients is to be sent over Hark, as delays in escalation of care are associated with adverse outcomes [25]. Second, as the ABCS requires a mobile phone for use, all health care workers must have access to a mobile phone or tablet device. Although most health care workers already carry personal mobile phones, provision of mobile phones or tablets should be considered for staff who do not have access to a device [38]. Third, the app must be constantly running on the mobile phone device, which will challenge the battery life of most popular devices. As such, hospitals that use Hark must ensure widespread access to mobile phone charging facilities.

Figure 4. Information about a simulated patient displayed in messages sent over Hark and a pager.

Shift St Mary's Hospital

Ward Charles Pannett

Patient Bauch A

Task Patient Review

back into bed. Did not hit her head able to stand back up herself and get back into bed. Moving all limbs without pain. Oriented to time and

Indication Other

Admitted from nursing home today with fevers and dysuria. Possible UTI Background of high cholesterol, previous fracture of L Neck of femur

2 NEWS: 2 Today 21:20

Limitations

This study has certain limitations, which future research should seek to address. First, this was a single-center study in a simulated setting. Hark should be tested across other sites in different geographic areas to determine whether the results obtained in this study can be replicated. As the study used 6 simulated cases as opposed real patient situations, participants may have evaluated the devices differently from how they would have if we had used in real clinical scenarios. In particular, users may have found Hark to be less disruptive when using these 6 scenarios than they would have when considering more cases with real clinical care.

Doctors' awareness before responding that all cases were about simulated patients may have influenced how quickly they responded to messages during our study. They may have deprioritized their response to messages about simulated patients or may have responded more quickly when using a preferred technology. To minimize this preference bias, we did not tell participants that we would be monitoring their response time. Future studies could aim to minimize this bias further by blinding participants between messages about simulated and actual patients.

It should also be noted that the sample of doctors who participated in this study consisted only of junior doctors. Although our results did not demonstrate a relationship between age and evaluation of the devices used, future research should include all grades of doctor to minimize any potential bias introduced by the age of the participant. It should also be noted that, although the study had a limited sample size of 22, the crossover design of the study allowed a larger number of data points to be collected from a smaller number of participants. The sample size was therefore sufficient to provide the required number of data points. The crossover design provided the additional benefit of minimizing bias between groups.

The implication of this study is that teams using Hark in clinical practice can transfer information more effectively so that clinicians can potentially prioritize their patients using objective, physiological parameters. This may result in fewer treatment delays and prevent avoidable harm [39-41]. Future research should aim to assess how the ABCS performs in a clinical environment when in use as a principal modality of communication and should include analysis of whether use of Hark has an impact on error rates and avoidable adverse events. Using the ABCS first in a pilot ward and then expanding from there may facilitate these efforts.

Conclusions

This study has investigated an alternative to the pager system for information transfer and task management. Hark has been demonstrated to improve the quality of information transfer and has been rated by users as more effective on several important measures, without any reduction in user satisfaction in a simulated environment. Using this device resulted in no delay in patient care. As one of the vital components of safe clinical care, improving interprofessional communication is a priority, and systems such as Hark can support this aim.

Conflicts of Interest

Drs Johnston, King, Arora, and Darzi are associated with the National Institute for Health Research (NIHR) Imperial Patient Safety Translational Research Centre (grant number 40490).
Hark was recently acquired by Google DeepMind, to which Dr King is affiliated.
The views expressed are those of the author(s). The authors declare no other conflicts of interest or other sources of funding for this work. This paper presents independent research supported by the NIHR Imperial Patient Safety Translational Research Centre. The views expressed are those of the authors and not necessarily those of the National Health Service, the NIHR, or the Department of Health of the United Kingdom.

References

1. Bion JF, Heffner JE. Challenges in the care of the acutely ill. Lancet 2004 Mar 20;363(9413):970-977. [doi: 10.1016/S0140-6736(04)15793-0] [Medline: 15043966]
2. Sutcliffe KM, Lewton E, Rosenthal MM. Communication failures: an insidious contributor to medical mishaps. Acad Med 2004 Feb;79(2):186-194. [Medline: 14744724]
3. Pucher P, Johnston M, Aggarwal R, Arora S, Darzi A. Effectiveness of interventions to improve patient handover in surgery: a systematic review. Surgery 2015 Jul;158(1):85-95. [doi: 10.1016/j.surg.2015.02.017] [Medline: 25999255]
4. Solet DJ, Norvell JM, Rutan GH, Frankel RM. Lost in translation: challenges and opportunities in physician-to-physician communication during patient handoffs. Acad Med 2005 Dec;80(12):1094-1099. [Medline: 16306279]
5. Greenberg CC, Regenbogen SE, Studdert DM, Lipsitz SR, Rogers SO, Zinner MJ, et al. Patterns of communication breakdowns resulting in injury to surgical patients. J Am Coll Surg 2007 Apr;204(4):533-540. [doi: 10.1016/j.jamcollsurg.2007.01.010] [Medline: 17382211]
6. Joint Commission. Sentinel Event Data Summary. 2015. URL: http://www.jointcommission.org/sentinel_event_statistics_quarterly/ [accessed 2016-03-23] [WebCite Cache ID 6gEKFS9J6]
7. Smith F. Who invented that bleeping thing? BMJ 2003 Sep;327:719. [doi: 10.1136/bmj.327.7417.719]
8. Wu RC, Lo V, Morra D, Wong BM, Sargeant R, Locke K, et al. The intended and unintended consequences of communication systems on general internal medicine inpatient care delivery: a prospective observational case study of five teaching hospitals. J Am Med Inform Assoc 2013;20(4):766-777 [FREE Full text] [doi: 10.1136/amiajnl-2012-001160] [Medline: 23355461]
9. Coiera E, Tombs V. Communication behaviours in a hospital setting: an observational study. BMJ 1998 Feb 28;316(7132):673-676 [FREE Full text] [Medline: 9522794]
10. Espino S, Cox D, Kaplan B. Alphanumeric paging: a potential source of problems in patient care and communication. J Surg Educ 2011;68(6):447-451. [doi: 10.1016/j.jsurg.2011.07.006] [Medline: 22000529]

11. Johnston MJ, King D, Arora S, Cooper K, Panda NA, Gosling R, et al. Requirements of a new communication technology for handover and the escalation of patient care: a multi-stakeholder analysis. J Eval Clin Pract 2014 Aug;20(4):486-497. [doi: 10.1111/jep.12186] [Medline: 24902627]
12. Reason J. Human Error. Cambridge: Cambridge University Press; 1990.
13. Baddeley AD. Working Memory. Oxford: Oxford University Press; 1986.
14. Wu R, Tran K, Lo V, O'Leary KJ, Morra D, Quan S, et al. Effects of clinical communication interventions in hospitals: a systematic review of information and communication technology adoptions for improved communication between clinicians. Int J Med Inform 2012 Nov;81(11):723-732. [doi: 10.1016/j.ijmedinf.2012.05.014] [Medline: 22727613]
15. Przybylo JA, Wang A, Loftus P, Evans K, Chu I, Shieh L. Smarter hospital communication: secure smartphone text messaging improves provider satisfaction and perception of efficacy, workflow. J Hosp Med 2014 Sep;9(9):573-578 [FREE Full text] [doi: 10.1002/jhm.2228] [Medline: 25110991]
16. Seddon M, Hay D. Task Manager: an innovative approach to improving hospital communication after hours. N Z Med J 2010 Oct 15;123(1324):57-66. [Medline: 20953223]
17. Locke KA, Duffey-Rosenstein B, De LG, Morra D, Hariton N. Beyond paging: building a web-based communication tool for nurses and physicians. J Gen Intern Med 2009 Jan;24(1):105-110 [FREE Full text] [doi: 10.1007/s11606-008-0827-2] [Medline: 18958533]
18. O'Connor C, Friedrich J, Scales D, Adhikari Neill K J. The use of wireless e-mail to improve healthcare team communication. J Am Med Inform Assoc 2009;16(5):705-713 [FREE Full text] [doi: 10.1197/jamia.M2299] [Medline: 19567803]
19. Rogers EM. Diffusion of Innovations. New York, NY: The Free Press; 1962.
20. Putzer G, Park Y. Are physicians likely to adopt emerging mobile technologies? Attitudes and innovation factors affecting smartphone use in the Southeastern United States. Perspect Health Inf Manag 2012;9:1b [FREE Full text] [Medline: 22737094]
21. Khanna V, Sambandam S, Gul A, Mounasamy V. "WhatsApp"ening in orthopedic care: a concise report from a 300-bedded tertiary care teaching center. Eur J Orthop Surg Traumatol 2015 Jul;25(5):821-826. [doi: 10.1007/s00590-015-1600-y] [Medline: 25633127]
22. Johnston MJ, King D, Arora S, Behar N, Athanasiou T, Sevdalis N, et al. Smartphones let surgeons know WhatsApp: an analysis of communication in emergency surgical teams. Am J Surg 2015 Jan;209(1):45-51. [doi: 10.1016/j.amjsurg.2014.08.030] [Medline: 25454952]
23. Joseph B, Pandit V, Khreiss M, Aziz H, Kulvatunyou N, Tang A, et al. Improving communication in level 1 trauma centers: replacing pagers with smartphones. Telemed J E Health 2013 Mar;19(3):150-154. [doi: 10.1089/tmj.2012.0114] [Medline: 23384333]
24. Lo V, Wu RC, Morra D, Lee L, Reeves S. The use of smartphones in general and internal medicine units: a boon or a bane to the promotion of interprofessional collaboration? J Interprof Care 2012 Jul;26(4):276-282. [doi: 10.3109/13561820.2012.663013] [Medline: 22482742]
25. Johnston M, Arora S, King D, Bouras G, Almoudaris A, Davis R, et al. A systematic review to identify the factors that affect failure to rescue and escalation of care in surgery. Surgery 2015 Apr;157(4):752-763. [doi: 10.1016/j.surg.2014.10.017] [Medline: 25794627]
26. Johnston MJ, Arora S, Pucher PH, Reissis Y, Hull L, Huddy JR, et al. Improving Escalation of Care: Development and Validation of the Quality of Information Transfer Tool. Ann Surg 2016 Mar;263(3):477-486. [doi: 10.1097/SLA.0000000000001164] [Medline: 25775058]
27. Johnston M, Mobasheri M, King D, Darzi A. The Imperial Clarify, Design and Evaluate (CDE) approach to mHealth app development. BMJ Innov 2015 Mar 31;1(2):39-42 [FREE Full text] [doi: 10.1136/bmjinnov-2014-000020]
28. Pucher P, Aggarwal R, Qurashi M, Singh P, Darzi A. Randomized clinical trial of the impact of surgical ward-care checklists on postoperative care in a simulated environment. Br J Surg 2014 Dec;101(13):1666-1673. [doi: 10.1002/bjs.9654] [Medline: 25350855]
29. Pucher PH, Aggarwal R, Srisatkunam T, Darzi A. Validation of the simulated ward environment for assessment of ward-based surgical care. Ann Surg 2014 Feb;259(2):215-221. [doi: 10.1097/SLA.0b013e318288e1d4] [Medline: 23470580]
30. Bonrath EM, Fritz M, Mees ST, Weber BK, Grantcharov TP, Senninger N, et al. Laparoscopic simulation training: does timing impact the quality of skills acquisition? Surg Endosc 2013 Mar;27(3):888-894. [doi: 10.1007/s00464-012-2529-4] [Medline: 23052509]
31. Berg R, Inaba K, Sullivan M, Okoye O, Siboni S, Minneti M, et al. The impact of heat stress on operative performance and cognitive function during simulated laparoscopic operative tasks. Surgery 2015 Jan;157(1):87-95. [doi: 10.1016/j.surg.2014.06.012] [Medline: 25482467]
32. Calatayud D, Arora S, Aggarwal R, Kruglikova I, Schulze S, Funch-Jensen P, et al. Warm-up in a virtual reality environment improves performance in the operating room. Ann Surg 2010 Jun;251(6):1181-1185. [doi: 10.1097/SLA.0b013e3181deb630] [Medline: 20485133]
33. King D, Darzi AL, Heys S, Thomsit M. Hark: Transformative Clinical Task Management. London UK: Imperial College London; 2016. URL: http://www.harkapp.co.uk/ [accessed 2016-03-23] [WebCite Cache ID 6gEMmY1aE]

34. Ehsanullah J, Ahmad U, Solanki K, Healy J, Kadoglou N. The surgical admissions proforma: Does it make a difference? Ann Med Surg (Lond) 2015 Mar;4(1):53-57 [FREE Full text] [doi: 10.1016/j.amsu.2015.01.004] [Medline: 25750727]
35. Scholder Ellen P, Bearden WO, Sharma S. Resistance to technological innovations: an examination of the role of self-efficacy and performance satisfaction. J Acad Marketing Sci 1991;19(4):297-307.
36. Wu R, Lo V, Morra D, Appel E, Arany T, Curiale B, et al. A smartphone-enabled communication system to improve hospital communication: usage and perceptions of medical trainees and nurses on general internal medicine wards. J Hosp Med 2015 Feb;10(2):83-89. [doi: 10.1002/jhm.2278] [Medline: 25352429]
37. Wu R, Appel L, Morra D, Lo V, Kitto S, Quan S. Short message service or disService: issues with text messaging in a complex medical environment. Int J Med Inform 2014 Apr;83(4):278-284. [doi: 10.1016/j.ijmedinf.2014.01.003] [Medline: 24495802]
38. Payne KB, Wharrad H, Watts K. Smartphone and medical related App use among medical students and junior doctors in the United Kingdom (UK): a regional survey. BMC Med Inform Decis Mak 2012;12:121 [FREE Full text] [doi: 10.1186/1472-6947-12-121] [Medline: 23110712]
39. Downey AW, Quach JL, Haase M, Haase-Fielitz A, Jones D, Bellomo R. Characteristics and outcomes of patients receiving a medical emergency team review for acute change in conscious state or arrhythmias. Crit Care Med 2008 Feb;36(2):477-481. [doi: 10.1097/01.CCM.0000300277.41113.46] [Medline: 18091535]
40. Ghaferi AA, Birkmeyer JD, Dimick JB. Complications, failure to rescue, and mortality with major inpatient surgery in medicare patients. Ann Surg 2009 Dec;250(6):1029-1034. [Medline: 19953723]
41. Ghaferi AA, Birkmeyer JD, Dimick JB. Hospital volume and failure to rescue with high-risk surgery. Med Care 2011 Dec;49(12):1076-1081. [doi: 10.1097/MLR.0b013e3182329b97] [Medline: 22002649]

Abbreviations

ABCS: app-based communication system
QUIT: QUality of Information Transfer Tool
SMS: short message service

Edited by G Eysenbach; submitted 23.06.15; peer-reviewed by R Wu, D Vollmer dahlke, T Poder; comments to author 19.10.15; revised version received 05.12.15; accepted 04.01.16; published 06.04.16.

Please cite as:
Patel B, Johnston M, Cookson N, King D, Arora S, Darzi A
Interprofessional Communication of Clinicians Using a Mobile Phone App: A Randomized Crossover Trial Using Simulated Patients
J Med Internet Res 2016;18(4):e79
URL: http://www.jmir.org/2016/4/e79/
doi:10.2196/jmir.4854
PMID:27052694

Original Paper

Comparison Between a Self-Administered and Supervised Version of a Web-Based Cognitive Test Battery: Results From the NutriNet-Santé Cohort Study

Karen E Assmann[1], MPH, MSc; Marion Bailet[1], MSc; Amandine C Lecoffre[1], MSc; Pilar Galan[1], MD, PhD; Serge Hercberg[1,2], MD, PhD; Hélène Amieva[3], PhD; Emmanuelle Kesse-Guyot[1], PhD

[1]Université Paris 13, Equipe de Recherche en Epidémiologie Nutritionnelle (EREN), Centre d'Epidémiologie et Statistiques Sorbonne Paris Cité, Inserm (U1153), Inra (U1125), Cnam, COMUE Sorbonne Paris Cité, Bobigny, France

[2]Département de Santé Publique, Hôpital Avicenne, Bobigny, France

[3]Inserm U897, Université Bordeaux, Bordeaux, France

Corresponding Author:
Karen E Assmann, MPH, MSc
Université Paris 13, Equipe de Recherche en Epidémiologie Nutritionnelle (EREN)
Centre d'Epidémiologie et Statistiques Sorbonne Paris Cité
Inserm (U1153), Inra (U1125), Cnam, COMUE Sorbonne Paris Cité
UFR SMBH
74 rue Marcel Cachin
Bobigny, 93017
France
Phone: 33 1 48 38 73 78
Fax: 33 1 48 38 89 31
Email: k.assmann@eren.smbh.univ-paris13.fr

Abstract

Background: Dementia is a major public health problem, and repeated cognitive data from large epidemiological studies could help to develop efficient measures of early prevention. Data collection by self-administered online tools could drastically reduce the logistical and financial burden of such large-scale investigations. In this context, it is important to obtain data concerning the comparability of such new online tools with traditional, supervised modes of cognitive assessment.

Objective: Our objective was to compare self-administration of the Web-based NutriNet-Santé cognitive test battery (NutriCog) with administration by a neuropsychologist.

Methods: The test battery included four tests, measuring, among others aspects, psychomotor speed, attention, executive function, episodic memory, working memory, and associative memory. Both versions of the cognitive battery were completed by 189 volunteers (either self-administered version first, n=99, or supervised version first, n=90). Subjects also completed a satisfaction questionnaire. Concordance was assessed by Spearman correlation.

Results: Agreement between both versions varied according to the investigated cognitive task and outcome variable. Spearman correlations ranged between .42 and .73. Moreover, a majority of participants responded that they “absolutely” or “rather” agreed that the duration of the self-administered battery was acceptable (184/185, 99.5%), that the tasks were amusing (162/185, 87.6%), that the instructions were sufficiently detailed (168/185; 90.8%) and understandable (164/185, 88.7%), and that they had overall enjoyed the test battery (182/185, 98.4%).

Conclusions: The self-administered version of the Web-based NutriCog cognitive test battery provided similar information as the supervised version. Thus, integrating repeated cognitive evaluations into large cohorts via the implementation of self-administered online versions of traditional test batteries appears to be feasible.

(J Med Internet Res 2016;18(4):e68) doi:10.2196/jmir.4862

KEYWORDS

cognition; executive function; internet; cohort studies

Introduction

More than one third of the population in more developed world regions will be aged 60 years and older by the year 2050, according to estimates published in a United Nations report [1]. This population aging is taking place worldwide and is associated with a significant increase in the burden of age-related cognitive decline [2]. According to the 2009 World Alzheimer Report [3], the number of dementia cases is estimated to double every 20 years. Given the major social and economic consequences of dementia, age-related cognitive decline is one of the key public health problems of our time.

Current evidence consistently indicates that Alzheimer's disease diagnoses are preceded by a very long pre-dementia phase [4-8]. Moreover, various studies suggest that this phase begins relatively early in life, indicating that studies on determinants of cognitive aging should focus on midlife factors [9-15]. Hence, longitudinal studies with available data on midlife risk factors and with repeated cognitive data are well adapted for the investigation of potential targets for early preventive measures [11]. Yet, the high logistical and personnel-related costs of traditional neuropsychological evaluations (ie, neuropsychological interviews or paper-and-pencil tests with supervision by trained personnel [16]) present major feasibility problems for epidemiological studies with large sample sizes. Thus, there has been a growing interest in the development of computerized cognitive evaluations as these allow for standardized, repeated procedures, systematic scoring, and automated saving or data management [17,18].

Two recent reviews of the literature have counted a total of 13 different computerized cognitive test batteries that are adapted to the context of age-related cognitive impairment or dementia [17,18]. Of these, only four are fully self-administered: the Computer-Administered Neuropsychological Screen for Mild Cognitive Impairment (CANS-MCI) battery, the Central Nervous System Vital Signs (CNS Vital Signs) battery, the MicroCog: Assessment of Cognitive Functioning (MicroCog) battery, and the Computer Assessment of Mild Cognitive Impairment (CAMCI) battery. One further battery, the Cambridge Neuropsychological Test Automated Battery (CANTAB) was described as "largely independent of verbal instructions." Such fully self-administered batteries that do not rely on examiners to explain instructions or to supervise the completion of tasks are of interest for multiple reasons. Logistical and personnel costs can be further reduced, effort and time investment for participants are lower (since testing can be done at home), and an integration of cognitive evaluations into large-scale epidemiological studies is possible. However, it is important to assess the concordance of such a fully self-administered mode of administration with a "supervised" mode of administration. A supervised mode of administration is characterized by the presence of a trained examiner who gives test instructions, supervises the completion of the battery, and is available to help in the case of comprehension problems.

The objective of this study was to compare these two types of administration (fully self-administered versus supervised) with respect to a cognitive test battery (ie, NutriCog) that was specifically developed for the French NutriNet-Santé cohort.

Methods

Choice of the Cognitive Tasks to Include in the NutriCog Battery

The selection of the cognitive tasks included in our battery was based on a literature review of available computerized neuropsychological batteries [17,18]. Three important considerations guided our choices. First, the tasks had to be feasible in the context of an entirely Web-based battery (in relation to connection speed limitations, the use of mouse and keyboard instead of stylus pens, etc). Second, given our objective to study long-term age-related cognitive decline starting from midlife, the tasks had to reflect cognitive processes that have been shown to show slight functional changes in early aging (ie, episodic memory, attention/working memory, and highly integrated executive functions) [16,19]. Third, the selected tasks had to be derived from robust, previously validated neuropsychological paradigms. The tasks that were finally chosen to be part of our battery are described in Table 1. Briefly, the first task ("Click") consists of connecting numbers (in ascending order) as fast as possible, the second task ("Maze") consists of discovering a labyrinth route while respecting a certain number of rules, the third task ("Cards") is a matching-to-sample task presented as a card game, and the fourth task ("Marbles") consists of memorizing and recalling the color pattern of several marbles with a specific location on the screen. The specific realizations of all tasks except the Click-task are randomly generated at the beginning of each session and were thus not identical across the two modes of administration (fully self-administered/supervised). For example, participants were confronted with distinct labyrinth routes for each of the two completions of the test battery (fully self-administered/supervised).

Table 1. Description of the cognitive test battery, NutriNet-Santé Study.

Task name	Cognitive processes involved	Task description	Outcome variables
Click	Visual-motor coordination, psychomotor speed, selective attention	Numbers (1-5) are connected in ascending order (as fast as possible). A set of three consecutive trials is completed. The task is similar to part A of the Trail Making Test[a].	Mean time taken (seconds)
Maze (Maze A, Maze B)	Episodic memory, procedural memory, working memory, visuospatial attention, executive function (planning, inhibition, mental flexibility)	A labyrinth path is discovered by the participant, who has to respect several different rules. First, a set of three consecutive trials is completed. Next, after performing the Cards task, the Marbles task, and a supplemental unrelated task (with a duration of a few minutes), the participant is presented with a final (fourth) trial. The task was adapted from the Groton Maze Learning test[b].	Total time taken (seconds); number of total clicks (n); number of total errors (n)—These variables are calculated separately for the initial 3 trials (Maze A) and the final trial (Maze B)
Cards	Working memory, executive function (inhibition, mental flexibility), sustained attention	Matching-to-sample (one-back type) task: cards are continuously presented to the participant, who must decide whether or not a presented card is the same as the one shown just before.	([1/ (incorrect answers+1000)] /time) * 100,000
Marbles	Visual memory, episodic memory, associative memory	The participant is presented with a central marble and several colored peripheral marbles. First, the color patterns of each marble have to be memorized. Second, the colors of the peripheral marbles disappear and the central marble continuously takes on different colors, and the participant has to click on the peripheral marble corresponding to the color shown.	([1/ (incorrect answers+1000)] /time) * 100,000

[a] Details on the Trail Making Test have been published elsewhere [16].

[b] Details on the Groton Maze Learning Test have been published elsewhere [20].

Development of an Operational Version of the NutriCog Battery

In order to obtain our final Web-based instrument, the NutriCog battery, the following working steps were undertaken: (1) adaptation of the task instructions to a self-administered computerized framework, (2) prototype development, (3) pilot testing, and (4) revision of the prototype. Pilot testing consisted of assessing comprehensibility of the instructions, timing, and potential technical glitches. These pilot tests have been conducted in our research institute, within a sample of individuals with varying age, education level, and sex.

Lessons that we have learned during the development of the operational version of our battery include the importance of constructive exchanges within an interdisciplinary team of neuropsychologists, epidemiologists, and computer scientists in order to find solutions that are relevant in terms of neuropsychological paradigms, the planned epidemiological investigations, and feasibility in terms of software development. Moreover, multiple rounds of pilot testing within samples of individuals with diversified characteristics were necessary to identify software errors and comprehensibility problems, and to calibrate the display times, for example, messages during the Maze task or cards during the Cards task.

Design of the Comparison Study

In order to compare the self-administered mode of the NutriCog Web-based cognitive battery to a supervised mode of this same battery, we conducted a comparison study in a subsample of the NutriNet-Santé cohort. Each participant of this subsample was asked to perform the test battery twice (ie, in both the self-administered mode and the supervised mode), with an intermission of about 2 months (mean 73.2, SD 17.2 days) in between the two realizations. In order to account for learning effects that have to be expected when repeatedly administering cognitive tests [16], we randomized participants into two study groups, differing in the order of administration of the test battery version. The respective study groups were named SA-SU (self-administered version first, supervised version second) and SU-SA (supervised version first, self-administered version second). The NutriCog test battery was available on the NutriNet-Santé personal Internet page of each participant. The volunteers were asked to use at least 12-inch monitors for optimal visualization of the tests, to use a mouse, and to be in a quiet place without any disruption.

Prior to administering the battery of cognitive tests, the participants were asked to complete a short questionnaire assessing their current mood. At the end of each round of assessment (ie, self-administered version and supervised version), participants were presented with a process evaluation questionnaire in order to provide feedback on the test battery. The expected time to complete the battery in full was 20-25 minutes (15-20 minutes for the cognitive tests and 5 minutes for the questionnaires). Finally, after the participants had completed both versions of the test battery, they were presented with a satisfaction questionnaire, designed to collect information on the acceptability of the duration of each version of the battery, the difficulty of the tasks, the presentation and comprehensibility of the instructions, and the overall appreciation of the test battery.

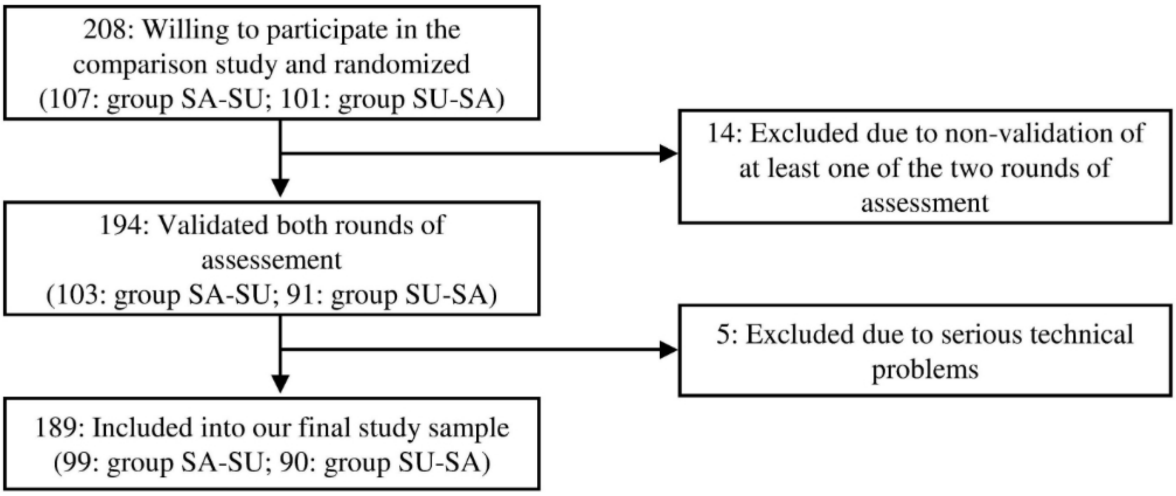
Administration of the Two Versions of the Test Battery

The self-administered version was completed by the volunteers while alone, following the instructions given on the webpage. The supervised version was completed with assistance by a trained neuropsychologist, who made a home-visit appointment with each volunteer. The instructions were given orally by the neuropsychologist, who was also available to answer any upcoming questions concerning the instructions. All administrations of the supervised version of the battery were realized by the same neuropsychologist. There were no other differences between the two assessment rounds.

Selection of Participants for the Comparison Study

The selection process for our study sample is presented in Figure 1. We selected a subsample of 1416 participants of the NutriNet-Santé cohort, via a stratified randomization procedure, using sex, age group (<50 years, ≥50 years), and educational level (≤2 years of higher education, >2 years of higher education) as stratification variables. Among those subjects who agreed to participate, we selected 208 individuals with varying sex, age, and educational level (our pre-defined objective was to obtain a final sample of approximately 200 participants) and randomly attributed them to the two study groups. Of these 208 subjects, 14 were excluded because they did not validate both versions of the test battery, and 5 individuals were excluded because serious technical (ie, computer-related) problems had occurred as they completed the battery. Thus, our final study sample consisted of 189 participants (group SA-SU: n=99; group SU-SA: n=90).

Figure 1. Selection of participants and attribution to the study groups SA-SU (self-administered version first, supervised version second) and SU-SA (self-administered version first, supervised version second).



Statistical Analyses

Creation of Composite Variables

For the Cards and Marbles tasks, composite variables relating the number of incorrect answers to the time taken were created, as both the quality of the responses and the rapidity of task completion are important for the evaluation of performances on these tasks. The following formula was applied in order to relate incorrect answers to time while obtaining readable (ie, not too small) values and improving the normality of the distribution: ([1/(incorrect answers+1000)] /time) *100,000. For these composite variables, higher scores indicate higher performances, while for all other variables, lower scores indicate higher performances.

Descriptive Analysis

Participant characteristics were presented as n and % or p50 and p25, p75 (ie, 50th and 25th, 75th percentile) and differences across study groups were tested via chi-square tests and (nonparametric) Wilcoxon rank-sum tests, respectively. In order to detect potential learning effects, median performances on the first administration and on the second administration of the battery were tabulated and compared via nonparametric paired tests (Wilcoxon signed-rank tests). Furthermore, we tabulated median performances according to test battery version (self-administered, supervised) as p50 and p25, p75. Finally, Spearman correlations between the different cognitive outcomes variables were tabulated.

Concordance Analysis

Concordance between the two versions of the test battery was estimated via Spearman correlation coefficients (presented with their *P* values). We decided to use Spearman correlations as our main statistical indicators in order to facilitate a comparison of our results with other methodological articles on computerized cognitive test batteries, which mostly report Pearson or Spearman correlations [18]. Spearman correlations were more adapted to our data than Pearson correlations as the cognitive outcome variables were not normally distributed.

Spearman correlations were presented for the whole study sample and stratified by age (<50 years, ≥50 years), educational level (≤2 years of post–high school education, >2 years of post–high school education), sex, and Web experience (Web novice, Web expert). Participants reporting that they were “inexperienced” or “beginners” in terms of Internet knowledge were designated as “Web novices” and participants report “advanced” or “expert” Internet knowledge were designated as “Web experts.”

Finally, participants’ responses to the satisfaction questionnaire were tabulated as n and %. All analyses were conducted with

SAS, version 9.3, and all statistical tests were two-sided with a significance level of .05.

Results

Our study sample had a median age of 57 years and 61.9% of the participants were women. Participant characteristics according to study group are presented in Table 2. Individuals in the SA-SU group were younger than participants in the SU-SA group and more frequently reported an advanced or expert level in Internet knowledge. In terms of gender, educational level, occupational category, and type of Internet connection, no significant differences between groups were observed.

Table 3 presents performances on the different cognitive tasks according to time of administration of the cognitive battery (first administration versus second administration). Except for the time taken on the “Click” task and for the “Marbles” task composite variable, the performances on the cognitive tasks were systematically better for the second administration.

For illustrative purposes, performances on cognitive tests according to version of the test battery (self-administrated, supervised) are presented in Multimedia Appendix 1. Multimedia Appendix 2 shows Spearman correlations among the different cognitive test variables for the self-administered version of the test battery.

Table 4 shows Spearman correlation coefficients as measures of concordance between cognitive performances according to version. Spearman correlations ranged between .42 and .73. The lowest correlations were observed for the outcome variables “mean number of clicks” and “mean number of total errors” of the Maze task (Maze A and Maze B): here, Spearman correlation coefficients ranged between .42 and .43.

Multimedia Appendix 3 shows Spearman correlation coefficients stratified by age group, educational level, sex, and Web experience. The coefficients ranged from .21-.80 and tended to be highest among participants who were higher educated and who had better Internet knowledge. However, these trends were not consistent across all outcome variables.

Tables 5 and 6 present the responses of participants to the satisfaction questionnaire. Notably, a majority of participants responded that they “absolutely” or “rather” agreed that the duration of the self-administered battery was acceptable (184/185, 99.5%), that the tasks were amusing (162/185, 87.6%), that the instructions were sufficiently detailed (168/185, 90.8%) and understandable (164/185, 88.7%), and that they had globally enjoyed the test battery (182/185, 98.4%). On the other hand, 81.1% (150/185) of participants reported to have preferred the supervised version of the self-administered version.

Table 2. Participant characteristics (N=189).

	Self-administered version first (n=99)		Supervised version first (n=90)		
	n or p50[a]	% or p25; p75[a]	n or p50[a]	% or p25; p75[a]	*P* [b]
Age	55.0	40.0; 65.0	59.0	51.0; 67.0	.02
Gender					.27
Male	34	34.3	38	42.2	
Female	65	65.7	52	57.8	
Education level					.74
<2 yrs post–high school education	34	34.3	33	36.7	
≥2 yrs post–high school education	65	65.7	57	63.3	
Occupational category					.50
Unemployed	13	13.1	7	7.8	
Employee	8	8.1	10	11.1	
Intermediate profession	12	12.1	8	8.9	
Managerial staff[c]	38	38.4	32	35.6	
Retired	28	28.3	33	36.7	
Self-evaluated Web knowledge [d,e]					.049
Web novice	6	6.1	13	14.9	
Web expert	92	93.9	74	85.1	
Type of connection [d]					.94
<512k	9	9.2	9	10.3	
≥512 and <1024k	18	18.4	16	18.4	
≥1024k	48	49.0	39	44.8	
Do not know	23	23.5	23	26.4	

[a] Values for age are not n and %, but median and 25th; 75th percentile.

[b] *P* value for the difference between both administration order groups, from Wilcoxon rank-sum nonparametric tests for age and chi-square tests for other variables.

[c] Or intellectual profession.

[d] 185 subjects (98 for the self-administered first and 87 for the supervised version first) returned the satisfaction questionnaire, thereby providing this information.

[e] Novice: “inexperienced” or “beginner” level; expert: “advanced” or “expert” level.

Table 3. Performance on cognitive tests: first versus second administration (N=189)[a].

		First administration		Second administration		
Test[b]	Variable	p50	p25; p75	p50	p25; p75	*P* [c]
Click	Time in seconds, mean	4.29	3.64; 5.49	4.32	3.68; 5.36	.70
Maze A						
	Time in seconds, mean	118.66	86.95; 168.67	101.16	79.28; 129.76	<.001
	Clicks, mean n	59.00	53.00; 70.33	54.00	48.33; 60.67	<.001
	Total errors, mean n	17.33	13.33; 25.00	14.00	10.67; 18.00	<.001
Maze B						
	Time in seconds	85.35	62.87; 112.64	72.11	51.72; 92.99	<.001
	Clicks, n	48.00	42.00; 55.00	44.00	37.00; 51.00	<.001
	Total errors, n	10.00	7.00; 15.00	8.00	4.00; 12.00	<.001
Cards	Composite variable[d]	1.26	1.09; 1.36	1.28	1.15; 1.40	<.001
Marbles	Composite variable[d]	2.21	1.48; 3.31	2.40	1.77; 3.33	.18

[a] Objective of this comparison: identification of potential learning effects. Lower scores indicate better performances, except for the Cards and Marbles composite variables, where higher scores indicate better performances.

[b] Maze A: Sum of the initial three rounds of the Maze task. Maze B: Final (fourth) round of the Maze task.

[c] Wilcoxon signed-rank test (nonparametric paired test).

[d] ([1/(incorrect answers+1000)] /time) * 100,000.

Table 4. Concordance of cognitive tests performances according to version (N=189)[a].

Test[b]	Variable	Spearman correlation	
		r	*P*
Click	Time in seconds, mean	.73	<.001
Maze A			
	Time in seconds, mean	.57	<.001
	Clicks, mean n	.43	<.001
	Total errors, mean n	.43	<.001
Maze B			
	Time in seconds	.53	<.001
	Clicks, n	.43	<.001
	Total errors, n	.42	<.001
Cards	Composite variable[c]	.64	<.001
Marbles	Composite variable[c]	.51	<.001

[a] Self-administered version versus supervised version.

[b] Maze A: Sum of the initial three rounds of the Maze task. Maze B: Final (fourth) round of the Maze task.

[c] ([1/ (incorrect answers+1000)] /time) * 100,000.

Table 5. Participant satisfaction with the two versions of the cognitive test battery, questions concerning the self-administered version (N=185)[a].

Questions	n	%
The duration is acceptable		
I absolutely agree	145	78.4
I rather agree	39	21.1
I rather don’t agree	1	0.5
I absolutely don’t agree	0	0.0
The exercises can be rapidly finished		
I absolutely agree	64	34.6
I rather agree	112	60.5
I rather don’t agree	8	4.3
I absolutely don’t agree	1	0.5
The exercises are too simple		
I absolutely agree	4	2.2
I rather agree	89	48.1
I rather don’t agree	87	47.0
I absolutely don’t agree	5	2.7
The exercises are amusing		
I absolutely agree	51	27.6
I rather agree	111	60.0
I rather don’t agree	23	12.4
I absolutely don’t agree	0	0.0
The instructions are sufficiently detailed		
I absolutely agree	80	43.2
I rather agree	88	47.6
I rather don’t agree	17	9.2
I absolutely don’t agree	0	0.0
The instructions are understandable		
I absolutely agree	69	37.3
I rather agree	95	51.4
I rather don’t agree	20	10.8
I absolutely don’t agree	1	0.5
Overall, I have appreciated the test battery		
I absolutely agree	83	44.9
I rather agree	99	53.5
I rather don’t agree	3	1.6
I absolutely don’t agree	0	0.0
I have had problems visualizing the graphics		
I absolutely agree	2	1.1
I rather agree	7	3.8
I rather don’t agree	49	26.5
I absolutely don’t agree	127	68.7
If I was proposed to repeat the test battery, I would agree		
I absolutely agree	155	83.8

Questions	n	%
I rather agree	30	16.2
I rather don’t agree	0	0.0
I absolutely don’t agree	0	0.0

[a] Only 185 out of the 189 participants of the study completed the satisfaction questionnaire.

Table 6. Participant satisfaction with the two versions of the cognitive test battery, questions concerning the supervised version, and comparison of the two versions (N=185)[a].

Questions	n	%
Questions concerning the supervised version		
The duration is acceptable		
I absolutely agree	157	84.9
I rather agree	28	15.1
I rather don’t agree	0	0.0
I absolutely don’t agree	0	0.0
I have appreciated completing the tests with a professional		
I absolutely agree	128	69.2
I rather agree	56	30.3
I rather don’t agree	0	0.0
I absolutely don’t agree	1	0.5
Comparison of the two versions		
Overall, which of the two versions have you preferred?		
The Web version	35	18.9
The supervised version	150	81.1

[a] Only 185 out of the 189 participants of the study completed the satisfaction questionnaire.

Discussion

Principal Findings

The aim of this comparison study was to assess the concordance of the fully self-administered mode of the Web-based NutriCog test battery with a supervised mode of the same battery, characterized by the presence of a trained neuropsychologist. Nine different outcome variables were evaluated to measure the performance of participants on four different tasks (Click, Maze, Cards, and Marbles). Spearman coefficients for the correlation between the two modes of administrations, in terms of cognitive test performances, ranged between .42 and .73. Correlation coefficients were lower for the Maze task (Mazes A and B) than for the other tasks. This could indicate potential problems concerning the comprehensibility of the Maze task instructions in the absence of a supervisor. However, based on suggestions made by the participants and by the neuropsychologist who was present during the supervised version of the test battery, these instructions have been slightly revised (after the comparison study was completed), in order to enhance understanding.

The observed correlation coefficients varied according to educational level and Web knowledge. Correlations tended to be higher among higher educated participants and among those with higher Internet knowledge. However, as can be expected in the context of an entirely Web-based study, there were only very few individuals who reported being “inexperienced” or having a “beginner level” with respect to Web knowledge.

Overall, in the context of this comparison study, very high values of concordance could not be expected due to multiple circumstances. First, we observed a clear learning effect, with test results that were almost systematically better at the second administration of the test battery, independently of the mode of administration. If such learning effects were of similar magnitude for the whole study sample, this would not affect the calculated Spearman correlations. On the other hand, differential learning effects would have lowered the observed concordance values in our sample. Further, cognitive performances are subject to a rather large amount of intra-individual variation [21].

To the best of our knowledge, no other study has yet investigated the concordance of a fully self-administered version of a cognitive test battery with a supervised version of the same test battery. However, the context of our study is similar to studies investigating test-retest reliability (ie, the correlation of performances on a first and a second administration of the same test battery), which have been conducted for multiple computerized cognitive batteries. A systematic review of the literature by Tierney et al identified 11 computerized cognitive test batteries [18], of which nine had information on the test-retest reliability for each subtest. The respective ranges of (Pearson or Spearman) correlation coefficients were .30-.74 (CAMCI) [22], .65-.88 (CANS-MCI) [23], .56-.90 (CNS Vital Signs) [24], .53-.93 (Cognitive Drug Research Computerized Assessment System for Dementia, COGDRAS-D) [25], .23-.79 (Cogstate) [26], .68-.80 (CSI) [27], .59-.98 (short form of the MicroCog battery) [28], and .40-.84 (Mindstreams Mild Impairment Battery) [29]. For the CANTAB battery, intraclass correlation coefficients were reported instead of Pearson or Spearman correlations, with a range of .09-.86 [30]. Of note, in our study, intraclass correlation coefficients (which were calculated after applying transformations to improve normality) ranged from .36-.65 (data not shown).

In this study, lower correlation coefficients than for such test-retest investigations had to be expected as we compared two different modes of administration. Although very high values of ≥.9 for specific subtests (as observed for the CNS Vital Signs, COGDRAS-D, and Microcog batteries) were not obtained in this investigation, the range of correlation coefficients found in our study (.42-.73) is roughly comparable to the ranges found in other studies.

Another element that supports the use of our Web-based cognitive battery in a fully self-administered mode is the fact that a majority (88.7%) of participants evaluated the test instructions as sufficiently detailed and understandable. As stated above, in order to further improve the comprehensibility of the test battery, the instructions concerning the Maze task have been slightly modified by taking into account the suggestions of participants of this comparison study. Besides, 81.1% of participants preferred the supervised version of the battery to the self-administered version. However, this is probably largely due to the fact that social interaction with a health professional was perceived as a more pleasant situation than completing the test battery alone.

Strengths and Limitations

A certain number of limitations to this study have to be considered. First, the observed learning effect between the first and second administration of the test battery is difficult to separate from differences in cognitive performances that are related to the mode of administration. Second, this comparison study aimed only to compare the fully self-administered mode of the battery to a supervised mode of administration. Data on the ability of the battery to accurately discriminate normal cognitive function from impaired cognitive function are not available. An important strength of our study is its originality, as to the best of our knowledge, no other study has yet compared full self-administration to supervised administration of cognitive batteries. Moreover, the study was conducted within a rather large sample of 189 participants of varying age, sex, and educational level. Finally, the concordance between the two modes of administrations was assessed for the whole population as well as for specific subgroups.

Conclusions

The concordance of the self-administered version and the supervised version of the Web-based NutriCog cognitive test battery was roughly similar to that observed for test-retest investigations of other test batteries. This indicates that these two different modes of administration provide similar information. In large epidemiological studies like the NutriNet-Santé cohort, the objective of cognitive evaluations is not to provide data with high validity in a clinical context, but to measure cognitive performances in a rapid and simple manner, with sufficient quality to permit valid conclusions on the population level. Given the drastic reduction of the logistical and financial burden that can be obtained by using fully self-administered tools, Web-based cognitive test batteries such as NutriCog provide interesting alternatives to supervised tools in the context of large cohort studies.

Acknowledgments

The NutriNet-Santé study is supported by the French Ministry of Health, the Institut de Veille Sanitaire, the Institut National de Prévention et d'Education pour la Santé, the Fondation pour la Recherche Médicale, the Institut National de la Santé et de la Recherche Médicale, the Institut National de la Recherche Agronomique, the Conservatoire National des Arts et Métiers, and the Paris 13 University.

None of the funding bodies had any involvement in study design, in the collection, analysis and interpretation of the data, in writing the manuscript, or the decision to submit.

KEA was supported by a doctoral fellowship from the University of Paris 13.

We thank all the scientists, dietitians, technicians, and assistants who helped carry out the NutriNet-Santé study. We especially thank the computer scientists, Younes Esseddik, Yasmina Chelghoum, Mohand Ait Oufella, Paul Flanzy, and Thi Hong Van Duong, and the statisticians, Veronique Gourlet, Charlie Menard, Fabien Szabo, Nathalie Arnault, Laurent Bourhis, and Stephen Besseau. We are grateful to the volunteers of the NutriNet-Santé study.

Authors' Contributions

EKG and HA designed the comparison study (development of the concept, design and protocol of the study and coordination of data collection) and the investigated cognitive test battery (choice of tests to be included, coordination of the technical development, and the pilot testing). MB was responsible for the practical coordination and realization of the comparison study. KEA performed the statistical analysis and wrote the article. EKG, HA, and MB provided methodological guidance. SH is the principal investigator of the NutriNet-Santé study. KEA, MB, ACL, PG, SG, HA, and EKG were involved in interpreting the results and editing the manuscript for important intellectual content. EKG, SH and HA had full responsibility for the final content, including the presentation of the NutriCog test battery and its development. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Performance on cognitive tests according to version.

[PDF File (Adobe PDF File), 31KB - jmir_v18i4e68_app1.pdf]

Multimedia Appendix 2

Spearman correlations between the different cognitive test variables for the self-administered version of the cognitive test battery NutriCog.

[PDF File (Adobe PDF File), 37KB - jmir_v18i4e68_app2.pdf]

Multimedia Appendix 3

Concordance of cognitive tests performances according to version: Spearman correlation coefficients, stratified by participant characteristics.

[PDF File (Adobe PDF File), 136KB - jmir_v18i4e68_app3.pdf]

References

1. Catalogue of United Nations Population Division Publications, Databases and Software 2003.: United Nations; 2002. World Population Ageing: 1950-2050, Chapter III: The changing balance between age groups URL: http://www.un.org/esa/population/publications/worldageing19502050/ [accessed 2016-03-07] [WebCite Cache ID 6ZYD7Ob0T]
2. McGuire LC, Ford ES, Ajani UA. Cognitive functioning as a predictor of functional disability in later life. Am J Geriatr Psychiatry 2006 Jan;14(1):36-42. [doi: 10.1097/01.JGP.0000192502.10692.d6] [Medline: 16407580]
3. Alzheimer's Disease International. 2009. World Alzheimer Report URL: http://www.alz.co.uk/research/files/WorldAlzheimerReport.pdf [accessed 2016-03-07] [WebCite Cache ID 6ZYDNprSM]
4. Mortimer JA, Borenstein AR, Gosche KM, Snowdon DA. Very early detection of Alzheimer neuropathology and the role of brain reserve in modifying its clinical expression. J Geriatr Psychiatry Neurol 2005 Dec;18(4):218-223 [FREE Full text] [doi: 10.1177/0891988705281869] [Medline: 16306243]
5. Amieva H, Le GM, Millet X, Orgogozo JM, Pérès K, Barberger-Gateau P, et al. Prodromal Alzheimer's disease: successive emergence of the clinical symptoms. Ann Neurol 2008 Nov;64(5):492-498. [doi: 10.1002/ana.21509] [Medline: 19067364]
6. Amieva H, Jacqmin-Gadda H, Orgogozo J, Le CN, Helmer C, Letenneur L, et al. The 9 year cognitive decline before dementia of the Alzheimer type: a prospective population-based study. Brain 2005 May;128(Pt 5):1093-1101 [FREE Full text] [doi: 10.1093/brain/awh451] [Medline: 15774508]
7. DeKosky ST, Marek K. Looking backward to move forward: early detection of neurodegenerative disorders. Science 2003 Oct 31;302(5646):830-834. [doi: 10.1126/science.1090349] [Medline: 14593169]
8. Linn RT, Wolf PA, Bachman DL, Knoefel JE, Cobb JL, Belanger AJ, et al. The 'preclinical phase' of probable Alzheimer's disease. A 13-year prospective study of the Framingham cohort. Arch Neurol 1995 May;52(5):485-490. [Medline: 7733843]
9. Coley N, Andrieu S, Gardette V, Gillette-Guyonnet S, Sanz C, Vellas B, et al. Dementia prevention: methodological explanations for inconsistent results. Epidemiol Rev 2008;30:35-66. [doi: 10.1093/epirev/mxn010] [Medline: 18779228]
10. Middleton LE, Yaffe K. Promising strategies for the prevention of dementia. Arch Neurol 2009 Oct;66(10):1210-1215 [FREE Full text] [doi: 10.1001/archneurol.2009.201] [Medline: 19822776]
11. Singh-Manoux A, Kivimäki M. The importance of cognitive aging for understanding dementia. Age (Dordr) 2010 Dec;32(4):509-512 [FREE Full text] [doi: 10.1007/s11357-010-9147-7] [Medline: 20454932]
12. Lee KS, Eom J, Cheong H, Oh BH, Hong CH. Effects of head circumference and metabolic syndrome on cognitive decline. Gerontology 2010;56(1):32-38. [doi: 10.1159/000236028] [Medline: 19729879]
13. Fratiglioni L, Paillard-Borg S, Winblad B. An active and socially integrated lifestyle in late life might protect against dementia. Lancet Neurol 2004 Jun;3(6):343-353. [doi: 10.1016/S1474-4422(04)00767-7] [Medline: 15157849]
14. Hughes T, Ganguli M. Modifiable midlife risk factors for late-life dementia. Rev Neurol 2010 Sep 1;51(5):259-262 [FREE Full text] [Medline: 20669124]

15. Rooney RFE. Preventing dementia: how lifestyle in midlife affects risk. Curr Opin Psychiatry 2014 Mar;27(2):149-157. [doi: 10.1097/YCO.0000000000000045] [Medline: 24441418]
16. Lezak M, Howieson D, Bigler ED, Tranel DD. Neuropsychological Assessment. New York: Oxford University Press; 2004.
17. Wild K, Howieson D, Webbe F, Seelye A, Kaye J. Status of computerized cognitive testing in aging: a systematic review. Alzheimers Dement 2008 Nov;4(6):428-437 [FREE Full text] [doi: 10.1016/j.jalz.2008.07.003] [Medline: 19012868]
18. Tierney MC, Lermer MA. Computerized cognitive assessment in primary care to identify patients with suspected cognitive impairment. J Alzheimers Dis 2010;20(3):823-832 [FREE Full text] [doi: 10.3233/JAD-2010-091672]
19. Depp C, Harmell A, Vahia I. Successful cognitive aging. Curr Top Behav Neurosci 2012;10:35-50. [doi: 10.1007/7854_2011_158] [Medline: 22275006]
20. Pietrzak RH, Maruff P, Mayes LC, Roman SA, Sosa JA, Snyder PJ. An examination of the construct validity and factor structure of the Groton Maze Learning Test, a new measure of spatial working memory, learning efficiency, and error monitoring. Arch Clin Neuropsychol 2008 Jul;23(4):433-445 [FREE Full text] [doi: 10.1016/j.acn.2008.03.002] [Medline: 18448309]
21. Salthouse TA, Nesselroade JR, Berish DE. Short-term variability in cognitive performance and the calibration of longitudinal change. J Gerontol B Psychol Sci Soc Sci 2006 May;61(3):P144-P151 [FREE Full text] [Medline: 16670183]
22. Saxton J, Morrow L, Eschman A, Archer G, Luther J, Zuccolotto A. Computer assessment of mild cognitive impairment. Postgrad Med 2009 Mar;121(2):177-185 [FREE Full text] [doi: 10.3810/pgm.2009.03.1990] [Medline: 19332976]
23. Tornatore JB, Hill E, Laboff JA, McGann ME. Self-administered screening for mild cognitive impairment: initial validation of a computerized test battery. J Neuropsychiatry Clin Neurosci 2005;17(1):98-105 [FREE Full text] [doi: 10.1176/jnp.17.1.98] [Medline: 15746489]
24. Gualtieri CT, Johnson LG. Reliability and validity of a computerized neurocognitive test battery, CNS Vital Signs. Arch Clin Neuropsychol 2006 Oct;21(7):623-643 [FREE Full text] [doi: 10.1016/j.acn.2006.05.007] [Medline: 17014981]
25. Simpson P, Surmon D, Wesnes K, Wilcock G. The cognitive drug research computerized assessment system for demented patients: A validation study. Int J Geriat Psychiatry 1991 Feb;6(2):95-102. [doi: 10.1002/gps.930060208]
26. Collie A, Maruff P, Darby DG, McStephen M. The effects of practice on the cognitive test performance of neurologically normal individuals assessed at brief test-retest intervals. J Int Neuropsychol Soc 2003 Mar;9(3):419-428. [doi: 10.1017/S1355617703930074] [Medline: 12666766]
27. Erlanger DM, Kaushik T, Broshek D, Freeman J, Feldman D, Festa J. Development and validation of a web-based screening tool for monitoring cognitive status. J Head Trauma Rehabil 2002 Oct;17(5):458-476. [Medline: 12802255]
28. Elwood RW. MicroCog: assessment of cognitive functioning. Neuropsychol Rev 2001 Jun;11(2):89-100. [Medline: 11572473]
29. Schweiger A, Doniger GM, Dwolatzky T, Jaffe D, Simon ES. Reliability of a novel computerized neuropsychological battery for mild cognitive impairment. Acta Neuropsychologica 2003;1(4):407-413.
30. Lowe C, Rabbitt P. Test/re-test reliability of the CANTAB and ISPOCD neuropsychological batteries: theoretical and practical issues. Cambridge Neuropsychological Test Automated Battery. International Study of Post-Operative Cognitive Dysfunction. Neuropsychologia 1998 Sep;36(9):915-923. [Medline: 9740364]

Abbreviations

CAMCI: Computer Assessment of Mild Cognitive Impairment
CANS-MCI: Computer-Administered Neuropsychological Screen for Mild Cognitive Impairment
CANTAB: Cambridge Neuropsychological Test Automated Battery
CNS Vital Signs: Central Nervous System Vital Signs
COGDRAS-D: Cognitive Drug Research Computerized Assessment System for Dementia
MicroCog: MicroCog: Assessment of Cognitive Functioning
SA-SU: self-administered version first, supervised version second
SU-SA: supervised version first, self-administered version second

Edited by G Eysenbach; submitted 01.07.15; peer-reviewed by R Pamphlett, J Parkin Kullmann; comments to author 29.07.15; revised version received 12.08.15; accepted 15.08.15; published 05.04.16.

Please cite as:
Assmann KE, Bailet M, Lecoffre AC, Galan P, Hercberg S, Amieva H, Kesse-Guyot E
Comparison Between a Self-Administered and Supervised Version of a Web-Based Cognitive Test Battery: Results From the NutriNet-Santé Cohort Study
J Med Internet Res 2016;18(4):e68
URL: http://www.jmir.org/2016/4/e68/
doi:10.2196/jmir.4862
PMID:27049114

Original Paper

Awareness and Use of the After-Visit Summary Through a Patient Portal: Evaluation of Patient Characteristics and an Application of the Theory of Planned Behavior

Srinivas Emani[1], MA, PhD; Michael Healey[1,2], MD; David Y Ting[3], MD; Stuart R Lipsitz[1], ScD; Harley Ramelson[1,4], MD; Vladimir Suric[1], BA; David W Bates[1,5], MSc, MD

[1]Division of General Internal Medicine, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, United States
[2]Brigham and Women's Physicians Organization, Brigham and Women's Hospital, Boston, MA, United States
[3]Massachusetts General Physicians Organization, Massachusetts General Hospital, Boston, MA, United States
[4]Information Services, Partners HealthCare, Boston, MA, United States
[5]Department of Healthcare Policy and Management, Harvard School of Public Health, Boston, MA, United States

Corresponding Author:
Srinivas Emani, MA, PhD
Division of General Internal Medicine, Department of Medicine
Brigham and Women's Hospital
Harvard Medical School
1620 Tremont Street, 3rd Floor
Boston, MA, 02120
United States
Phone: 1 617 732 7063
Fax: 1 617 732 7072
Email: semani1@partners.org

Abstract

Background: Patient portals are being used to provide a clinical summary of the office visit or the after-visit summary (AVS) to patients. There has been relatively little research on the characteristics of patients who access the AVS through a patient portal and their beliefs about the AVS.

Objective: The aim was to (1) assess the characteristics of patients who are aware of and access the AVS through a patient portal and (2) apply the Theory of Planned Behavior (TPB) to predict behavioral intention of patients toward accessing the AVS provided through a patient portal.

Methods: We developed a survey capturing the components of TPB (beliefs, attitude, perceived norm, and perceived behavioral control). Over a 6-month period, patients with a patient portal account with an office visit in the previous week were identified using our organization’s scheduling system. These patients were sent an email about the study and a link to the survey via their portal account. We applied univariate statistical analysis (Pearson chi-square and 1-way ANOVA) to assess differences among groups (aware/unaware of AVS and accessed/did not access AVS). We reported means and standard deviations to depict belief strengths and presented correlations between beliefs and attitude, perceived norm, and perceived behavioral control. We used hierarchical regression analysis to predict behavioral intention toward accessing the AVS through the patient portal.

Results: Of the 23,336 patients who were sent the survey, 5370 responded for a response rate of 23.01%. Overall, 76.52% (4109/5370) were aware that the AVS was available through the patient portal and 54.71% of those (2248/4109) accessed the AVS within 5 days of the office visit. Patients who accessed the AVS had a greater number of sessions with the portal (mean 119, SD 221.5) than those who did not access the AVS (mean 79.1, SD 123.3, $P<.001$); the difference was not significant for awareness of the AVS. The strongest behavioral beliefs with accessing the AVS were being able to track visits and tests (mean 2.53, SD 1.00) followed by having medical information more readily accessible (mean 2.48, SD 1.07). In all, 56.7% of the variance in intention to access the AVS through the portal was accounted for by attitude, perceived norm, and perceived behavioral control.

Conclusions: Most users of a patient portal were aware that the AVS was accessible through the portal. Patients had stronger beliefs about accessing the AVS with the goal of timely and efficient access of information than with engaging in their health care. Interventions to improve patient access of the AVS can focus on providers promoting patient beliefs about the value of the AVS for tracking tests and visits, and timely and efficient access of information.

(J Med Internet Res 2016;18(4):e77) doi:10.2196/jmir.5207

KEYWORDS

patient portal; after-visit summary (AVS); meaningful use; electronic health records (EHRs); beliefs

Introduction

The adoption and use of patient portals tethered to electronic health records (EHRs) has accelerated in the last decade. A primary driver of this growth has been the Medicare and Medicaid EHR Incentive Program, widely referred to as the EHR Meaningful Use (MU) program, introduced in the Health Information Technology for Economic and Clinical Health (HITECH) provision of the American Recovery and Reinvestment Act of 2009 [1,2]. The objectives of the MU program are to increase the adoption of EHRs and the meaningful use of EHRs to improve delivery of care, decrease medical errors, improve efficiency of care, and enhance patient centeredness of care [2]. The MU program is being implemented in three stages with the criteria for achieving meaningful use of the EHR becoming more rigorous with each stage. Patient portals are expected to play a key role in the MU program by providing patients with timely and efficient access to information, engaging patients in their care, and enhancing patient centeredness of care [3,4]. One of the core objectives of the MU program is to allow patients to view online and download their health information, such as test results, problem and medication lists, and medication allergies. For example, the Blue Button initiative has been implemented by a number of organizations to allow patients to download a copy of their health information by clicking on a blue circle on the patient portal page [5].

Patient portals are also being used to provide a clinical summary of the office visit or the after-visit summary (AVS) to patients. The Centers for Medicare and Medicaid Services (CMS) has defined the AVS as a clinical summary that “provides a patient with relevant and actionable information and instructions” such as the provider’s office contact information, date and location of visit, an updated medication list, updated vitals, reason(s) for visit, procedures and other instructions based on clinical discussions that took place during the office visit, any updates to a problem list, immunizations or medications administered during visit, summary of topics covered/considered during visit, and time and location of next appointment/testing, if scheduled [6]. Stage 1 of the MU program specified that the AVS should be provided to patients for more than 50% of all office visits within three business days.

The AVS requirement was controversial in the physician community and, in spite of the Stage 1 requirement of the MU program for the provision of the AVS, there has been relatively little research on how patients view the AVS. In a survey of the printed version of the AVS provided at an office visit, Neuberger and colleagues [7] reported that 88% of respondents said the information on the AVS was easy to understand and 84% said that the AVS was helpful. Chung and colleagues [8] reported similar results in their survey of a printed version of the AVS: 93% of patients agreed that they understood the information on the AVS and 93% agreed that having the AVS was helpful. Ralston and colleagues [9] reported that the AVS was the fastest growing use of their organization’s patient portal and may reflect the patient’s desire for information about their care plan and needs. Pavlik and colleagues [10] found patient satisfaction with the MU version of the AVS did not differ significantly from other content versions of the AVS. In this study, 30% of the patients reported that they plan to keep the AVS for their next appointment.

Although these studies provide some understanding of patient opinions about the AVS, we know relatively little about predictors of patient access to the AVS through the patient portal, such as do younger and more highly educated patients access (ie, retrieve) the AVS more through the patient portal compared to older and less educated patients, and what role does patient experience with the patient portal play in their accessing the AVS through the portal? There is also a lack of evidence on patient beliefs and attitude toward accessing the AVS. Do patients believe that accessing the AVS through the patient portal will provide information in a timely manner, allow them to track their visits and tests, and reinforce their provider’s instructions? This study contributes to the sparse literature on patient portals and the AVS by addressing the following objectives: (1) assess the characteristics of patients who are aware of and access (retrieve) the AVS through a patient portal and (2) apply the Theory of Planned Behavior (TPB) to evaluate beliefs, attitude, perceived norm, perceived behavioral control, and predict behavioral intention of patients toward accessing the AVS provided through a patient portal.

Methods

Theoretical Model

In a previous study, we pointed to the lack of application of theoretical models in the study of patient adoption and use of patient portals [11]. In that study, Rogers’ Diffusion of Innovation model was successfully applied to assess patient perceptions of a patient portal. The perceived attributes of ease of use and relative advantage of the portal emerged as significant predictors of portal adoption and value. Another technology adoption model, the Unified Theory of Adoption and Use of Technology, was successfully applied by Turvey and colleagues [12] in their study of use of the Blue Button at the Department of Veterans Affairs patient portal, MyHealtheVet. In that study, factors such as knowledge and usability of the Blue Button emerged as significant barriers to the use of the Blue Button associated with the patient portal. However, we continued to find a lack of application of theoretical models in the study of patient portals. Furthermore, there was a need for theoretical models that could predict behavioral intention and behavior toward the use of patient portals and the specific functionality associated with portals such as the AVS. The application of such models could also yield useful prescriptive implications for practitioners who are interested in improving patient use of patient portals.

In this study, we applied a prominent theoretical model, the Theory of Planned Behavior (TPB), to patient portals with a focus on predicting patients' behavioral intention toward accessing the AVS provided through the portal. According to TPB, three major factors lead to the formation of a patient's intention to perform a behavior: attitude toward the behavior, perceived norm, and perceived behavioral control [13]. Underlying each of these factors is a set of beliefs: behavioral beliefs about the positive or negative consequences of accessing the AVS (determines attitude), normative beliefs that important people would approve or disapprove of the patient accessing the AVS or that these referents themselves are accessing the AVS (determines perceived norm), and control beliefs that facilitate or impede the patient accessing the AVS through the patient portal (determines perceived behavioral control). Behavioral intention is a direct predictor of the behavior of accessing the AVS through the patient portal. However, factors such as skills and environmental factors may influence the relationship between intention and behavior. For example, a patient may lack the computer skills to access the AVS through the portal even if they form an intention to access the AVS. Finally, background factors such as education, age, and race may indirectly influence beliefs about accessing the AVS. Over the last three decades, TPB has been successfully applied to predict intention across a wide range of health and risk behaviors, including exercise, breast self-examination, eating a low-fat diet, condom use, alcohol consumption, smoking, and using drugs [13,14].

Survey Instrument

To apply TPB to predict patients' intention to access the AVS through the patient portal, we developed and implemented a cross-sectional survey that captured the different components of TPB. We followed the approach recommended by Fishbein and Ajzen [13] in developing our survey instrument. First, we conducted a pilot study to gather data on the beliefs related to patients' accessing the AVS through the patient portal. In TPB, the behavior of interest is defined by four elements through the principle of compatibility: the action performed, the target at which the action is directed, the context in which it is performed, and the time at which it is performed [13]. In our pilot study, we defined the behavior of interest as accessing (action) the AVS (target) through the patient portal (context) within 7 days of the visit (time). Our pilot survey consisted of three questions related to this behavior of interest: (1) asking patients to list the advantages and disadvantages of accessing the AVS through the patient portal within 7 days of the visit to identify behavioral beliefs, (2) asking patients to list individuals or groups who would approve or disapprove of their accessing the AVS through the patient portal to identify injunctive normative beliefs, and (3) asking patients to list the factors or circumstances that would make it easy or difficult for them to access the AVS through the patient portal to identify control beliefs. Based on the responses to the pilot survey, we created the items for our survey. Additionally, we changed the time component for accessing the AVS to within 5 days of the visit. Multimedia Appendix 1 lists the survey items on TPB we developed for our study.

Recruitment

The study was implemented in the ambulatory care practices of an academic medical center affiliated with Partners HealthCare, an integrated delivery system located in Eastern Massachusetts. Partners developed its own patient portal, Patient Gateway, following its strategy of developing and implementing its own EHR, the Longitudinal Medical Record. The patient portal has functionality similar to other vendor portals, including requests for appointments, prescription refills and referrals, access to certain components of the EHR (eg, laboratory results), and secure messaging with the practice and provider. The AVS is made available to patients through the portal. Patient portal transactions are stored permanently in the Partners clinical information systems and can be accessed for research purposes after institutional review board approval. Over a 6-month period, patients with a patient portal account and an office visit in the previous week were identified using the Partners scheduling system. These patients were sent an email about the study through their portal account. The email included a link to the survey. After 7 days, patients were sent a reminder email with another link to the survey. Patients were not compensated for the survey. All study materials and methods were approved by the Partners Health Care Institutional Review Board.

Statistical Analysis

We present frequencies and means of sociodemographic characteristics and factors related to portal experience for the different groups: aware of AVS / unaware of AVS and accessed AVS / did not access AVS. To assess for differences between the groups, we conducted chi-square tests for categorical data (Pearson chi-square for dichotomous and nominal variables) and robust 1-way ANOVA for continuous variables. We employed multiple regression analysis using a forced entry method to assess predictors of patient satisfaction with the AVS. To test the application of TPB, we computed Cronbach alpha for the major factors (attitude, perceived norm, and perceived behavioral control). We then created scales for each factor using the mean of the scores of the items for each scale. We also computed Cronbach alpha and created a scale for behavioral intention to access the AVS through the patient portal. We present means and standard deviations for the belief items captured through our survey and correlations of the belief items with respective factors. Finally, we conducted hierarchical regression analysis to predict behavioral intention from the major factors of TPB as well as external factors such as sociodemographics and portal experience.

Results

Response Rate

Of the 23,336 patients who received the online survey, 5370 responded for a response rate of 23.01%. Table 1 shows the characteristics of the responders and nonresponders. Overall, 61.79% (3318/5370) of responders and 62.51% (11,231/17,966) of nonresponders were female; 90.76% (4874/5370) of responders were white compared to 86.44% (15,530/17,966) of nonresponders (*P*<.001). Responders were older (mean 56.6, SD 14.0 years) than nonresponders (mean 50.4, SD 15.4 years, *P*<.001). Responders also had a portal account for a longer time (mean 3.5, SD 3.3 years) compared to nonresponders (mean 2.8, SD 2.9 years) and sent a greater number of messages (mean 7.7, SD 20.1) compared to nonresponders (mean 4.3, SD 13.8). Given the large sample sizes of responders and nonresponders, the statistically significant differences should be viewed with caution.

Table 1. Characteristics of responders and nonresponders (N=23,336).

Characteristics	Responders n=5370	Nonresponders n=17,966	*P*
Gender (female), n (%)	3318 (61.79)	11,231 (62.51)	.34
Age (years), mean (SD)	56.6 (14.0)	50.4 (15.4)	<.001
Race (white), n (%)	4874 (90.76)	15,530 (86.44)	<.001
Selected problems on problem list,[a] mean (SD)	2.1 (1.2)	2.0 (1.1)	<.001
Years with patient portal account, mean (SD)	3.5 (3.3)	2.8 (2.9)	<.001
Sessions with patient portal, mean (SD)	100.4 (173.5)	57.9 (109.6)	<.001
Messages sent via patient portal, mean (SD)	7.7 (20.1)	4.3 (13.8)	<.001

[a] The selected problems included hypertension, hyperlipidemia, diabetes, cancer (any), coronary artery disease, congestive heart failure, asthma, osteoarthritis, rheumatoid arthritis, and depression.

Awareness and Access of the After-Visit Summary

Among the 5370 responders of the survey, 4109 (76.52%) reported that they were aware of the availability of the AVS through the patient portal and 1169 (21.77%) reported that they were not aware of the availability of the AVS through the portal (92/5370, 1.71% did not respond to this question). Table 2 shows characteristics of patients who were aware and unaware of the AVS; 61.35% (2521/4109) of female patients were aware of the AVS and 63.99% (748/1169) of female patients were unaware of the AVS. Patients who were aware of the AVS had a mean age of 56.4 (SD 14.0) years compared to a mean age of 57.1 (SD 13.7) years for patients who were unaware of the AVS. In all, 90.58% (3722/4109) and 91.79% (1073/1169) of patients were white among patients who were aware and unaware of the AVS, respectively. In all, 67.78% (2452/3618) of patients who were aware of the AVA and 67.7% (674/995) who were unaware of the AVS reported their marital status as married or an unmarried couple; 51.11% (1873/3655) of patients who were aware of the AVS and 50.65% (508/1003) who were unaware of the AVS reported their health status as very good or excellent. The proportion of patients who reported that they were a 4-year college graduate or more was lower in the aware group (67.54%, 2461/3644) compared to the unaware group (79.42%, 795/1001, *P*<.001). The proportion of patients who reported total household income from all sources as US $75,000 or more was also lower in the aware group (62.55%, 2036/3255) compared to the unaware group (66.4%, 570/858, *P*=.04). Patients who were aware of the AVS had a portal account for a mean of 3.4 (SD 3.3) years compared to a mean of 4.0 (SD 3.4) years for patients who were not aware of the AVS (*P*<.001). The two groups did not differ on the number of sessions with the portal or the number of messages sent via the portal. Satisfaction with the portal was significantly higher in the AVS aware group (45.28%, 164/3653 reporting excellent) compared to the unaware group (30.1%, 299/993, *P*<.001).

Table 2. Characteristics of respondents who were aware/unaware of the after-visit summary (AVS).

Characteristics	Aware of AVS n=4109	Unaware of AVS n=1169	*P*
Gender (female), n (%)	2521 (61.35)	748 (63.99)	.10
Age (years), mean (SD)	56.4 (14.0)	57.1 (13.7)	.16
Race (white), n (%)	3722 (90.58)	1073 (91.79)	.21
Education (≥4-year college degree), n (%)	2461/3644 (67.54)	795/1001 (79.42)	.001
Income (≥US$75,000),[a] n (%)	2036/3255 (62.55)	570/858 (66.4)	.04
Marital status (married or unmarried couple), n (%)	2452/3618 (67.78)	674/995 (67.7)	.98
Health status (very good or excellent),[b] n (%)	1873/3665 (51.11)	508/1003 (50.65)	.80
Selected problems on problem list,[c] mean (SD)	2.1 (1.2)	2.2 (1.2)	.36
Years with patient portal account, mean (SD)	3.4 (3.3)	4.0 (3.4)	.001
Sessions with patient portal, mean (SD)	101.8 (185.4)	97.4 (128.6)	.44
Messages sent via patient portal, mean (SD)	7.5 (19.7)	8.5 (20.9)	.14
Satisfaction with patient portal (excellent),[d] n (%)	1654/3653 (45.28)	299/993 (30.1)	.001

[a] In total household income from all sources before taxes.

[b] Rating of overall health was captured as excellent (5), very good (4), good (3), fair (2), or poor (1).

[c] The selected problems included hypertension, hyperlipidemia, diabetes, cancer (any), coronary artery disease, congestive heart failure, asthma, osteoarthritis, rheumatoid arthritis, and depression.

[d] Satisfaction with patient portal was captured as excellent (5), very good (4), good (3), fair (2), or poor (1).

Of the 4109 patients who were aware of the AVS, 2248 (54.71%) reported that they accessed the AVS through the patient portal within 5 days of the visit and 1805 (43.93%) did not access the AVS (56/4109, 1.4% did not respond). The top two reasons provided for not accessing the AVS through the portal were did not have a need for the AVS (45.43%, 820/1805) and did not remember that AVS was available through the portal (31.63%, 571/1805). Another 14.68% (265/1805) reported that they had received a copy of the AVS from their doctor’s office. Only 3.49% (63/1805) reported that they did not know how to access the AVS through the patient portal.

In total, 61.30% (1378/2248) of patients who accessed the AVS through the patient portal were female and 61.61% (1112/1805) of patients who did not access the AVS through the portal were female (Table 3). Mean age of patients who accessed the AVS was 56.5 (SD 14.1) years and mean age was 56.3 (SD 13.8) years for patients who did not access the AVS. In all, 90.52% (2035/2248) of patients who accessed the AVS and 90.47% (1633/1805) of patients who did not access the AVS were white. Of patients who accessed the AVS, 66.08% (1325/2005) had a 4-year college degree or more compared to 69.22% (1102/1592) of patients who did not access the AVS. Also, 62.85% (1132/1801) of patients who accessed the AVS and 62.18% (878/1412) of patients who did not access the AVS had a total household income of US $75,000 or more. Patients who accessed the AVS had a greater proportion reporting a status as a married or unmarried couple (69.69%, 1391/1996) compared to patients who did not access the AVS (65.44%, 1032/1577, *P*=.007). Patients reporting health status of very good or excellent was similar in both groups (50.59%, 1022/202 in access group and 51.75%, 827/1598 in did not access group). Patients who accessed the AVS had a portal account for a greater number of years (mean 3.6, SD 3.3) compared to patients who did not access the AVS (mean 3.1, SD 3.1, *P*<.001). Patients who accessed the AVS also used the portal more than patients who did not access the AVS (mean 119, SD 221.5 sessions vs mean 79.1, SD 123.3 sessions, *P*<.001; mean 8.2, SD 20.8 messages sent via the portal vs mean of 6.6, SD 18.2 messages sent via the portal, *P*=.01). Patients who accessed the AVS reported greater satisfaction with the portal (excellent: 51.03%, 1032/2022) compared to patients who did not access the AVS (excellent: 38.00%, 602/1584, *P*<.001).

Table 3. Characteristics of respondents who accessed or did not access the after-visit summary (AVS).[a]

Characteristics	Accessed AVS n=2248	Did not access AVS n=1805	*P*
Gender (female), n (%)	1378 (61.30)	1112 (61.61)	.84
Age (years), mean (SD)	56.5 (14.1)	56.3 (13.8)	.73
Race (white), n (%)	2035 (90.52)	1633 (90.47)	.95
Education (≥4-year college degree), n (%)	1325/2005 (66.08)	1102/1592 (69.22)	.046
Income (≥US $75,000),[b] n (%)	1132/1801 (62.85)	878/1412 (62.18)	.70
Marital status (married or unmarried couple), n (%)	1391/1996 (69.69)	1032/1577 (65.44)	.007
Health status (very good or excellent),[c] n (%)	1022/2020 (50.59)	827/1598 (51.75)	.49
Selected problems on problem list,[d] mean (SD)	2.2 (1.2)	2.1 (1.1)	.03
Years with patient portal account, mean (SD)	3.6 (3.3)	3.1 (3.1)	.001
Sessions with patient portal, mean (SD)	119.0 (221.5)	79.1 (123.3)	.001
Messages sent via patient portal, mean (SD)	8.2 (20.8)	6.6 (18.2)	.01
Satisfaction with patient portal (excellent),[e] n (%)	1032/2022 (51.03)	602/1584 (38.00)	.001

[a] Data presented for patients who were aware that the AVS was available through the patient portal.

[b] In total household income from all sources before taxes.

[c] Rating of overall health was captured as excellent (5), very good (4), good (3), fair (2), or poor (1).

[d] The selected problems included hypertension, hyperlipidemia, diabetes, cancer (any), coronary artery disease, congestive heart failure, asthma, osteoarthritis, rheumatoid arthritis, and depression.

[e] Satisfaction with patient portal was captured as excellent (5), very good (4), good (3), fair (2), or poor (1).

Satisfaction With After-Visit Summary

We assessed patient satisfaction with the AVS by asking patients to rate the AVS on a scale from 1 (poor) to 5 (excellent). The mean satisfaction with the AVS was 3.9 (SD 1.12). Patients 65 years and older reported greater satisfaction with the AVS (mean 3.9, SD 1.1) than patients younger than 65 years of age (mean 3.8, SD 1.1, *P*=.02). Nonwhite patients reported greater satisfaction (mean 4.0, SD 1.0) than white patients did (mean 3.9, SD 1.1, *P*=.04). Patients who reported very good or excellent health status reported greater satisfaction with the AVS (mean 4.0, SD 1.1) than patients who reported other (poor/fair/good) health status (mean 3.8, SD 1.2, *P*=.02). Patients who had less than a 4-year college degree reported greater satisfaction (mean 4.1, SD 1.0) than patients who had a 4-year college degree or more (mean 3.8, SD 1.2, *P*<.001). We fitted a multiple regression (forced entry) with satisfaction with AVS as the dependent variable and sociodemographics and portal-related variables as predictors (Table 4). Satisfaction with the patient portal was the most significant predictor of satisfaction with the AVS (beta=.679, *P*<.001). The number of portal sessions was also a significant predictor of the AVS (beta=–.095, *P*<.001). Among sociodemographic variables, age, gender, race, and education were significant predictors of satisfaction with AVS. Finally, patient-reported health status was a significant predictor of satisfaction with AVS.

The Theory of Planned Behavior

Components of Theory of Planned Behavior

Table 5 shows results of reliability analyses for the three main factors in TPB and for the outcome variable of behavioral intention. Cronbach alpha was very good for all the factors. We created scales for each of the factors using a mean of the scores of the items pertaining to each factor.

Table 4. Multiple regression with satisfaction with after-visit summary as the dependent variable.

Predictors	Beta (standardized coefficient)	*P*
(Constant)		<.001
Age	.039	.002
Gender	–.036	.003
Race	–.031	.008
Health status	.037	.003
Education	–.069	<.001
Income	–.018	.19
Marital status	–.017	.20
Years with patient portal account	–.003	.82
Sessions with patient portal	–.095	<.001
Messages sent via patient portal	.002	.89
Satisfaction with patient portal	.679	<.001

Table 5. Reliability analysis for attitude, perceived norm, perceived behavioral control, and behavioral intentions.

TPB factor	Items	Cronbach alpha
Attitude	My accessing the visit summary report via patient gateway within 5 days of the visit is bad/good, pleasant/unpleasant, harmful/beneficial, useless/useful	.87
Perceived norm	Most people who are important to me think that I should access the visit summary report via patient gateway within 5 days of the visit; most people whose opinions I value would approve of my accessing the visit summary report via patient gateway within 5 days of the visit; most people I respect and value will access the visit summary report via patient gateway within 5 days of the visit; most people like me will access the visit summary report via patient gateway within 5 days of the visit	.86
Perceived behavioral control	I am confident that I can access the visit summary report via patient gateway within 5 days of the visit; my accessing the visit summary report via patient gateway within 5 days of the visit is completely up to me; if I really wanted to, I can access the visit summary report via patient gateway within 5 days of the visit; I have complete control over whether or not I access the visit summary report via Patient Gateway within 5 days of the visit	.88
Behavioral intention	I intend to access the visit summary report via patient gateway within 5 days of the visit; I will access the visit summary report via patient gateway within 5 days of the visit; how likely or unlikely is it that you will access the visit summary report via patient gateway within 5 days of the visit; I plan to access the visit summary report via patient gateway within 5 days of the visit	.96

Table 6 shows results for behavioral belief strength, outcome evaluation, and the product of behavioral belief strength and outcome evaluation for the behavioral beliefs included in this study. Mean belief strength was on the positive side for all the behavioral beliefs. The most strongly held beliefs were patients' ability to track visits and tests, and patients' having medical information more readily accessible. The two beliefs that were lowest in strength were clarifying issues with their doctor and reinforcing doctor's instructions. These two beliefs also had the lowest outcome evaluation and the greatest variation in belief strength and outcome evaluation. The mean behavioral belief strength×outcome evaluation products show that the three beliefs with the strongest positive impact on attitude were being able to track visits and tests, having medical information more readily accessible, and a more efficient way to obtain medical information. The correlations between behavioral belief strength×outcome evaluation and attitude were positive and significant and ranged from .45 to .52.

Table 7 shows results for normative belief strengths, motivation to comply, and the product of normative belief strengths and motivation to comply for the three referents included in the study. Mean normative belief strength was positive for all three referents. The strongest belief strength was associated with the patients' doctor. Motivation to comply was also positive with the strongest motivation to comply associated with the patient's doctor. Correlations between normative belief strength×motivation to comply and injunctive norm were all positive and significant.

Table 8 shows results for control beliefs, power of factor, and the product of control beliefs and power of factor. The strongest control belief was that the patient will have access to the Internet within 5 days of the visit. This was followed by the control belief "it will be easy for me to access the AVS via the patient portal" and the belief "I will remember the user ID and password for the patient portal." Correlations between the product of control belief and power of factor and perceived behavioral control were all positive and significant. The strongest correlation was for the control belief “it will be easy for me to access the AVS via the patient portal.”

Table 6. Behavioral beliefs, outcome evaluation, and correlations with attitude.

My accessing the visit summary report of my office visit via patient gateway within 5 days of the visit will result in:	Belief strength, mean (SD)	Outcome evaluation, mean (SD)	Belief strength×outcome evaluation		
			Mean (SD)	r [a]	P
My obtaining medical information (laboratory results, test results) in a more timely manner	2.38 (1.18)	2.55 (0.93)	6.62 (3.60)	.48	<.001
My having up to date medical information	2.40 (1.16)	2.57 (0.94)	6.74 (3.50)	.50	<.001
My having medical information more readily accessible	2.48 (1.07)	2.58 (0.92)	6.93 (3.36)	.50	<.001
Being able to clarify issues with my doctor	2.21 (1.27)	2.42 (1.10)	6.17 (3.73)	.45	<.001
My being able to track my visits and tests	2.53 (1.00)	2.60 (0.87)	7.06 (3.23)	.52	<.001
My being able to view all my medical information in one location	2.38 (1.21)	2.53 (1.04)	6.74 (3.58)	.45	<.001
Reinforcing my doctor’s instructions	2.18 (1.30)	2.37 (1.12)	6.11 (3.76)	.45	<.001
A more efficient way to obtain my medical information	2.41 (1.14)	2.54 (0.98)	6.84 (3.41)	.49	<.001

[a] Correlation between behavioral belief strength×outcome evaluation and attitude.

Table 7. Injunctive normative beliefs, motivation to comply, and correlations with injunctive norm.

My _____ thinks that I should access the visit summary report of my office visit via patient gateway within 5 days of the visit	Normative belief strength, mean (SD)	Motivation to comply, mean (SD)	Normative belief strength×motivation to comply		
			Mean (SD)	r [a]	P
Spouse/Partner	0.92 (1.77)	0.57 (1.77)	2.49 (3.92)	.31	<.001
Doctor	1.66 (1.55)	2.12 (1.24)	4.38 (4.27)	.45	<.001
Nurse	1.19 (1.59)	1.49 (1.46)	2.87 (3.97)	.38	<.001

[a] Correlation between normative belief strength×motivation and injunctive norm.

Table 8. Control beliefs, power of factors, and correlations with perceived behavioral control.

Belief	Control belief strength, mean (SD)	Power of factor,[a] mean (SD)	Control belief strength×power of factor		
			Mean (SD)	r [b]	P
I will receive a reminder email after the office visit that the visit summary report is available via patient gateway	2.08 (1.42)	6.07 (1.42)	15.02 (7.94)	.21	<.001
I will have all my medical information in the visit summary report after the office visit	2.08 (1.33)	6.08 (1.33)	14.57 (8.08)	.27	<.001
I will have access to the Internet within 5 days of the office visit	2.76 (0.70)	6.76 (0.70)	17.85 (6.59)	.29	<.001
I will remember my patient gateway user ID and password	2.51 (0.98)	6.51 (0.98)	16.62 (7.28)	.27	<.001
I will understand the information in the visit summary report	2.39 (0.93)	6.39 (0.93)	14.07 (9.26)	.21	<.001
It will be easy for me to access the visit summary report via patient gateway	2.64 (0.81)	6.64 (0.81)	16.95 (6.90)	.33	<.001

[a] Power of factor measured on a scale from 1 (likely) to 7 (unlikely) and then reverse coded.

[b] Correlation between control belief × power of factor and perceived behavioral control.

Predictors of Behavioral Intention

We employed hierarchical multiple regression to assess predictors of behavioral intention with respect to accessing the AVS through the patient portal and to test the TPB model (Table 9). Given the TPB model, we entered attitude, perceived norm, and perceived behavioral control in the first model. In the second and third models, we entered variables external to the TPB model: variables related to patient portal use were entered in the second model and sociodemographic variables and health status were entered in the third model. The first model consisting of the direct predictors of behavioral intention, attitude, perceived norm, and perceived behavioral control was significant and accounted for 56.7% of the variance in behavioral intention. The second model with portal variables was also significant but added only 5%. Years with portal account, number of sessions, and satisfaction with the AVS were significant predictors in this model. The third model was also significant and added another 3%. Age was the only sociodemographic variable that was significant in this model (beta=.049, *P*<.001*)*.

Table 9. Hierarchical multiple regression analysis predicting behavioral intention.

Factors	Beta (standardized coefficient)	R^2 change	*P*
TPB factors		.567	<.001
(Constant)			<.001
Attitude	.432		<.001
Perceived norm	.291		<.001
Perceived behavioral control	.206		<.001
TPB factors and patient portal factors		.005	<.001
(Constant)			<.001
Attitude	.415		<.001
Perceived norm	.285		<.001
Perceived behavioral control	.197		<.001
Number of years with patient portal account	.040		.001
Number of sessions with patient portal	.039		.003
Number of messages sent via patient portal	–.008		.54
Satisfaction with patient portal	.003		.87
Satisfaction with AVS	.044		.006
TPB factors, patient portal factors, sociodemographics, and self-reported health status		.003	.004
(Constant)			<.001
Attitude	.412		<.001
Perceived norm	.284		<.001
Perceived behavioral control	.197		<.001
Number of years with patient portal account	.030		.02
Number of sessions with patient portal	.035		.008
Number of messages sent via patient portal	–.004		.75
Satisfaction with patient portal	.003		.85
Satisfaction with AVS	.042		.009
Age	.049		<.001
Gender	.019		.09
Race	.012		.28
Health status	–.001		.91
Education	–.011		.33
Income	–.008		.53
Marital status	.010		.40

Discussion

In this study, we assessed patient awareness and access of the AVS provided through a patient portal. A large majority of users of the portal were aware that the AVS was available through the portal. Of those who were aware, just over half reported that they accessed the AVS through the portal within 5 days of an office visit. There were no differences between the groups with respect to sociodemographics (eg, age, gender, race, and self-reported health status). Education and income were related to awareness of the AVS but in a reverse direction than expected: users of the portal with more education and higher income were more likely to be unaware of the availability of the AVS through the portal. However, these differences did not carry over to the access of the AVS through the portal. Given these findings on race, income, and education, our study does not find a digital divide in the case of both awareness and access of the AVS through the patient portal. In a previous study on the digital divide associated with the adoption and use of the patient portal, we found that the digital divide did not carry over to portal use, specifically the relationship between income and the frequency of secure messaging through the portal [15]. In this study, accessing the AVS through the patient portal yields a similar finding: once patients have adopted the portal and are using it, issues of digital divide may not persist at least with respect to some of the functionality of the portal. However, there is a need for additional research on the digital divide with respect to use of portal functionality.

Although previous studies have reported high levels of satisfaction overall with patient portals, there has been little research on satisfaction with the use of specific portal functionality such as the AVS. Our study found a high level of satisfaction with the AVS similar to the finding by Pavlik and colleagues [10]. At the same time, overall satisfaction with the patient portal was positively and significantly related to both awareness and access of the AVS through the patient portal. Satisfaction with the portal was also the most important predictor of satisfaction with the AVS. We agree with the need for empirically measuring a quality indicator such as patient satisfaction with the portal because it appears to be a driver of patient satisfaction with specific functionality of the portal such as the AVS [4].

We do not know of other work that has previously applied TPB to assess patients' beliefs and predict behavioral intention toward accessing their AVS through a patient portal and our findings support the use of the theoretical model in this area. The correlations between beliefs about accessing the AVS through the patient portal and the major determinants of the TPB model (attitude, perceived norm, and perceived behavioral control) were positive and significant, and similar to those reported in studies on other health behaviors. Ajzen [16] reported mean correlations between the expectancy-value index of beliefs and a direct attitude measure ranging from .50 to .53 based on the findings of two meta-analyses of studies applying TPB to health behaviors. The correlations in our study (Table 5) ranged from .45 to .50 with a mean of .48. McEachan and colleagues [14] conducted a meta-analysis of 237 prospective tests of the TPB applied to health behaviors. They reported a corrected (for sampling and measurement error) mean correlation of .57 between attitude and intention, a mean correlation of .54 between perceived behavioral control and intention, and a mean correlation of .40 between subjective norm and intention. In our study, correlation between attitude and intention was .65, perceived behavioral control and intention was .47, and perceived norm and intention was .58**.** The prediction of behavioral intention to access the AVS through the patient portal from attitude, perceived norm, and perceived behavioral control yielded similar findings to other applications of TPB in the health arena. Ajzen [16] reported a meta-analysis showing the mean multiple correlation between the three major determinants of TPB and behavioral intention to range from .59 and .66. In terms of prediction of intention, McEachan and colleagues [14] reported that 44.3% of the variance in intention was accounted by attitude, subjective norm, and perceived behavioral control. Attitude was the strongest predictor (beta=.35) followed by perceived behavioral control (beta=.34) and subjective norm (beta=.15). In our study, 56.7% of the variance in intention was accounted by attitude, perceived norm, and perceived behavioral control. Attitude was the strongest predictor of intention (beta=.43) followed by perceived norm (beta=.29) and perceived behavioral control (beta=.21).

In terms of specific beliefs, our study found that behavioral beliefs related to patient access of information through the AVS, specifically the ability to track visits and tests, have medical information more readily accessible, and obtain medical information more efficiently, were more important than beliefs about patient engagement in their health care, such as clarifying issues with their doctor or reinforcing instructions. This finding is similar to the finding obtained in studies on patients accessing their doctor's notes that patients value access to their information [17-20]. In the VA Open Notes study, patients accessed their doctor's notes to be better prepared for clinic visits, remember their care plan better, and feel more in control of their health [17,18]. Delbanco and colleagues [19,20] also found patients accessed their doctor's notes to be better prepared for future visits and have a greater sense of control over their health. In their organization, Ralston and colleagues [9] noted the AVS may serve patient's information and care needs better because it provides a focused plan of care combined with educational materials hyperlinked to other sources. The question of when patients prefer to access their AVS compared to their doctor's notes through a patient portal, and the relative value of these uses of the portal, are important topics for future research.

In the case of normative and control beliefs, the strongest normative belief and motivation to comply were associated with the patient's doctor. Patients believe that their doctor thinks they should access the AVS and they want to do what their doctor thinks. The importance of clinician encouragement of patients using online tools such as patient portals has been identified in several studies [9,11,21]. Our study finds that clinicians also have an important role in encouraging patients to access specific functionality of portals such as the AVS. In the case of control beliefs, the ease of accessing the AVS through the patient portal was identified as an important belief. The importance of ease of access and use of patient portals has been well documented in empirical and scoping studies

[11,12,22,23]. Our study finds that even after patients have adopted patient portals, ease of access of specific functionality of the portals can facilitate or impede intention and behavior to use the functionality.

This study had two objectives: (1) assess the characteristics of patients who are aware of and access the AVS through a patient portal and (2) apply TPB to evaluate beliefs, attitude, perceived norm, perceived behavioral control, and predict behavioral intention of patients toward accessing the AVS through a patient portal. A majority of users of a patient portal selected for this study were aware that the AVS was accessible through the portal, but almost a third did not remember that the AVS was available and therefore did not access it. Patients may need to be reminded that the AVS is accessible through the portal at the time they leave their office visit, especially if patient preference is to receive information through the patient portal. In terms of patient characteristics, we did not find evidence of a digital divide with respect to income or education in either awareness or access of the AVS in our portal users. On the other hand, portal users in one setting who accessed their doctor's notes were more educated than those who did not access the notes [18]. Additional research is needed on issues of digital divide with respect to different uses of the patient portal, such as accessing the AVS or doctor's notes.

With respect to the meaningful use of EHR incentive program, two goals were envisioned related to EHR functionality such as the AVS [1,2]: (1) provide patients with timely and efficient access to their health information and (2) motivate patients to engage in their health care. Other studies have also identified patient engagement in their health care (eg, shared decision making with their doctor) as an important purpose of the AVS [24]. Our study found that patients had stronger beliefs about the AVS with respect to timely and efficient access of information than with engaging in their health care. This finding may reflect patients' value of the AVS as a permanent personal record to review whenever the need arises [25]. On the other hand, it is possible that the use of the AVS to engage patients in their health care is not being promoted. Pavlik and colleagues [10] noted the need for concerted efforts to remind patients of important information available through the AVS than simply providing the AVS to patients. Such efforts can lead to patient activation and the use of information by patients to undertake recommended treatment plans and self-management, both of which are important goals for the AVS [24]. Although Stage 1 of the MU program included a core requirement for the provision of the AVS, Stage 2 of the MU program no longer includes this core requirement. This is unfortunate given the value of the AVS for providing timely and efficient access of information as reported by patients in our study.

We found TPB to be a suitable theoretical model to predict behavioral intention of patients toward accessing the AVS through a patient portal. Our findings match applications of TPB for predicting intention with respect to a variety of other health behaviors. Thus, this study provides an important contribution to the application of theoretical models to the study of patient portals and extends some of the prior theoretical work on this topic [11,12]. Beyond its theoretical contribution, the application of TPB can suggest interventions that are relevant to practitioners. For example, Fishbein and Ajzen [13] recommend interventions that target and change relevant salient beliefs or make new beliefs salient in support of recommended behavior. Our study found that the strongest behavioral beliefs related to accessing the AVS through the patient portal are those related to tracking visits and tests, and having medical information more readily and efficiently accessible. Our study also found that doctors are an important social agent for patients with respect to accessing the AVS. For those patients who are not accessing the AVS through the patient portal, a simple intervention that organizations can implement would be to encourage doctors and support staff to discuss with patients the advantages of accessing the AVS through the portal, such as tracking visits and tests. This intervention would also help those patients who do not remember that the AVS is available through the portal. Similarly, doctors can ask patients to use the AVS for clarifying instructions and engaging in shared decision making. Engaging patients in a dialog about the use of the AVS may also help facilitate two important factors identified by us in a survey of physician beliefs about the AVS: (1) enhancing physician satisfaction with the AVS and (2) promoting positive beliefs about the effect of the AVS on patient outcomes and the care the physician personally delivers [26].

Although this study yielded valuable insights into awareness and access of the AVS through a patient portal and the application of TPB to this area, it is associated with some limitations. The response rate in our study was low. We relied on a self-report of patients accessing the AVS through the patient portal. With respect to the time component, we chose 5 days from the office visit as the period within which the patient accessed the AVS. There is a need to assess other time periods after an office visit in which the patient could access the AVS. The study was conducted in the setting of a Northeast academic medical center and the results may not be generalizable to other regions and patient populations. However, we have no reason to suspect that the beliefs identified in this study and their respective strengths would differ across different institutional settings in which the AVS is available through a patient portal (academic medical center vs other) or across different platforms (homegrown vs vendor patient portals).

Acknowledgments

The authors thank the Partners Siemens Research Council (PSRC) for funding the study, and Mohan Ganasekaran, Partners IS, for help with the programming required for the study.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Survey Items.

[PDF File (Adobe PDF File), 44KB - jmir_v18i4e77_app1.pdf]

References

1. Blumenthal D, Tavenner M. The "meaningful use" regulation for electronic health records. N Engl J Med 2010 Aug 5;363(6):501-504. [doi: 10.1056/NEJMp1006114] [Medline: 20647183]
2. Buntin MB, Jain SH, Blumenthal D. Health information technology: laying the infrastructure for national health reform. Health Aff (Millwood) 2010 Jun;29(6):1214-1219 [FREE Full text] [doi: 10.1377/hlthaff.2010.0503] [Medline: 20530358]
3. Kannry J, Beuria P, Wang E, Nissim J. Personal health records: meaningful use, but for whom? Mt Sinai J Med 2012;79(5):593-602. [doi: 10.1002/msj.21334] [Medline: 22976365]
4. Kruse CS, Bolton K, Freriks G. The effect of patient portals on quality outcomes and its implications to meaningful use: a systematic review. J Med Internet Res 2015;17(2):e44 [FREE Full text] [doi: 10.2196/jmir.3171] [Medline: 25669240]
5. Hogan TP, Nazi KM, Luger TM, Amante DJ, Smith BM, Barker A, et al. Technology-assisted patient access to clinical information: an evaluation framework for blue button. JMIR Res Protoc 2014;3(1):e18 [FREE Full text] [doi: 10.2196/resprot.3290] [Medline: 24675395]
6. Centers for Medicare and Medicaid Services. Stage 2 Eligible Professional Meaning Use Core Measures Measure 8 of 17. 2014 Nov. URL: https://www.cms.gov/Regulations-and-Guidance/Legislation/EHRIncentivePrograms/downloads/Stage2_EPCore_8_ClinicalSummaries.pdf [accessed 2015-10-07] [WebCite Cache ID 6c6eLnWa3]
7. Neuberger M, Dontje K, Holzman G, Corser B, Keskimaki A, Chant E. Examination of office visit patient preferences for the after-visit summary (AVS). Perspect Health Inf Manag 2014;11:1d [FREE Full text] [Medline: 25593570]
8. Chung A, Shea C. After-visit clinical summaries: what is meaningful to patients? 2013 Nov 18 Presented at: American Medical Informatics Association Symposium; Nov 18, 2013; Washington, DC.
9. Ralston JD, Carrell D, Reid R, Anderson M, Moran M, Hereford J. Patient web services integrated with a shared medical record: patient use and satisfaction. J Am Med Inform Assoc 2007;14(6):798-806 [FREE Full text] [doi: 10.1197/jamia.M2302] [Medline: 17712090]
10. Pavlik V, Brown AE, Nash S, Gossey JT. Association of patient recall, satisfaction, and adherence to content of an electronic health record (EHR)-generated after visit summary: a randomized clinical trial. J Am Board Fam Med 2014;27(2):209-218 [FREE Full text] [doi: 10.3122/jabfm.2014.02.130137] [Medline: 24610183]
11. Emani S, Yamin CK, Peters E, Karson AS, Lipsitz SR, Wald JS, et al. Patient perceptions of a personal health record: a test of the diffusion of innovation model. J Med Internet Res 2012;14(6):e150 [FREE Full text] [doi: 10.2196/jmir.2278] [Medline: 23128775]
12. Turvey C, Klein D, Fix G, Hogan TP, Woods S, Simon SR, et al. Blue Button use by patients to access and share health record information using the Department of Veterans Affairs' online patient portal. J Am Med Inform Assoc 2014;21(4):657-663. [doi: 10.1136/amiajnl-2014-002723] [Medline: 24740865]
13. Fishbein M, Ajzen I. Predicting and Changing Behavior: The Reasoned Action Approach. New York: Psychology Press; 2010.
14. McEachan RRC, Conner M, Taylor NJ, Lawton RJ. Prospective prediction of health-related behaviours with the Theory of Planned Behaviour: a meta-analysis. Health Psychology Review 2011 Sep;5(2):97-144. [doi: 10.1080/17437199.2010.521684]
15. Yamin CK, Emani S, Williams DH, Lipsitz SR, Karson AS, Wald JS, et al. The digital divide in adoption and use of a personal health record. Arch Intern Med 2011 Mar 28;171(6):568-574. [doi: 10.1001/archinternmed.2011.34] [Medline: 21444847]
16. Ajzen I. Attitudes and persuasion. In: Deaux K, editor. The Oxford Handbook of Personality and Social Psychology (Oxford Library of Psychology). New York: Oxford University Press; 2012.
17. Woods SS, Schwartz E, Tuepker A, Press NA, Nazi KM, Turvey CL, et al. Patient experiences with full electronic access to health records and clinical notes through the My HealtheVet Personal Health Record Pilot: qualitative study. J Med Internet Res 2013;15(3):e65 [FREE Full text] [doi: 10.2196/jmir.2356] [Medline: 23535584]
18. Nazi KM, Turvey CL, Klein DM, Hogan TP, Woods SS. VA OpenNotes: exploring the experiences of early patient adopters with access to clinical notes. J Am Med Inform Assoc 2015 Mar;22(2):380-389. [doi: 10.1136/amiajnl-2014-003144] [Medline: 25352570]
19. Walker J, Leveille SG, Ngo L, Vodicka E, Darer JD, Dhanireddy S, et al. Inviting patients to read their doctors' notes: patients and doctors look ahead: patient and physician surveys. Ann Intern Med 2011 Dec 20;155(12):811-819 [FREE Full text] [doi: 10.7326/0003-4819-155-12-201112200-00003] [Medline: 22184688]
20. Delbanco T, Walker J, Bell SK, Darer JD, Elmore JG, Farag N, et al. Inviting patients to read their doctors' notes: a quasi-experimental study and a look ahead. Ann Intern Med 2012 Oct 2;157(7):461-470 [FREE Full text] [doi: 10.7326/0003-4819-157-7-201210020-00002] [Medline: 23027317]

21. Ralston JD, Coleman K, Reid RJ, Handley MR, Larson EB. Patient experience should be part of meaningful-use criteria. Health Aff (Millwood) 2010 Apr;29(4):607-613 [FREE Full text] [doi: 10.1377/hlthaff.2010.0113] [Medline: 20368589]
22. Kaelber DC, Jha AK, Johnston D, Middleton B, Bates DW. A research agenda for personal health records (PHRs). J Am Med Inform Assoc 2008;15(6):729-736 [FREE Full text] [doi: 10.1197/jamia.M2547] [Medline: 18756002]
23. Archer N, Fevrier-Thomas U, Lokker C, McKibbon KA, Straus SE. Personal health records: a scoping review. J Am Med Inform Assoc 2011;18(4):515-522 [FREE Full text] [doi: 10.1136/amiajnl-2011-000105] [Medline: 21672914]
24. Hummel J, Evans P. Providing Clinical Summaries to Patients After Each Office Visit: A Technical Guide. Seattle: Qualis Health; 2012. URL: https://www.healthit.gov/sites/default/files/measure-tools/avs-tech-guide.pdf [accessed 2016-03-10] [WebCite Cache ID 6fuJsyJTI]
25. Tang PC, Newcomb C. Informing patients: a guide for providing patient health information. J Am Med Inform Assoc 1998;5(6):563-570 [FREE Full text] [Medline: 9824803]
26. Emani S, Ting DY, Healey M, Lipsitz SR, Ramelson H, Suric V, et al. Physician perceptions and beliefs about generating and providing a clinical summary of the office visit. Appl Clin Inform 2015;6(3):577-590. [doi: 10.4338/ACI-2015-04-RA-0043] [Medline: 26448799]

Abbreviations

AVS: after-visit summary
EHR: electronic health record
HITECH: Health Information Technology for Economic and Clinical Health
MU: Meaningful Use
TPB: Theory of Planned Behavior

Edited by A Moorhead; submitted 28.10.15; peer-reviewed by K Nazi, C Kruse, G Vergeire-Dalmacion; comments to author 20.11.15; revised version received 12.01.16; accepted 17.01.16; published 13.04.16.

Please cite as:
Emani S, Healey M, Ting DY, Lipsitz SR, Ramelson H, Suric V, Bates DW
Awareness and Use of the After-Visit Summary Through a Patient Portal: Evaluation of Patient Characteristics and an Application of the Theory of Planned Behavior
J Med Internet Res 2016;18(4):e77
URL: http://www.jmir.org/2016/4/e77/
doi:10.2196/jmir.5207
PMID:27076485

Original Paper

Is Participatory Design Associated with the Effectiveness of Serious Digital Games for Healthy Lifestyle Promotion? A Meta-Analysis

Ann DeSmet[1], PhD; Debbe Thompson[2], PhD; Tom Baranowski[2], PhD; Antonio Palmeira[3], PhD; Maïté Verloigne[1], PhD; Ilse De Bourdeaudhuij[1], PhD

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

[2]Baylor College of Medicine, United States Department of Agriculture/Agricultural Research Service (USDA/ARS) Children's Nutrition Research Center, Houston, TX, United States

[3]University of Lisbon, Universidade Lusófona de Humanidades e Tecnologias, Faculty of Physical Education and Sports and Interdisciplinary Centre for the Study of Human Performance (CIPER) – Faculty of Human Movement, Lisbon, Portugal

Corresponding Author:
Ilse De Bourdeaudhuij, PhD
Ghent University
Department of Movement and Sport Sciences
Watersportlaan 2
Ghent, 9000
Belgium
Phone: 32 92646311
Fax: 32 92646484
Email: ilse.debourdeaudhuij@ugent.be

Abstract

Background: Serious digital games can be effective at changing healthy lifestyles, but large differences in their effectiveness exist. The extent of user involvement in game design may contribute to game effectiveness by creating a better fit with user preferences. Participatory design (PD), which represents active user involvement as informant (ie, users are asked for input and feedback) or codesigner (ie, users as equal partners in the design) early on and throughout the game development, may be associated with higher game effectiveness, as opposed to no user involvement or limited user involvement.

Objective: This paper reports the results of a meta-analysis examining the moderating role of PD in the effectiveness of serious digital games for healthy lifestyle promotion.

Methods: Four databases were searched for peer-reviewed papers in English that were published or in press before October 2014, using a (group-) randomized controlled trial design. Effectiveness data were derived from another meta-analysis assessing the role of behavior change techniques and game features in serious game effectiveness.

Results: A total of 58 games evaluated in 61 studies were included. As previously reported, serious digital games had positive effects on healthy lifestyles and their determinants. Unexpectedly, PD (g=0.075, 95% CI 0.017 to 0.133) throughout game development was related to lower game effectiveness on behavior (Q=6.74, *P*<.05) than when users were only involved as testers (g=0.520, 95% CI 0.150 to 0.890, *P*<.01). Games developed with PD (g=0.171, 95% CI 0.061 to 0.281, *P*<.01) were also related to lower game effectiveness on self-efficacy (Q=7.83, *P*<.05) than when users were not involved in game design (g=0.384, 95% CI 0.283 to 0.485, *P*<.001). Some differences were noted depending on age group, publication year of the study, and on the specific role in PD (ie, informant or codesigner), and depending on the game design element. Games developed with PD were more effective in changing behavioral determinants when they included users in design elements on game dynamics (beta=.215, 95% CI .075 to .356, *P*<.01) and, more specifically, as an informant (beta=.235, 95% CI .079 to .329, *P*<.01). Involving users as informants in PD to create game levels was also related to higher game effectiveness (Q=7.02, *P*<.01). Codesign was related to higher effectiveness when used to create the game challenge (Q=11.23, *P*<.01), but to lower game effectiveness when used to create characters (Q=4.36, *P*<.05) and the game world (Q=3.99, *P*<.05).

Conclusions: The findings do not support higher effectiveness of games developed with PD. However, significant differences existed among PD games. More support was found for informant roles than for codesign roles. When PD was applied to game dynamics, levels, and game challenge, this was associated with higher effectiveness than when it was applied to game aesthetics.

Since user involvement may have an important influence on reach, adoption, and implementation of the intervention, further research and design efforts are needed to enhance effectiveness of serious games developed with PD.

(J Med Internet Res 2016;18(4):e94) doi:10.2196/jmir.4444

KEYWORDS

serious games; video games; computer games; games; health promotion; meta-analysis; review; design; community-based participatory research

Introduction

Serious digital games are a form of organized play, using a digital device, intended to be both entertaining and educational [1]. Serious games have shown promising effects in promoting healthy lifestyles [2,3]. Healthy lifestyles can prevent a wide range of diseases, such as some cancers, cardiovascular diseases, stroke, dementia, mental illness, and diabetes [4-7], and having effective interventions to promote these lifestyles is therefore of great public health importance. A recent meta-analysis of serious games for healthy lifestyle promotion revealed that serious games were effective at changing (1) behavior (eg, number of steps taken per day), (2) individual determinants of this behavior (eg, knowledge, attitudes, social norms, self-efficacy, skills, and perceived environmental barriers or facilitators), and (3) clinical outcomes (eg, body mass index). Although the effects were small, they were in line with the magnitude of effects found for other computer-delivered interventions [8,9]. These effects applied across health behaviors (eg, physical activity, illness self-management, and social behavior). Games were also equally effective for both men and women and for all age groups [2]. Apart from being effective, serious games are also well liked by their target group: adolescent users preferred serious games to traditional educational approaches, such as classroom teaching [10]. Being enjoyable, absorbing, and intrinsically motivating [1,11], serious games may overcome motivational barriers that health promotion programs often encounter [12-14].

Despite their potential, large variations exist in serious game effectiveness that are not well understood [15]. The recent meta-analysis of serious games for healthy lifestyle promotion, which showed games were effective, also investigated the role of theory, individual tailoring, and sample and study characteristics in game effectiveness [2]. A game was considered effective if it reached its goal of improving healthy lifestyles (eg, being more physically active), improving the health outcomes related to these lifestyles (eg, reducing obesity), or improving determinants predictive of a healthy lifestyle (eg, having a positive attitude toward physical activity). Randomized controlled trials were included in the meta-analysis. To be deemed effective, the gain in these health-related outcomes had to be higher in the condition receiving the intervention (ie, the game) than in the control condition. Independent variables included theory, individual tailoring, study quality, health behavior, and study and sample characteristics. Theoretical foundation reflected the theory used to guide intervention development (eg, behavior change theories and game-based learning theories) and was a significant moderator of game effectiveness. Individual tailoring reflected the degree to which the intervention content or appearance was adjusted to match the individual user's characteristics (eg, girls would receive a different game challenge than boys) and was also a significant moderator of effectiveness. Of the investigated study characteristics (eg, average time during which they were exposed to the game), only the time duration between the end of the intervention and the measurement of effects affected game effectiveness.

Other characteristics, such as specific type of health behavior (eg, physical activity, preventive behavior and illness-self-management, and mental health promotion and social behavior) and study quality (ie, study validity and reliability, such as sampling method, measures, and blinding) did not affect game effectiveness.

Heterogeneity in effectiveness, however, remained that could not be fully explained by these moderators, pointing to the need to explore other potential moderators of serious game effectiveness. The extent to which serious games were developed in participatory design (PD) between users and professionals may be an important moderator, as appreciation of game features may differ by target group [16-18]. This may be particularly important when professionals do not share the same characteristics as the target group [19]. Involving end users in design should increase game fit with user preferences, which has been hypothesized to enhance game effectiveness [20].

Participatory design is a broad term that comprises several purposes, methods, and intensities of user involvement. An overarching PD principle is that users should be involved in the development of tools designed to benefit their quality of life [21]. In a pragmatic perspective, PD is considered beneficial via its potential to improve program effectiveness and user adoption [22]; our study reflects this perspective. Other perspectives on user participation consider PD to be useful regardless of its potential contribution to program effectiveness. PD is also regarded as a moral imperative motivated by its potential effects on community empowerment, giving communities (ie, groups of people with shared interests, needs, or identities) more control over their lives by active involvement in program design [22].

PD has similarities to other concepts of user involvement in game design. It is closely related to the concept of user-centered design (UCD), which emphasizes continuous user involvement at every design stage with iterative testing. UCD considers PD to be one of its approaches [23]. Formative research is also frequently mentioned as a way to involve users in health game development [24]. Formative research generally refers to collecting data among target users to ensure the intervention is acceptable and appropriate for them, for example, at a cultural or cognitive level [25]. This type of formative research resembles the UCD-specified role for the user as informant. However, power imbalances between users and professionals may be maintained in formative research [26]. Formative evaluation may be one method used within PD, but does not necessarily constitute PD [27].

In this paper, *user involvement* is an umbrella term describing the degree to which the end user influences game design. User involvement can take several forms: users, testers, informants, and codesign partners [28]. As users or testers, the target group is observed during game play. They are asked for acceptability and/or usability of an early version of the game developed without user input—also known as alpha testing [29,30]. These forms of user involvement are not considered PD, since the experts initiated the design and made the final decisions without user involvement. In the mode of informant—users are asked for input and feedback—or codesigner—users as equal partners in the design—here considered forms of PD, users are actively involved and are asked for input starting at an early stage of design, prior to product development [28].

Active user involvement in the development of health promotion interventions has been advocated to ensure user concerns are adequately reflected in the program [31], to integrate user and professional expertise for mutual learning [32], and to increase community acceptability and adoption of the intervention [19,32]. The benefits of user involvement on program effectiveness are rarely assessed [33]. Collecting hard evidence on the value of PD is difficult, as its fuzzy processes are considered irreconcilable with commonly used research methods for outcome measurements [22]. An experimental design with and without user participation may not be practically feasible, since insights into which factors to control in a comparison condition may surface only during the PD process [22]. Funding bodies, nevertheless, often require evidence of the contribution that PD has made to effectiveness [19,22].

A meta-analysis may advance our understanding of how PD relates to game effectiveness by quantifying and comparing differences across studies and by overcoming small sample sizes in individual studies [34]. This can add to the limited evidence of the value of PD in serious game design and effectiveness.

Several game design elements (eg, levels, challenge, feedback, and tailoring) may contribute to game effectiveness [12,35]. PD in educational games showed that users were focused on design elements, such as narrative, sounds, setting, and characters, but struggled with integrating other design elements in gameplay, such as educational content [36]. Since a recent meta-analysis on serious games indicated that only certain design elements (eg, levels and adaptive game challenge) were associated with game effectiveness (data available upon request from the authors), we could expect user involvement in these design elements to be more strongly associated with game effectiveness.

This study conducted a meta-analysis and meta-regression analysis to investigate the role of user involvement and PD in the effectiveness of serious games for healthy lifestyle promotion. The research questions included the following: (1) Does game effectiveness differ by user involvement (ie, no involvement, tester only, or PD)?, (2) Does the influence of user involvement on effectiveness differ by age group (ie, average participant age <18 years or ≥18 years) or by publication year of the study (ie, game studies published before 2010 or between 2010 and 2014)?, (3) Does PD relate differentially to game effectiveness depending on the design role for users (ie, informant or codesigner)?, and (4) Does PD relate differentially to game effectiveness depending on the design element (eg, educational content, game challenge, or narrative) in which it was applied?

Methods

Overview

The research protocol was inspired by Cochrane guidelines [37]. Meta-analysis reporting was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement [38], with the exception of sensitivity analyses and effectiveness data per study, which are reported elsewhere. Effectiveness data for this study were obtained from a meta-analysis evaluating the role of game design elements in serious game effectiveness. This meta-analysis was an update—search was updated for papers between 2013 and 2014—of a previous meta-analysis. For background information on effectiveness calculations, we refer the readers to an earlier meta-analysis by DeSmet et al [2].

Search Strategy and Study Selection

Inclusion and Exclusion Criteria

The following were the inclusion and exclusion criteria for the studies:

1. Studies were included if they investigated serious digital games, defined as organized play with a set of rules by which to play and a goal, which created a challenge, provided feedback or showed outcomes, entailed interaction, and had a topic [1]. Studies were excluded if they investigated commercial off-the-shelf games, multimedia programs with no interaction (eg, only watching a video without a challenge), and board games.

2. Studies were included if they were designed for healthy lifestyle promotion aiming to improve health behaviors, such as healthy diet (eg, fruit and vegetable consumption), physical activity (eg, number of steps taken per day), social behavior (eg, emotion recognition and not bullying others), health responsibility and maintenance (eg, illness self-management and not smoking), and stress management or self-actualization (eg, personal growth and mental health promotion) [39]. Studies were excluded if they investigated games that only targeted an increased skill level but did not target lifestyle change (eg, athletic performance), that were only used in a therapeutic context and with no intent to create a lifestyle change (eg, treatment support), or that were used for professional education (eg, teaching medical skills).

3. Studies were included if they reported outcomes on behavior or its determinants (eg, knowledge and attitudes). Studies were excluded if they only consisted of usability evaluations, player

experiences, or case studies or that only reported effects on clinical outcomes, not healthy lifestyles.

4. Studies were included if they provided data that allowed the computation of effect sizes. Studies were excluded if there were no data available in the article or after consulting authors allowing an effect size to be calculated.

5. Studies were included if they used a research design with a control condition to which either individuals or groups of individuals were randomly assigned. Studies were excluded if they consisted of a one-group, pretest post-test design or a one-group, post-test-only design.

6. Studies were included if they were reported in English.

Search Strategy

Four databases were searched for peer-reviewed publications since the creation of the journal databases until October 2014: PubMed (1966), Web of Science (1926), Cumulative Index to Nursing and Allied Health Literature (CINAHL) (1937), and PsycINFO (1887). The search was conducted using the keywords *games*, *video games*, or *interactive multimedia* and *health*. Search results were complemented with hand-searching studies reported in reviews, examining the table of contents of relevant specialized journals and databases (ie, Computers in Human Behavior, CyberPsychology, Behavior and Social Networking, Games for Health Journal, JMIR Serious Games, Telemedicine and E-Health, and Health Games Research database), and by requesting qualifying manuscripts from the local Digital Games Research Association (DiGRA) chapter. Authors were contacted for more information when data for coding or effect size calculation were lacking. When unclear, the coding frame was presented to the authors for completion or review and correction.

Coding Frame

Primary and Secondary Outcomes

The following primary outcomes were studied: behavior and behavioral determinants (ie, knowledge, attitudes, subjective norm, perceived barriers, skills, self-efficacy, and behavioral intention). These behavioral determinants are considered the key changeable influences on behaviors [40]. Whatever the authors identified as attitudes, skills, etc, were counted in those categories. As secondary outcomes, clinical effects (eg, weight and depression score) were included, when applicable. These outcomes could relate to several healthy lifestyles, in accordance with the Health-Promoting Lifestyle Profile scale [39]: healthy diet, physical activity, social skills and behavior, health responsibility and maintenance, and stress management and self-actualization.

User Preferences

A coding frame for PD of serious games was based on the items listed in Textbox 1 [12,28,41-43].

To provide an index of PD for mechanics, dynamics, and aesthetics, the number of aspects in which users were involved as informant or as codesigner were summed per area. User involvement could relate to the choice to include certain features and/or to how these features were designed. The user group was defined as the end users for whom the intervention was intended, not as the stakeholders who were important in facilitating the implementation of the intervention (eg, teachers and health professionals). Stakeholder involvement in PD was coded separately.

Also coded were the number of consultations with the users (ie, exact number or constant involvement) and stakeholders (ie, exact number or constant involvement), the method used to involve users (eg, focus groups), and the sampling strategy to recruit the user and stakeholder groups (see Multimedia Appendix 1). These process measures may indicate reliability and validity of the findings from consulting the target group [19]. The full coding frame can be found in Multimedia Appendix 1.

Two coders (ADS, Wendy Van Lippevelde) independently coded user involvement for a random selection of one-third of the game studies, showing good coder agreement (κ=.83). All authors of the included PD studies were offered an opportunity to review and, if necessary, adjust the more detailed coding sheet for their study to guarantee coding accuracy.

Textbox 1. Coding frame for participatory design upon which serious games were based.

A coding frame for PD of serious games was based on the following:

1. The degree of user involvement. Users were either (1) not involved, (2) involved as testers only (of a finalized version), or (3) involved in PD, as either informants at an early stage (ie, preferences and suggestions are elicited prior to and during development) or as codesigners who were part of a design team and in charge of one or more parts of game development [28].
2. When users were involved in PD as informants or codesigners, we coded the following aspects:
    - The design elements in which they were involved, inspired by the Mechanics, Dynamics, Aesthetics (MDA) framework [41].
        1. Mechanics included controls (eg, push buttons and mouse) and actions or interactivity (eg, jump and run).
        2. Dynamics included rewards (eg, points and badges) and levels (eg, number of levels and how to level up or down).
        3. Aesthetics included narrative (eg, story), challenge (eg, game type), character looks (eg, clothing style), game world looks (eg, colors and setting), and language (eg, jargon and phrases).
    - The involvement of users in choosing behavioral change strategies employed in serious games (eg, individual tailoring [12] and feedback [42,43]).
    - The involvement of users as informants or codesigners of educational content.

Meta-Analytic Procedure

Game effectiveness was defined as an improvement in healthy lifestyles, behavioral determinants, or clinical outcomes associated with these lifestyles. Reported effects on *behavior* reflect average effects on the entire category of healthy lifestyles (eg, physical activity and healthy diet). Reported effects on *behavioral determinants* also show average effects on all behavioral determinants considered as one category (eg, knowledge and attitudes). Similarly, all clinical outcomes are considered as one category across health areas (eg, body mass index and depression scores). Hedges' g was calculated as the effect size index, which corrects for small sample sizes [44]. A positive or a negative Hedges' g indicated that the serious game respectively increased or reduced adoption of a healthy lifestyle or its determinants. In cases where the intervention targeted a reduction of unhealthy lifestyles, the computed sign of the effect size was reversed so all positive differences reflected an improvement in healthy lifestyles for the treatment group compared to the control group.

Study results were pooled using a random effects model. Tests for significance of the effect size are indicated in Table 1 by the mention "within category." Two *P* values per row are reported: one for Hedges' g effect sizes (considered significant at *P*<.05) and one for Cochran's Q values, testing for heterogeneity among studies in this category (considered significant at *P*<.05). A significant *P* value for Hedges' g indicated a significant effect of the games in this category, whereas a significant *P* value for Cochran's Q indicated that large variations existed between these studies' effectiveness.

Moderator analyses were conducted to explain differences in effect sizes between studies. For all moderator analyses, a mixed-effects model was used, and Cochran's Q test and I^2[45] were reported to investigate the degree of heterogeneity in effect sizes. Moderator analyses test whether heterogeneity can be explained by differences between several categories on the moderating variable, indicated in Table 1 by the mention "between categories." A significant *P* value associated with the Q test indicated the moderator was able to explain some heterogeneity between the studies' effectiveness. Moderator analyses were only conducted with the categories that contained a sufficient sample. A sufficient sample was defined as having at least three studies in a category, or a minimum sample size of 250 participants per category. A sample size smaller than this was shown not to have sufficient power to detect even large, homogeneous moderator effects [46]. This minimum sample size was thus chosen to make meaningful moderator analyses, without being overly restrictive. It should be emphasized, however, that this sample size may still be underpowered to detect small or heterogeneous moderator effects, and not finding significant moderator effects should hence not be interpreted as evidence for no effect [46]. When a category did not have a sufficient sample, it was excluded from the moderator analysis, which was then performed only on the remaining categories where possible—two or more categories. When a moderator was not relevant for a study (eg, narratives are not used, as they do not fit within certain game types), the study was not included in that particular moderator analysis.

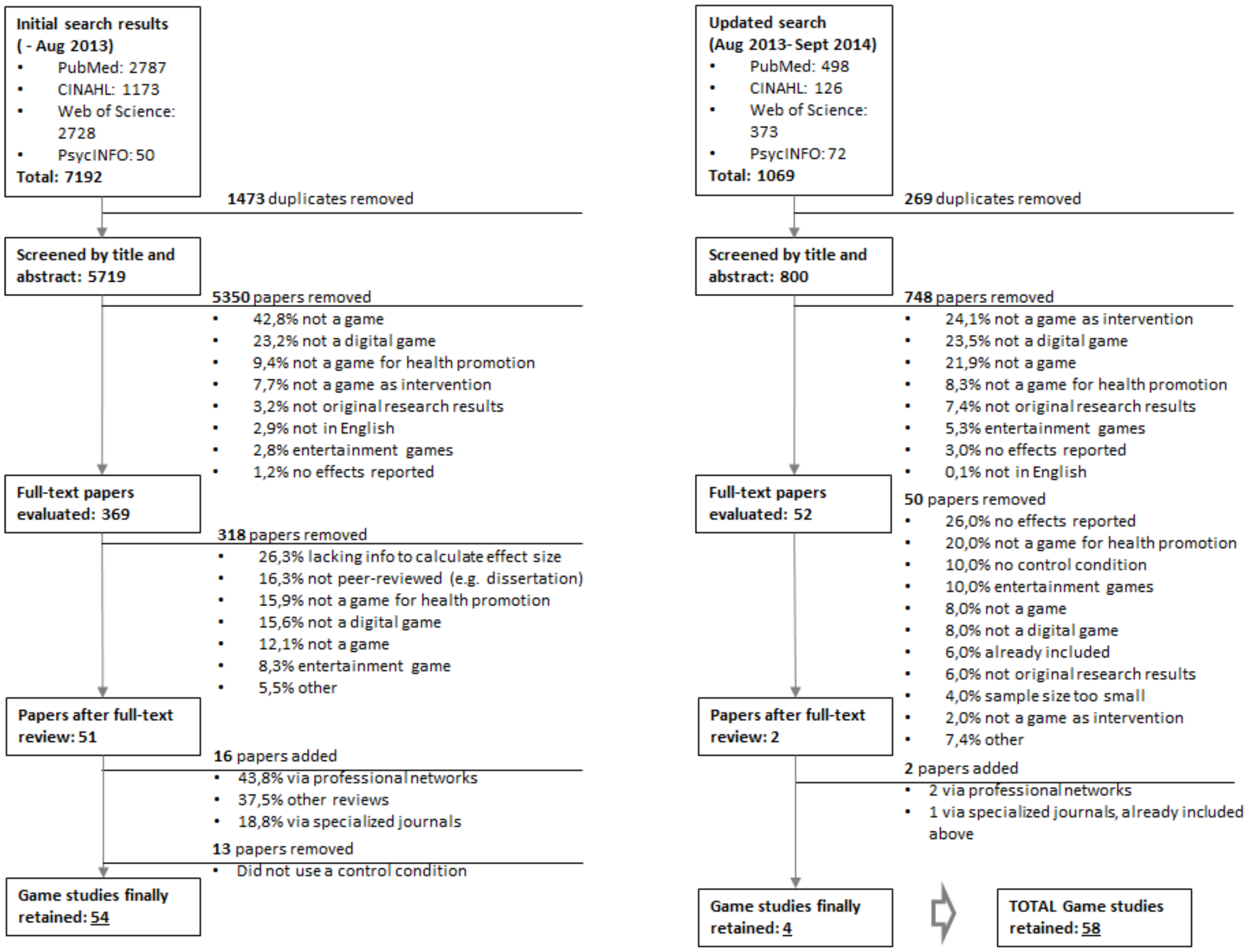
Meta-regression (ie, method-of-moments procedure) was performed for continuous moderators [47], where the slope (beta) and its *P* value indicated the importance of this moderator in understanding linear changes in effect sizes. To maintain the independence of the data, whenever necessary, effect sizes were averaged across different outcomes. All analyses were performed with Comprehensive Meta-Analysis software, version 3 (Biostat Inc, Englewood, NJ, USA). Effect sizes of 0.20-0.49 were considered small effects, 0.50-0.79 were considered moderate, and ≥0.80 were considered large [48]. The research protocol for this study is available in Multimedia Appendix 2.

Results

Overview

The database search was conducted at two time points: once for papers published or accepted until July 2013, and once updated for papers published or accepted until October 2014 (see Figure 1). In total, the search yielded 8261 hits, from which 1742 (21.09%) duplicates were removed. Next, 6098 articles out of 6519 (93.54%) were deleted after reading the abstract and title. After reading the full texts of the remaining articles (n=421) and adding studies from other sources (n=18), such as a search in specialized journals and databases and requests via professional networks, 58 out of 439 (13.2%) game papers were retained that fit our inclusion/exclusion criteria.

Figure 1. Flowchart of game study selection.



Descriptive Information

A total of 58 games were studied in a total of 61 evaluations—3 papers studied more than one game and 3 games were also evaluated in more than one paper. Most studies came from North America (43/61, 71%), with the remainder originating from Europe (16/61, 26%) and Asia Pacific (2/61, 3%), including Australia and New Zealand. None came from Africa or Central or South America, possibly related to the English language restriction in our inclusion criteria. Most evaluated games for children (32/61, 53%) or adolescents (30/61, 49%). Around a third (21/61, 34%) evaluated games for adults, whereas only one game targeted the elderly. Over half the game studies were published in 2010 or later (33/61, 54%).

Moderating Role of User Involvement—None, Tester, or Participatory Design—in Game Effectiveness

For 8 out of 61 studies (13%), no information on target group involvement could be obtained. Of the remaining 53 evaluations, more than half of the game evaluations involved the target group in the game design (36/53, 68%). Of these 36 studies involving the target group, 11 (31%) game evaluations only pilot-tested a finished version, and 25 out of 36 (69%) involved the target group in PD. As previously reported, serious game studies were effective in changing healthy lifestyle behavior, behavioral determinants, and clinical outcomes [2].

Table 1 lists game study effect sizes and the role of user involvement as a moderator in explaining differences between these effect sizes. Results were organized by type of outcome: behavior, behavioral determinants, and clinical outcomes. Table 1 contains several analyses per outcome. The *Total* row reports the effect sizes for all included game studies that measured effects on a particular outcome. The row, *User involvement (between categories)*, shows the results of a moderator analysis testing for differences in effect sizes between game studies of different categories of user involvement (eg, for behavior, user involvement was a significant moderator in explaining heterogeneity between game studies [Q=6.74, *P*=.03]). The ensuing rows show the average effect sizes of game studies within each category of user involvement (eg, for game studies where users were not involved, effects on behavior were not significant [g=0.540, *P*=.17], but large differences existed in effect sizes of game studies belonging to this category [Q=324.39, *P*<.001]).

Table 1. User involvement as moderator in game study effectiveness.

Outcome	Sample size[a], n	k[b], n	Hedges' g[c] (95% CI)	*P* g	Q[d]	*P* Q	I[2e], %
Behavior							
Total[f]	317,582	24	0.216 (0.113 to 0.319)	<.001	408.37	<.001	94
User involvement (between categories)	11,684	22			6.74	.03	
User involvement (within categories)							
No involvement	3238	4	0.540 (-0.230 to 1.311)	.17	324.29	<.001	99
Involvement only as tester	430	4	0.520 (0.150 to 0.890)	.01	10.21	.02	71
Involvement as informant or codesigner	8016	14	0.075 (0.017 to 0.133)	.01	14.67	.33	11
Behavioral determinants							
Total[f]	22,366	51	0. 317 (0.244 to 0.391)	<.001	219.72	<.001	77
User involvement (between categories)	13,000	44			3.24	.20	
User involvement (within categories)							
No involvement	4649	16	0.288 (0.161 to 0.414)	<.001	39.26	.001	62
Involvement only as tester	1335	8	0.420 (0.278 to 0.562)	<.001	9.97	.19	30
Involvement as informant or codesigner	7016	20	0.265 (0.170 to 0.361)	<.001	49.37	<.001	62
Self-efficacy							
Total[f]	14,564	21	0.227 (0.130 to 0.324)	<.001	86.14	<.001	77
User involvement (between categories)	6398	20			7.83	.02	
User involvement (within categories)							
No involvement	1545	8	0.384 (0.283 to 0.485)	<.001	2.73	.91	0
Involvement only as tester	434	3	0.305 (-0.044 to 0.654)	.09	6.80	.03	71
Involvement as informant or codesigner	4419	9	0.171 (0.061 to 0.281)	.002	22.60	.004	65
Clinical outcomes							
Total[f]	9488	11	0. 071 (0.031 to 0.111)	.001	8.08	.62	0
User involvement (between categories)	1084	7			0.05	.82	
User involvement (within categories)							
No involvement	119	2	-0.036 (-0.402 to 0.330)	.85	0.20	.66	0
Involvement only as tester	384	3	0.034 (-0.168 to 0.236)	.74	0.63	.73	0
Involvement as informant or codesigner	700	4	-0.001 (-0.217 to 0.216]	.99	5.63	.13	47

[a] Combined participant sample size.

[b] k: number of studies.

[c] Hedges' g (random effects).

[d] Cochran's Q test: homogeneity statistic (mixed effects).

[e] I^2 index: inconsistency, a second measure of heterogeneity.

[f] The total includes the studies with missing values on the user involvement variable. For measures on behavior, one very large study (n=297,737) was included here, which was not involved in moderator analyses on user involvement, due to missing information.

There were significant differences (Q=6.74, *P*=.03) between the game studies' effects on behavior depending on the nature of target group involvement. Game studies for which the target group was involved in PD (g=0.075, *P*=.01) were significantly less effective in changing behavior than game studies that were only pilot-tested among the target group after design (g=0.520, *P*=.01). Effects of game studies developed in PD with the target group showed no significant differences with game studies that had not involved the target group (g=0.540, *P*=.17) (see Table 1). We assessed whether this applied to both older (ie, published before 2010) and more recent (ie, published between 2010 and 2014) game studies. When analyzing older and recent game studies separately, the difference in effectiveness on behavior by levels of active user involvement was only significant for recent game studies (Q=3.87, *P*=.049), but not for older game studies (Q=1.15, *P*=.28). As in the full sample analysis, games where users were involved merely as testers were more effective (g=0.577, *P*=.03) than games where users were involved in PD (g=0.066, *P*=.09). There were insufficient observations available here for comparison with studies without user involvement (see Multimedia Appendix 3, Tables A3-1 and A3-2).

We also assessed whether this applied both to game studies evaluated among a younger audience (ie, average age of participants <18 years) and among an adult audience (ie, average age of participants ≥18 years). This difference was only significant in studies with, on average, adult participants (Q=4.39, *P*=.04). Game studies among adult participants where the target group had only been involved as testers were more effective (g=0.577, *P*=.03) than studies developed in PD (g=0.009, *P*=.92). There were insufficient observations available here for comparison with studies without user involvement (see Multimedia Appendix 3, Tables A3-3 and A3-4).

There were no significant differences between the effects on behavioral determinants. When examining determinants separately, user involvement only significantly moderated effects on self-efficacy (Q=7.83, *P*=.02). Game studies for which the target group had been involved in PD (g=0.171, *P*=.002) were less effective in increasing self-efficacy than game studies for which the target group had not been involved in the design (g=0.384, *P*<.001), but showed no significant differences with game studies where the target group had been involved in pilot testing (g=0.305, *P*=.09) (see Table 1). When examining this by the publication year of the studies, this difference only appeared significant (Q=21.14, *P*=.01) for recent game studies (ie, published between 2010 and 2014) and not for studies published prior to 2010 (Q=2.63, *P*=.11). Recent game studies (ie, published 2010-2014) for which the target group had been involved in PD (g=0.098, *P*=.17) were less effective in increasing self-efficacy than game studies for which the target group had only been involved as testers (g=0.483, *P*<.001), but showed no significant difference with game studies where users were not involved in the design (g=0.281, *P*=.04) (see Multimedia Appendix 3, Tables A3-1 and A3-2).

We also assessed whether this applied both to game studies evaluated among a younger audience (ie, average age of participants <18 years) and among an adult audience (ie, average age of participants ≥18 years). This difference was only significant in studies with, on average, younger participants (Q=9.66, *P*=.002). Game studies among children and adolescents where the target group had not been involved were more effective (g=0.377, *P*<.001) than studies in this age group developed in PD (g=0.127, *P*=.03). There were insufficient observations available here for comparison with studies where users were only involved as testers (see Multimedia Appendix 3, Tables A3-3 and A3-4). Differences in effects on clinical outcomes could only be assessed between two levels of involvement. This difference in effects between levels of user involvement was not significant (see Table 1).

In summary, PD was associated with less health behavior change than when users only pilot-tested a finalized game version—all game studies and those with average sample age ≥18 years. There were no significant differences in game studies where users had not been involved in the design process. Games developed in PD also related to lower effects on self-efficacy than when users were not involved in any phase of the game design—all game studies and those with average sample age <18 years—or when users were only involved in pilot testing—recent game studies.

Differential Moderating Role of Participatory Design in Game Study Effectiveness by Design Elements

Descriptive Information

A total of 25 out of 36 (69%) game studies involved users in PD, either as informants or codesigners, and were available for a more in-depth exploration of the role of PD in game effectiveness. For 3 studies reported in 2 papers [49,50], no detailed information could be obtained from the authors; therefore, they were not included in the moderator analyses (see Table 2). This resulted in a total of 22 game studies for which detailed information was available.

Focus groups were the most frequently used method to consult users in the serious game development (14/22, 64%) [51-64]. A range of 2-16 focus groups informed game design. Out of 22 studies, 4 (18%) involved users via interviews [65-68], 2 (9%) via classroom discussions [58,69], and 1 (5%) via a survey [70]. In 3 studies out of 22 (14%), a user technical advisory board was established [55,57,62], and in 3 studies (14%) reported in 2 papers, target group users were recruited as members of the design team [70,71].

On average, 97.17 (SD 127.72) users per study were involved in PD of the game studies. Users were recruited via purposive sampling in 12 out of 22 studies (55%) [51-54,58,59,61,65,68,69,71], via a self-selected convenience sample (eg, flyers) in 8 studies out of 22 (36%) [55-57,60,62,63,66,70], and/or via snowball sampling in 4 out of 22 studies (18%) [55,57,62,64]. For one study, no information could be obtained on sampling method [67]. For some studies, the number of stages in which users were involved was unclear. A total of 6 out of 22 studies (27%)—discussed in 5 papers—mentioned continuous user involvement [59,70,71] or involvement in several stages [62,64]. For other studies, there was a range of 2-20 consultations, but it was unclear if these related to different design stages. Other stakeholders (eg, health professionals, parents, and teachers) were consulted in 13 out of 22 game studies (59%) [52,53,55-64,68]; on average, 24.91 (SD 22.24) other stakeholders were consulted.

Table 2. User involvement in game design elements.

Study[a]	N[b]	Ch[c]	CLV[d]	GL[e]	LW[f]	R[g]	L[h]	Co[i]	IA[j]	F[k]	T[l]	LC[m]
Campbell et al, 1999 [65]												I[n]
Yawn et al, 2000 [64]	C[o]	C	I	I				I	I			I
Baranowski et al, 2003 [69]		I	I	I	I	I						I
Campbell et al, 2004 [66]					I				I			I
Jago et al, 2006 [53]			I			I						I
Kato et al, 2008 [68]			I	I								I
Trepka et al, 2008 [63]			I	I								I
Thompson et al, 2009 [61]			I			I						I
Tortolero et al, 2010 [62]												I
Baranowski et al, 2011 [51]	I	I	I				I					
Dias and Aganti, 2011 [67]			I, C									
Sapouna et al, 2011 [58]	I, C	I, C	I	I	I, C			I, C	C			C
Swartz et al, 2011 [60]												I
Markham et al, 2012 [55]												I
Molnar and Kostkova, 2012 [56]		I										
Brown et al, 2012 [52]	I		I, C									
Schotland and Littman, 2012 [59]		I, C	I			I	I					
Christensen et al, 2013 [70]	I, C		I, C	I, C	I, C				I			
Majumdar et al, 2013 [54]	I	I	I			I		I	I			I
Song et al, 2013 (a-b[p]) [71]	C	C		C	C	C	I, C			C	C	C
Peskin et al, 2014 [57]												I

[a] Studies are chronologically organized by publication year.

[b] N: narrative.

[c] Ch: challenge.

[d] CLV: character looks/voice.

[e] GL: game world looks.

[f] LW: language, wording.

[g] R: rewards.

[h] L: levels.

[i] Co: controls.

[j] IA: interactivity/action.

[k] F: feedback.

[l] T: tailoring.

[m] LC: learning content.

[n] I: user involvement as informant.

[o] C: user involvement as codesigner.

[p] Two different types of games (a-b) were evaluated in this paper.

Moderating Role of Involvement as Informant or Codesigner in Game Study Effectiveness

As these analyses only related to some of the studies mentioned above (ie, 25 studies using PD), too few observations were available to analyze several of the moderators. This was the case for the moderating role of PD in game study effectiveness on clinical outcomes and for the moderating role of PD in game effectiveness on specific behavioral determinants (eg, attitudes and self-efficacy). Detailed analyses to assess the effectiveness of PD characteristics among PD studies by groups of publication year or average sample age were also not conducted due to insufficient observations.

Moderator analyses for PD as one category, considering informant and codesign together, are described in the following text. Results from detailed moderator analyses for the role as informants or as codesigner separately are summarized in the following text and described in detail in Multimedia Appendix 3 (Table A3-5).

The total number of aspects in which users were involved as informant or as codesigner was summed. There was no significant relationship between the number of areas in which users were involved as informants or codesigner and effectiveness on behavior (beta=.002, 95% CI -.026 to .029, *P*=.91, k=14, n=8016), nor on behavioral determinants (beta=.045, 95% CI -.009 to .098, *P*=.10, k=16, n=6155). Similar results were noted when analyzing the relationship between PD in these areas with effectiveness, separately for involvement as informant or as codesigner (see Multimedia Appendix 3, Table A3-5).

Moderating Role of Involvement as Informant or Codesigner in Mechanics, Dynamics, Aesthetics, or Educational Content in Game Study Effectiveness

There were insufficient studies (k=1) available to perform moderator analyses where users were involved in feedback or tailoring decisions.

Summed Mechanics, Dynamics, Aesthetics Areas

There were no significant relationships between degree of PD in mechanics (beta=-.030, 95% CI -.117 to .056, *P*=.20, k=14, n=8016), dynamic aspects (beta=.074, 95% CI -.040 to .187, *P*=.20, k=14, n=8016), or aesthetic aspects (beta=-.001, 95% CI -.038 to .035, *P*=.94, k=14, n=8016) and effectiveness on behavior.

There were also no significant relationships between degree of PD in mechanics (beta=-.078, 95% CI -.343 to .188, *P*=.57, k=16, n=6155) or aesthetic aspects (beta=.046, 95% CI -.030 to .121, *P*=.24, k=16, n=6155) and effectiveness on behavioral determinants. There was, however, a positive, significant relationship between the degree of PD in dynamics and effectiveness on behavioral determinants (beta=.215, 95% CI .075 to .356, *P*=.003, k=16, n=6155). Similar results were noted when analyzing the relationship between PD in these areas with effectiveness, separately for involvement as informant (beta=.235, 95% CI .079 to .329, *P*=.003, k=16, n=6155), but not for involvement as codesigner (see Multimedia Appendix 3, Table A3-5).

Specific Mechanics, Dynamics, Aesthetics Areas

We examined the role of PD as informant or codesigner in specific MDA areas of game design in explaining game study effectiveness. Significance of moderators is summarized in Table 3. Full information on these moderator analyses is provided in Multimedia Appendix 3 (Tables A3-7 to A3-15).

Table 3. Overview of significance of moderator analyses for the role of participatory design in specific Mechanics, Dynamics, Aesthetics design elements in game study effectiveness.

Comparisons between categories of involvement and no involvement	Informant/ codesigner		Informant		Codesigner	
	Q[a]	*P*	Q	*P*	Q	*P*
Behavior (between categories)						
Narrative	0.05	.82	0.05	.82	0.03	.87
Challenge	1.07	.30	1.07	.30	1.04	.31
Characters	0.36	.55	1.80	.18	4.38	.04
Game world	0.25	.62	0.25	.62	3.99	.046
Language	0.79	.37	0.79	.37	0.10	.75
Rewards	0.85	.36	0.85	.36	N/A[b]	
Levels	N/A		N/A		N/A	
Controls	0.37	.54	0.37	.54	N/A	
Action/interactivity	0.11	.74	0.54	.47	N/A	
Behavioral determinants (between categories)						
Narrative	0.43	.51	1.35	.24	0	.99
Challenge	2.93	.09	N/A	N/A	11.23	.001
Characters	1.17	.28	1.27	.26	0.21	.65
Game world	0.32	.57	0.05	.83	0.59	.44
Language	0.01	.93	0.46	.50	0.59	.44
Rewards	2.07	.15	0.20	.66	N/A	
Levels	7.02	.008	7.02	.008	N/A	
Controls	N/A		N/A		N/A	
Action/interactivity	0.48	.49	0.48	.49	N/A	

[a] Cochran’s Q test: homogeneity statistic (mixed effects).

[b] N/A: not applicable/not available due to insufficient observations.

PD in designing narratives, language, rewards, controls, and actions or interactivity did not relate significantly to game study effectiveness on behavior or its determinants. Significant differences were found for PD for creating the challenge, levels, character looks, and game world design. Involving users in PD on the challenge and levels related to higher game study effectiveness, whereas involvement in PD in character looks and game world design was associated with lower game study effectiveness. These differences in effects are discussed more in detail below.

When creating the challenge, there was a significant difference when considering codesign separately. Game studies where the challenge was codesigned with users (g=0.791, 95% CI 0.447 to 1.135, *P*<.001, k=4, n=318) were significantly more effective (Q=11.23, *P*=.001) in changing behavioral determinants than those where users had not been involved in codesign (g=0.192, 95% CI 0.128 to 0.257, *P*<.001, k=12, n=5837).

A significant relationship was also noted between user involvement as either informant or codesigner and game study effectiveness on behavioral determinants. Game studies where users were involved in the creation of levels as either informants or codesigners (g=0.771, 95% CI 0.347 to 1.196, *P*<.001, k=3, n=231) were significantly more effective (Q=7.02, *P*=.008) than game studies where users were not involved as informants or codesigners in level design (g=0.191, 95% CI 0.130 to 0.253, *P*<.001, k=13, n=5294). The same significant finding was noted when analyzing user involvement as informants separately, but could not be assessed for codesigners separately.

User involvement as codesigner, in the creation of characters, related to a significant difference in game study effectiveness on behavior. Game studies where users were involved in codesign of the characters (g=-0.022, 95% CI -0.128 to 0.083, *P*=.68, k=4, n=1485) were significantly less effective (Q=4.38, *P*=.04) at changing behavior than game studies where users were not involved in codesign (g=0.107, 95% CI 0.047 to 0.168, *P*<.001, k=10, n=6531). Differences in game study effectiveness on behavior were not significant when users were involved only as informants (Q=1.70, *P*=.18). Similarly, there was a significant difference in game effectiveness on behavior when only considering codesign.

Game studies where users had been involved in codesign of the game world (g=-0.075, 95% CI -0.232 to 0.082, *P*=.35, k=1, n=624) were significantly less effective (Q=3.99, *P*=.046) at changing behavior than game studies where users had not been involved in codesign (g=0.095, 95% CI 0.039 to 0.150, *P*=.001, k=13, n=7392). Note that this finding is based on only one study, albeit with a sufficient sample size, where codesign had been applied in creating the game world. Differences in game study effectiveness on behavior were not statistically significant when users were involved only as informants (Q=0.05, *P*=.83).

Educational Content

There were no significant differences (Q=0.01, *P*=.94) between the effectiveness of game studies on behavior when users were involved in PD of the educational content (g=0.084, 95% CI 0.030 to 0.144, *P*=.003, k=11, n=7027) or where users were not involved (g=0.078, 95% CI -0.144 to 0.301, *P*=.49, k=3, n=989). There were also no significant differences (Q=1.56, *P*=.21) between the effectiveness on behavioral determinants where users were (g=0.196, 95% CI 0.113 to 0.278, *P*<.001, k=11, n=4269) or were not involved in PD of educational content (g=0.381, 95% CI 0.102 to 0.659, *P*=.007, k=5, n=1886). Similar results were noted when analyzing the relationship between PD in educational content with effectiveness, separately for involvement as informant or as codesigner (see Multimedia Appendix 3, Table A3-6).

Discussion

Principal Findings

This study assessed whether user involvement PD in serious games for healthy lifestyle promotion related to game study effectiveness, and if this relationship varied by target group and study characteristics, by the design role and areas in which users were active participants. To our knowledge, this is the first study that conducted a meta-analysis of the relationship of user involvement and PD with serious game effectiveness.

More than half of the game studies involved users to some extent, with 25 out of 61 (41%) game studies having involved users as either informants or codesigners. User involvement in early stages of game design was previously reported to be low [72]; we did not find this in our study. However, this discrepancy may be influenced by the large number of recent game studies included in our study, since user involvement in serious game design was suggested to be on the rise recently [72].

Our first research question examined whether user involvement related significantly to game study effectiveness. Findings showed that serious game studies where the target group was involved in PD (ie, as informants or as codesigners) were, surprisingly, less effective in changing health behavior than when the target group was only involved as *testers* in the game. This was especially so for recent game studies and for game studies evaluated among an, on average, adult audience. Games developed with PD were also less effective at increasing self-efficacy (ie, how competent users perceive themselves to be in adopting a healthy lifestyle) than games developed without user involvement, in recent game studies and in studies evaluated among an, on average, younger audience. These results deviate from conclusions from a systematic review on nongame interventions, where higher community participation in the design process was associated with a higher achievement of health outcomes [73]. These conclusions, however, were obtained from *vote counting* (ie, correlation between number of outcomes achieved and degree of community engagement) and were not meta-analytic findings, which would have taken sample sizes and strength of effects across studies into account.

Several hypotheses can be advanced to explain these unexpected findings. First, users may need to be more than just a member of the target group; for example, they need sufficient subject-domain or design expertise to create a successful partnership with game designers [74]. This expertise may help avoid the risk of users focusing on irrelevant aspects, or on ideas that conflict with the pedagogical goals of the game [28,36]. Information on expertise level could be surmised for only two studies, where users were full members of the design team. Future teams intending to build a serious game with PD may test selecting users with content or design expertise, adjusting for lack of expertise, for example, by choosing appropriate techniques [75], or taking time to create user expertise [28,75]. Lack of experience with PD on the professional side, for example, not providing clear instructions or expectations, insufficient positive reinforcement, and not succeeding in creating a shared goal, could also lower the success of PD [76,77], which may have been operating in the studies included in this meta-analysis. Studies using PD in a school context noticed that without clear guidance, instructions, or encouragement from educators, children only did the minimum required, did not use all functionalities, and were reluctant to experiment and revise. When users did the bare minimum, they seemed to not know or share the overarching goal of the design process [76,77]. These findings also applied to adults where, additionally, not feeling recognized for their intellectual contribution could hinder the success of PD [77]. Information on these issues was not available for this review.

As a second hypothesis, codesign techniques (eg, storyboarding and paper prototyping) may need to be adjusted to the level of user design experience, users' cognitive abilities, and stage of game development [74,75]. This was not described in most included studies, but may be associated with the quality of idea generation [75], and may overcome several barriers influencing codesign, such as users tending to think along familiar lines instead of exploring new avenues, or forgetting or being afraid to bring up ideas when generating ideas in a group [75]. Future serious game studies using PD should more clearly describe the process and techniques used [22], to further advance our understanding of how PD may relate to game effectiveness.

Some age differences were noted. Although PD related to lower effectiveness in both age groups, this lower effectiveness was shown in different outcomes. PD related to a lower effectiveness on self-efficacy for a younger audience, and a lower effectiveness on behavior for an, on average, adult audience. While informant design and codesign are classifications that apply both to PD with children and adults [75], the specific techniques used within these approaches (eg, comic boarding and sticky notes) may not be equally effective among children and adults. It could be hypothesized that younger participants may be tempted to add game features that enhanced the game experience and enjoyment. However, these may lead to an increased cognitive load, ultimately decreasing learning outcomes [78] and perceived competence in performing the target behavior. Adults involved in PD may not have been aware of a gap between positive behavioral intentions and behavioral performance and the methods needed to bridge this gap, such as goal-setting techniques [79]. Further research and possibly adjustment of techniques by age group are therefore warranted. More detailed descriptions of the process and techniques used in PD may also aid in understanding the lower effectiveness of PD in recent game studies. It seems from Table 2 that older game studies used more informant design, whereas newer studies used more codesign. These codesign methods may not yet be as well established as informant design methods or, given a recent proliferation of PD during game development [72], may not be adequately applied or used after insufficient training in PD. These are hypotheses that remain to be tested in future research.

Third, since our findings showed significant effectiveness differences by forms of PD, further examination of differences between roles of informants and codesigners might provide an explanation for differences in effectiveness. Case studies have suggested involving users as informants is more beneficial than involving them as codesigners [80,81], since users are often unable to consider game characteristics in relation to the learning objectives [81]. Our meta-analysis showed that the number of design elements where users were involved as informant or as codesigner did not significantly relate to game effectiveness on behavior or behavioral determinants.

Differences were found between involvement as informant or as codesigner in dynamics and game effectiveness on behavioral determinants. When examining user roles separately, a positive association was only significant when users were involved as informants and not when they were codesigners. This confirms observations from case studies that involving users as informants may be more effective than as codesigners in certain design elements [80,81].

User involvement as codesigner related to higher game effectiveness when involved in creating the game challenge, but to lower effectiveness if users were involved as codesigners in creating characters or in designing the looks of the game world. Different methods in PD, such as Informant Design [80], Bonded Design [75], and Co-design or Cooperative Inquiry [74], vary in the roles participants take in the process [75]. In Informant Design, users are asked to provide input and feedback at some, but not all, development stages. They provide suggestions, but do not carry out the design itself. In both Bonded Design and Cooperative Inquiry, users are involved in the creation of the material itself, and are considered integral design partners. In Bonded Design, user involvement may span a shorter time or may be restricted to certain development stages, whereas in Cooperative Inquiry, users are partners throughout the whole development process [75]. Our findings seem to provide more support for Informant Design than for Bonded Design (ie, shorter periods as codesigner) or Cooperative Inquiry (ie, continuous codesigning). Involving the target group as codesigners requires a lengthy iterative process compared to Informant Design, and compared to only pilot-testing a finalized version [28]. Possibly, there was insufficient time available in serious game design projects to properly execute codesign. Only 6 games using PD in our meta-analysis involved users throughout the entire design process, whereas 7 studies involved users in only one design element. This may indicate that most studies using codesign partners opted for Bonded Design rather than full and continuous cooperation.

Lastly, when users were involved in design decisions on game dynamics, specifically on creating levels, or were involved in creating the challenge, effectiveness at changing behavioral determinants was higher. Most game studies, however, involved users in the aesthetic parts of the design, such as character, narrative, or game world creation. These did not relate to game study effectiveness, or even related negatively to effectiveness. The positive relationship of PD in creating the game challenge, game levels, and dynamics with game study effectiveness may derive from the fact that these elements are more important to game study effectiveness in general. Involving users in trivial aspects not associated with effectiveness may be counterproductive and difficult to defend from a cost-effectiveness point of view.

The importance of user involvement in game dynamics may also be explained by the need to adjust the game demands to the cognitive and technical abilities of the users, since PD may help in creating better tools by assessing the cognitive processes of how users interact with technology [82].

In summary, our findings demonstrated that informant roles may have a stronger link to effectiveness than codesign, depending on the design element. Users should be involved in decisions that relate to the game dynamics (eg, levels) and to the challenge (eg, game type). Although PD in serious games for healthy lifestyle promotion was associated with lower effectiveness than only pilot-testing a version developed by professionals, several recommendations were made to increase effectiveness of PD in serious game design. Since other health interventions indicated that early user involvement was important to achieve reach, adoption, and sustained implementation of the intervention [83], it is essential to more fully understand the role of PD in serious games and how to optimize it, as all these elements together—see, for example, the Reach, Effectiveness, Adoption, Implementation, and Maintenance (RE-AIM) framework—determine the public health impact of health-promoting interventions [84,85].

Limitations and Directions for Future Research

Some limitations to this meta-analysis need to be noted. First, "no evidence for an effect" does not equal "evidence for no effect." In some areas, analyses were likely statistically underpowered or were not available due to insufficient observations (eg, codesign roles per design element). Second, moderator analyses are equivalent to correlational analyses. There is always the possibility of a hidden, third variable. For example, a lower effectiveness for game studies developed in codesign may be influenced by lack of time in the project for optimal execution of codesign, insufficient user expertise, and a mismatch between techniques and user roles, expertise, and design phase. A more rigorous description of PD processes and of the application of recommended methods and techniques is needed in future PD game research, which would also enable a uniform coding based on what is reported in papers. Third, publication bias on the effects for behavioral determinants was noted, as reported elsewhere [2]. The reported effect sizes for these outcomes may be overestimated as a result. Lastly, no information was available on the impact of the input provided by users on the eventual game. Given often-conflicting opinions and goals between users and professionals [23], it is possible users' input was not integrated into final versions of the games included in the meta-analysis.

Conclusions

Our findings indicate that serious game studies for healthy lifestyle promotion developed with PD were less effective than game studies where users only pilot-tested a version designed by professionals. However, significant differences existed between the effectiveness of game studies developed with PD, suggesting certain types of PD may be more effective. This was the case for game studies where users were involved as informants in the design of game dynamics (eg, game levels). Involving users as codesigners for character or game world creation was less effective than not involving users in codesign for these game features. User involvement in designing the game challenge as informant or codesigner, on the other hand, related to higher game study effectiveness. These findings suggest PD should be mostly focused on areas crucial for general game study effectiveness, such as game dynamics and the challenge, whereas involvement in more trivial aspects, such as other aesthetic components, may be counterproductive. Involving users in user testing and informant roles may be more beneficial than as codesigners. Further research is needed to more fully explore how to incorporate PD into serious game design.

Acknowledgments

This study received a grant (No. 110051) from the Flemish Agency for Innovation by Science and Technology. MV is supported by the Research Foundation Flanders (FWO) with a postdoctoral research fellowship (FWO13/PDO/191). DT and TB are supported by the US Department of Agriculture, funded in part with federal funds from the United States Department of Agriculture/Agricultural Research Service (USDA/ARS) under Cooperative Agreement number 3092-5-001-058. The authors wish to thank Wendy Van Lippevelde for her assistance in double coding, and all authors of the included studies for providing supplementary information when requested.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Coding frame.

[PDF File (Adobe PDF File), 41KB - jmir_v18i4e94_app1.pdf]

Multimedia Appendix 2

Research protocol.

[PDF File (Adobe PDF File), 99KB - jmir_v18i4e94_app2.pdf]

Multimedia Appendix 3

Tables on moderator analyses.

[PDF File (Adobe PDF File), 135KB - jmir_v18i4e94_app3.pdf]

References

1. Prensky M. Digital Game-Based Learning. 2nd edition. St. Paul, MN: Paragon House; 2007.
2. DeSmet A, Van Ryckeghem D, Compernolle S, Baranowski T, Thompson D, Crombez G, et al. A meta-analysis of serious digital games for healthy lifestyle promotion. Prev Med 2014 Dec;69:95-107. [doi: 10.1016/j.ypmed.2014.08.026] [Medline: 25172024]
3. DeSmet A, Shegog R, Van Ryckeghem D, Crombez G, De Bourdeaudhuij I. A systematic review and meta-analysis of interventions for sexual health promotion involving serious digital games. Games Health J 2015 Apr;4(2):78-90. [doi: 10.1089/g4h.2014.0110]
4. Fratiglioni L, Paillard-Borg S, Winblad B. An active and socially integrated lifestyle in late life might protect against dementia. Lancet Neurol 2004 Jun;3(6):343-353. [doi: 10.1016/S1474-4422(04)00767-7] [Medline: 15157849]
5. Kim YS, Park YS, Allegrante JP, Marks R, Ok H, Ok CK, et al. Relationship between physical activity and general mental health. Prev Med 2012 Nov;55(5):458-463. [doi: 10.1016/j.ypmed.2012.08.021] [Medline: 22981733]

6. Lopez AD, Mathers CD, Ezzati M, Jamison DT, Murray CJ. Global and regional burden of disease and risk factors, 2001: Systematic analysis of population health data. Lancet 2006 May 27;367(9524):1747-1757. [doi: 10.1016/S0140-6736(06)68770-9] [Medline: 16731270]
7. Peel NM, McClure RJ, Bartlett HP. Behavioral determinants of healthy aging. Am J Prev Med 2005 Apr;28(3):298-304. [doi: 10.1016/j.amepre.2004.12.002] [Medline: 15766620]
8. Krebs P, Prochaska JO, Rossi JS. A meta-analysis of computer-tailored interventions for health behavior change. Prev Med 2010;51(3-4):214-221 [FREE Full text] [doi: 10.1016/j.ypmed.2010.06.004] [Medline: 20558196]
9. Portnoy DB, Scott-Sheldon LA, Johnson BT, Carey MP. Computer-delivered interventions for health promotion and behavioral risk reduction: A meta-analysis of 75 randomized controlled trials, 1988-2007. Prev Med 2008 Jul;47(1):3-16 [FREE Full text] [doi: 10.1016/j.ypmed.2008.02.014] [Medline: 18403003]
10. Vogel JJ, Vogel DS, Cannon-Bowers J, Bowers CA, Muse K, Wright M. Computer gaming and interactive simulations for learning: A meta-analysis. J Educ Comput Res 2006 Sep 1;34(3):229-243. [doi: 10.2190/FLHV-K4WA-WPVQ-H0YM]
11. Graesser A, Chipman P, Leeming F, Biedenbach S. Deep learning and emotion in serious games. In: Ritterfeld U, Cody M, Vorderer P, editors. Serious Games: Mechanisms and Effects. New York, NY: Routledge; 2009:83-102.
12. Baert V, Gorus E, Mets T, Geerts C, Bautmans I. Motivators and barriers for physical activity in the oldest old: A systematic review. Ageing Res Rev 2011 Sep;10(4):464-474. [doi: 10.1016/j.arr.2011.04.001] [Medline: 21570493]
13. McGuire AM, Anderson DJ, Fulbrook P. Perceived barriers to healthy lifestyle activities in midlife and older Australian women with type 2 diabetes. Collegian 2014;21(4):301-310. [Medline: 25632727]
14. Toobert DJ, Strycker LA, Glasgow RE, Bagdade JD. If you build it, will they come? Reach and adoption associated with a comprehensive lifestyle management program for women with type 2 diabetes. Patient Educ Couns 2002;48(2):99-105. [Medline: 12401412]
15. Ritterfeld U, Cody M, Vorderer P. Introduction. In: Ritterfeld U, Cody M, Vorderer P, editors. Serious Games: Mechanisms and Effects. New York, NY: Routledge; 2009:3-9.
16. Homer BD, Hayward EO, Frye J, Plass JL. Gender and player characteristics in video game play of preadolescents. Comput Human Behav 2012 Sep;28(5):1782-1789. [doi: 10.1016/j.chb.2012.04.018]
17. Wang H, Shen C, Ritterfeld U. Enjoyment of digital games. What makes them "seriously" fun? In: Ritterfeld U, Cody M, Vorderer P, editors. Serious Games: Mechanisms and Effects. New York, NY: Routledge; 2009:25-47.
18. Yee N. Motivations for play in online games. Cyberpsychol Behav 2006 Dec;9(6):772-775. [doi: 10.1089/cpb.2006.9.772] [Medline: 17201605]
19. DeSmet A, Palmeira A, Beltran A, Brand L, Davies VF, Thompson D. The yin and yang of formative research in designing serious (exer-)games. Games Health J 2015 Feb;4(1):63-66. [doi: 10.1089/g4h.2014.0123] [Medline: 26181683]
20. Guy S, Ratzki-Leewing A, Gwadry-Sridhar F. Moving beyond the stigma: Systematic review of video games and their potential to combat obesity. Int J Hypertens 2011;2011:179124 [FREE Full text] [doi: 10.4061/2011/179124] [Medline: 21629863]
21. Halskov K, Hansen NB. The diversity of participatory design research practice at PDC 2002-2012. Int J Hum Comput Stud 2015 Feb;74:81-92. [doi: 10.1016/j.ijhcs.2014.09.003]
22. Frauenberger C, Good J, Fitzpatrick G, Iversen OS. In pursuit of rigour and accountability in participatory design. Int J Hum Comput Stud 2015 Feb;74:93-106. [doi: 10.1016/j.ijhcs.2014.09.004]
23. Baek E, Cagiltay K, Boling E, Frick T. User-centered design and development. In: Spector M, Merrill MD, van Merrienboer J, Driscoll MP, editors. Handbook of Research on Educational Communications and Technology. 3rd edition. New York, NY: Lawrence Erlbaum Associates; 2008:659-670.
24. Thompson D. Talk to me, please!: The importance of qualitative research to games for health. Games Health J 2014 Jun;3(3):117-118. [doi: 10.1089/g4h.2014.0023] [Medline: 26196170]
25. Gittelsohn J, Steckler A, Johnson CC, Pratt C, Grieser M, Pickrel J, et al. Formative research in school and community-based health programs and studies: "State of the art" and the TAAG approach. Health Educ Behav 2006 Feb;33(1):25-39 [FREE Full text] [doi: 10.1177/1090198105282412] [Medline: 16397157]
26. Bogart LM, Uyeda K. Community-based participatory research: Partnering with communities for effective and sustainable behavioral health interventions. Health Psychol 2009 Jul;28(4):391-393 [FREE Full text] [doi: 10.1037/a0016387] [Medline: 19594261]
27. Wang F, Hannafin MJ. Design-based research and technology-enhanced learning environments. Educ Technol Res Dev 2005 Dec;53(4):5-23. [doi: 10.1007/BF02504682]
28. Druin A. The role of children in the design of new technology. Behav Inf Technol 2002 Jan;21(1):1-25. [doi: 10.1080/01449290110108659]
29. Beltran A, O'Connor T, Hughes S, Baranowski J, Nicklas TA, Thompson D, et al. Alpha test of a videogame to increase children's vegetable consumption. Games Health J 2012 Jun;1(3):219-222 [FREE Full text] [doi: 10.1089/g4h.2011.0027] [Medline: 24761315]
30. Thompson D, Baranowski T, Buday R, Baranowski J, Juliano M, Frazior M, et al. In pursuit of change: Youth response to intensive goal setting embedded in a serious video game. J Diabetes Sci Technol 2007 Nov;1(6):907-917 [FREE Full text] [Medline: 19885165]

31. Israel BA, Parker EA, Rowe Z, Salvatore A, Minkler M, López J, et al. Community-based participatory research: Lessons learned from the Centers for Children's Environmental Health and Disease Prevention Research. Environ Health Perspect 2005 Oct;113(10):1463-1471 [FREE Full text] [Medline: 16203263]
32. Cargo M, Mercer SL. The value and challenges of participatory research: Strengthening its practice. Annu Rev Public Health 2008;29:325-350. [Medline: 18173388]
33. Sandoval JA, Lucero J, Oetzel J, Avila M, Belone L, Mau M, et al. Process and outcome constructs for evaluating community-based participatory research projects: A matrix of existing measures. Health Educ Res 2012 Aug;27(4):680-690 [FREE Full text] [doi: 10.1093/her/cyr087] [Medline: 21940460]
34. Borenstein M, Hedges L, Higgins J, Rothstein HR. Introduction to Meta-Analysis. West Sussex, UK: John Wiley & Sons; 2009.
35. Cugelman B. Gamification: What it is and why it matters to digital health behavior change developers. JMIR Serious Games 2013;1(1):e3 [FREE Full text] [doi: 10.2196/games.3139] [Medline: 25658754]
36. Ke F. An implementation of design-based learning through creating educational computer games: A case study on mathematics learning during design and computing. Comput Educ 2014 Apr;73:26-39. [doi: 10.1016/j.compedu.2013.12.010]
37. Higgings J, Green S. Cochrane Handbook for Systematic Reviews of Interventions, version 5.1.0. London, UK: The Cochrane Collaboration; 2011.
38. 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]
39. Walker SN, Sechrist KR, Pender NJ. The health-promoting lifestyle profile: Development and psychometric characteristics. Nurs Res 1987;36(2):76-81. [Medline: 3644262]
40. Bartholomew L, Parcel G, Kok G, Gottlieb N, Fernàndez ME. Planning Health Promotion Programs: An Intervention Mapping Approach. 3rd edition. San Francisco, CA: Jossey-Bass; 2011.
41. Hunicke R, LeBlanc M, Zubek R. MDA: A formal approach to game design and game research. In: Proceedings of the AAAI Workshop on Challenges in Game AI. 2004 Presented at: AAAI Workshop on Challenges in Game AI; July 25-26, 2004; San Jose, CA p. 1.
42. Baranowski MT, Bower PK, Krebs P, Lamoth CJ, Lyons EJ. Effective feedback procedures in games for health. Games Health J 2013 Dec;2(6):320-326. [doi: 10.1089/g4h.2013.1328]
43. Lyons EJ. Cultivating engagement and enjoyment in exergames using feedback, challenge, and rewards. Games Health J 2015 Feb;4(1):12-18. [doi: 10.1089/g4h.2014.0072] [Medline: 26181675]
44. Hedges LV. Distribution theory for Glass's estimator of effect size and related estimators. J Educ Behav Stat 1981 Jan 01;6(2):107-128. [doi: 10.3102/10769986006002107]
45. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ 2003 Sep 6;327(7414):557-560 [FREE Full text] [doi: 10.1136/bmj.327.7414.557] [Medline: 12958120]
46. Hempel S, Miles JN, Booth MJ, Wang Z, Morton SC, Shekelle PG. Risk of bias: A simulation study of power to detect study-level moderator effects in meta-analysis. Syst Rev 2013;2:107 [FREE Full text] [doi: 10.1186/2046-4053-2-107] [Medline: 24286208]
47. Thompson SG, Higgins JP. How should meta-regression analyses be undertaken and interpreted? Stat Med 2002 Jun 15;21(11):1559-1573. [doi: 10.1002/sim.1187] [Medline: 12111920]
48. Cohen J. Statistical Power Analysis for the Behavioral Sciences. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.
49. Mustanski B, Garofalo R, Monahan C, Gratzer B, Andrews R. Feasibility, acceptability, and preliminary efficacy of an online HIV prevention program for diverse young men who have sex with men: The keep it up! intervention. AIDS Behav 2013 Nov;17(9):2999-3012 [FREE Full text] [doi: 10.1007/s10461-013-0507-z] [Medline: 23673793]
50. Paperny DM, Starn JR. Adolescent pregnancy prevention by health education computer games: Computer-assisted instruction of knowledge and attitudes. Pediatrics 1989 May;83(5):742-752. [Medline: 2654867]
51. Baranowski T, Baranowski J, Thompson D, Buday R, Jago R, Griffith MJ, et al. Video game play, child diet, and physical activity behavior change: A randomized clinical trial. Am J Prev Med 2011 Jan;40(1):33-38 [FREE Full text] [doi: 10.1016/j.amepre.2010.09.029] [Medline: 21146765]
52. Brown K, Newby K, Bayley J, Joshi P, Judd B, Baxter A. Development and Evaluation of a Serious Game for Relationships and Sex Education Using Intervention Mapping. Coventry, UK: Coventry University; 2012:1-112.
53. Jago R, Baranowski T, Baranowski JC, Thompson D, Cullen KW, Watson K, et al. Fit for Life Boy Scout badge: Outcome evaluation of a troop and Internet intervention. Prev Med 2006 Mar;42(3):181-187. [doi: 10.1016/j.ypmed.2005.12.010] [Medline: 16458955]
54. Majumdar D, Koch PA, Lee H, Contento IR, Islas-Ramos AD, Fu D. "Creature-101": A serious game to promote energy balance-related behaviors among middle school adolescents. Games Health J 2013 Oct;2(5):280-290 [FREE Full text] [doi: 10.1089/g4h.2013.0045] [Medline: 24761326]
55. Markham CM, Tortolero SR, Peskin MF, Shegog R, Thiel M, Baumler ER, et al. Sexual risk avoidance and sexual risk reduction interventions for middle school youth: A randomized controlled trial. J Adolesc Health 2012 Mar;50(3):279-288. [doi: 10.1016/j.jadohealth.2011.07.010] [Medline: 22325134]

56. Molnar A, Kostkova P. On effective integration of educational content in serious games. In: Proceedings of 13th IEEE International Conference on Advanced Learning Technologies. 2013 Presented at: 13th IEEE International Conference on Advanced Learning Technologies; July 15-18, 2013; Beijing, China p. 15-18. [doi: 10.1109/ICALT.2013.94]
57. Peskin MF, Shegog R, Markham CM, Thiel M, Baumler ER, Addy RC, et al. Efficacy of It's Your Game-Tech: A computer-based sexual health education program for middle school youth. J Adolesc Health 2015 May;56(5):515-521. [doi: 10.1016/j.jadohealth.2015.01.001] [Medline: 25739520]
58. Sapouna M, Wolke D, Vannini N, Watson S, Woods S, Schneider W, et al. Virtual learning intervention to reduce bullying victimization in primary school: A controlled trial. J Child Psychol Psychiatry 2010 Jan;51(1):104-112. [doi: 10.1111/j.1469-7610.2009.02137.x] [Medline: 19703096]
59. Schotland M, Littman K. Using a computer game to teach young children about their brains. Games Health J 2012 Dec;1(6):442-448 [FREE Full text] [doi: 10.1089/g4h.2012.0039] [Medline: 24761319]
60. Swartz LH, Sherman CA, Harvey SM, Blanchard J, Vawter F, Gau J. Midlife women online: Evaluation of an Internet-based program to prevent unintended pregnancy & STIs. J Women Aging 2011;23(4):342-359. [doi: 10.1080/08952841.2011.613255] [Medline: 22014222]
61. Thompson D, Baranowski T, Baranowski J, Cullen K, Jago R, Watson K, et al. Boy Scout 5-a-Day Badge: Outcome results of a troop and Internet intervention. Prev Med 2009 Dec;49(6):518-526. [doi: 10.1016/j.ypmed.2009.09.010] [Medline: 19765608]
62. Tortolero SR, Markham CM, Peskin MF, Shegog R, Addy RC, Escobar-Chaves SL, et al. It's Your Game: Keep It Real: Delaying sexual behavior with an effective middle school program. J Adolesc Health 2010 Feb;46(2):169-179 [FREE Full text] [doi: 10.1016/j.jadohealth.2009.06.008] [Medline: 20113923]
63. Trepka MJ, Newman FL, Davila EP, Matthew KJ, Dixon Z, Huffman FG. Randomized controlled trial to determine the effectiveness of an interactive multimedia food safety education program for clients of the special supplemental nutrition program for women, infants, and children. J Am Diet Assoc 2008 Jun;108(6):978-984. [doi: 10.1016/j.jada.2008.03.011] [Medline: 18502229]
64. Yawn BP, Algatt-Bergstrom PJ, Yawn RA, Wollan P, Greco M, Gleason M, et al. An in-school CD-ROM asthma education program. J Sch Health 2000 Apr;70(4):153-159. [Medline: 10790839]
65. Campbell MK, Honess-Morreale L, Farrell D, Carbone E, Brasure M. A tailored multimedia nutrition education pilot program for low-income women receiving food assistance. Health Educ Res 1999 Apr;14(2):257-267 [FREE Full text] [Medline: 10387505]
66. Campbell MK, Carbone E, Honess-Morreale L, Heisler-Mackinnon J, Demissie S, Farrell D. Randomized trial of a tailored nutrition education CD-ROM program for women receiving food assistance. J Nutr Educ Behav 2004;36(2):58-66. [Medline: 15068753]
67. Dias M, Agante L. Can advergames boost children's healthier eating habits? A comparison between healthy and non-healthy food. J Consum Behav 2011 Jun 28;10(3):152-160. [doi: 10.1002/cb.359]
68. Kato PM, Cole SW, Bradlyn AS, Pollock BH. A video game improves behavioral outcomes in adolescents and young adults with cancer: A randomized trial. Pediatrics 2008 Aug;122(2):e305-e317. [doi: 10.1542/peds.2007-3134] [Medline: 18676516]
69. Baranowski T, Baranowski J, Cullen KW, Marsh T, Islam N, Zakeri I, et al. Squire's Quest! Dietary outcome evaluation of a multimedia game. Am J Prev Med 2003 Jan;24(1):52-61. [Medline: 12554024]
70. Christensen JL, Miller LC, Appleby PR, Corsbie-Massay C, Godoy CG, Marsella SC, et al. Reducing shame in a game that predicts HIV risk reduction for young adult MSM: A randomized trial delivered nationally over the Web. J Int AIDS Soc 2013;16(3 Suppl 2):18716 [FREE Full text] [Medline: 24242264]
71. Song H, Kim J, Kwon RJ, Jung Y. Anti-smoking educational game using avatars as visualized possible selves. Comput Human Behav 2013 Sep;29(5):2029-2036. [doi: 10.1016/j.chb.2013.04.008]
72. Khaled R, Vasalou A. Bridging serious games and participatory design. Int J Child Comput Interact 2014 May;2(2):93-100. [doi: 10.1016/j.ijcci.2014.03.001]
73. Krishnaswami J, Martinson M, Wakimoto P, Anglemeyer A. Community-engaged interventions on diet, activity, and weight outcomes in U.S. schools: A systematic review. Am J Prev Med 2012 Jul;43(1):81-91. [doi: 10.1016/j.amepre.2012.02.031] [Medline: 22704752]
74. Yip J, Clegg T, Bonsignore E, Gelderblom H, Rhodes E, Druin A. Brownies or bags-of-stuff? Domain expertise in cooperative inquiry with children. In: Proceedings of the 12th International Conference on Interaction Design and Children. 2013 Presented at: 12th International Conference on Interaction Design and Children; June 24-27, 2013; New York, NY p. 201-210. [doi: 10.1145/2485760.2485763]
75. Walsh G, Foss E, Yip J, Druin A. FACIT PD: A framework for analysis and creation of intergenerational techniques for participatory design. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2013 Presented at: SIGCHI Conference on Human Factors in Computing Systems; April 27-May 2, 2013; Paris, France p. 2893-2902. [doi: 10.1145/2470654.2481400]
76. Denner J, Werner L, Ortiz E. Computer games created by middle school girls: Can they be used to measure understanding of computer science concepts? Comput Educ 2012 Jan;58(1):240-249. [doi: 10.1016/j.compedu.2011.08.006]

77. Joubert M, Wishart J. Participatory practices: Lessons learnt from two initiatives using online digital technologies to build knowledge. Comput Educ 2012 Aug;59(1):110-119. [doi: 10.1016/j.compedu.2011.09.024]
78. Schrader C, Bastiaens TJ. The influence of virtual presence: Effects on experienced cognitive load and learning outcomes in educational computer games. Comput Human Behav 2012;28:648-658.
79. Gollwitzer P, Sheeran P. Implementation intentions and goal achievement: A meta-analysis of effects and processes. Adv Exp Soc Psychol 2006;38:69.
80. Scaife M, Rogers Y, Aldrich F, Davies L. Designing for or designing with? Informant design for interactive learning environments. In: Proceedings of the ACM SIGCHI Conference on Human factors in Computing Systems. 1997 Presented at: ACM SIGCHI Conference on Human factors in Computing Systems; March 22-27, 1997; Atlanta, GA p. 343-350. [doi: 10.1145/258549.258789]
81. Vasalou A, Ingram G, Khaled R. User-centered research in the early stages of a learning game. In: Proceedings of the Designing Interactive Systems Conference. 2012 Presented at: Designing Interactive Systems Conference; June 11-15, 2012; Newcastle Upon Tyne, UK p. 116-125. [doi: 10.1145/2317956.2317976]
82. Sharples M, Jeffery N, du Boulay J, Teather D, Teather B, du Boulay G. Socio-cognitive engineering: A methodology for the design of human-centred technology. Eur J Oper Res 2002 Jan;136(2):310-323. [doi: 10.1016/S0377-2217(01)00118-7]
83. Compernolle S, De Cocker K, Lakerveld J, Mackenbach JD, Nijpels G, Oppert JM, et al. A RE-AIM evaluation of evidence-based multi-level interventions to improve obesity-related behaviours in adults: A systematic review (the SPOTLIGHT project). Int J Behav Nutr Phys Act 2014;11:147 [FREE Full text] [doi: 10.1186/s12966-014-0147-3] [Medline: 25480391]
84. Glasgow RE, Vogt TM, Boles SM. Evaluating the public health impact of health promotion interventions: The RE-AIM framework. Am J Public Health 1999 Sep;89(9):1322-1327. [Medline: 10474547]
85. Glasgow RE, Lichtenstein E, Marcus AC. Why don't we see more translation of health promotion research to practice? Rethinking the efficacy-to-effectiveness transition. Am J Public Health 2003 Aug;93(8):1261-1267. [Medline: 12893608]

Abbreviations

C: user involvement as codesigner
Ch: challenge
CINAHL: Cumulative Index to Nursing and Allied Health Literature
CIPER: Interdisciplinary Centre for the Study of Human Performance
CLV: character looks/voice
Co: controls
DiGRA: Digital Games Research Association
F: feedback
FWO: Research Foundation-Flanders
GL: game world looks
I: user involvement as informant
IA: interactivity/action
k: number of studies
L: levels
LC: learning content
LW: language, wording
MDA: Mechanics, Dynamics, Aesthetics
N: narrative
N/A: not applicable
PD: participatory design
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
R: rewards
RE-AIM: Reach, Effectiveness, Adoption, Implementation, and Maintenance
T: tailoring
UCD: user-centered design
USDA/ARS: United States Department of Agriculture/Agricultural Research Service

Edited by G Eysenbach; submitted 19.03.15; peer-reviewed by D Majumdar, F Blumberg; comments to author 19.07.15; revised version received 28.07.15; accepted 14.12.15; published 29.04.16.

Please cite as:
DeSmet A, Thompson D, Baranowski T, Palmeira A, Verloigne M, De Bourdeaudhuij I
Is Participatory Design Associated with the Effectiveness of Serious Digital Games for Healthy Lifestyle Promotion? A Meta-Analysis
J Med Internet Res 2016;18(4):e94
URL: http://www.jmir.org/2016/4/e94/
doi:10.2196/jmir.4444
PMID:27129447

Original Paper

What is the Profile of Individuals Joining the KNEEguru Online Health Community? A Cross-Sectional Mixed-Methods Study

Philip Bright[1,2], HBSc, PGCRM, PGCHE; Karen Hambly[1], HBSc, PhD; Sandra Tamakloe[1], BSc (Hons), MSc

[1]School of Sport and Exercise Sciences, University of Kent, Chatham, United Kingdom

[2]Research Department, European School of Osteopathy, Maidstone, United Kingdom

Corresponding Author:
Philip Bright, HBSc, PGCRM, PGCHE
School of Sport and Exercise Sciences
University of Kent
Medway Building
Chatham Maritime
Chatham,
United Kingdom
Phone: 44 (0)1634 88 (880
Fax: 44 (0)1634 88 (880
Email: philbright@eso.ac.uk

Abstract

Background: The use of the Internet for seekers of health-related information provides convenience and accessibility to diverse sources (of variable quality) for many medical conditions. There is a suggestion that patients may find empowerment by engaging with Internet health care strategies and communities. The profile of consumers of online health information on knee pain has not been explored.

Objective: Our objective was to identify the characteristics and motivations of online health information-seekers accessing the online health community, KNEEguru (KG). The study was designed to obtain the respondents' sociodemographic profile, together with their main reasons and motivations for joining such a community, their health information-seeking behavior, the extent of their knee problems, and their general Internet usage.

Methods: We undertook an online questionnaire survey, offered to users of the KG website from June to July 2012. A mix of open and closed questions was used to facilitate inductive enquiry. Quantitative responses were analyzed using univariate analysis; qualitative thematic analysis of the open responses was completed and a conceptual model was developed.

Results: One-hundred and fifty-two respondents took part (11.56% response rate, 152/1315), with a mean age of 40.1 years. Of this cohort, 61.2% were female, 68.4% were in domestic partnerships, 57.2% were employed, 75.0% had higher education qualifications, and 80.3% were of white/Caucasian ethnicity. Females were associated with joining KG in order to get emotional support from other users (OR 2.11, 95% CI 1.04 - 4.27, *P*=.04). Respondents' self-perception of health was associated with reported quality of life (OR 10.86, 95% CI 3.85 - 30.43, *P*<.001). Facebook users were associated with joining KG to share experiences (OR 2.34, 95% CI 1.04 - 5.56, *P*=.03). Post-surgery respondents were associated with joining KG to compare symptoms with other users (OR 7.31, 95% CI 2.06 - 39.82, *P*<.001). Three key themes were induced: condition, emotion and support. Respondents expressed distress and frustration at uncertainty of prognosis around various knee conditions, with some users preferring to initially observe rather than engage. Conversely, a strong desire to inform and support other community members was stated with reciprocation of ideas and experiences. KG was conceptualized as a filter that takes an individual's condition and emotional response to that condition as basis for support; this filter facilitated validation as the outcome of engagement.

Conclusions: This study, in line with wider literature, suggests that users of an online knee-specific community are typically female, middle-aged, white/Caucasian, married, employed, and have attained a level of higher education. These users demonstrate a pragmatic approach to health care information with altruistic motivations and a desire to share experiences as a means of validation. This finding emphasizes a means of promoting efficient and appropriate online health care, and demonstrates the benefits of the Internet as a viable complement to clinical engagement.

(J Med Internet Res 2016;18(4):e84) doi:10.2196/jmir.5374

KEYWORDS

Information seeking behaviour; Internet; Nonverbal communication; Knee-pain

Introduction

The use of the Internet for seekers of health-related information provides convenience and accessibility to diverse sources of variable quality [1]. There is a suggestion that patients may find empowerment by engaging with Internet health care strategies [2]. There is also some perceived skepticism in seeking medical information online due to doubts about accuracy, reliability and bias [3]; this is further compounded with the potential danger that Internet health provision medicalizes the trivial and engenders the *sick* role [4]. Despite concerns regarding potential misinformation, online health communities (OHCs) continue to thrive with growing clinician moderation [5] to add credibility to the health-related information generated via social media [6]. This clinician-validated approach, alongside adherence to the Health on the Net Foundation code of conduct [7] and online assessment tools such as DISCERN instrument [8], are establishing quality benchmarks for online health care information [3].

OHCs and Internet-based health care strategies are as varied as the specific conditions they represent [9,10,11,12] and the multi-media aspects of the Internet are also being explored and assessed [13]. There are a number of joint-replacement and osteoarthritis (OA) resources online [10,14], which are purported to have a beneficial impact on patient/practitioner shared-decision making. Knee-related Internet resources and attitudes of the online communities of knee pain sufferers are not widely reported; this is despite self-care programs demonstrating efficacy for controlling pain and maintaining function [15]. Fifty percent of people aged 50 and over will report knee pain during any one year, with one quarter describing this joint pain as severe and disabling [16]. Increasing age, gender, and obesity are identified as risk factors for progression of knee OA in people older than 50, contributing to OA as the sixth most disabling condition globally [17]; younger individuals are more likely to suffer knee pain as a result of acute injury, repetitive strain, or rare juvenile onset of OA [18].

KNEEguru (KG) is an OHC with over 33,000 members. KG is stated as a resource for the general public with knee problems, particularly those who have had or are contemplating knee surgery, and is overseen by a range of clinical experts [19]. Previous studies have investigated activity levels of consumers on the KG website with regard to articular cartilage repair procedures and suitability of specific knee outcome measures to patients [20,21]. While the profile of general online health care consumers has been reported in adult populations [22,23,24], the profile and experiences of individuals who would selectively engage with knee-specific OHC are not known.

Aims & Objectives

This study sought to explore the expressed motivations of participants seeking specific online health information regarding the knee. The extent to which the perceived benefits and quantifiable motives were related to characteristics of respondents was also assessed. Relating this to theorized benefits and challenges of Internet health could potentiate further perspectives on knee-pain sufferers and how their profiles compare with other OHC users.

Methods

Design

A self-administered, cross-sectional survey of individuals registering on the KG website was undertaken from June to July 2012. Participants were self-selecting and opportunity sampling was deployed; invitation was via a *pop-up window* that appeared upon navigating to the KG registration page. The sole exclusion criterion was participants under 18 years of age. Informed consent was given by participants explicitly indicating agreement to complete the survey, and no incentive for participation was offered.

The questionnaire was hosted on the Bristol Online Survey (University of Bristol, UK) software platform and initially piloted for face validity. The instrument was designed to identify the characteristics and motivations of users of the website both quantitatively and qualitatively. The survey consisted of 30 main questions, four of which were open responses, and the remainder were closed or Likert scale questions (74 items including sub-questions). Anonymized participants' demographic and health status characteristics, extent of knee pain, reasons for registering on the website, and questions related to health information-seeking behavior were captured. There was no adaptive or conditional logic in the response processing, and the open qualitative questioning allowed respondents to directly elaborate on their experiences and motivations for engaging with KG (the instrument is included in Multimedia Appendix 1).

The procedures for handling, processing, storage, and destruction of the data were compliant with the Data Protection Act 1998. The University of Kent ethics committee provided approval for this study.

Analysis

A mix of open and closed questions was used to facilitate inductive enquiry. Summary statistics were calculated to report sociodemographic data, reasons and motivations for joining KG, Internet and social media usage, knee problem demographics, and participants' perception of health and quality of life. Cross-tabulations for quantitative responses were analyzed using a χ^2, Fisher's Exact test, and odds ratios to examine differences in proportions by responders' characteristics. Significance levels were set at $P<.05$ for the Pearson $\chi^2$and Fisher's exact tests; for all odds ratio calculations, a 95% confidence interval was calculated. Qualitative thematic analysis of the open responses was completed using a framework approach and iterative open coding. This technique was used to create an initial descriptive representation of themes and sub-themes encountered in the participants' narrative. Triangulation of inducted themes was completed by two independent researchers. Further refinement of thematic content engendered a conceptual model of how participants rationalized engagement.

Results of the study were analyzed in a mixed-methods approach using Excel version 14 (Microsoft Corporation, Redmond, WA, USA), SPSS version 20 (SPSS Inc., Chicago, IL, USA) and Analyse-it version 3.76 (Analyse-it Software, Ltd., Leeds, UK). Excel was used to store and analyze open-text, facilitating the coding framework and thematic analysis. Summary and inferential statistics were calculated using a combination of Excel, Analyse-it and SPSS.

Results

Qualitative Questionnaire Responses

One-hundred and fifty-two respondents took part (11.56 % response rate from 1315 registrants approached) with a mean age of 40.1 years. Of the 152 respondents, 61.2% (93/152) were female, 68.4% (104/152) were in domestic partnerships, 57.2% (87/152) were employed, 75.0% (114/152) had higher education qualifications and 80.3% (122/152) were of white/Caucasian ethnicity. The United States was the most represented domicile (55.3%, 84/152) followed by the United Kingdom (21.7%, 33/152), alongside a global selection of other nations. The highest proportion of responders (57.9%, 88/152) reported the sharing of experience as the important motivation for engaging with KG (Table 1).

Table 1. Responses to reasons for engagement questions.

Question of motivation for engagement	Percentage rating as important[a]
To get emotional support from others	38
To vent out emotions related to the knee problem	31
To validate my experience	43
To seek recognition	12
To offer emotional support to others	42
To share my experience with others	58

a: “Important” and “Very Important” grouped together compared to “Neither important or non-important”, “Not important at all”, “Not relevant”, and “Not very important”.

Gender was not typically statistically significant as a determinant of response; females were associated with joining KG in order to get emotional support from other users (OR 2.11, 95% CI 1.04 - 4.27, *P*=.04) but no difference existed when looking for information about health or use of social media. Respondents’ self-perception of health was significantly associated with reported quality of life (OR 10.86, 95% CI 3.85 - 30.43, *P*<.001). Facebook users demonstrated an association with joining KG to share experiences (OR 2.34, 95% CI 1.04 - 5.56, *P*=.03). Post-surgery respondents were associated with joining KG to compare symptoms with other users (OR 7.31, 95% CI 2.06 - 39.82, *P*< .001) rather than compare recovery (OR 2.34, 95% CI 0.75 - 8.72, *P*=.14). Education to a minimum of graduate level was seen as an indicator of high daily Internet usage when compared to secondary level attainment only (OR 13.29, 95% CI 1.26 - 67.28, *P*=.01).

Thematic Analysis of Qualitative Responses

Four themes and 43 sub-themes were initially derived from all 152 responses to the mandatory question: *Why are you registering with KNEEguru?* These responses were rarefied into three overarching themes and 24 sub-themes, outlined in Textbox 1: condition (8 sub-themes), emotion (9 sub-themes), and support (7 sub-themes). Inter-rater agreement on overarching- and sub-themes was 100% and 64% respectively. The thematic content will be discussed in turn with reference to illustrative quotes in Multimedia Appendix 2.

Textbox 1. Major themes and grouped sub-themes.

- Condition – relating to reported situation and extenuating circumstances
    - Prognosis/progression
    - Procedure/treatment
    - Symptom
    - Diagnosis
    - Resolution/recovery
    - Cost
    - Quality of life/debilitation
    - Quality of practitionership
- Emotion – relating to emotional impact on the lives of the responders
    - Confusion
    - Anxiety/frustration
    - Pragmatism/stoicism
    - Altruistic
    - Empathy
    - Empowerment/inspiration
    - Trust/confidence
    - Validation
    - Expectation
- Support – relating to perceived merit of engaging with the OHC
    - Shared experience
    - Surrogate
    - Guidance/contextualization/informed decision-making
    - Proactivity/self-management/locus of control
    - Voyeuristic
    - Future of healthcare
    - Beneficence

Condition

Participants were compelled to describe their predisposing knee-related issues, as a rationale for engagement. A major motivational factor reported was the issue surrounding prognosis or progression; individuals were either concerned at potential outcomes of their condition or recounted the prognostic information gained from medics or their own research. Sequelae of traumatic events alternated between positive and negative experiences (I.a) with potentially distressing outcomes also described (I.b). The rate of progression was closely monitored by some individuals and posted as a potential measure for comparison (I.c); limitation of specific procedures was then reported within the context of resolution (I.d).

Perceptions of condition effect and progress were intimately bound with an underlying causative incident or procedure often aligned to a specific diagnosis. Participants were erudite and well-versed in medical terminology from an informed and critical stance (I.e). Further context was provided by individual descriptions of symptoms both prior to intervention and in chronic situations (I.f). A rich thread of narrative illustrated participants' perspectives on perceived effects of their complaint. These physical manifestations were often cited as a primary reason for seeking guidance (I.g).

Resolution and recovery of participants' knee issues were key motivations for engagement with the KG forum. Many respondents expressed a strong desire to expedite a return to full function, or had regained appropriate functional status (I.h). Some individuals presented positive outcomes, potentially related to their prior standing (I.i). This finding was a counterpoint to the overarching cost, both financially and in terms of the quality of life, that participants emphasized. Individuals depicted insidious, limiting effects of their condition and resultant anxiety (I.j) leading to further distress, despondency and isolation (I.k).

The final concepts informing perception around participants' knee conditions were the quality of practitioner and consistency of patient-handling. The reported standard of care was highly variable relative to individual experience (I.l). Others stated satisfaction with the level of guidance around treatment (I.m) but competency was seen as an issue in the context of surgery, rehabilitation, and expectation (l.n, l.o).

Emotion

A strong emotional response to injury, treatment, and follow-up care was professed by most participants with varying degrees of impact. The sub-themes embodied were *confusion around conflicting advice* and *anxiety and frustration* at uncertainty of their situation, which was occasionally offset by pragmatism and stoicism. A strong altruistic tendency with empathetic reciprocation of experience was regularly articulated. The experience of engaging with the OHC was seen as empowering and inspirational, feeding off the legitimate shared experience of participants. This result engendered trust and confidence, which led to validation of the experience. The management of participants' expectations was then informed by this validation.

Participants expressed confusion with regard to their situation and the guidance provided from health care resources (II.a), exacerbated by the lack of support material found elsewhere on the Internet (II.b). The issue of uncertainty of diagnosis, when compounded by conflicting information, was also voiced (II.c). This confusion was seen to underpin anxiety and frustration, which prompted engagement with the community; standards of care and lack of progress incited further exasperation (II.d).

Specific technical issues around medical procedures were cited as causes of distress and concern by a number of participants (II.e). The general uncertainty or lack of clarity around impending procedures and their outcomes motivated some individuals (II.f); similarly, nuanced response to surgery prompted further need for counsel (II.g). Individuals offset these issues of anxiety and uncertainty with a pragmatic and stoic response. Experience provided a resigned attitude to outcome for some (II.h), while others were keen to avoid surgery with a reserved approach (II.i). Pragmatism and resignation were also described with a sense of personal responsibility regarding knee health (II.j) and resultant psychological impact (II.k).

A strong desire to inform and support other community members was stated with reciprocation of ideas and perspective. The altruistic desire to help others as a result of sharing the benefit of individual experience was expressed (II.l), and reciprocation of experience was expected (II.m). This altruism was seen as a determinant of empowerment and inspiration. Participants clearly described the motivation derived from engaging in the OHC as mitigating the effects of their knee problems (II.n). This result was framed by issues of trust and confidence influenced by internal and external factors. Internal factors were expressed as the uncertainty of the medical prognosis or rationalization of participants' condition (II.o, II.p). External influences were felt to be the direct consequences of medical staff and, as previously stated, variable standard of care (II.q, II.r).

Participants entrusted the authenticity of experiences described, often in counterpoint to their mistrust of practitioners. A common outcome described was validation of experience based around exposure to the OHC. The community mentality facilitated sharing and rationalization of experiences of knee pain via a self-determined process (II.s, II.t). This validation was explicit in terms of palliation of fear (II.u), while others saw a direct need for affirmation of their predicament (II.v). Many participants described their expectations of outcomes from KG interaction or previously unmet expectations. Generally increasing awareness and achieving an informed perspective were described (II.w). Participants anticipated management of their own expectations via collaboration with KG users (II.x), potentially avoiding prognostic changes eliciting concern (II.y).

Support

The emotional response to individuals' knee conditions engendered various concepts of support. Responses commonly manifested as descriptions of shared experience with the outcome of validation and awareness (III.a). Sharing information was seen as a pathway to substantiate participants' experience (III.b) and this reciprocity was anticipated as a consequence of involvement (III.c). Engagement was often undertaken by surrogates demonstrating concern and exploring outcomes for close relatives; the individual's enquiries were often necessitated as a primary carer (III.d). The process of support and guidance was emphasized in respect to trauma (III.e); these complications of events around others were often the cause for concern that prompted action (III.f).

The sub-theme of guidance and contextualization was readily expressed as part of the information-seeking behavior. Participants were avid consumers of knee health care (III.g). Others were motivated by existing discussion material and suitably consoled to pursue further support (III.h). Guidance sought was often tempered by the progress reported by others (III.i). The expectation expressed was that the process of guidance would lead to informed decision-making regarding procedures or prognoses (III.j). The participants rationalized this advice and guidance as a means for reassurance and affirmation (III.k).

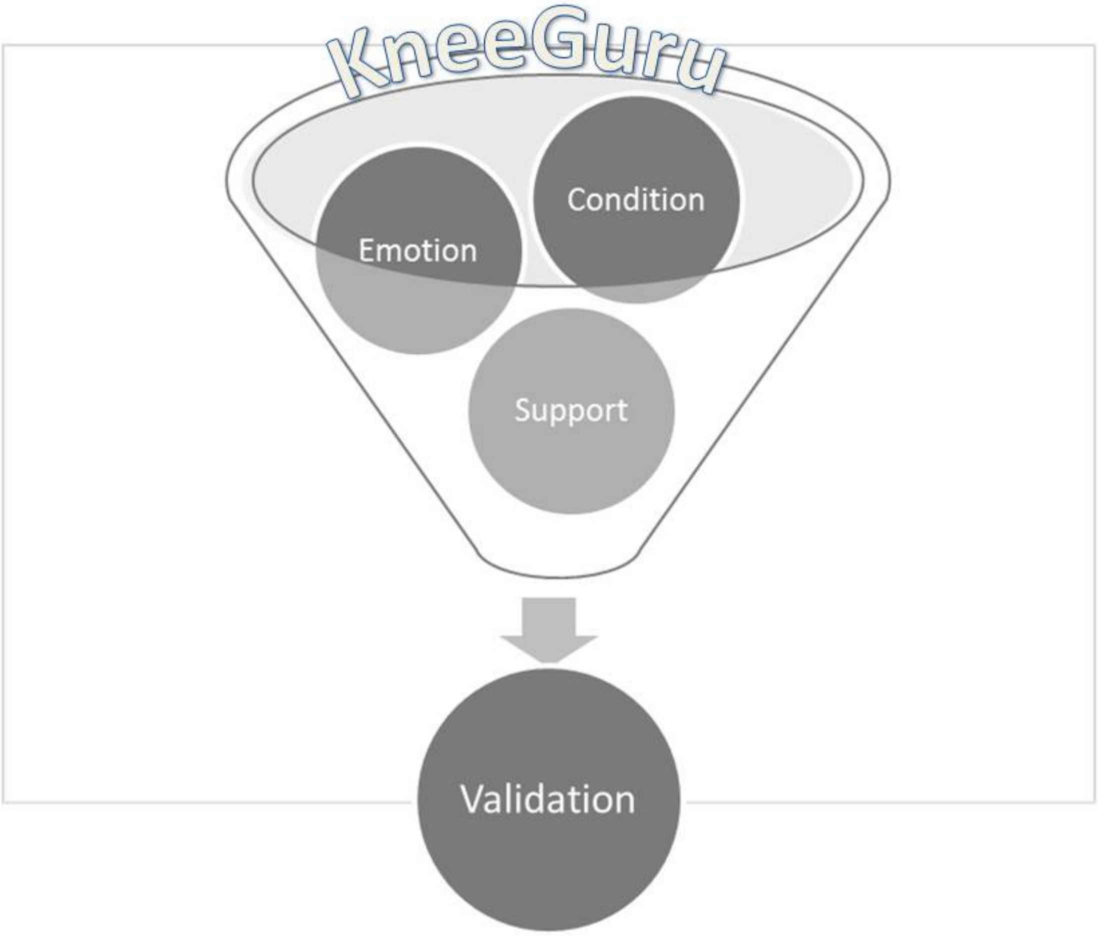
A key element of support was seen as facilitating proactivity via a forum for self-management and autonomy (III.l, III.m). Respondents declared a growing need for establishing a locus of control through the community (III.n). The need to achieve a sense of authority over their knee condition was important to some participants (III.o). Certain individuals adopted a voyeuristic approach to engagement and chose to peruse material without full access to the OHC (III.p). Participants declared a history of observation with burgeoning extenuating circumstances dictating a course of action (III.q), while others simply declared a curiosity around fellow OHC consumers, stating that the sole reason for interaction was verification of users (III.r).

Interaction with web-based technology was identified as the future of health care by some respondents (III.s), and seen as being vital and trail-blazing (III.t). The general perception of an accessible, informed, and knowledgeable community (underpinned with expert advice) was seen as highly beneficial. This sub-theme of beneficence was described in terms of assistance and well-being (III.u). Mitigation of fear, distress, and symptom-response was also volunteered (III.v) with immersion within the OHC seen to establish a true community spirit (III.w).

The interlinked themes of condition, emotion, and support were seen to be related within the context of KG. Participants declared a condition-based knee issue and their consequent emotional response which demanded support. This framework led to the development of a final conceptual model (Figure 1):

The personal experience of engagement with the OHC is viewed with KG as a filter that takes an individual’s condition, and emotional response to that condition, that drives the need for support. Processing through this filter facilitates validation as the outcome of engagement. This validation is established through the community nature of KG and is seen to have a major beneficial effect for participants.

Figure 1. Conceptual model of engagement with online health communities.



Discussion

Principal Results

This study sought to explore the characteristics and expressed motivations for participants seeking specific online health information regarding complaints of the knee. The extent to which the perceived benefits and quantifiable motives were related to characteristics of respondents was also established. The participants were seen to have an emotional response to their knee conditions that prompted support through KG; this engagement proved to be a validatory experience.

While females were more represented in the cohort of responders, in line with other reports of OHC participants [22], gender was not always significant as a determinant of response. Female participants were associated with joining KG in order to get emotional support from other users. A higher incidence of females has been seen to engage with online support communities for combating depression [25]. This gender-related tendency is seemingly supported in anxiety-inducing behavior reported across various physical conditions such as cancer, flu, and respiratory disorders [26,27,28,29]. Qualitative emotional responses were described in detail by both our male and female respondents, potentially mitigating the gender selection bias commonly reported [23].

Our study respondents also demonstrated that no differences existed between genders when searching directly for information about health. This result may relate to the specificity of the OHC and musculoskeletal focus offered by KG. Musculoskeletal pain frequency is reportedly higher in females [30], alongside incidence of knee OA [31]. This tendency is mirrored by severity of knee pain reported for certain female populations [32], potentially mediated by biomechanics and progressive decline in estrogen [33]. Females have less functionality and activity following knee replacement in Western countries [34], while Asian populations seemingly have less gender-specific outcomes post-surgery [35]. Our demographic did not describe explicit issues experienced around gender as a motivational factor for engaging with KG. Community support for their presenting condition was highly regarded and accessible, but seemingly lacked recognition of the latest evidence describing the characteristics that influence knee pathology [36]. This trend may suggest that the decisions influenced by OHC are not always clinically rational, and females may be more likely to prevaricate in seeking a resolution for joint-related morbidity [37].

There may be a perceived inevitability regarding the condition of OA that marks this as a particularly nuanced area of health care [38,39]. The descriptions of being resigned to the outcome of the disease process reported by our participants may be an indication of awareness and expectations being influenced by wide-ranging sources [40]. Specific patient decision aids, akin to OHC, have been seen to have positive effects on patient choice and awareness, but have not led to significant differences in surgical outcomes [41]. Long-term patient expectations for OA may lead to the contemplation of surgery, but pain management and functional outcomes are more revered; generalized optimism for long-term outcomes prevail over short-term response [42]. Potential conflicts between informed patients' and clinicians' expectations, where the former value symptom relief and the latter prioritize safety [43], may also account for our study's dissonant theme of dissatisfaction with variable standards of health care.

This finding of criticality of clinical health encounters may be further supported by our finding of association between higher education and greater Internet usage, and wider implications of health-seeking information [44]. Further studies reporting on online behavior demographics show mixed issues regarding influences and participation with social media and subsequent outcomes [45,46]. The context and necessity of engagement would seem to be crucial with uptake of technology, and social networking, demonstrably related to age and generational cohort. The perceived ubiquity of technology in developed cultures is presented as both beneficent and maleficent in equal measure [40,47]. The disenfranchised, technologically-challenged individual may adopt a deterministic view that has no locus of control [48]. Our study's indication regarding education and online activity within Generation X (mean age 40 years) suggests a utilitarian adaption to keep pace with the digital natives of Generation Y born after 1980 [48].

Facebook users demonstrated an association with joining KG to share experiences; previous studies demonstrate the frequency of social networking site use was not a significant predictor of supportive interaction [49]. Facebook users have previously been shown to be more willing to engage with student and community activities [50,51]. Facebook's platform has also been successfully explored as a potential medium to disseminate knowledge-transfer of health care information regarding OA [52]. As Facebook has developed as an *intranet* within the Internet, it is quickly facilitating information exchange through selective sharing, interaction, and self-monitoring of activities [53]. The implications for general health care are still to be fully understood or widely adopted [54,55], but the facilitation of patient empowerment is a major development [56]. Arguably, as supported in our study, social networks acting as introducers for secure OHCs is a model that can authenticate patient experience and mitigate concerns surrounding privacy and social anxiety [57,58].

Participants' emotional responses were well-described, although this was not directly supported in our quantitative findings. Emotional support is reported across a range of conditions, with various blogging platforms and communities specifically created for provision of guidance and advice [59]. Emotional support is seen as more valuable, and likely to engender and prolong engagement, than informational support [60]. The outpouring of emotion in our thematic content suggested a catharsis borne out by the validatory statements. Online communities would seem to provide an outlet for greater unfettered expression, and exchange of sympathy, unrivalled by the clinical encounter alone [61]. The ideas of relatedness, mutual respect, and engendering competency that are purported to underpin OHC [62] can be seen as antecedents of shared-decision making, influencing primary health care and challenging paternalism [63]. The burgeoning OHC are informing patients' decisions and their impact is being felt across multiple conditions and scenarios [64,65,66].

Respondents' self-perception of health was significantly associated with reported quality of life (QOL). While seemingly obvious, concepts of health between patients and practitioners are rarely reported; it would appear that there is congruence but patients describe how they value the professional over the profession they represent [67]. This attitude was reported by our respondents with stated predilection for supporting clinicians based on personal preference. With relation to knee and hip OA, QOL has been influenced by attitudes to health and social support transactions outside of clinical encounters [68]. Our study's findings of the validatory experience offered by OHC participation elucidates the wider finding of social support components mitigating effects of OA and the negative impact on QOL [69,70].

Post-surgery respondents were associated with joining KG to compare symptoms with other users rather than compare recovery, which may be supported by psychological impact of symptoms on post-surgical knee outcomes [71]. The implications of anxiety and pain catastrophization around surgical procedures can spur further self-motivated desire to engage in social activity [72]. The descriptions of validating experience from our study potentiate the mitigation of postoperative pain predicted by catastrophizing [73]. Wider quantitative findings suggest the level of education, tangible support, problem-solving, coping, and internal locus of control reported in our study are predictive of functional outcome following knee surgery [74].

The qualitative responses provided further evidence of surgical outcome denoting condition as a motivation for engagement. The emotional impact of this was well-documented in our study and reflects wider reports of pre-surgical anxiety [75]. Self-efficacy measures are indicated as vital to postoperative psychological and functional outcomes [76]; the use of OHC as part of this self-determination demands greater scrutiny. The full package of care around knee conditions needs to be further developed to integrate recommended use of validated online communities that are proving to be viable resources to complement clinical rehabilitation and patient autonomy.

Limitations

Only 11.56% (152/1315) of registrants agreed to take part in the survey, which may limit generalization of the quantitative findings. The richness of the qualitative responses may be subject to a Pygmalion Effect [77]; individuals believing that appeasing expectations of the pedagogue (or researcher/clinician in this case) would provide them with greater subsequent consideration. The low response rate may indicate bias, but closer scrutiny suggests the respondent characteristics are representative of samples reported in similar studies. There is also evidence of concordance between the quantitative and qualitative findings.

Conclusions

This study, in line with wider literature, suggests that users of an online knee-specific community are typically female, middle-aged, white/Caucasian, married, employed, and have attained a level of higher education. Respondents demonstrate a pragmatic approach to health care information with altruistic motivations and a desire to share experiences as a means of validation. This finding emphasizes a means of promoting efficient and appropriate online health care, and demonstrates the benefits of the Internet as a viable complement to clinical engagement. Consideration of integrated packages of care around knee health should include the recommendation of OHC support in future.

Acknowledgments

PB was responsible for preparation of the manuscript and thematic analysis. PB and ST were responsible for statistical analyses. KH conceived and developed the questionnaire. All authors reviewed and approved the final version of the manuscript. The authors would like to thank Hannah Kasiri-Martino and Oliver Bright for assistance with the thematic triangulation.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Bristol Online Survey Questionnaire.

[PDF File (Adobe PDF File), 179KB - jmir_v18i4e84_app1.pdf]

Multimedia Appendix 2

Themes and illustrative quotes.

[PDF File (Adobe PDF File), 53KB - jmir_v18i4e84_app2.pdf]

References

1. Atkinson NL, Saperstein SL, Pleis J. Using the Internet for health-related activities: findings from a national probability sample. J Med Internet Res 2009;11(1):e4 [FREE Full text] [doi: 10.2196/jmir.1035] [Medline: 19275980]
2. Samoocha D, Bruinvels DJ, Elbers NA, Anema JR, van der Beek AJ. Effectiveness of web-based interventions on patient empowerment: a systematic review and meta-analysis. J Med Internet Res 2010;12(2):e23 [FREE Full text] [doi: 10.2196/jmir.1286] [Medline: 20581001]
3. Khazaal Y, Chatton A, Zullino D, Khan R. HON label and DISCERN as content quality indicators of health-related websites. Psychiatr Q 2012 Mar;83(1):15-27. [doi: 10.1007/s11126-011-9179-x] [Medline: 21547515]
4. Clark J. Medicalization of global health 1: has the global health agenda become too medicalized? Glob Health Action 2014;7:23998 [FREE Full text] [Medline: 24848659]
5. Huh J, Pratt W. Weaving clinical expertise in online health communities. Proc SIGCHI Conf Hum Factor Comput Syst 2014;2014:1355-1364 [FREE Full text] [doi: 10.1145/2556288.2557293] [Medline: 26413582]
6. Hajli MN, Sims J, Featherman M, Love PE. Credibility of information in online communities. J Strat Market 2014:1-16.
7. Health On the Net Foundation. 2013. URL: http://www.hon.ch/HONcode/ [accessed 2015-11-26] [WebCite Cache ID 6dKDOSfVW]
8. The DISCERN Instrument. 1999. URL: http://www.discern.org.uk/ [accessed 2015-11-26] [WebCite Cache ID 6dKDa6v3c]
9. Wong L, Yan H, Margel D, Fleshner NE. Urologists in cyberspace: a review of the quality of health information from American urologists' websites using three validated tools. Can Urol Assoc J 2013;7(3-4):100-107 [FREE Full text] [doi: 10.5489/cuaj.501] [Medline: 23671523]

10. Nassiri M, Bruce-Brand RA, O'Neill F, Chenouri S, Curtin PT. Surfing for hip replacements: has the internet tidal wave led to better quality information. J Arthroplasty 2014 Jul;29(7):1339-1344.e1. [doi: 10.1016/j.arth.2014.01.009] [Medline: 24559520]
11. Chumber S, Huber J, Ghezzi P. A methodology to analyze the quality of health information on the internet: the example of diabetic neuropathy. Diabetes Educ 2015 Feb;41(1):95-105. [doi: 10.1177/0145721714560772] [Medline: 25480397]
12. Stienen JJ, Ottevanger PB, Wennekes L, Dekker HM, van der Maazen RW, Mandigers CM, et al. Development and evaluation of an educational e-Tool to help patients with non-Hodgkin's lymphoma manage their personal care pathway. JMIR Res Protoc 2015;4(1):e6 [FREE Full text] [doi: 10.2196/resprot.3407] [Medline: 25575019]
13. Gabarron E, Fernandez-Luque L, Armayones M, Lau AY. Identifying measures used for assessing quality of YouTube videos with patient health information: a review of current literature. Interact J Med Res 2013;2(1):e6 [FREE Full text] [doi: 10.2196/ijmr.2465] [Medline: 23612432]
14. Pietrzak E, Cotea C. The University of Queensland. 2012. Can the Internet improve health outcomes in osteoarthritis patients and provide useful information about the disease? URL: http://www.camvh.org.au/docs/ehealth/Telemanagement_of_osteoarthritis_26_04_12.pdf [accessed 2015-11-26] [WebCite Cache ID 6dKEruGzY]
15. Mazzuca SA, Brandt KD, Katz BP, Hanna MP, Melfi CA. Reduced utilization and cost of primary care clinic visits resulting from self-care education for patients with osteoarthritis of the knee. Arthritis Rheum 1999 Jun;42(6):1267-1273 [FREE Full text] [doi: 10.1002/1529-0131(199906)42:6<1267::AID-ANR25>3.0.CO;2-E] [Medline: 10366121]
16. Blagojevic M, Jinks C, Jeffery A, Jordan KP. Risk factors for onset of osteoarthritis of the knee in older adults: a systematic review and meta-analysis. Osteoarthr Cartil 2010 Jan;18(1):24-33 [FREE Full text] [doi: 10.1016/j.joca.2009.08.010] [Medline: 19751691]
17. Silverwood V, Blagojevic-Bucknall M, Jinks C, Jordan JL, Protheroe J, Jordan KP. Current evidence on risk factors for knee osteoarthritis in older adults: a systematic review and meta-analysis. Osteoarthr Cartil 2015 Apr;23(4):507-515 [FREE Full text] [doi: 10.1016/j.joca.2014.11.019] [Medline: 25447976]
18. Arden N, Nevitt MC. Osteoarthritis: epidemiology. Best Pract Res Clin Rheumatol 2006 Feb;20(1):3-25. [doi: 10.1016/j.berh.2005.09.007] [Medline: 16483904]
19. KNEEguru. 2015. URL: http://www.kneeguru.co.uk/ [accessed 2015-11-26] [WebCite Cache ID 6dKFmXBx8]
20. Hambly K. Activity profile of members of an online health community after articular cartilage repair of the knee. Sports Health 2011 May;3(3):275-282 [FREE Full text] [doi: 10.1177/1941738111402151] [Medline: 23016018]
21. Hambly K, Griva K. IKDC or KOOS? Which measures symptoms and disabilities most important to postoperative articular cartilage repair patients? Am J Sports Med 2008 Sep;36(9):1695-1704. [doi: 10.1177/0363546508317718] [Medline: 18577582]
22. Powell J, Inglis N, Ronnie J, Large S. The characteristics and motivations of online health information seekers: cross-sectional survey and qualitative interview study. J Med Internet Res 2011 Feb;13(1):e20 [FREE Full text] [doi: 10.2196/jmir.1600] [Medline: 21345783]
23. Bianco A, Zucco R, Nobile CG, Pileggi C, Pavia M. Parents seeking health-related information on the Internet: cross-sectional study. J Med Internet Res 2013;15(9):e204 [FREE Full text] [doi: 10.2196/jmir.2752] [Medline: 24047937]
24. Klemenc-Ketis Z, Kersnik J. Seeking health advice on the Internet in patients with health problems: a cross-sectional population study in Slovenia. Inform Health Soc Care 2013 Sep;38(3):280-290. [doi: 10.3109/17538157.2013.764306] [Medline: 23514043]
25. Houston TK, Cooper LA, Ford DE. Internet support groups for depression: a 1-year prospective cohort study. Am J Psychiatry 2002 Dec;159(12):2062-2068. [Medline: 12450957]
26. Brooks-Pollock E, Tilston N, Edmunds WJ, Eames KT. Using an online survey of healthcare-seeking behaviour to estimate the magnitude and severity of the 2009 H1N1v influenza epidemic in England. BMC Infect Dis 2011;11:68 [FREE Full text] [doi: 10.1186/1471-2334-11-68] [Medline: 21410965]
27. Uddin MJ, Alam N, Koehlmoos TP, Sarma H, Chowdhury MA, Alam DS, et al. Consequences of hypertension and chronic obstructive pulmonary disease, healthcare-seeking behaviors of patients, and responses of the health system: a population-based cross-sectional study in Bangladesh. BMC Public Health 2014;14:547 [FREE Full text] [doi: 10.1186/1471-2458-14-547] [Medline: 24888580]
28. Haggerty JL, Roberge D, Lévesque J, Gauthier J, Loignon C. An exploration of rural-urban differences in healthcare-seeking trajectories: implications for measures of accessibility. Health Place 2014 Jul;28:92-98. [doi: 10.1016/j.healthplace.2014.03.005] [Medline: 24793139]
29. Hvidberg L, Wulff CN, Pedersen AF, Vedsted P. Barriers to healthcare seeking, beliefs about cancer and the role of socio-economic position. A Danish population-based study. Prev Med 2015 Feb;71:107-113 [FREE Full text] [doi: 10.1016/j.ypmed.2014.12.007] [Medline: 25524610]
30. Fillingim RB, King CD, Ribeiro-Dasilva MC, Rahim-Williams B, Riley JL. Sex, gender, and pain: a review of recent clinical and experimental findings. J Pain 2009 May;10(5):447-485 [FREE Full text] [doi: 10.1016/j.jpain.2008.12.001] [Medline: 19411059]

31. Srikanth VK, Fryer JL, Zhai G, Winzenberg TM, Hosmer D, Jones G. A meta-analysis of sex differences prevalence, incidence and severity of osteoarthritis. Osteoarthr Cartil 2005 Sep;13(9):769-781 [FREE Full text] [doi: 10.1016/j.joca.2005.04.014] [Medline: 15978850]
32. Han H, Lee J, Kang S, Chang CB. The relationship between the presence of depressive symptoms and the severity of self-reported knee pain in the middle aged and elderly. Knee Surg Sports Traumatol Arthrosc 2015 May 17:1-9. [doi: 10.1007/s00167-015-3628-2] [Medline: 25982621]
33. Karsdal MA, Bay-Jensen AC, Lories RJ, Abramson S, Spector T, Pastoureau P, et al. The coupling of bone and cartilage turnover in osteoarthritis: opportunities for bone antiresorptives and anabolics as potential treatments? Ann Rheum Dis 2014 Feb;73(2):336-348. [doi: 10.1136/annrheumdis-2013-204111] [Medline: 24285494]
34. Cherian JJ, O'Connor MI, Robinson K, Jauregui JJ, Adleberg J, Mont MA. A prospective, longitudinal study of outcomes following total knee arthroplasty stratified by gender. J Arthroplasty 2015 Aug;30(8):1372-1377. [doi: 10.1016/j.arth.2015.03.032] [Medline: 25865810]
35. Gen LY, Bin Abd Razak HR, Chi CH, Chye TH. No gender-based differences in outcomes after conventional total knee arthroplasty in Asians. J Arthroplasty 2015 Sep;30(9):1548-1550. [doi: 10.1016/j.arth.2015.03.021] [Medline: 25869589]
36. Stubbs B, Hurley M, Smith T. What are the factors that influence physical activity participation in adults with knee and hip osteoarthritis? A systematic review of physical activity correlates. Clin Rehabil 2015 Jan;29(1):80-94. [doi: 10.1177/0269215514538069] [Medline: 24917590]
37. Mora M, Shell JE, Thomas CS, Ortiguera CJ, O'Connor MI. Gender differences in questions asked in an online preoperative patient education program. Gend Med 2012 Dec;9(6):457-462. [doi: 10.1016/j.genm.2012.10.002] [Medline: 23141296]
38. Collier R. A joint effort to fight arthritis. Can Med Assoc J 2012 Oct 16;184(15):E787-E788 [FREE Full text] [doi: 10.1503/cmaj.109-4306] [Medline: 23008485]
39. Gignac MA, Davis AM, Hawker G, Wright JG, Mahomed N, Fortin PR, et al. What do you expect? You're just getting older: a comparison of perceived osteoarthritis-related and aging-related health experiences in middle- and older-age adults. Arthritis Rheum 2006 Dec 15;55(6):905-912 [FREE Full text] [doi: 10.1002/art.22338] [Medline: 17139636]
40. Vance K, Howe W, Dellavalle RP. Social Internet sites as a source of public health information. Dermatol Clin 2009 Apr;27(2):133-6, vi. [doi: 10.1016/j.det.2008.11.010] [Medline: 19254656]
41. Stacey D, Taljaard M, Dervin G, Tugwell P, O'Connor AM, Pomey MP, et al. Impact of patient decision aids on appropriate and timely access to hip or knee arthroplasty for osteoarthritis: a randomized controlled trial. Osteoarthr Cartil 2015 Aug 4:1-9. [doi: 10.1016/j.joca.2015.07.024] [Medline: 26254238]
42. Dwek M, McBain H, Cleanthous S, Shipley M, Newman S. Osteoarthritis: patient expectations about future pain, stiffness and function. Musculoskeletal Care 2015 Jun;13(2):84-92. [doi: 10.1002/msc.1089] [Medline: 25470575]
43. Cordero-Ampuero J, Darder A, Santillana J, Caloto MT, Nocea G. Evaluation of patients' and physicians' expectations and attributes of osteoarthritis treatment using Kano methodology. Qual Life Res 2012 Oct;21(8):1391-1404. [doi: 10.1007/s11136-011-0058-6] [Medline: 22134806]
44. AlGhamdi KM, Moussa NA. Internet use by the public to search for health-related information. Int J Med Inform 2012 Jun;81(6):363-373. [doi: 10.1016/j.ijmedinf.2011.12.004] [Medline: 22217800]
45. Bolton RN, Parasuraman A, Hoefnagels A, Migchels N, Kabadayi S, Gruber T, et al. Understanding Generation Y and their use of social media: a review and research agenda. J Serv Manage 2013 Jun 14;24(3):245-267. [doi: 10.1108/09564231311326987]
46. Koteyko N, Hunt D, Gunter B. Expectations in the field of the Internet and health: an analysis of claims about social networking sites in clinical literature. Sociol Health Illn 2015 Mar;37(3):468-484. [doi: 10.1111/1467-9566.12203] [Medline: 25847533]
47. Lober WB, Flowers JL. Consumer empowerment in health care amid the Internet and social media. Semin Oncol Nurs 2011 Aug;27(3):169-182. [doi: 10.1016/j.soncn.2011.04.002] [Medline: 21783008]
48. Surry DW, Baker FW. The co-dependent relationship of technology and communities. Br J Educ Technol 2015 Sep 08;47(1):13-28. [doi: 10.1111/bjet.12349]
49. Oh HJ, Ozkaya E, LaRose R. How does online social networking enhance life satisfaction? The relationships among online supportive interaction, affect, perceived social support, sense of community, and life satisfaction. Comput Human Behav 2014 Jan;30:69-78. [doi: 10.1016/j.chb.2013.07.053]
50. Junco R. The relationship between frequency of Facebook use, participation in Facebook activities, and student engagement. Comput Educ 2012 Jan;58(1):162-171. [doi: 10.1016/j.compedu.2011.08.004]
51. Tufekci Z. Grooming, gossip, Facebook and MySpace: what can we learn about these sites from those who won't assimilate? Inform Commun Soc 2008 Jun;11(4):544-564. [doi: 10.1080/13691180801999050]
52. Brosseau L, Wells GA, Brooks S, De Angelis G, Bell M, Egan M, et al. People getting a grip on arthritis II: an innovative strategy to implement clinical practice guidelines for rheumatoid arthritis and osteoarthritis patients through Facebook. Health Educ J 2013 Jan 25;73(1):109-125. [doi: 10.1177/0017896912471031]
53. Abell L, Brewer G. Machiavellianism, self-monitoring, self-promotion and relational aggression on Facebook. Comput Human Behav 2014 Jul;36:258-262. [doi: 10.1016/j.chb.2014.03.076]

54. Hawn C. Take two aspirin and tweet me in the morning: how Twitter, Facebook, and other social media are reshaping health care. Health Aff (Millwood) 2009;28(2):361-368 [FREE Full text] [doi: 10.1377/hlthaff.28.2.361] [Medline: 19275991]
55. Maher CA, Lewis LK, Ferrar K, Marshall S, De BI, Vandelanotte C. Are health behavior change interventions that use online social networks effective? A systematic review. J Med Internet Res 2014;16(2):e40 [FREE Full text] [doi: 10.2196/jmir.2952] [Medline: 24550083]
56. Greene JA, Choudhry NK, Kilabuk E, Shrank WH. Online social networking by patients with diabetes: a qualitative evaluation of communication with Facebook. J Gen Intern Med 2011 Mar;26(3):287-292 [FREE Full text] [doi: 10.1007/s11606-010-1526-3] [Medline: 20945113]
57. Pedersen ER, Kurz J. Using Facebook for health-related research study recruitment and program delivery. Curr Opin Psychol 2016 Jun;9:38-43. [doi: 10.1016/j.copsyc.2015.09.011]
58. Shaw A, Timpano K, Tran T, Joorman J. Correlates of Facebook usage patterns: the relationship between passive Facebook use, social anxiety symptoms, and brooding. Comput Human Behav 2015;Correlates of Facebook usage patterns:575-580. [doi: 10.1016/j.chb.2015.02.003]
59. Ploderer B, Reitberger W, Oinas-Kukkonen H, van Gemert-Pijnen J. Social interaction and reflection for behaviour change. Pers Ubiquit Comput 2014 Jul 4;18(7):1667-1676. [doi: 10.1007/s00779-014-0779-y]
60. Wang YC, Kraut R, Levine J. To stay or leave?: the relationship of emotional and informational support to commitment in online health support groups. In: Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work. 2012 Feb 11 Presented at: Computer Supported Cooperative Work; Feb 11 2012; Seattle, WA p. 833-842. [doi: 10.1145/2145204.2145329]
61. Yao T, Zheng Q, Fan X. The impact of online social support on patients' quality of life and the moderating role of social exclusion. J Serv Res 2015 May 08;18(3):369-383. [doi: 10.1177/1094670515583271]
62. Zhang Y. Understanding the sustained use of online health communities from a self-determination perspective. J Assn Inf Sci Tec 2015 Jul 07:n/a-n/a. [doi: 10.1002/asi.23560]
63. Barry MJ, Edgman-Levitan S. Shared decision making -pinnacle of patient-centered care. N Engl J Med 2012 Mar 1;366(9):780-781. [doi: 10.1056/NEJMp1109283] [Medline: 22375967]
64. Rood JA, Van Zuuren FJ, Stam F, van der Ploeg T, Huijgens PC, Verdonck-de Leeuw IM. Cognitive coping style (monitoring and blunting) and the need for information, information satisfaction and shared decision making among patients with haematological malignancies. Psychooncology 2015 May;24(5):564-571. [doi: 10.1002/pon.3699] [Medline: 25333875]
65. Kehl KL, Landrum MB, Arora NK, Ganz PA, van Ryn M, Mack JW, et al. Association of actual and preferred decision roles with patient-reported quality of care: shared decision making in cancer care. JAMA Oncol 2015 Apr;1(1):50-58. [doi: 10.1001/jamaoncol.2014.112] [Medline: 26182303]
66. Hageman MG, Reddy R, Makarawung DJ, Briet JP, van Dijk CN, Ring D. Do upper extremity trauma patients have different preferences for shared decision-making than patients with nontraumatic conditions? Clin Orthop Relat Res 2015 Nov;473(11):3542-3548. [doi: 10.1007/s11999-015-4375-x] [Medline: 26040968]
67. Papp R, Borbas I, Dobos E, Bredehorst M, Jaruseviciene L, Vehko T, et al. Perceptions of quality in primary health care: perspectives of patients and professionals based on focus group discussions. BMC Fam Pract 2014;15:128 [FREE Full text] [doi: 10.1186/1471-2296-15-128] [Medline: 24974196]
68. Ethgen O, Vanparijs P, Delhalle S, Rosant S, Bruyère O, Reginster JY. Social support and health-related quality of life in hip and knee osteoarthritis. Qual Life Res 2004 Mar;13(2):321-330. [Medline: 15085904]
69. Oh J, Yi M. Structural equation modeling on quality of life in older adults with osteoarthritis. J Korean Acad Nurs 2014 Feb;44(1):75-85 [FREE Full text] [doi: 10.4040/jkan.2014.44.1.75] [Medline: 24637288]
70. French HP, Galvin R, Horgan NF, Kenny RA. Prevalence and burden of osteoarthritis amongst older people in Ireland: findings from The Irish LongituDinal Study on Ageing (TILDA). Eur J Public Health 2015 Jun 23:1-7. [doi: 10.1093/eurpub/ckv109] [Medline: 26105959]
71. Sullivan M, Tanzer M, Stanish W, Fallaha M, Keefe FJ, Simmonds M, et al. Psychological determinants of problematic outcomes following total knee arthroplasty. Pain 2009 May;143(1-2):123-129. [doi: 10.1016/j.pain.2009.02.011] [Medline: 19304392]
72. Somers TJ, Keefe FJ, Godiwala N, Hoyler GH. Psychosocial factors and the pain experience of osteoarthritis patients: new findings and new directions. Curr Opin Rheumatol 2009 Sep;21(5):501-506. [doi: 10.1097/BOR.0b013e32832ed704] [Medline: 19617836]
73. Khan RS, Ahmed K, Blakeway E, Skapinakis P, Nihoyannopoulos L, Macleod K, et al. Catastrophizing: a predictive factor for postoperative pain. Am J Surg 2011 Jan;201(1):122-131. [doi: 10.1016/j.amjsurg.2010.02.007] [Medline: 20832052]
74. Lopez-Olivo MA, Landon GC, Siff SJ, Edelstein D, Pak C, Kallen MA, et al. Psychosocial determinants of outcomes in knee replacement. Ann Rheum Dis 2011 Oct;70(10):1775-1781. [doi: 10.1136/ard.2010.146423] [Medline: 21791452]
75. Pouli N, Das NR, Lincoln NB, Walsh D. The experience of living with knee osteoarthritis: exploring illness and treatment beliefs through thematic analysis. Disabil Rehabil 2014;36(7):600-607. [doi: 10.3109/09638288.2013.805257] [Medline: 23789746]

76. Magklara E, Burton CR, Morrison V. Does self-efficacy influence recovery and well-being in osteoarthritis patients undergoing joint replacement? A systematic review. Clin Rehabil 2014 Sep;28(9):835-846. [doi: 10.1177/0269215514527843] [Medline: 24668361]
77. Dauvrin M, Lorant V. Leadership and cultural competence of healthcare professionals: a social network analysis. Nurs Res 2015;64(3):200-210 [FREE Full text] [doi: 10.1097/NNR.0000000000000092] [Medline: 25871625]

Abbreviations

KG: KNEEguru
OA: Osteoarthritis
OHC: online health communities
QOL: quality of life

Edited by G Eysenbach; submitted 01.12.15; peer-reviewed by M O'Connor, P Ghezzi; comments to author 20.01.16; revised version received 24.01.16; accepted 25.01.16; published 18.04.16.

Please cite as:
Bright P, Hambly K, Tamakloe S
What is the Profile of Individuals Joining the KNEEguru Online Health Community? A Cross-Sectional Mixed-Methods Study
J Med Internet Res 2016;18(4):e84
URL: http://www.jmir.org/2016/4/e84/
doi:10.2196/jmir.5374
PMID:27089531

Original Paper

The Pros and Cons of Getting Engaged in an Online Social Community Embedded Within Digital Cognitive Behavioral Therapy for Insomnia: Survey Among Users

Neil S Coulson[1], PhD; Richard Smedley[1], PhD; Sophie Bostock[2,3], PhD; Simon D Kyle[4], PhD; Rosie Gollancz[2], BSc; Annemarie I Luik[2,4], PhD; Peter Hames[2], MA; Colin A Espie[2,4], PhD

[1]Division of Rehabilitation & Ageing, School of Medicine, University of Nottingham, Nottingham, United Kingdom

[2]Big Health, London, United Kingdom

[3]University of Southampton, Faculty of Medicine, Southampton, United Kingdom

[4]Sleep and Circadian Neuroscience Institute (SCNi), Nuffield Department of Clinical Neurosciences, Oxford, United Kingdom

Corresponding Author:
Neil S Coulson, PhD
Division of Rehabilitation & Ageing
School of Medicine
University of Nottingham
Medical School, Queen's Medical Centre
University of Nottingham
Nottingham, NG7 2UH
United Kingdom
Phone: 44 1158466642
Fax: 44 115 8230231
Email: neil.coulson@nottingham.ac.uk

Abstract

Background: Sleepio is a proven digital sleep improvement program based on cognitive behavioral therapy techniques. Users have the option to join an online community that includes weekly expert discussions, peer-to-peer discussion forums, and personal message walls.

Objective: The aim of this study was to conduct an online survey to (1) explore the reasons for deciding to engage with the Sleepio online community, (2) explore the potential benefits arising from engagement with the online community, and (3) identify and describe any problematic issues related to use of the online community.

Methods: We developed an online survey and posted an invitation to the community discussion forum inviting users to participate. In addition, we sent an email invitation to 970 individuals who had previously or were currently working through the Sleepio program to participate in this study.

Results: In total, 100 respondents (70/100, 70% female; mean age 51 years, range 26–82 years) completed the online survey. Most respondents had started Sleepio with chronic sleep problems (59/100, 59% up to 10 years; 35/100, 35% >10 years) and had actively engaged with the online community (85/100, 85%) had made a discussion or wall post). At the time of the survey, respondents had used Sleepio for a median of 12 weeks (range from 3 weeks to 2 years). We analyzed responses to the open-ended questions using thematic analysis. This analysis revealed 5 initial drivers for engagement: (1) the desire to connect with people facing similar issues, (2) seeking personalized advice, (3) curiosity, (4) being invited by other members, and (5) wanting to use all available sleep improvement tools. Advantages of engagement included access to continuous support, a reduced sense of isolation, being part of a nonjudgmental community, personalized advice, positive comparisons with others, encouragement to keep going, and altruism. We found 5 potential disadvantages: design and navigation issues, uncertain quality of user-generated content, negative comparisons with others, excessive time commitments, and data privacy concerns. Participants related their community experiences to engagement with the Sleepio program, with many stating it had supported their efforts to improve their sleep, as well as helping with adherence and commitment to the program. Despite some concerns, members regarded the Sleepio community as a valuable resource.

Conclusions: Online communities may be a useful means through which to support long-term engagement with Web-based therapy for insomnia.

(J Med Internet Res 2016;18(4):e88) doi:10.2196/jmir.5654

KEYWORDS

engagement; sleep; online community; discussion forum; insomnia; cognitive behavioral therapy

Introduction

Epidemiological studies suggest that the prevalence of clinical insomnia disorder ranges from 10% to 12% worldwide [1,2], and problems are often long lasting [3]. Insomnia disorder is defined as a combination of difficulty initiating and maintaining sleep, but it also has a significant negative impact on daytime functioning [4]. A range of negative daytime consequences are typically linked to insomnia, including increased fatigue, diminished work productivity, lower quality of life, and lower relationship satisfaction, as well as increased incidence of poor health [5]. Furthermore, the development of mental and physical health problems has been linked to chronic insomnia [6,7].

While pharmacological treatments have an evidence base for improving insomnia, patients often report a preference for cognitive behavioral therapy (CBT) [8]. Indeed, the evidence for CBT as an effective treatment is persuasive [9-11]; however, access is limited due to resource and expertise constraints [12-15]. As a result, attention has turned to the Internet as a means to deliver treatment for chronic insomnia [16]. Furthermore, accumulating evidence in recent years suggests that such interventions can provide effective treatment for patients [17-19]. For example, one of the most rigorous studies of Web-based CBT used a randomized, placebo-controlled trial design to examine the impact of participation in the Sleepio intervention. The findings revealed that CBT delivered through a media-rich Web app with automated support and a community forum was effective in improving the sleep and associated daytime functioning of adults with insomnia disorder [19,20].

Sleepio is a subscription-based, fully automated digital sleep intervention that delivers CBT-based content to patients with chronic insomnia via the Web and mobile phones [21]. The course is delivered through 6 weekly sessions, facilitated by an animated character, “The Prof.” As each session begins, The Prof reviews progress, and examines the diary data submitted during the week, current sleep status and pattern, and progress made toward goals previously set. Automated algorithms tailor CBT content, incorporating behavioral (eg, stimulus control), cognitive (eg, thought restructuring), and lifestyle (sleep hygiene) components, to individual participant’s responses. Throughout the course, participants are also provided with the opportunity to engage with an optional community that comprises weekly discussions with a sleep expert (live text-based question-and-answer sessions), peer-to-peer discussion forums, and personal message walls. Comments on the Sleepio community can be rated by other users according to helpfulness, and color coding is used to signpost the most helpful users. On completion of the first 6 core sessions with The Prof, users become “graduates” and gain access to a graduate-only discussion forum. Active and consistently helpful graduates are identified by an external moderator and invited to volunteer to become “greeters,” which means they are able to send a welcome greeting to new users when they start Sleepio.

Several factors have been associated with greater levels of engagement with and lower rates of attrition from Web-based interventions, including interactivity, tailored content, social networking, and reminders [22-24]. However, there has been little evaluation of the role of online communities that are embedded within specific intervention programs, such as the Sleepio community. Understanding the user experience of such communities may provide useful insights for researchers, clinicians, and developers keen to maximize engagement with Web-based programs and reduce dropout.

The aim of this study was to explore, using a qualitative methodology, the experiences of individuals undertaking the Sleepio intervention who chose to engage with the optional online community. In particular, the study aimed to elicit the reasons for deciding to engage with the community, as well as any perceived benefits and disadvantages.

Methods

Procedure

We developed an online survey (see Textbox 1 for questions) using Bristol Online Surveys (University of Bristol, Bristol, UK). Next, we posted an invitation to the community discussion forum inviting users to participate. In addition, we sent an email invitation to 970 individuals who had previously or were currently working through the Sleepio program to participate in this study. Included within each invitation (and reminder) was a link to the study website, which explained the nature of the research and that ethical approval had been obtained from the University of Nottingham, Nottingham, UK. Each user, after indicating their consent, also provided their Sleepio username and date of birth in order to link their usage of the program with the responses they provided in this study.

Textbox 1. Interview questions on an online survey regarding usage of the Sleepio digital sleep intervention.

1. The Sleepio community is an optional part of the program; why did you decide to access it?
2. What are the benefits to you from accessing the community?
3. What do you value most about the Sleepio community?
4. How does it help? Could you give some examples of how it has helped you?
5. Can you describe any disadvantages? If so, what are they and do you have an idea how they can be resolved?

Data Analysis

We analyzed the responses to each question according to the principles of inductive thematic analysis, as described by Braun and Clarke [25]. Each set of responses was read multiple times by 2 authors (NC and RS) in order to gain familiarity with the data and to identify potential emerging themes. Next, these 2 authors discussed their preliminary analysis and together created a thematic framework that was then used to analyze all the responses independently. After further discussion, the authors (NC and RS) confirmed a set of themes, and identified and extracted relevant data to represent each of these themes. The remaining authors then reviewed the themes for coherence.

Results

In total, 100 respondents (70/100, 70% female; mean age 51 years, range 26–82 years) completed the online survey. Most respondents had started Sleepio with chronic sleep problems (59/100, 59% up to 10 years; 35/100, 35% >10 years) and had actively engaged with the online community (85/100, 85%) had made a discussion or wall post). At the time of the survey, respondents had used Sleepio for a median of 12 weeks (range from 3 weeks to 2 years).

Reasons for Engagement With the Sleepio Online Community

Respondents described 5 key motivations that led them to engage with the online community: (1) connecting with people facing similar issues, (2) seeking personalized advice, (3) curiosity, (4) being invited by other members, and (5) wanting to use all available sleep improvement tools.

Connecting With People Facing Similar Issues

For many, the motivation to engage with the online community stemmed from a need to connect with others facing similar issues. Users described how they wanted to compare their experiences of sleep disturbance with others’, and the community would provide a platform to do this. For many, the need to connect with others who “understand” what they were going through was of critical importance. For example, one user stated, “It was important for me to connect with the community as they were the people who were actually going through the program with me” (female, 60 years).

Seeking Personalized Advice

For some, the decision to access the online community arose from the need to ask a question, clarify how to do something, or use it as an additional source of information to supplement what they had learned during the CBT sessions. This helped individuals to obtain as much information as possible to help them overcome their sleep problems, provided them with advice on the practical aspects of implementing what they had learned during the course, and clarified what may or may not work. In the words of one participant, it allowed her “to ask detailed questions about how to implement certain aspects of the Sleepio course, and to ask advice about techniques” (female, 53 years).

Curiosity

For some users, curiosity was their motivation to access the online community. In particular, some described having no previous experience of online communities, discussion forums, or social media. As a consequence, they were interested to learn more about the Sleepio community, how it works, and the content of messages posted by other users. After having browsed the discussion forum to determine its potential relevance, most recognized that it was a potentially valuable and integral part of the Sleepio course: “It was there and I was curious so went for a look, and immediately realized how potentially valuable it was to me” (female, 43 years).

Being Invited by Other Members

A small number of users mentioned receiving an email from a graduate greeter inviting them to join the online community. While some had initially thought these invitations were automatically generated, on realizing it was indeed a real person they accepted the invitation. For others, there was no initial uncertainty and the invitation motivated them to access the community: “I received a message off someone and decided to have a look” (female, 32 years).

Wanting to Use All Available Sleep Improvement Tools

Some saw the online community as an integral part of the treatment program. As a result, users described their willingness to engage with the community in order to maximize their chances of sleep improvement: “It was part of the program offered and it was so important to me to try and improve my sleeping pattern I wanted to access all that was offered to give myself the best opportunity to achieve my goals” (no gender given, 56 years).

Perceived Benefits of Engagement With the Sleepio Online Community

The benefits derived from engaging with the online community were captured in 7 themes: continuous access, reduced sense of isolation, being part of a supportive and nonjudgmental community, personalized guidance and reassurance, positive comparison with others, encouragement to keep going, and altruism.

Continuous Access

The availability of the online community, 24 hours per day and 7 days per week, was appreciated, with many users describing how they benefited by simply knowing the community was accessible whenever they needed it. Beyond this, users could access the community and engage with it, “in my own time” (female, 45 years), with some describing how this was very helpful when engaging with sleep restriction tasks (ie, one of the tasks within the therapy part of the program), for example, “fill some of the hours when I have to stay awake” (female, 63 years). Moreover, easy access to the community meant that users could quickly solicit information or advice: “The immediacy of being able to connect and get answers” (female, 61 years).

Reduced Sense of Isolation

On accessing the online community, one of the most immediate benefits that users described was realizing that they were not the only person to be experiencing sleep-related problems: “It made me realize the extent to which sleep deprivation is affecting so many people and that you are not unique” (female, 64 years). Furthermore, the knowledge that they were no longer alone appeared to be of comfort for several: “Knowing you are not alone in your experiences means a lot” (female, 56 years). This was particularly noted when they felt isolated and without an adequate support network: “Sleep problems and insomnia is the loneliest place in the world” (female, 44 years).

Being Part of a Supportive and Nonjudgmental Community

The online community was considered a safe venue through which users could obtain much needed support: “It’s a safe place to talk about sleeping problems without them being dismissed as trivial” (female, 41 years). Indeed, users described how they felt that they could ask a question, discuss a personal issue, or simply vent without fear of ridicule or judgment. As a result, engaging with the online community appeared to be as important as the actual program itself in terms of supporting their efforts to improve their sleep. Several noted the “willingness of the community to engage, especially with newcomers to the program” (female, 68 years), and how their engagement with a community of “respect,” “understanding,” and “empathy” was in itself a great comfort.

Personalized Guidance and Reassurance

The Sleepio community appeared to be an important source of information and advice, and many users described examples of specific questions they had, that were then answered by other users: “It gave me answers that I needed to know and how to handle different situations, such as when being away on a trip” (female, 68 years). This was particularly evident during the early stages of the course, where many users described initial concerns or uncertainties: “Helped when I didn’t understand the format especially during the first couple of weeks” (female, 56 years).

Positive Comparisons With Others

By reading the stories, experiences, and updates posted by others, users were able to directly compare their own progress with that of others. In so doing, users described 2 distinct types of comparison. First, some users described how they compared themselves with others who were doing better, and that this was a source of inspiration: “seeing that others have succeeded” (female, 32 years). Second, and conversely, some described how they felt better after having read that others were in a much worse position than they were: “I...realized that there were many people with much worse issues” (male, 55 years).

Encouragement to Keep Going

Several users described how engagement with the online community motivated them to keep going with the course. This was particularly evident during challenging aspects of the course, such as the sleep restriction task: “I value being able to read other people’s experiences, how they have coped, especially when some weeks are challenging, such as the sleep restriction part of the program. I value feedback to my comments because they have been positive and motivating, it encourages me to persevere” (female, 56 years).

Altruism

For several users, it appeared important to be able to provide information, advice, and support, as well as to receive it. In so doing, these users clearly wished to give something back to the community: “The opportunity to help others which I like doing, is satisfying and enables me to return some of the benefit I have gained” (female, 77 years). For some, the act of giving back also helped them address their own sleep problems: “I’ve also found that helping others gives me insights into my own sleep problems” (male, 67 years).

Perceived Disadvantages of Engagement With the Sleepio Online Community

The disadvantages of using the online community focused on 5 areas of concern: design and navigation issues, uncertain quality of user-generated content, negative comparisons with others, excessive time commitments, and data privacy concerns.

Design and Navigation Issues

Some community users felt that some aspects of the onscreen user interface could be improved and described problems they had experienced while using it. Specifically, some found the interface difficult to navigate, especially if they did not have any prior experience of using forums or other social media: “I’ve never use a community chat before so was a bit nervous and unsure of how it worked” (female, 45 years). Some felt that navigating their way through conversations was difficult: “I often found the discussions hard to follow as it was sometimes difficult to see which questions were being answered and to follow a sequence of questions and answers through when other questions and answers popped up in the middle” (female, 68 years). Additionally, it wasn’t always clear to users where to post messages or how to search for information that was relevant to their particular problem: “Maybe a better search function would be a great way to find the posts that are relevant to my problem” (male, 68 years).

Uncertain Quality of User-Generated Content

Within the community, a vast amount of information and advice was shared online and some users appeared critical of the scientific evidence underpinning user-generated content: “Sometimes people post suggestions (about diet, natural remedies...) that are not supported by scientific evidence which I find confusing” (female, 38 years). This disadvantage was rooted in the fact that “the community aren’t sleep research experts” (male, 39 years) and, while often the information or advice may be based on experience, “their information is not reliable and you don’t know who they are, to be able to trust what they say” (female, 57 years). This issue was considered especially problematic if users were posting to help someone clearly in distress: “So the possibility of ill advice for someone who needs a doctor was the only possible downside I saw, not for me, but for another member” (female, 58 years).

Negative Comparisons With Others

For some users, reading about other people's experiences was on occasion problematic. In some instances, reading about the stress, anxiety, and distress being experienced by others was in itself a source of anxiety and stress: "I saw how others struggled, found it too tough, and gave up: this just scared me, made me anxious, tense and probably worked against what the program is meant to be doing" (female, 29 years). However, some users described similar feelings when others posted about their success: "Reading how wonderful people feel now they are improving after a few days of the program!!!, Not so nice to see this if it isn't working for you after a few weeks longer" (female, 59 years).

Excessive Time Commitments

Engaging with the online community was time consuming, and some users described feeling "overwhelmed" by the volume of messages being posted. As a result, some felt they were spending too much time in the community: "I can find myself spending way too much time or trying to make the time to see all" (female, 59 years). For others, reading messages was "addictive" and something that could "eat up time unnecessarily." However, time spent was meaningful, as users undertook to offer advice and support through posting messages or replying to questions: "It takes a lot of time to compose a meaningful helpful message, particularly to someone who is struggling" (female, 61 years).

Data Privacy Concerns

A small number of users expressed concerns regarding the perceived privacy of the discussion forum. In particular, some felt it was not sufficiently clear that it was openly accessible, and therefore any message posted to the forum could, in theory, be read by anyone. Additional concerns were focused on the belief that user-generated content could be accessed through Internet search engines. As a consequence, some users described feeling upset and explained how this had led them to disengage from making an active contribution (ie, posting) or to be more discerning about the content they uploaded to the forum: "Even though the community is anonymous and members' names are not public, the content of the messages is public and that was quite upsetting to discover that... I have been more careful since discovering that in terms of what I post" (female, 61 years).

Discussion

The problem of attrition within digital interventions has been widely documented [26]. Despite this, there has been little work exploring how online communities can support long-term engagement with Web-based therapy. In this study, we wished to explore Sleepio program participants' decisions to engage with and their experiences in the optional community. What is evident from their responses is that their reasons for engaging with the community varied but that their overall experiences were broadly positive and appeared to contribute to their sustained engagement with the Sleepio program and, for some, sleep improvement.

Arguably, one of the most important features of the community is that it is available at any point of the day or night, 7 days a week, and regardless of time zone. Furthermore, it is convenient and provides an immediate source of support to individuals who, according to their own stories, may feel isolated, exhausted, frustrated, and despondent as a result of their sleep problems [27].

Our results also revealed that the Sleepio program was considered demanding, especially the sleep restriction component. Previous work has indeed identified important challenges in the delivery of sleep restriction therapy to patients with insomnia [28,29], particularly in relation to adverse events and likely implementation or adherence challenges. Therefore, as an adjunct tool for individuals engaging with Web-based therapy for insomnia, the presence of a community, and in particular an asynchronous discussion forum, may be particularly helpful in terms of providing support.

For many, the community offered a means to connect with other people who not only had sleep problems but also were undertaking the Sleepio program. Indeed, this notion of connection was also reflected in responses concerning perceived benefits of the community. These findings resonate with previous research that has demonstrated how online communities can act as important venues through which individuals facing challenging health problems can soon feel less isolated and more supported [30-32], regardless of how much they actively engage with the community. Indeed, evidence suggests that community users typically vary in their level of engagement, ranging from those who post messages regularly through to those who simply lurk (ie, read messages only), but their experience may be equally rewarding [33].

The Sleepio community was certainly an important source of information, advice, and ideas for users, and this is also consistent with previous literature [30]. In particular, users sought and received a high volume of experiential information, which appeared to be well received by others, not least because it was understandable, relevant, and credible. Previous research has found information provision to be one of the key functions within online support communities [34,35]. We found a range of information being exchanged, including factual information, advice, and personal suggestions for coping strategies and management. Going forward, it would be useful to undertake a more fine-grained analysis of the specific types of informational support requested and exchanged within such an online community and how this relates to engagement with the various components of the Sleepio program.

In addition to the aforementioned examples of how the community benefited users, we also identified a clear motivational function. It was evident from responses that participation in the community provided the impetus to keep going, and users offered messages of support and encouragement, especially during difficult parts of the program (eg, sleep restriction). Indeed, previous research has identified poor adherence to CBT interventions for insomnia, particularly in the context of implementing behavioral advice [36]. Our results suggest that engagement with the online community may be one avenue through which user motivation and adherence can be supported. Such findings are comparatively novel within the broader literature, since the majority of online communities are either standalone or embedded within a program, but such online communities embedded within a program have typically not been evaluated in isolation. In future, researchers may wish to examine more explicitly the means through which engagement with online communities motivates individuals to persist with Web-based therapy for insomnia; however, our findings present some useful starting points for investigation (eg, provision of information and advice, and social comparison).

Our analysis also revealed some disadvantages arising from the use of the online community. While some of these (ie, design and navigation issues, and privacy concerns) were clearly associated with operational aspects of the Sleepio community and therefore can be easily addressed, others have been reported elsewhere in the literature. For example, some users were concerned about the accuracy of information posted by other members of the community. However, our analysis did not consider the accuracy of any information exchanged between users. In the future, it would be beneficial to ascertain the extent to which user-posted information may be inaccurate, not applicable, or perhaps unsafe. Despite these concerns, other studies that have examined the accuracy of information posted by users of discussion forums suggest that this is not a common problem but may only be more problematic in communities with low levels of activity [37]. Indeed, other users swiftly correct any inaccurate information that is posted [38]. The Sleepio community, although not continuously monitored by staff, has several mechanisms in place to create a safe environment. First, posts can be flagged by other users when considered not in line with the terms of use. Second, keywords that, for example, display a medical risk or personal information, are flagged automatically. In addition, expert sessions take place on a weekly basis to provide support by a clinical psychologist.

While we found that reading about the experiences of other users could be helpful, we also found that, for some, this actually had a negative impact on their well-being. In particular, some considered reading about the “horror stories” posted by others to be unhelpful, upsetting, and distressing. Such findings have also been reported elsewhere in the literature and may reflect the fact that many individuals use discussion forums to vent [32] or offload their daily hassles and bad experiences [39]. However, for those who read such messages, the impact can be problematic. While this is a challenging issue to address, one solution may be in the design and organization of forums, where subforums may be used to channel different types of user-generated content into specific locations, for example, sleep problems while being at college, or success stories. Similar mechanisms are seen in face-to-face versions of group therapy; while groups are helpful, they can sometimes be experienced as negative.

One final challenge facing users was the amount of time it took to engage with the online community and in particular all the discussions taking place. This issue was further exacerbated when users wished to provide help and support to others, as they viewed composing a helpful and tailored message to be time consuming. As online communities, such as this, continue to flourish, this issue is likely to persist and may indeed worsen. Restructuring forums into subforums may help to address this issue as well, such that users can choose which content they wish to engage with each time they access the community.

Study Limitations

We acknowledge some limitations to this study. First, despite our best efforts to recruit participants, the sample size is modest, and therefore the extent to which the views expressed by respondents represents the entire pool of Sleepio users is debatable. One obvious concern may be that only those who held particularly positive or negative views toward the community chose to participate in this study. Similarly, since the majority of our respondents were active users of the community, future research needs to consider the experiences of those who either engaged in a limited way or not at all. Second, the data for this study were captured by a single online survey, and it could be argued that experiences, views, and attitudes toward the Sleepio community may change over time, as users move through the program. While this is arguably true, we did note a broad range of users (ie, newcomers, users enrolled on the course, and graduates of the course), as well as evidence of respondents recalling specific good or bad experiences within the community in the past. All that said, it would be useful to engage in more longitudinal work to follow the experiences of those engaging with the Sleepio program from the point of entry and to consider how participation in the community contributes to relevant outcomes and indicators of adherence to the Sleepio intervention content. Third, since the focus of this study was on users’ experiences, we know little about those who chose not to engage with the community and their reasons behind this decision.

Conclusion

Despite some concerns, members regarded the Sleepio community as a valuable resource. Online communities may be a useful means through which to support long-term engagement with Web-based therapy for insomnia.

Acknowledgments

The authors would like to thank all the participants who gave their time freely to support this work.

Conflicts of Interest

This study was supported by the Sleepio team; however, all data analysis was conducted by NC and RS.

References

1. Ohayon MM. Epidemiology of insomnia: what we know and what we still need to learn. Sleep Med Rev 2002 Apr;6(2):97-111. [Medline: 12531146]
2. Durrence HH, Riedel BW, Taylor DJ, Bush A, Lichstein K. Epidemiology of Sleep: Age, Gender, and Ethnicity. Mahwah, NJ: Lawrence Erlabaum; 2004.
3. Morin CM, Bélanger L, LeBlanc M, Ivers H, Savard J, Espie CA, et al. The natural history of insomnia: a population-based 3-year longitudinal study. Arch Intern Med 2009 Mar 9;169(5):447-453. [doi: 10.1001/archinternmed.2008.610] [Medline: 19273774]
4. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, 5th edition. Arlington, VA: American Psychiatric Association Publishing; 2013.
5. Kyle SD, Morgan K, Espie CA. Insomnia and health-related quality of life. Sleep Med Rev 2010 Feb;14(1):69-82. [doi: 10.1016/j.smrv.2009.07.004] [Medline: 19962922]
6. Vgontzas AN, Liao D, Pejovic S, Calhoun S, Karataraki M, Bixler EO. Insomnia with objective short sleep duration is associated with type 2 diabetes: A population-based study. Diabetes Care 2009 Nov;32(11):1980-1985 [FREE Full text] [doi: 10.2337/dc09-0284] [Medline: 19641160]
7. Baglioni C, Battagliese G, Feige B, Spiegelhalder K, Nissen C, Voderholzer U, et al. Insomnia as a predictor of depression: a meta-analytic evaluation of longitudinal epidemiological studies. J Affect Disord 2011 Dec;135(1-3):10-19. [doi: 10.1016/j.jad.2011.01.011] [Medline: 21300408]
8. Vincent N, Lionberg C. Treatment preference and patient satisfaction in chronic insomnia. Sleep 2001 Jun 15;24(4):411-417. [Medline: 11403525]
9. Irwin MR, Cole JC, Nicassio PM. Comparative meta-analysis of behavioral interventions for insomnia and their efficacy in middle-aged adults and in older adults 55+ years of age. Health Psychol 2006 Jan;25(1):3-14. [doi: 10.1037/0278-6133.25.1.3] [Medline: 16448292]
10. Morin CM, Bootzin RR, Buysse DJ, Edinger JD, Espie CA, Lichstein KL. Psychological and behavioral treatment of insomnia:update of the recent evidence (1998-2004). Sleep 2006 Nov;29(11):1398-1414. [Medline: 17162986]
11. Riemann D, Perlis ML. The treatments of chronic insomnia: a review of benzodiazepine receptor agonists and psychological and behavioral therapies. Sleep Med Rev 2009 Jun;13(3):205-214. [doi: 10.1016/j.smrv.2008.06.001] [Medline: 19201632]
12. Dyas JV, Apekey TA, Tilling M, Ørner R, Middleton H, Siriwardena AN. Patients' and clinicians' experiences of consultations in primary care for sleep problems and insomnia: a focus group study. Br J Gen Pract 2010 May;60(574):e180-e200 [FREE Full text] [doi: 10.3399/bjgp10X484183] [Medline: 20423574]
13. Falloon K, Arroll B, Elley CR, Fernando A. The assessment and management of insomnia in primary care. BMJ 2011 May 27;342:2899. [Medline: 21622505]
14. Espie CA. "Stepped care": a health technology solution for delivering cognitive behavioral therapy as a first line insomnia treatment. Sleep 2009 Dec;32(12):1549-1558 [FREE Full text] [Medline: 20041590]
15. Espie C, Hames P, McKinstry B. Use of the internet and mobile media for delivery of cognitive behavioral insomnia therapy. Sleep Med Clin Sep 30 2013;8(3):407-419. [doi: 10.1016/j.jsmc.2013.06.001]
16. Ritterband LM, Thorndike FP. The further rise of internet interventions. Sleep 2012 Jun;35(6):737-738 [FREE Full text] [doi: 10.5665/sleep.1850] [Medline: 22654185]
17. Ritterband LM, Thorndike FP, Gonder-Frederick LA, Magee JC, Bailey ET, Saylor DK, et al. Efficacy of an Internet-based behavioral intervention for adults with insomnia. Arch Gen Psychiatry 2009 Jul;66(7):692-698 [FREE Full text] [doi: 10.1001/archgenpsychiatry.2009.66] [Medline: 19581560]
18. Vincent N, Lewycky S. Logging on for better sleep: RCT of the effectiveness of online treatment for insomnia. Sleep 2009 Jun;32(6):807-815 [FREE Full text] [Medline: 19544758]
19. Espie CA, Kyle SD, Williams C, Ong JC, Douglas NJ, Hames P, Brown June S L. A randomized, placebo-controlled trial of online cognitive behavioral therapy for chronic insomnia disorder delivered via an automated media-rich web application. Sleep 2012 Jun;35(6):769-781 [FREE Full text] [doi: 10.5665/sleep.1872] [Medline: 22654196]
20. Espie CA, Kyle SD, Miller CB, Ong J, Hames P, Fleming L. Attribution, cognition and psychopathology in persistent insomnia disorder: outcome and mediation analysis from a randomized placebo-controlled trial of online cognitive behavioural therapy. Sleep Med 2014 Aug;15(8):913-917. [doi: 10.1016/j.sleep.2014.03.001] [Medline: 24791643]
21. Espie C, Kyle S, Sheaves B. Sleepio. 2016. URL: https://www.sleepio.com/ [accessed 2016-04-04] [WebCite Cache ID 6gWQg1Vme]
22. Davies C, Corry K, Van Itallie A, Vandelanotte C, Caperchione C, Mummery WK. Prospective associations between intervention components and website engagement in a publicly available physical activity website: the case of 10,000 Steps Australia. J Med Internet Res 2012;14(1):e4 [FREE Full text] [doi: 10.2196/jmir.1792] [Medline: 22260810]
23. Schneider F, de Vries H, Candel M, van de Kar A, van Osch L. Periodic email prompts to re-use an internet-delivered computer-tailored lifestyle program: influence of prompt content and timing. J Med Internet Res 2013;15(1):e23 [FREE Full text] [doi: 10.2196/jmir.2151] [Medline: 23363466]
24. Richardson CR, Buis LR, Janney AW, Goodrich DE, Sen A, Hess ML, et al. An online community improves adherence in an internet-mediated walking program. Part 1: results of a randomized controlled trial. J Med Internet Res 2010;12(4):e71 [FREE Full text] [doi: 10.2196/jmir.1338] [Medline: 21169160]

25. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol Jan 1 2006;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
26. Eysenbach G. The law of attrition. J Med Internet Res 2005;7(1):e11 [FREE Full text] [doi: 10.2196/jmir.7.1.e11] [Medline: 15829473]
27. Kyle SD, Espie CA, Morgan K. "...Not just a minor thing, it is something major, which stops you from functioning daily": quality of life and daytime functioning in insomnia. Behav Sleep Med 2010;8(3):123-140. [doi: 10.1080/15402002.2010.487450] [Medline: 20582756]
28. Kyle SD, Morgan K, Spiegelhalder K, Espie CA. No pain, no gain: an exploratory within-subjects mixed-methods evaluation of the patient experience of sleep restriction therapy (SRT) for insomnia. Sleep Med 2011 Sep;12(8):735-747. [doi: 10.1016/j.sleep.2011.03.016] [Medline: 21907616]
29. Kyle SD, Miller CB, Rogers Z, Siriwardena AN, Macmahon KM, Espie CA. Sleep restriction therapy for insomnia is associated with reduced objective total sleep time, increased daytime somnolence, and objectively impaired vigilance: implications for the clinical management of insomnia disorder. Sleep 2014 Feb;37(2):229-237 [FREE Full text] [doi: 10.5665/sleep.3386] [Medline: 24497651]
30. White M, Dorman SM. Receiving social support online: implications for health education. Health Educ Res 2001 Dec;16(6):693-707 [FREE Full text] [Medline: 11780708]
31. Buchanan H, Coulson N, Malik S. Health-related internet support groups and dental anxiety: the fearful patient's online journey. Int J Web Based Communities 2010 Jan 01;6(4):362-375. [doi: 10.1504/IJWBC.2010.035839]
32. Coulson NS. How do online patient support communities affect the experience of inflammatory bowel disease? An online survey. JRSM Short Rep 2013;4(8):2042533313478004 [FREE Full text] [doi: 10.1177/2042533313478004] [Medline: 24040493]
33. Mo P, Coulson N. Empowering processes in online support groups among people living with HIV/AIDS: A comparative analysis of ‘lurkers’ and ‘posters’. Comput Hum Behav 2010 Sep;26(5):1183-1193. [doi: 10.1016/j.chb.2010.03.028]
34. Braithwaite DO, Waldron VR, Finn J. Communication of social support in computer-mediated groups for people with disabilities. Health Commun 1999;11(2):123-151. [doi: 10.1207/s15327027hc1102_2] [Medline: 16370973]
35. Coulson NS, Buchanan H, Aubeeluck A. Social support in cyberspace: a content analysis of communication within a Huntington's disease online support group. Patient Educ Couns 2007 Oct;68(2):173-178. [doi: 10.1016/j.pec.2007.06.002] [Medline: 17629440]
36. Matthews EE, Schmiege SJ, Cook PF, Berger AM, Aloia MS. Adherence to cognitive behavioral therapy for insomnia (CBTI) among women following primary breast cancer treatment: a pilot study. Behav Sleep Med 2012;10(3):217-229. [doi: 10.1080/15402002.2012.666220] [Medline: 22742439]
37. Hwang KO, Farheen K, Johnson CW, Thomas EJ, Barnes AS, Bernstam EV. Quality of weight loss advice on internet forums. Am J Med 2007 Jul;120(7):604-609 [FREE Full text] [doi: 10.1016/j.amjmed.2007.04.017] [Medline: 17602934]
38. Esquivel A, Meric-Bernstam F, Bernstam EV. Accuracy and self correction of information received from an internet breast cancer list: content analysis. BMJ 2006 Apr 22;332(7547):939-942 [FREE Full text] [doi: 10.1136/bmj.38753.524201.7C] [Medline: 16513686]
39. Malik S, Coulson NS. 'They all supported me but I felt like I suddenly didn't belong anymore': an exploration of perceived disadvantages to online support seeking. J Psychosom Obstet Gynaecol 2010 Sep;31(3):140-149. [doi: 10.3109/0167482X.2010.504870] [Medline: 20718585]

Abbreviations

CBT: cognitive behavioral therapy

Edited by G Eysenbach; submitted 15.02.16; peer-reviewed by H Semper, C Williams; comments to author 09.03.16; revised version received 15.03.16; accepted 16.03.16; published 25.04.16.

Please cite as:
Coulson NS, Smedley R, Bostock S, Kyle SD, Gollancz R, Luik AI, Hames P, Espie CA
The Pros and Cons of Getting Engaged in an Online Social Community Embedded Within Digital Cognitive Behavioral Therapy for Insomnia: Survey Among Users
J Med Internet Res 2016;18(4):e88
URL: http://www.jmir.org/2016/4/e88/
doi:10.2196/jmir.5654
PMID:27113540

Original Paper

Crowdsourcing and the Accuracy of Online Information Regarding Weight Gain in Pregnancy: A Descriptive Study

Tammy Chang[1,2], MD, MPH, MS; Bianca A Verma[3], MPH; Trevor Shull[1], MD; Michelle H Moniz[2,4], MD, MSc; Lauren Kohatsu[5]; Melissa A Plegue[1], MA; Kevyn Collins-Thompson[6], PhD

[1]Department of Family Medicine, University of Michigan, Ann Arbor, MI, United States
[2]Institute for Healthcare Policy and Innovation, University of Michigan, Ann Arbor, MI, United States
[3]University of Michigan Medical School, Ann Arbor, MI, United States
[4]Department of Obstetrics & Gynecology, University of Michigan, Ann Arbor, MI, United States
[5]Optimity Advisors, Washington, DC, United States
[6]School of Information, University of Michigan, Ann Arbor, MI, United States

Corresponding Author:
Tammy Chang, MD, MPH, MS
Department of Family Medicine
University of Michigan
2800 Plymouth Rd
Building 14- Room G107
Ann Arbor, MI,
United States
Phone: 1 734 647 3305
Fax: 1 734 763 7770
Email: tachang@med.umich.edu

Abstract

Background: Excess weight gain affects nearly half of all pregnancies in the United States and is a strong risk factor for adverse maternal and fetal outcomes, including long-term obesity. The Internet is a prominent source of information during pregnancy; however, the accuracy of this online information is unknown.

Objective: To identify, characterize, and assess the accuracy of frequently accessed webpages containing information about weight gain during pregnancy.

Methods: A descriptive study was used to identify and search frequently used phrases related to weight gain during pregnancy on the Google search engine. The first 10 webpages of each query were characterized by type and then assessed for accuracy and completeness, as compared to Institute of Medicine guidelines, using crowdsourcing.

Results: A total of 114 queries were searched, yielding 305 unique webpages. Of these webpages, 181 (59.3%) included information regarding weight gain during pregnancy. Out of 181 webpages, 62 (34.3%) contained no specific recommendations, 48 (26.5%) contained accurate but incomplete recommendations, 41 (22.7%) contained complete and accurate recommendations, and 22 (12.2%) were inaccurate. Webpages were most commonly from for-profit websites (112/181, 61.9%), followed by government (19/181, 10.5%), medical organizations or associations (13/181, 7.2%), and news sites (12/181, 6.6%). The largest proportion of for-profit sites contained no specific recommendations (44/112, 39.3%). Among pages that provided inaccurate information (22/181, 12.2%), 68% (15/22) were from for-profit sites.

Conclusions: For-profit websites dominate the online space with regard to weight gain during pregnancy and largely contain incomplete, inaccurate, or no specific recommendations. This represents a significant information gap regarding an important risk factor for obesity among mothers and infants. Our findings suggest that greater clinical and public health efforts to disseminate accurate information regarding healthy weight gain during pregnancy may help prevent significant morbidity and may support healthier pregnancies among at-risk women and children.

(J Med Internet Res 2016;18(4):e81) doi:10.2196/jmir.5138

KEYWORDS

Internet; crowdsourcing; weight gain; pregnancy

Introduction

Appropriate weight gain during pregnancy has important health implications for both mothers and infants [1]. Excess weight gain in pregnancy is associated with adverse obstetric and fetal/neonatal outcomes, including hypertensive disorders of pregnancy, gestational diabetes, excessive fetal growth, prolonged labor, birth injury, and cesarean delivery [2,3]. Furthermore, excess weight gain in pregnancy may be associated with long-term increased risk of obesity in both mother and child [4-7].

In 2009, the Institute of Medicine (IOM) updated its guidelines for weight gain in pregnancy, recommending more stringent weight gain with increasing prepregnancy body mass index (BMI) [1]. However, many women are unaware of these guidelines [8,9]. In the United States, only 32% of women gain the appropriate amount of weight during pregnancy based on the IOM guidelines, while 47% of women gain excessive weight and 21% gain inadequate weight [10]. Although the American Congress of Obstetricians and Gynecologists (ACOG) recommends discussing appropriate weight gain, diet, and exercise at the initial visit and periodically throughout pregnancy, clinicians caring for pregnant women often do not follow this recommendation [11,12]. In one recent study of over 300 women, only 12% reported being counseled correctly by their health care provider regarding how much they should gain during pregnancy [11]. Clinicians may be unaware, unfamiliar, or unaccepting of the guidelines, or reluctant to discuss the sensitive topic of weight gain with their overweight or obese patients [13-15].

The Internet is a common source of health information among pregnant women, with nearly all women (94%) using the Internet for pregnancy-related information [16-19]. Among these women, nearly all (98%) begin their online health inquiries using search engines such as Google [18]. Most women (99%) search online because they want to find out more information on their own, beyond the information that is provided to them by their health care provider [18]. Pregnant women commonly seek information about topics such as antenatal complications, childbirth, pregnancy symptoms, and health promotion/lifestyle issues, including nutrition and physical activity during pregnancy [18-20]. In general, pregnant women consider online health information to be of reasonable quality and reliable, and information found online plays a significant role in women's decision-making during pregnancy [17-19].

Despite widespread online health information seeking during pregnancy, relatively little is known about whether online sources accurately reflect the latest IOM guidelines for weight gain in pregnancy. The objective of this study was to characterize the type of website and accuracy of information of the most frequently accessed websites containing information about weight gain in pregnancy.

Methods

Identification of Commonly Accessed Webpages

The authors identified 13 initial search queries that included keywords and phrases related to weight gain during pregnancy based on clinical experiences with pregnant patients (ie, "Am I gaining enough weight during pregnancy?," "Healthy weight gain during pregnancy," and "Am I gaining too much weight during pregnancy?"). To help ensure the queries used by our study reflected how people actually search on this topic, we expanded the initial set of 13 queries into 114 variants based on query suggestions produced by Google's autocomplete feature (see Multimedia Appendix 1). The autocomplete feature suggests similar queries that have been typed before by Google users, or that occur as content on the Web [21].

Each query was entered into the Google search engine to yield a list of webpages on August 8, 2014. A website is defined as a set of webpages typically served from a single Web domain [22]. A Windows 8 Pro (64-bit) machine using the US English locale and running Google Chrome (v46) from Ann Arbor, Michigan, was used for all search results retrieval and processing. A new Google account was used with personalization not activated, so that search history did not influence the results. Safe search was also not activated.

The first 10 webpages that resulted from each search were included in this study, as these represent the websites most likely to be viewed by actual users; webpages ranked 11 or greater on a Google search received less than 5% of overall traffic [23]. Paid advertisements/webpages at the top of the Google search result were not included. Duplicate websites were removed to create the list of the most commonly accessed websites resulting from Google searching for information on weight gain during pregnancy.

Data Collection

Crowdsourcing as Data Collection Method

Crowdsourcing was used in our study to characterize the type and accuracy of the commonly accessed webpages. Crowdsourcing is the process of "obtaining needed services, ideas, or content by soliciting contributions from a large group of people, and especially from the online community, rather than from traditional employees or suppliers" [24]. This online data collection method is increasingly used in health-related research and is an ideal method for this study for two reasons: (1) crowdsourcing obtains the judgments of individuals who are likely more similar to patients than medical experts and (2) crowdsourcing employs several layers of validation to ensure a high level of accuracy in data collection [25-28]. In our case, Crowdflower.com [29] was used to identify US residents to complete *microtasks*, or small tasks online that involved identifying simple ideas or phrases from webpages. People who complete microtasks online are called contributors and receive a small reimbursement for the tasks they complete. Contributors were instructed to evaluate only the body of the webpage ("corpus") and to disregard any advertisements or pop-ups. Contributors were paid between US $0.05 and $0.15 for each microtask completed, depending on the difficulty of the task.

Crowdflower Tasks

Each webpage was analyzed independently by three Crowdflower contributors. Contributors were drawn from the pool of millions of Crowdflower contributors and webpages were assigned randomly. Thus, the three contributors assigned to each webpage varied. Each of the following three questions was a separate *task* used to evaluate each of the webpages:

1. Task 1: Is the webpage about weight gain during pregnancy?
2. Task 2: What type of website is this (eg, for-profit company, government, news, medical organization, blog, university, nonprofit/foundation, or medical journal)?
3. Task 3: Did this webpage include the Institute of Medicine guidelines information for total weight gain for each prepregnancy weight group [1]?

These three tasks were completed between September 20, 2014, and December 3, 2014. Textbox 1 contains the abstraction questionnaire with the exact questions and answer choices used by contributors to evaluate each webpage.

Textbox 1. Abstraction questionnaire, including questions and answer choices.

Task 1: Does this webpage include information about healthy weight gain DURING pregnancy?

- Yes
- No

Task 2: Which description below best describes the type of webpage this is?

- Government (.gov)
- News (primary goal of site is to report news)
- University or academic center with associated hospitals/clinics (examples: Mayo Clinic, Kaiser Permanente)
- Company (for-profit business, examples: babycenter.com, WebMD)
- Blog
- Medical organization or associations (examples: Institute of Medicine [IOM], the American College of Obstetricians and Gynecologists [ACOG])
- Medical journal (example: American Journal of Obstetrics and Gynecology)
- Nonprofit/foundation (example: March of Dimes)
- Other

Task 3: Which of the following SPECIFIC recommendations for TOTAL weight gain during pregnancy were mentioned, if any? Select all that apply.

- Underweight women (or women with body mass index [BMI] less than 18.5 kg/m^2) are recommended to gain 28-40 pounds
- Normal-weight women (or women with BMI 18.5-24.9 kg/m^2) are recommended to gain 25-35 pounds
- Overweight women (or women with BMI 25-29.9 kg/m^2) are recommended to gain 15-25 pounds
- Obese women (or women with BMI ≥30 kg/m^2) are recommended to gain 11-20 pounds
- Recommendations were given in pounds (lbs), but were different than what is listed above for one or more groups of women
- Recommendations were only given in kilograms (kg)
- No SPECIFIC recommendations for total weight gain during pregnancy were given
- No SPECIFIC recommendations for total weight gain during pregnancy were given in the text, though there is an ONLINE CALCULATOR
- Recommendations were only given for multiple pregnancies (twins)

Quality Control of Crowdflower Responses

The authors used three methods to ensure the accuracy of contributors' work. First, Crowdflower itself provides a quality-control mechanism by maintaining historical trust levels for each contributor, based on their performance from previous jobs. These trust levels are used to maintain a high-quality pool of potential contributors. Second, before they could contribute to the study, potential contributors were required to successfully complete a training session in which they correctly evaluated at least two of three gold standard test webpages. These test webpages were a subset of the total webpages, for which gold standard responses to the three abstraction form questions were previously agreed upon by two investigators (TC, TS, BV, or LK). Third, additional gold standard test webpages—unidentified to contributors—were interspersed randomly throughout the full sample of webpages for each task. Contributors were required to maintain at least 65% accuracy of these test questions to be considered *trusted* contributors. The process of determining accuracy was automated by Crowdflower and reported to contributors in real time. If a contributor's accuracy decreased below this threshold, they were removed from the job and their prior responses were disregarded.

Task Flow for Crowdsourced Data Collection

Despite the use of specific search terms, some webpages were determined *not* to be about weight gain during pregnancy and were excluded from further analyses using the question in Task 1. For this task, if at least two of three contributors determined the webpage to be about weight gain in pregnancy, it was considered a relevant webpage for further analysis.

The remaining webpages that did include information about weight gain during pregnancy were then evaluated using questions in Tasks 2 and 3. For Task 2, if at least two of three contributors agreed on the website type, this was considered the correct answer. If three contributors answered differently in Task 2, the webpage was reviewed by two investigators (TC, TS, BV, or LK) to determine the correct website type.

For Task 3, each contributor could select more than one response and each response selected by at least two of three contributors was deemed correct. Two investigators (TC, TS, BV, or LK) also reviewed any webpages that had no majority consensus in order to determine a final answer. Weight gain recommendations that were reported in kilograms were evaluated for accuracy using the technique described above. Webpages that only included recommendations for multiple gestations (eg, twins) were removed from our sample (n=1).

Researcher Assessment of Crowdflower Results

Each webpage's recommendations for total weight gain were classified as *complete and accurate* (consistent with IOM recommendations for each prepregnancy BMI category), *incomplete but accurate* (consistent with IOM recommendations but exclusive of one or more of the BMI categories), *inaccurate* (inconsistent with IOM recommendations, such as an incorrect range; recommendation listed was outside of the range; or incorrect BMI categories), *no recommendation* (no specific recommendations of weight gain ranges based on prepregnancy BMI), or *no recommendation but calculator* (for webpages that include a calculator to determine a user's recommended weight gain, but do not include specific recommendations in the text of the page). Inaccurate webpages were reviewed by one investigator (TC) to determine whether the webpage was reporting recommendations that were higher or lower than recommended by the IOM, or if the inaccuracy was related to errors or omission in reporting prepregnancy BMI ranges for their recommendation.

Descriptive statistics were used to describe the distribution of the *type of sites*, and the *accuracy* and *completeness* of webpages. This article was developed using publicly available information. No human participants took part in any protocol and, as such, this study was deemed exempt from review by the University of Michigan Institutional Review Board.

Results

After querying the 114 search terms and aggregating the top 10 website results of each query and removing any duplicates, 305 unique webpages remained. Of these 305 unique webpages, 181 (59.3%) included information regarding weight gain during pregnancy. Webpages were most commonly from for-profit websites (112/181, 61.9%), followed by government (19/181, 10.5%), medical organizations or associations (13/181, 7.2%), and news sites (12/181, 6.6%) (see Table 1).

Overall, one-third (62/181, 34.3%) of all websites that contained information regarding weight gain during pregnancy gave no specific recommendations for total weight gain during pregnancy. The largest proportion of webpages overall were from for-profit sites that contained no specific recommendations (44/181, 24.3%). Among pages that provided inaccurate information (22/181, 12.2%), 68% (15/22) were from for-profit sites (see Table 2). Among pages that provided accurate and complete information (41/181, 22.7%), 41% (17/41) were from for-profit sites, 22% (9/41) from government sites, and 17% (7/41) from medical organizations. Please see Multimedia Appendix 2 for a list of the top 10 accurate webpages by frequency of appearance in Google searching.

Table 1. Frequency of each type of website (n=181).

Type of website	Frequency, n (%)
Company (for-profit business)	112 (61.9)
Government	19 (10.5)
Medical organization or association	13 (7.2)
News	12 (6.6)
Personal blog	10 (5.5)
Nonprofit/foundation	8 (4.4)
University or academic medical center	6 (3.3)
Medical journal	1 (0.6)

Table 2. Accuracy and completeness by type of website used; first page (top 10) results only (n=181[a]).

Type of domain/website of the selected webpages	Complete and accurate, n (%)	Incomplete but accurate, n (%)	Inaccurate, n (%)	No recommendation, n (%)	No recommendation, but weight gain calculator, n (%)
Government (n=19)	9 (47)	4 (21)	2 (11)	3 (16)	1 (5)
News (n=12)	4 (33)	4 (33)	0 (0)	4 (33)	0 (0)
Personal blog (n=10)	0 (0)	2 (20)	2 (20)	6 (60)	0 (0)
University or academic medical center (n=6)	2 (33)	2 (33)	0 (0)	2 (33)	0 (0)
Medical organization or association (n=13)	7 (54)	2 (15)	1 (8)	2 (15)	1 (8)
Nonprofit/foundation (n=8)	2 (25)	4 (50)	2 (25)	0 (0)	0 (0)
Company (for-profit business) (n=112)	17 (15.2)	30 (26.8)	15 (13.4)	44 (39.3)	6 (5.4)
Medical journal (n=1)	0 (0)	0 (0)	0 (0)	1 (100)	0 (0)
Overall (n=181)	41 (22.7)	48 (26.5)	22 (12.2)	62 (34.3)	8 (4.4)

[a] One site excluded for providing recommendations for twins.

Of inaccurate webpages, 18% (4/22) recommended more than the recommended amount of weight, 14% (3/22) recommended less, and 64% (14/22) of webpages included errors or omissions in reporting prepregnancy BMI ranges for their recommendations. The 4 pages out of 22 (18%) that recommended more weight gain reported between 2 and 5 pounds more than the recommended amount. For example, askdrsears.com [30] reported, “If you begin pregnancy slightly above your ideal weight, a healthy weight gain is 20 to 25 pounds; if you are obese, less than 20 pounds,” while the correct recommendation for overweight women is 15-25 pounds. For the 3 pages out of 22 (14%) that recommended less weight gain, they ranged from 1 to 8 pounds less. For example, sharecare.com [31] reported, “The recommended weight gain is 25-30 lbs for average weight women. 15-20 lbs if overweight.” Again, the recommendation for overweight women is 15-25 pounds. Among the webpages that included errors or omissions in reporting BMI ranges, errors in the prepregnancy BMI category deviated from the correct definition by 1-2 BMI points (kg/m^2). Out of 22 websites, 4 (18%) combined the overweight and obese groups and combined the recommendations (ie, “Women with pre-pregnancy BMI >25 can gain between 11-25 pounds”).

Discussion

Principal Findings

The majority of webpages assessed in our study do not reflect the latest IOM guidelines and either present inaccurate or incomplete information, or provide no recommendations. Additionally, for-profit websites currently dominate this online sphere, and these were most likely to contain inaccurate, incomplete, or no specific recommendations compared to other website types. Our study suggests that despite established guidelines for healthy weight gain in pregnancy, many women may not access these guidelines when searching the Internet, potentially increasing their risk of adverse maternal or fetal outcomes from too much or too little weight gain.

Among inaccurate webpages, most recommendations either were within a few pounds of the IOM recommendations or incorrectly defined prepregnancy BMI categories by just 1-2 points (kg/m^2). However, some webpages combined recommendations for overweight and obese women. This is concerning because obese women may be at highest risk of obstetric complications from excessive gestational weight gain. More concerning may be that nearly one-third of webpages had no recommendations despite appearing in the top 10 results of a Google query for information regarding weight gain during pregnancy. This creates a significant knowledge gap for pregnant women and can have long-term health consequences for mother and baby.

Patients may receive no or inaccurate information on weight gain in pregnancy not just from the Internet, but also from their clinicians, friends, and family. Studies have shown that patients want specific information on weight gain during pregnancy, and they trust and rely on information given by clinicians [32,33]. Clinicians also have the advantage of providing face-to-face information to their patients. However, clinicians are often incomplete information sources regarding weight gain during pregnancy because they fail to address the topic or provide information that is not aligned with the IOM guidelines [34,35]. Perhaps the same reluctance among clinicians to discuss the sensitive subject of weight may also apply to online contributors who may not want to offend their pregnant readers by discussing specific weight gain guidelines on their websites. This lack of information, from both clinicians and the Internet, may be compounded by misinformation, such as the widespread cultural belief of “eating for two,” that is often perpetuated and reinforced by pregnant women’s friends and family members [36-39].

Our findings underscore the inadequacy of online resources—key sources of health information for pregnant women—in addressing weight gain in pregnancy [10]. These findings also suggest two specific ways to improve awareness of the IOM weight gain guidelines, for both pregnant women and the people who influence them:

1. Clinicians and caregivers need to fill the knowledge gap. At this time, clinicians and those that provide support and care for pregnant women cannot depend on the Internet to transmit accurate information about healthy weight gain during pregnancy. As a trusted and preferred source of health information, clinicians, nurses, and public health workers should anticipate a knowledge gap about this topic and strive to discuss healthy weight gain with every pregnant woman [32,33]. Robust data from systematic reviews and meta-analyses have demonstrated the efficacy of motivational interviewing, behavioral self-monitoring, goal setting, structured moderate physical exercise programs, and dietary counseling [40-44]. Ideally, providers of prenatal care should address healthy diet and exercise habits early in pregnancy, including discussions of individualized goals for total weight gain across the pregnancy. The frequency of prenatal visits allows for revisiting this topic and drawing on clinical support resources, such as nurses, nutritional counselors, and social workers, to augment the clinician's efforts.

In addition, because of the known challenges in patient-clinician communication regarding weight gain during pregnancy, research should also investigate effective communication approaches that work for both clinicians and at-risk patients. These approaches could include how the topic should be framed, what information should be presented, and the optimal timing of this counseling.

2. Accurate webpages must be more accessible. For-profit webpages dominate online search results because they know how to market their websites. Academic and government websites, on the other hand, do not seem to be using the techniques that ensure a webpage appears as a top result in Google searching. These techniques include optimizing the keywords used in the website, so that they are in the file name and headings and are used throughout the page, especially in the beginning of the first sentence; creating sitemaps (ie, a list of pages of a website accessible to crawlers or users); and including related links that can be detected and indexed to create greater visibility for the website and the health message [45]. By employing these techniques, websites with accurate information can become more accessible by appearing higher in Google searching, something shown to improve Web traffic to those sites [23]. These techniques can have an exponential effect as other sites can link to them and spread the accurate and usable knowledge. Furthermore, visible and easily searchable sites can help to educate not only pregnant women, but other influential people, such as family and friends who often search on behalf of their friends and loved ones [46].

Limitations

This study describes the text that is contained in the webpage, not whether a patient understood the text, as we did not evaluate the presentation or readability of the information. Also, the specific characteristics and motivations of contributors on Crowdflower are unknown and may bias their judgments, though we believe this bias is limited by the multiple layers of quality control described in the Methods. We also did not analyze the accuracy of online calculators, though they were a small proportion of webpages (8/181, 4.4%). Finally, only the Google search engine results were evaluated in this study, which may limit the generalizability of our findings.

Conclusions

In conclusion, frequently accessed online information regarding weight gain during pregnancy does not reflect the latest IOM guidelines, with a large proportion of webpages displaying inaccurate or incomplete information, or no recommendations. Accurate information regarding healthy weight gain during pregnancy is vital during the prenatal period to prevent long-term morbidity in mother and baby. Our study adds insight into a potential modifiable factor that may contribute to a large proportion of US women gaining an unhealthy amount of weight during pregnancy and the vital role of clinicians and medical organizations in educating and supporting women during this critical window for maternal and child health.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Original queries and query variants.

[PDF File (Adobe PDF File), 27KB - jmir_v18i4e81_app1.pdf]

Multimedia Appendix 2

Top 10 accurate webpages by frequency of appearance during Google searching.

[JPG File, 70KB - jmir_v18i4e81_app2.jpg]

References

1. Committee to Reexamine IOM Pregnancy Weight Guidelines. In: Rasmussen KM, Yaktine AL, editors. Weight Gain During Pregnancy: Reexamining the Guidelines. Washington, DC: National Academies Press; 2009.

2. Siega-Riz AM, Viswanathan M, Moos M, Deierlein A, Mumford S, Knaack J, et al. A systematic review of outcomes of maternal weight gain according to the Institute of Medicine recommendations: Birthweight, fetal growth, and postpartum weight retention. Am J Obstet Gynecol 2009 Oct;201(4):339.e1-339.e14. [doi: 10.1016/j.ajog.2009.07.002] [Medline: 19788965]
3. Crane JM, White J, Murphy P, Burrage L, Hutchens D. The effect of gestational weight gain by body mass index on maternal and neonatal outcomes. J Obstet Gynaecol Can 2009 Jan;31(1):28-35. [Medline: 19208280]
4. Rooney BL, Schauberger CW. Excess pregnancy weight gain and long-term obesity: One decade later. Obstet Gynecol 2002 Aug;100(2):245-252. [Medline: 12151145]
5. Schmitt NM, Nicholson WK, Schmitt J. The association of pregnancy and the development of obesity - Results of a systematic review and meta-analysis on the natural history of postpartum weight retention. Int J Obes (Lond) 2007 Nov;31(11):1642-1651. [doi: 10.1038/sj.ijo.0803655] [Medline: 17607325]
6. Olson CM, Strawderman MS, Dennison BA. Maternal weight gain during pregnancy and child weight at age 3 years. Matern Child Health J 2009 Nov;13(6):839-846. [doi: 10.1007/s10995-008-0413-6] [Medline: 18818995]
7. Oken E, Taveras EM, Kleinman KP, Rich-Edwards JW, Gillman MW. Gestational weight gain and child adiposity at age 3 years. Am J Obstet Gynecol 2007 Apr;196(4):322.e1-322.e8 [FREE Full text] [doi: 10.1016/j.ajog.2006.11.027] [Medline: 17403405]
8. Downs DS, Savage JS, Rauff EL. Falling short of guidelines? Nutrition and weight gain knowledge in pregnancy. J Womens Health Care 2014;3:1-14 [FREE Full text] [doi: 10.4172/2167-0420.1000184] [Medline: 25599012]
9. Shub A, Huning EY, Campbell KJ, McCarthy EA. Pregnant women's knowledge of weight, weight gain, complications of obesity and weight management strategies in pregnancy. BMC Res Notes 2013;6:278 [FREE Full text] [doi: 10.1186/1756-0500-6-278] [Medline: 23866845]
10. Deputy NP, Sharma AJ, Kim SY, Hinkle SN. Prevalence and characteristics associated with gestational weight gain adequacy. Obstet Gynecol 2015 Apr;125(4):773-781. [doi: 10.1097/AOG.0000000000000739] [Medline: 25751216]
11. McDonald SD, Pullenayegum E, Taylor VH, Lutsiv O, Bracken K, Good C, et al. Despite 2009 guidelines, few women report being counseled correctly about weight gain during pregnancy. Am J Obstet Gynecol 2011 Oct;205(4):333.e1-333.e6. [doi: 10.1016/j.ajog.2011.05.039] [Medline: 21784404]
12. American College of Obstetricians and Gynecologists. ACOG Committee opinion no. 548: Weight gain during pregnancy. Obstet Gynecol 2013 Jan;121(1):210-212. [doi: 10.1097/01.AOG.0000425668.87506.4c] [Medline: 23262962]
13. Stengel MR, Kraschnewski JL, Hwang SW, Kjerulff KH, Chuang CH. "What my doctor didn't tell me": Examining health care provider advice to overweight and obese pregnant women on gestational weight gain and physical activity. Womens Health Issues 2012;22(6):e535-e540 [FREE Full text] [doi: 10.1016/j.whi.2012.09.004] [Medline: 23122213]
14. Chang T, Llanes M, Gold KJ, Fetters MD. Perspectives about and approaches to weight gain in pregnancy: A qualitative study of physicians and nurse midwives. BMC Pregnancy Childbirth 2013;13:47 [FREE Full text] [doi: 10.1186/1471-2393-13-47] [Medline: 23433216]
15. Cabana MD, Rand CS, Powe NR, Wu AW, Wilson MH, Abboud PA, et al. Why don't physicians follow clinical practice guidelines? A framework for improvement. JAMA 1999 Oct 20;282(15):1458-1465. [Medline: 10535437]
16. Declercq E, Sakala C, Corry M, Applebaum S. Listening to Mothers II: Report of the Second National U.S. Survey of Women's Childbearing Experiences. New York, NY: Childbirth Connection; 2006 Oct. URL: http://www.childbirthconnection.org/pdfs/LTMII_report.pdf [accessed 2015-09-14] [WebCite Cache ID 6bXnOKUpr]
17. Larsson M. A descriptive study of the use of the Internet by women seeking pregnancy-related information. Midwifery 2009 Feb;25(1):14-20. [doi: 10.1016/j.midw.2007.01.010] [Medline: 17408822]
18. Huberty J, Dinkel D, Beets MW, Coleman J. Describing the use of the Internet for health, physical activity, and nutrition information in pregnant women. Matern Child Health J 2013 Oct;17(8):1363-1372. [doi: 10.1007/s10995-012-1160-2] [Medline: 23090284]
19. Lagan BM, Sinclair M, Kernohan WG. Internet use in pregnancy informs women's decision making: A Web-based survey. Birth 2010 Jun;37(2):106-115. [doi: 10.1111/j.1523-536X.2010.00390.x] [Medline: 20557533]
20. Szwajcer EM, Hiddink GJ, Maas L, Koelen MA, van Woerkum CM. Nutrition-related information-seeking behaviours of women trying to conceive and pregnant women: Evidence for the life course perspective. Fam Pract 2008 Dec;25 Suppl 1:i99-i104 [FREE Full text] [doi: 10.1093/fampra/cmn077] [Medline: 18974061]
21. Google Help. Autocomplete URL: https://support.google.com/websearch/answer/106230?hl=en [accessed 2015-09-14] [WebCite Cache ID 6bXnW7tAs]
22. Merriam-Webster. Definition of website URL: http://www.merriam-webster.com/dictionary/website [accessed 2015-09-14] [WebCite Cache ID 6bXpHOeCJ]
23. Chitika. 2013 Jun 07. The value of Google result positioning URL: https://chitika.com/google-positioning-value [accessed 2015-09-14] [WebCite Cache ID 6bXoeqIoe]
24. Merriam-Webster. Definition of crowdsourcing URL: http://www.merriam-webster.com/dictionary/crowdsourcing [accessed 2015-09-14] [WebCite Cache ID 6bXntHqST]
25. Cavazos-Rehg PA, Krauss M, Fisher SL, Salyer P, Grucza RA, Bierut LJ. Twitter chatter about marijuana. J Adolesc Health 2015 Feb;56(2):139-145 [FREE Full text] [doi: 10.1016/j.jadohealth.2014.10.270] [Medline: 25620299]

26. Gottlieb A, Hoehndorf R, Dumontier M, Altman RB. Ranking adverse drug reactions with crowdsourcing. J Med Internet Res 2015;17(3):e80 [FREE Full text] [doi: 10.2196/jmir.3962] [Medline: 25800813]
27. Harris JK, Mart A, Moreland-Russell S, Caburnay CA. Diabetes topics associated with engagement on Twitter. Prev Chronic Dis 2015;12:E62 [FREE Full text] [doi: 10.5888/pcd12.140402] [Medline: 25950569]
28. Kristan J, Suffoletto B. Using online crowdsourcing to understand young adult attitudes toward expert-authored messages aimed at reducing hazardous alcohol consumption and to collect peer-authored messages. Transl Behav Med 2015 Mar;5(1):45-52 [FREE Full text] [doi: 10.1007/s13142-014-0298-4] [Medline: 25729452]
29. Crowdflower. URL: http://www.crowdflower.com/ [accessed 2016-03-17] [WebCite Cache ID 6g4tbeD6G]
30. Ask Dr Sears. URL: http://www.askdrsears.com/ [accessed 2016-03-17] [WebCite Cache ID 6g5KJIRfd]
31. Sharecare. URL: https://www.sharecare.com/ [accessed 2016-03-17] [WebCite Cache ID 6g5KBwabo]
32. Hesse BW, Nelson DE, Kreps GL, Croyle RT, Arora NK, Rimer BK, et al. Trust and sources of health information: The impact of the Internet and its implications for health care providers: Findings from the first Health Information National Trends Survey. Arch Intern Med 2005;165(22):2618-2624. [doi: 10.1001/archinte.165.22.2618] [Medline: 16344419]
33. Ferrari RM, Siega-Riz AM, Evenson KR, Moos M, Carrier KS. A qualitative study of women's perceptions of provider advice about diet and physical activity during pregnancy. Patient Educ Couns 2013 Jun;91(3):372-377 [FREE Full text] [doi: 10.1016/j.pec.2013.01.011] [Medline: 23399436]
34. Stotland NE, Gilbert P, Bogetz A, Harper CC, Abrams B, Gerbert B. Preventing excessive weight gain in pregnancy: How do prenatal care providers approach counseling? J Womens Health (Larchmt) 2010 Apr;19(4):807-814 [FREE Full text] [doi: 10.1089/jwh.2009.1462] [Medline: 20078239]
35. Phelan S, Phipps MG, Abrams B, Darroch F, Schaffner A, Wing RR. Practitioner advice and gestational weight gain. J Womens Health (Larchmt) 2011 Apr;20(4):585-591 [FREE Full text] [doi: 10.1089/jwh.2010.2316] [Medline: 21413898]
36. Chuang CH, Stengel MR, Hwang SW, Velott D, Kjerulff KH, Kraschnewski JL. Behaviours of overweight and obese women during pregnancy who achieve and exceed recommended gestational weight gain. Obes Res Clin Pract 2014;8(6):e577-e583 [FREE Full text] [doi: 10.1016/j.orcp.2013.12.254] [Medline: 25434913]
37. Hammond C. BBC Future. 2012 Jun 26. Should pregnant women eat for two? URL: http://www.bbc.com/future/story/20120626-can-pregnant-women-eat-for-two [accessed 2015-09-14] [WebCite Cache ID 6bXo0fOcv]
38. Deierlein AL, Siega-Riz AM, Adair LS, Herring AH. Effects of pre-pregnancy body mass index and gestational weight gain on infant anthropometric outcomes. J Pediatr 2011 Feb;158(2):221-226 [FREE Full text] [doi: 10.1016/j.jpeds.2010.08.008] [Medline: 20863516]
39. Stuebe AM, Lyon H, Herring AH, Ghosh J, Wise A, North KE, et al. Obesity and diabetes genetic variants associated with gestational weight gain. Am J Obstet Gynecol 2010 Sep;203(3):283.e1-283.e17 [FREE Full text] [doi: 10.1016/j.ajog.2010.06.069] [Medline: 20816152]
40. Thangaratinam S, RogoziÃƒâ€¦Ã¢â‚¬Å¾ska E, Jolly K, Glinkowski S, Duda W, Borowiack E, et al. Interventions to reduce or prevent obesity in pregnant women: A systematic review. Health Technol Assess 2012 Jul;16(31):iii-iv, 1 [FREE Full text] [doi: 10.3310/hta16310] [Medline: 22814301]
41. Sanabria-Martínez G, García-Hermoso A, Poyatos-León R, Álvarez-Bueno C, Sánchez-López M, Martínez-Vizcaíno V. Effectiveness of physical activity interventions on preventing gestational diabetes mellitus and excessive maternal weight gain: A meta-analysis. BJOG 2015 Aug;122(9):1167-1174. [doi: 10.1111/1471-0528.13429] [Medline: 26036300]
42. Brown MJ, Sinclair M, Liddle D, Hill AJ, Madden E, Stockdale J. A systematic review investigating healthy lifestyle interventions incorporating goal setting strategies for preventing excess gestational weight gain. PLoS One 2012;7(7):e39503 [FREE Full text] [doi: 10.1371/journal.pone.0039503] [Medline: 22792178]
43. Choi J, Fukuoka Y, Lee JH. The effects of physical activity and physical activity plus diet interventions on body weight in overweight or obese women who are pregnant or in postpartum: A systematic review and meta-analysis of randomized controlled trials. Prev Med 2013 Jun;56(6):351-364 [FREE Full text] [doi: 10.1016/j.ypmed.2013.02.021] [Medline: 23480971]
44. Hill B, Skouteris H, Fuller-Tyszkiewicz M. Interventions designed to limit gestational weight gain: A systematic review of theory and meta-analysis of intervention components. Obes Rev 2013 Jun;14(6):435-450. [doi: 10.1111/obr.12022] [Medline: 23534901]
45. Search Engine Optimization Starter Guide. Mountain View, CA: Google; 2010. URL: http://static.googleusercontent.com/media/www.google.com/en/us/webmasters/docs/search-engine-optimization-starter-guide.pdf [WebCite Cache ID 6bXoGwwoh]
46. Pew Research Center. Washington, DC: Pew Internet & American Life Project Health fact sheet URL: http://www.pewinternet.org/fact-sheets/health-fact-sheet/ [WebCite Cache ID 6bXoLHEEq]

Abbreviations

ACOG: American Congress of Obstetricians and Gynecologists
BMI: body mass index
IOM: Institute of Medicine

Edited by G Eysenbach; submitted 14.09.15; peer-reviewed by M Bardus; comments to author 12.11.15; revised version received 02.12.15; accepted 23.01.16; published 07.04.16.

Please cite as:
Chang T, Verma BA, Shull T, Moniz MH, Kohatsu L, Plegue MA, Collins-Thompson K
Crowdsourcing and the Accuracy of Online Information Regarding Weight Gain in Pregnancy: A Descriptive Study
J Med Internet Res 2016;18(4):e81
URL: http://www.jmir.org/2016/4/e81/
doi:10.2196/jmir.5138
PMID:27056465

Original Paper

Patient Perspectives on Sharing Anonymized Personal Health Data Using a Digital System for Dynamic Consent and Research Feedback: A Qualitative Study

Karen Spencer[1*], PhD; Caroline Sanders[2*], PhD; Edgar A Whitley[3], PhD; David Lund[4], PhD; Jane Kaye[5], PhD; William Gregory Dixon[1,6,7], MRCP(UK), PhD

[1]Arthritis Research UK Centre for Epidemiology, Manchester Academic Health Sciences Centre, The University of Manchester, Manchester, United Kingdom

[2]National Institute for Health Research (NIHR) School for Primary Care Research, Manchester Academic Health Sciences Centre, The University of Manchester, Manchester, United Kingdom

[3]Department of Management, London School of Economics and Political Science, London, United Kingdom

[4]HW Communications Ltd, Lancaster, United Kingdom

[5]Centre for Health, Law and Emerging Technologies, Nuffield Department of Population Health, University of Oxford, Oxford, United Kingdom

[6]Health eResearch Centre, Farr Institute, Manchester Academic Health Sciences Centre, The University of Manchester, Manchester, United Kingdom

[7]Department of Rheumatology, Salford Royal National Health Service (NHS) Foundation Trust, Salford, United Kingdom

[*]these authors contributed equally

Corresponding Author:
William Gregory Dixon, MRCP(UK), PhD
Arthritis Research UK Centre for Epidemiology
Manchester Academic Health Sciences Centre
The University of Manchester
Room 2.900 Stopford Building
Oxford Road
Manchester,
United Kingdom
Phone: 44 0161 275 1642
Fax: 44 0161 275 5043
Email: will.dixon@manchester.ac.uk

Abstract

Background: Electronic health records are widely acknowledged to provide an important opportunity to anonymize patient-level health care data and collate across populations to support research. Nonetheless, in the wake of public and policy concerns about security and inappropriate use of data, conventional approaches toward data governance may no longer be sufficient to respect and protect individual privacy. One proposed solution to improve transparency and public trust is known as Dynamic Consent, which uses information technology to facilitate a more explicit and accessible opportunity to opt out. In this case, patients can tailor preferences about whom they share their data with and can change their preferences reliably at any time. Furthermore, electronic systems provide opportunities for informing patients about data recipients and the results of research to which their data have contributed.

Objective: To explore patient perspectives on the use of anonymized health care data for research purposes. To evaluate patient perceptions of a Dynamic Consent model and electronic system to enable and implement ongoing communication and collaboration between patients and researchers.

Methods: A total of 26 qualitative interviews and three focus groups were conducted that included a video presentation explaining the reuse of anonymized electronic patient records for research. Slides and tablet devices were used to introduce the Dynamic Consent system for discussion. A total of 35 patients with chronic rheumatic disease with varying levels of illness and social deprivation were recruited from a rheumatology outpatient clinic; 5 participants were recruited from a patient and public involvement health research network.

Results: Patients were supportive of sharing their anonymized electronic patient record for research, but noted a lack of transparency and awareness around the use of data, making it difficult to secure public trust. While there were general concerns about detrimental consequences of data falling into the wrong hands, such as insurance companies, 39 out of 40 (98%) participants generally considered that the altruistic benefits of sharing health care data outweighed the risks. Views were mostly positive about the use of an electronic interface to enable greater control over consent choices, although some patients were happy to share their data without further engagement. Participants were particularly enthusiastic about the system as a means of enabling feedback regarding data recipients and associated research results, noting that this would improve trust and public engagement in research. This underlines the importance of patient and public involvement and engagement throughout the research process, including the reuse of anonymized health care data for research. More than half of patients found the touch screen interface easy to use, although a significant minority, especially those with limited access to technology, expressed some trepidation and felt they may need support to use the system.

Conclusions: Patients from a range of socioeconomic backgrounds viewed a digital system for Dynamic Consent positively, in particular, feedback about data recipients and research results. Implementation of a digital Dynamic Consent system would require careful interface design and would need to be located within a robust data infrastructure; it has the potential to improve trust and engagement in electronic medical record research.

(J Med Internet Res 2016;18(4):e66) doi:10.2196/jmir.5011

KEYWORDS

eHealth; data sharing; public trust; consent

Introduction

The National Health Service (NHS) provides health care for over 60 million citizens throughout their lives, with vast amounts of information about patients' treatments and outcomes collected in their medical records. Such real-world data is an important asset for UK health research: patients' "cradle to grave" records are increasingly captured within electronic patient record (EPR) systems, providing the opportunity to anonymize patient-level health care data and collate across populations to support research. The importance and vast opportunity of sharing health care data for research is explicit within the UK government's Strategy for UK Life Sciences [1]. This has been supported by the recent cross-funder investment to establish the Farr Institute, a network of academic eHealth Centres of Excellence [2].

There is a reasonable expectation in society that the sensitive and personal nature of health care data requires that it should be carefully managed and access to it should be restricted only to those with a legitimate purpose. As a consequence, the UK legal regime has conditions for the use of health care data but at the same time allows certain exemptions for research carried out in the public interest. Under the Data Protection Act (1998), patient consent is not required when anonymized data are used for research, although there may be societal concerns that "go beyond compliance with the requirements of formal regulation" [3]. Health care data are highly personal and are usually of a sensitive nature, making it difficult to anonymize data effectively to maintain the privacy of patients. The Data Protection Act also has a fair processing obligation that requires patients to be informed about what happens to their data that applies to all kinds of data [4]. This aligns with a shared societal expectation that patients have a right to know how their data are being used and should be given the opportunity to consent but also object to their data being shared with others, even in the case of "anonymized" data.

The UK government's care.data initiative, a program intended to enable sharing of anonymized primary care health records with "researchers and organizations outside the NHS" [5] for research and service improvement, was paused in 2014 due to a loss of public trust [3]. Trust is often taken as the measure of an individual's willingness to be vulnerable to the actions of another person on the basis that the trustee will act according to the trustor's confident expectations [6]. Different forms of trust include deterrence-based trust, where the trustor is confident that the trustee will act as expected because sanctions for breach of trust are very high; calculus-based trust, where the trustor "evaluates" the reputation/certification of the trustee; relational trust that arises when repeated interactions have gone well; and institution-based trust, which combines calculus and relational trust through the proxy of the trusted institution [7]. Although health care institutions can normally assume a high level of institutional trust [8], if lost, it can be difficult to repair [9]. Reasons cited for this loss of public trust included concerns that personal health care data might be used inappropriately (eg, sharing with insurance companies or being sold for profit [10], as well as lack of clarity as to how patients should opt out). The population-level approach of the above campaign failed to reassure many patients about potential misuse of data and, although recent studies have shown that most patients support confidential reuse of health data, concerns have also arisen surrounding security, privacy, and control over access of EPRs [11-13]. Previous research has highlighted that the UK public has little knowledge of how their EPRs are used for medical research purposes [13] and the lack of transparency and engagement with patients is viewed to undermine public trust with implications for acceptable models of consent [14-16]. The Nuffield Council on Bioethics, in their review of the care.data plans, recommended that health authorities track the use of patient data, give people greater access, and say how their data is used [17]. This is important for maintaining trust and requires the opportunity and process for opting out of data sharing to be clear. Furthermore, an independent review of the care.data program by the National Data Guardian has asked of any future system, "How can patients check, update, or change their preferences and see that their choices have been respected?" [18].

One proposed solution for the problems outlined above is known as Dynamic Consent, which uses information technology (IT) to facilitate a more explicit and accessible opportunity to opt out, whereby patients can tailor preferences about whom they share their data with, and can change their preferences reliably at any time preventing any further data sharing [19,20]. This is achieved technically by binding patient information with consent expressions [19]. In addition, via the same digital interface, patients can be provided with information as to the recipients of their data plus other information, such as results of research derived from their data contribution. Demonstrating to patients how sharing their data has contributed to improved care within the population could build community trust, and show how patients are already contributing to research within the NHS: a pledge within the NHS Constitution [21]. A prototype Dynamic Consent interface has been developed by the Ensuring Consent and Revocation (EnCoRe) project [20]. This was initially designed in the context of biobanking to allow patients to consent for the collection of biobank tissue and data, but the same principles, architecture, and philosophy could be used to facilitate the trusted sharing of EPR data for research. Implementing such a system for this purpose, however, faces some unknowns. While previous surveys have suggested that patients are willing to share their EPR for research [11], would they wish to express consent preferences using a digital system? Would patients value feedback information about who the recipients are and the results of the research via such a system? There is also a need to address the feasibility and barriers for using such a system, and how it could best be implemented.

The purpose of this study was to undertake qualitative research to (1) explore patient perspectives on the use of anonymized health data for research purposes and (2) to evaluate patient perceptions of a Dynamic Consent model and electronic system to enable and implement ongoing communication and collaboration between patients and researchers in this context.

Methods

Participants and Methods

A combination of qualitative in-depth interviews and focus groups were used to first explore patients' perspectives on the use of anonymized personal health data for research, before introducing the model of Dynamic Consent and feedback and seeking patients' views. Interviews and focus groups were also conducted to seek patient views on an electronic prototype system to collect consent and provide feedback. Focus groups are considered a valuable approach for exploring a range of public and patient views in health research, especially where the goal is to explore and develop a new intervention or service [22]. Combining interviews and focus groups enabled us to maximize recruitment because people could choose whether they wanted to take part in an interview or focus group. Interviews were effective in allowing in-depth discussions related to personal views and experiences. The interaction within focus groups generated some level of debate and consensus, as well as creative ideas about data and information sharing and the potential use of an electronic system.

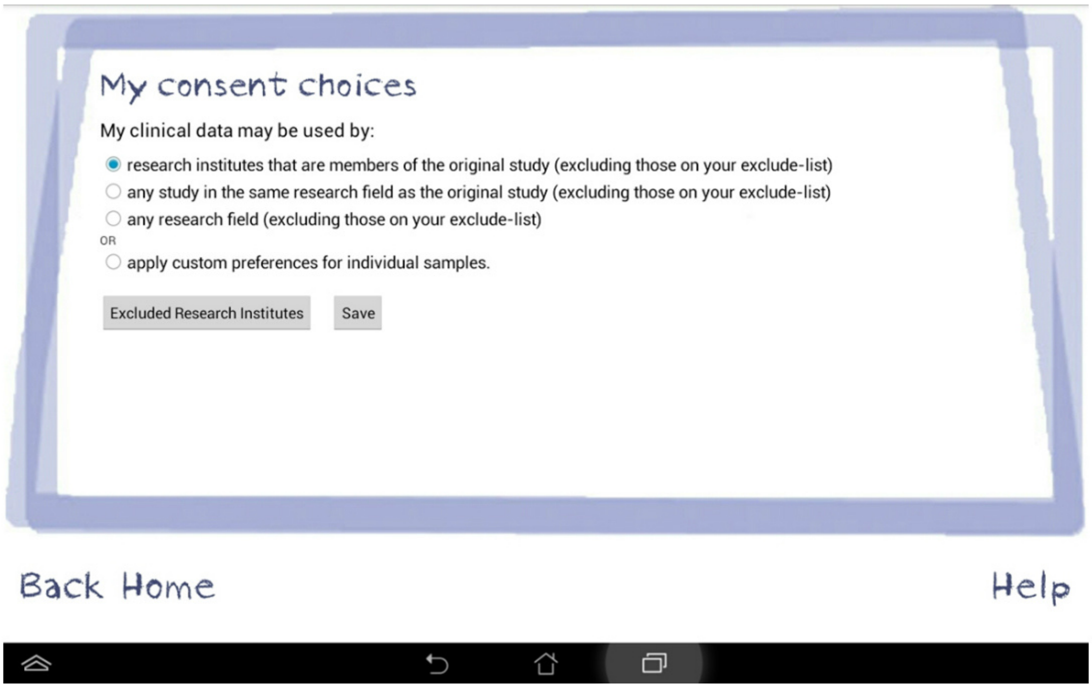
Participants were recruited from a rheumatology outpatient clinic in a large teaching hospital (n=35) and from a patient and public involvement (PPI) health research network (n=5), both based in Salford, Greater Manchester in the United Kingdom. Within the clinic, unselected patients were identified by members of the clinical team and directed to the research associate (KS) for further information. All participants were provided with a patient information sheet describing the study. The sample was to some extent a convenience sample based on who responded to advertisements via the PPI network. However, we were able to purposively sample via the outpatient clinic to ensure maximum variation, including a mixture of men and women of various ages with varied levels of illness and health care experiences [23]. The final sample also had varied occupational, educational, and social circumstances, which were referred to in interviews and focus group discussions.

Three focus groups were conducted, consisting of 4-6 participants along with a moderator and note taker, and lasted approximately 90 minutes. Participants were organized into focus groups pragmatically according to when they consented and were available. A total of 26 semistructured interviews were conducted with patients, each lasting between 45 and 60 minutes. The focus groups and interviews were audiotaped with permission from participants; written informed consent was obtained prior to the start of any discussion. The study received ethical approval from Liverpool East Research Ethics Committee (Ref: 13/NW/0722). A patient and public involvement group comprised of five members was established at the start of the project. This group convened quarterly to inform aspects of the study, such as the design of information and interview guides, and to discuss and refine emerging findings from the focus groups and interviews.

Procedure for Interviews/Focus Groups

An interview/focus group topic guide was developed initially from the literature and subsequent topics were added if they arose during data collection. Topics discussed during the interviews and focus groups included the following: previous knowledge and understanding of how health data are stored and shared beyond the NHS, views and concerns regarding the storage and sharing of EPRs for research purposes, willingness of participants to share their health data and with whom, views about a Dynamic Consent model for reuse of anonymized health data, and views about a prototype electronic system for Dynamic Consent using a tablet device. As we were unsure as to the level of knowledge that participants held relating to their electronic patient records and how they might be anonymized and collated to benefit research, we developed a short 5-minute film that informed participants of current practice within the United Kingdom [24]. Included were visual examples of a clinical consultation involving entry of patient-level data into an EPR system followed by large anonymized datasets being used by university researchers. This was presented on the tablet device during individual interviews and via a projector during focus group discussions. Following initial discussion focused on understanding and views about storage and use of health data, the moderator introduced the Dynamic Consent prototype on a tablet interface with touch screen technology. The interface screens included the ability for patients to state their willingness (or not) to share their anonymized records with specific groups, for example research institutes or private companies—entitled *My consent choices*. Additional screens provided details of which groups had accessed their shared data, research studies using participants' data, as well as links to published research and relevant news items (see Figures 1-3 for screenshots of the Dynamic Consent interface). The prototype interface was intended to provide sufficient detail to elicit patients' views about the concept of Dynamic Consent. Details within the interface, such as how best to categorize research groups or optimal methods for patient feedback, were not tested but will form the basis of future research.

Figure 1. Screenshot of the Dynamic Consent prototype interface: My consent choices.



My consent choices

My clinical data may be used by:

● research institutes that are members of the original study (excluding those on your exclude-list)

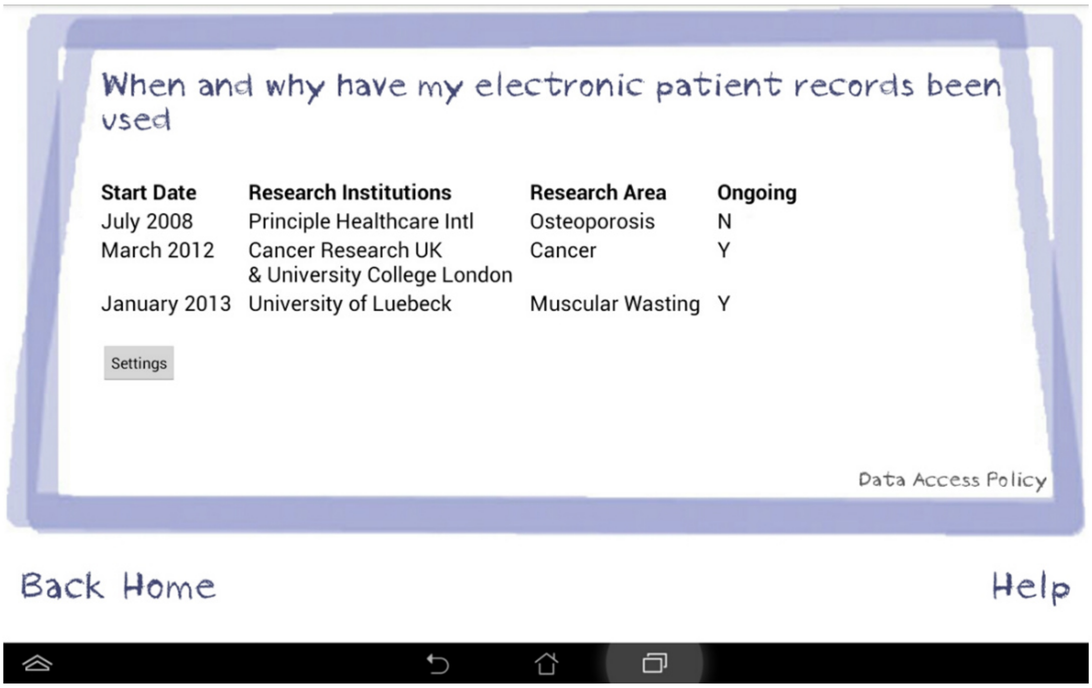
○ any study in the same research field as the original study (excluding those on your exclude-list)

○ any research field (excluding those on your exclude-list)

OR

○ apply custom preferences for individual samples.

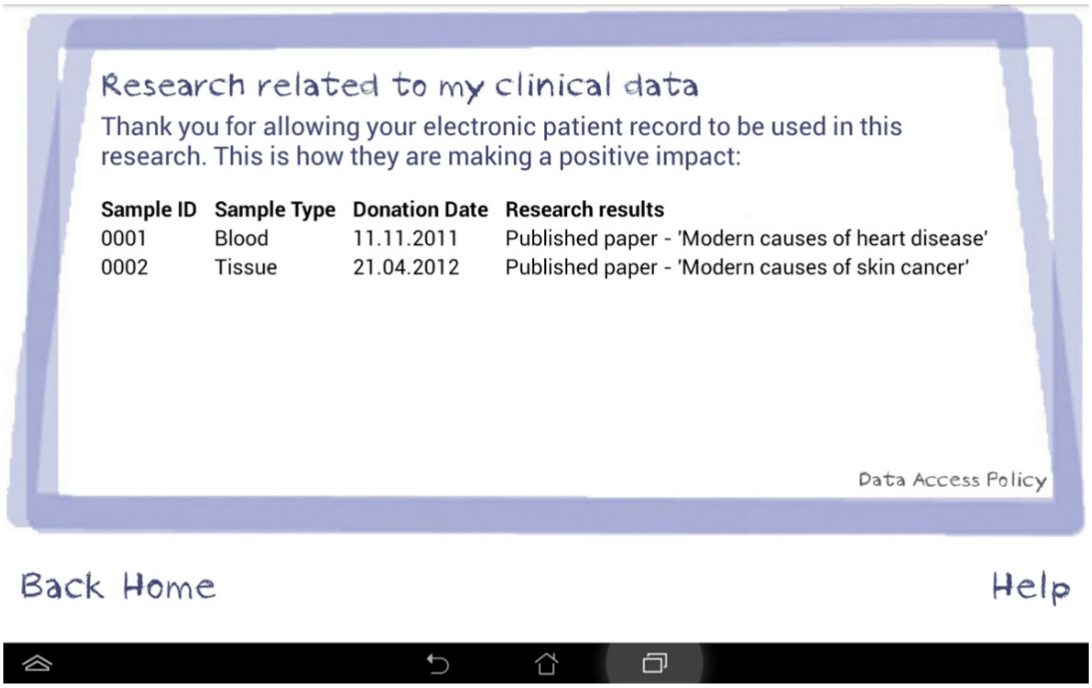
Figure 2. Screenshot of the Dynamic Consent prototype interface: When and why have my electronic patient records been used.



When and why have my electronic patient records been used

Start Date	Research Institutions	Research Area	Ongoing
July 2008	Principle Healthcare Intl	Osteoporosis	N
March 2012	Cancer Research UK & University College London	Cancer	Y
January 2013	University of Luebeck	Muscular Wasting	Y

Figure 3. Screenshot of the Dynamic Consent prototype interface: Research related to my clinical data.



Research related to my clinical data

Thank you for allowing your electronic patient record to be used in this research. This is how they are making a positive impact:

Sample ID	Sample Type	Donation Date	Research results
0001	Blood	11.11.2011	Published paper - 'Modern causes of heart disease'
0002	Tissue	21.04.2012	Published paper - 'Modern causes of skin cancer'

Data Analysis

Interviews and focus groups were transcribed verbatim and NVivo version 10 software (QSR International) was used to facilitate analysis. Data were analyzed thematically using some key techniques of a grounded theory approach, including open coding and constant comparison to identify key (ie, selective) codes [25]. An iterative and inductive approach to analysis was followed so that analysis started in parallel with the data collection; initial results informed subsequent data collection as themes and issues were identified and informed further questions and probing around these emerging themes. For example, a key focus for initial discussions had been the model of Dynamic Consent and the associated interface to specify preferences. However, initial findings demonstrated that while people valued the potential for an opportunity to opt out of specific research, they were more enthusiastic to discuss the research feedback components of the system. Further discussions and questions about this enabled understanding of why this component was considered a priority for a diverse group of participants. Memos and documents were written about emerging categories, to summarize a point, to critique information, and to relate emergent theories to existing literature [25]. Authors KS, CS, and WGD met on a regular basis to discuss the development of codes, themes, categories, and theories about the phenomenon being studied. Recruitment ceased once data saturation was established.

Results

Overview

Of the 40 participants, 23 (58%) were women and 17 (43%) were men. Ages ranged from 23 to 88 years (mean 61, SD 13). With the exception of one white Canadian participant, participants described themselves as white British. All were suffering from a chronic rheumatic disease. Three key themes characterized participants’ views on the use of anonymized EPRs for medical research and their perceptions of Dynamic Consent: (1) the role of trust and perceived social responsibility to share health data, (2) transparency through Dynamic Consent and patient feedback and the potential for enhanced control and patient engagement, and (3) operational and technological scope and challenges for participation.

The Role of Trust and Perceived Social Responsibility to Share Health Data

Individuals often indicated a high degree of trust in the NHS, for example stating, "I trust the NHS to store my information confidentially" (Participant #4) or "I'm generally quite trustful of hospitals and [general practitioners] GPs" (Participant #11). These and additional comments (see Textbox 1) indicate a sense of institutional trust in the health care system, as well as medical professionals. Most participants felt confident that electronic health records were managed securely and anonymity was preserved when used for research. Respondents tended to express a greater concern about security of financial data compared to health data. There was acknowledgement that there may have been some decline in public trust of the NHS and the medical profession in the wider population, but a number of people viewed the media to be responsible for overinflating a sense of public concern due to the "negative press coverage" (Participant #2) allowing "distractions from actual issues" (Participant #5). While the majority expressed satisfaction toward governance arrangements within the NHS, there was an expressed view that no system could be completely secure. A small minority of participants described concern about risk to their privacy, speculating that patients with more sensitive health conditions may be "more guarded of what happens with their health information" (Participant #11) due to fear of stigmatization. There was a general fear of detrimental consequences if data were to fall into the wrong hands, such as insurance companies, suggesting a more nuanced calculative sense of trust that went beyond trust in health care institutions. However, 39 out of 40 (98%) participants considered that the benefits of storing and sharing EPRs for medical research outweighed any perceived risk in terms of data security. Most participants appreciated the importance of medical research and the importance of sharing their EPR for the benefit of medical progress and the health of future generations (see Textbox 1). Only one participant held a contrary view, preferring not to share his data:

> *I would hate for my health details to be in there [national database]...It would be a good idea, but that's in a really nice ideal world, and it's not an ideal world. I would opt out. It's not that I don't trust the NHS, it's that I don't trust, you know, people…people make mistakes.* [Participant #15]

The discussions did not raise issues around different levels of anonymization or the potential for reidentification of patient identity through unique patterns of clinical history.

Textbox 1. Quotations representing the role of trust and perceived social responsibility to share health data.

> *If researchers or health care professionals, or anybody were to look at my own personal records...I trust that they [NHS] have those skills to keep it anonymized.* [Participant #2]
>
> *I think most of my electronic record is pretty safe in the NHS, I trust them to protect my identity and look after the information.* [Participant #18]
>
> *I would never do personal banking because I'm of the age group that doesn't trust things. But things like National Health, I would say yes.* [Participant #14]
>
> *There's always going to be the pros and cons with [storing and sharing health data]. However, for me personally, the pros outweigh the cons.* [Participant #5]
>
> *I don't care what people know about my health...I suppose for insurance for stuff and things like that, could bother some people.* [Participant #12]
>
> *I mean I can trust the doctors and all...but other people, no. Once it leaves the NHS, I'd be wondering where it's going and who's looking at it.* [Participant #19]
>
> *Once you have been in receipt of the excellent kind of care and treatment that I've had, I think you have a social responsibility that if you can help the next generation by having your information provided to the researchers to [do] some good.* [Focus group #3]
>
> *I am happy to share my health records. As long as it benefits other people.* [Participant #28]
>
> *I understand that advances can't be done in medical science, unless people like me and others are taking part in research.* [Participant #1]

Transparency Through Dynamic Consent and Patient Feedback, and the Potential for Enhanced Control and Patient Engagement

Despite the high level of institutional trust and sense of social responsibility, participants reported low levels of awareness about how their personal health data was currently stored and shared for medical research. Importantly, respondents highlighted that fear can come from the unknown causing people to be "very fearful, because you don't know what's going on, you don't know if it's identifiable" (Focus group #1). Some respondents referred to the need for greater information so that "you're dealing with the information, rather than all these things that might be not true" (Participant #29), again highlighting the limits in high-level and more abstract institutional trust. The desire for greater transparency and engagement regarding the use of their data was reflected in repeatedly positive responses regarding the potential use of the digital interface as a useful tool to enable insight into how data is used for specific research studies. Respondents were mostly positive about the potential use of the interface to enable greater control over consent for specific studies because "it gives you choices" (Focus group #1), although some were happy to share their data without wanting to engage further (see below). Despite a clear introduction, patients did not talk about the time-varying nature of consent preferences, instead talking about the value of using the system to make a one-off decision if they wanted to opt out. Most respondents did not raise concerns about changing their minds at a later point and this was not explicitly asked about by researchers. Where respondents did give an example of wanting to reverse inclusion of their data, they assumed this would be possible. Many respondents thought they would try out using the system if invited, and comments indicated that participants were particularly enthusiastic regarding the feedback component of the interface. As indicated in the previous section, respondents understood the need for using health data to enable medical progress; however, they had previously had very little insight into outcomes of research using health records. The electronic system was in this context considered to be especially valuable in providing a mechanism to enable detailed transparency and feedback on relevant research, which may also improve trust and public engagement in research. For example, the patient quoted in the previous section who said he would opt out viewed the Dynamic Consent interface as a useful tool to improve patient control:

> *If there's a trail and you can see where it's being used...you can find out who is using it, what it's being used for, and why it's being used. And then, you know, you could stop it being used.* [Participant #15]

Participants commented on the positive benefits of gaining feedback of where and when their health data had contributed to published articles or breaking news items (see Textbox 2).

Despite the many positive benefits of the system described and the value of feedback, there were a number of respondents who stated that they would not want to use the system for either consent or feedback. They instead reiterated views that they trusted the NHS and researchers to use their data appropriately:

> *Well, honestly, [laughingly] I don't think I would really bother [with the Dynamic Consent interface], but I don't mind anybody having the information to benefit, you know, other people.* [Focus group #1]
>
> *We've got to have research so we can make things better, I mean, what benefit would it be for me to check that feedback. Because they've got the information then, and then they know how to treat me...telling me wouldn't benefit me. Giving me the end product would benefit me.* [Participant #28]

Textbox 2. Quotations representing transparency through Dynamic Consent and patient feedback, and the potential for enhanced control and patient engagement.

> *I like the idea of the Dynamic Consent where you can opt in for bits of it, say, they sent me something online and said we want you to take part in this study, this is what it'll involve, X, Y and Z...I might say well, I'm happy to do X and Y but not Z...I quite like that.* [Participant #3]
>
> *Well, I think it's good, you know, to be able to get involved and to be able to track and control what is happening.* [Focus group #3]
>
> *I think a lot more people would like to know where their health information was being used. Some people might refuse getting involved [in research] because there's a fear of where the information is going.* [Focus group #2]
>
> *[It] lets you know what's happening. And you might find out it's [health data] somewhere where you don't want it to go, but at least you'd know about it.* [Participant #12]
>
> *I just love this idea [Dynamic Consent], the updates they're great. If I was involved with something [research] and it got published, I could go on the Internet and click on that [dynamic interface] and it would give me all the published papers on it.* [Participant #4]
>
> *I thought, oh, that's nice to see the actual papers that have been written on things that you contributed to.* [Participant #5]
>
> *I think this [Dynamic Consent] would encourage more people to get involved with research. Yes, definitely it will improve people's trust.* [Focus group #3]

Operational and Technological Scope and Challenges for Participation

Easy usability of the interface was another positive aspect of Dynamic Consent described by participants. In trying out the interface in the focus groups or interviews, many individuals commented that it was easy to use, describing it as "simple and quick" (Participant #16) with the touch screen viewed as "straightforward [for] people with a variety of conditions" (Participant #2). During focus groups and the PPI groups, participants demonstrated that it was easy to use for people with arthritis involving their hands. Some participants, who had no previous experience of using a tablet device, were able to navigate the prototype easily while being directed to various parts of the app. Participants expressed surprise at how easy it was to use, and said they would be enthusiastic to try a live version. However, a number of respondents also expressed a view that they or others, especially older people, may need initial support to be introduced to using the system. A minority of participants (10/40, 25%) described their potential inability to use the Dynamic Consent interface due to either lack of access to IT platforms at home or lack of confidence in utilizing new information technologies.

These less positive comments were mainly from participants that did not have access to a home computer, never used the Internet, and confessed to being less comfortable utilizing digital technology such as the touch screen interface (see Textbox 3). Out of these 10 participants, 4 (40%) further discussed their willingness to receive support (eg, from a volunteer or a member of staff who could talk through use of the system) to enable use of the system. The remaining few considered they were too old and/or ill—two with terminal cancer—to engage with the technology. A few participants did express they would be happy to complete an alternative paper copy to give consent.

Textbox 3. Quotations representing operational and technological scope and challenges for participation.

> *It is, you know, very straightforward, most people could use it, with a variety of conditions, so it's accessible in terms of that.* [Participant #2]
>
> *Well, it's very easy to use, isn't it, the touch screen, it's a lot easier than a computer.* [Focus group #1]
>
> *It's hard to take this new technology in, you know, when you get to our age you're thinking why bother.* [Focus group #1]

> *Well, if you show me what to do I'd use it because I've never used one, an iPad, you know. As I say, I do use a computer, but not an iPad.* [Focus group #1]
>
> *I wouldn't know where to start with that [Dynamic Consent interface], I can't even send a text...I don't have confidence there...It's out of my league.* [Participant #30]
>
> *Technology-wise, you know, I think it would be quite interesting if you were into that. I can give you an answer if I was, but, no. For me personally, no, but I'd fill a piece of paper in for you.* [Participant #31]

Discussion

Principal Findings

Patients in this study were highly supportive of sharing their anonymized electronic patient record for research and perceived a Dynamic Consent system for consent and feedback to be valuable if implemented. The three key themes characterized by the participants' views were as follows: trust and social responsibility play a major role in patients' views about sharing health care data; there is scope for a Dynamic Consent system to facilitate transparency and patient engagement in reuse of health care data that would be highly valued, and would help mitigate concerns about institutional trustworthiness by enhancing individual control and empowerment [26]; and there are some technological and operational challenges for implementing an electronic system for Dynamic Consent. The discussion is structured around these three core themes below.

The findings echo previous research that patients tend to be supportive of the use of their personal health data for research [11,27,28] and reflect a sense of social responsibility and altruism, as well as potential personal benefits associated with medical research [29]. However, while there was a high level of institutional trust in the NHS and health professionals, similar to other studies, there were concerns about security and potential recipients, especially private companies, who might use data inappropriately if it were exported outside of the NHS [14]. Some participants in this study expressed a view that the media were responsible for overinflating the degree of opposition to reuse health data. Nonetheless, views also demonstrated that trust is not universally assumed, and people want reassurance that the conditions underpinning trust are preserved. Such conditions, including values of reciprocity, nonexploitation, and the public good [3], go beyond the established legal framework; current arrangements mean that people lack necessary information and opportunities for greater control over consent and engagement with research based on reused EPRs.

The findings in this study regarding the value of increased transparency and engagement of patients in the reuse of anonymized health care data reflect recent recommendations of the Caldicott 2 review [4] and the recent Nuffield Council on Bioethics report on "The collection, linking, and use of data in biomedical research" [17]. The recommendations aim to "[provide] greater clarity for members of the public about ways that their biomedical data are used" by providing patient-level information about the recipients of their data and the results thereof. Although research is currently conducted using anonymized health care data without consent, few people are aware of this—a finding reinforced throughout our discussions with patients. Such transparency is deemed an important prerequisite for maintaining public trust [16], providing a rationale for greater openness and engagement with patients. The Dynamic Consent system was considered valuable in this respect and could be viewed to enable the black box around consent and the reuse of health care data to be opened.

The initial emphasis when designing the adapted Dynamic Consent prototype for the reuse of health care data, instead of its original purpose of biobanking, was for enabling patients to have greater control over the reuse of their data. However, during the course of the study and analysis of data it has been apparent that patients particularly valued the feedback components enabling greater transparency of how their data were used. They thought this would give insight into previously hidden research of relevance to their health care, which would make them feel valued as participants. There was much less emphasis from participants on the potential for the system to enable greater control regarding consent. This resonates with other research findings that patients valued explicit consent for use of health data and that this was associated mostly with an interest and a curiosity in the kind of research to which they were contributing [14], as well as the opportunity to engage more closely with the research environment [26]. This also aligns with the major emphasis placed on public and patient involvement in research and provision of health care, making patients feel like active participants in a *research active* nation [30]. In practice, provision of feedback on research using EPRs requires an infrastructure that can support an audit trail of which users have accessed the data. It also requires that the system collect lay summaries of the research findings with a link back to the patient participants. Data access agreements would require research groups to upload their results at the end of their studies. No such infrastructure currently exists and would be challenging to implement nationally. However, it aligns well with the recent investment in four national eHealth Centres of Excellence and plans for developing safe havens for health care data for research [2]. Establishing a patient view into such a research infrastructure could deliver a trusting relationship between the patient community and the data repository and its users.

When the touch screen interface was presented, there were examples of enthusiastic engagement, which was balanced against a notable minority who found the technology daunting. This is a common finding in various studies of IT-based interventions to support home monitoring [31] and some have found major barriers associated with nonadoption of IT-based initiatives [32]. Despite some concern about using the system discussed in this study, participants often stated that they would be willing to use the system in a clinical setting if support was provided. Additionally, patients with chronic rheumatic disease can have problems with dexterity and were thus a good population in which to test the touch screen interfaces. While this initial study provided the encouraging results that participants did not demonstrate limitations due to physical functioning, future implementation would need to consider other groups of patients with special needs, such as poor vision. A minority of participants expressed a view that they would not want to use the electronic system even with support because they had no experience or even preferred to avoid using IT devices. However, this is an important issue to consider in planning for implementation of this model of consent and feedback. Because they expressed willingness to engage with aspects of the system—expressing consent preferences and receiving feedback—use of alternative formats or methods for support need further consideration in refining the design of the system. We have since held an implementation workshop with 35 patients to consider further practical issues around implementation. During this positive and supportive meeting, similar and additional considerations were raised, including the need for enabling hands-on support, paper versions of lay research summaries, and tiered options for the depth of information provided to suit varied literacy levels and levels of interest.

The study was conducted among a specific population of patients with chronic rheumatic disease, and thus we need to consider the generalizability of the study. Patients with a chronic disease might be more motivated to share data compared to those with better general health; conversely, a more extensive medical history could make people reluctant to share personal information. Some have reported variations in views regarding requirements for consent that can be influenced by sociodemographic factors and medical history. For example, previous research has shown that patients may be reluctant to share other aspects of health care, such as sexual history or mental health history [33]. Indeed, one participant speculated that people with more sensitive health conditions might be "more guarded of what happens with their health information" (Participant #11). Depression is a common comorbidity in rheumatic diseases such as rheumatoid arthritis [34]. Consequently, our population might represent a group less willing to share their data than the general population. No patients raised concerns specifically about sharing their rheumatology records. Willingness to join the study may have been influenced by an underlying support for data sharing from participants, although it is equally possible that people opposed to data sharing may have been motivated to join. Our experience suggested that few of the participants understood how health care data were currently shared for research and this potential bias is likely to be small. All of our participants were white, reflecting the local demographic. This may bias the study toward more favorable results, as previous studies have suggested privacy concerns may be higher in black and minority ethnic groups [12]. Recruitment from within a clinical setting may have influenced responses toward higher levels of trust. However, views were similar between participants recruited from the clinic and participants from the PPI research network. Implementation of a Dynamic Consent system would need to consider how the setting (eg, touch screens in clinical settings with endorsement from clinical teams versus Web-based systems from home) might influence uptake, engagement, and consent preferences. The implementation plans would also need to extend testing into other population groups.

Conclusions

In conclusion, this study has generated promising results: a willingness for patients to share their anonymized EPR data for research and a favorable view of a technical solution to meet the needs of recent national recommendations to bring greater transparency and patient engagement in the reuse of EPRs. While uncertainty remains about the degree to which patients will specify consent options in practice, the system offers a potentially valuable technical solution to the challenges of maintaining public trust when sharing medical records for research. This work has provided important insights that will inform the future design of the intervention. We plan to include further codesign [34] in order to maximize the potential for successful implementation and piloting in practice. It represents a first step toward implementation that requires thoughtful development and evaluation, necessarily in a setting with supportive infrastructure. Nonetheless, recent commitment to eHealth research within the government and from funders makes this vision plausible and achievable.

Acknowledgments

KS was supported by a research-funding award from Arthritis Research UK. JK was supported by a Wellcome Trust grant (096599/2/11/Z). WGD was supported by a Medical Research Council (MRC) Clinician Scientist Fellowship (G0902272). Background work was partially supported by the Technology Strategy Board, the Engineering and Physical Sciences Research Council, and the Economic and Social Research Council (grant number EP/G002541/1) for JK, DL, and EAW.

Conflicts of Interest

None declared.

References

1. Strategy for UK Life Sciences. London, UK: Department for Business, Innovation and Skills; 2011 Dec. URL: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/32457/11-1429-strategy-for-uk-life-sciences.pdf [accessed 2016-03-06] [WebCite Cache ID 6fogF9CgQ]
2. The Farr Institute of Health Informatics Research. URL: http://www.farrinstitute.org/ [accessed 2016-03-04] [WebCite Cache ID 6fkyQF5Q3]

3. Carter P, Laurie GT, Dixon-Woods M. The social licence for research: Why care.data ran into trouble. J Med Ethics 2015 May;41(5):404-409 [FREE Full text] [doi: 10.1136/medethics-2014-102374] [Medline: 25617016]
4. Caldicott F. Information: To Share or Not to Share? The Information Governance Review. 2013 Mar. URL: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/192572/2900774_InfoGovernance_accv2.pdf [accessed 2016-03-06] [WebCite Cache ID 6fogm6dR4]
5. Better Information Means Better Care. London, UK: National Health Service; 2014. URL: https://www.england.nhs.uk/wp-content/uploads/2014/01/cd-leaflet-01-14.pdf [accessed 2016-03-04] [WebCite Cache ID 6fkymvGKh]
6. Mayer R, Davis J, Schoorman F. An integrative model of organizational trust. Acad Manage Rev 1995 Jul 01;20(3):709-734 [FREE Full text] [doi: 10.5465/AMR.1995.9508080335]
7. Rousseau DM, Sitkin SB, Burt RS, Camerer C. Not so different after all: A cross-discipline view of trust. Acad Manage Rev 1998 Jul 01;23(3):393-404. [doi: 10.5465/Amr.1998.926617] [Medline: WOS:000075127800005]
8. Gilson L. Trust and the development of health care as a social institution. Soc Sci Med 2003 Apr;56(7):1453-1468. [Medline: 12614697]
9. Bachmann R, Gillespie N, Priem R. Repairing trust in organizations and institutions: Toward a conceptual framework. Organ Stud 2015 Aug 21;36(9):1123-1142. [doi: 10.1177/0170840615599334]
10. Sheather J, Brannan S. Patient confidentiality in a time of care.data. BMJ 2013;347:f7042. [Medline: 24284340]
11. Ipsos MORI. 2011 Jun 09. Public support for research in the NHS URL: https://www.ipsos-mori.com/researchpublications/researcharchive/2811/Public-support-for-research-in-the-NHS.aspx [accessed 2016-03-04] [WebCite Cache ID 6fkzn4yhF]
12. Luchenski SA, Reed JE, Marston C, Papoutsi C, Majeed A, Bell D. Patient and public views on electronic health records and their uses in the United kingdom: Cross-sectional survey. J Med Internet Res 2013;15(8):e160 [FREE Full text] [doi: 10.2196/jmir.2701] [Medline: 23975239]
13. Kass N, Natowicz M, Hull R, Faden R, Plantinga L, Gostin L, et al. The use of medical records in research: What do patients want? J Law Med Ethics 2003;31(3):429-433. [Medline: 14626550]
14. Hill EM, Turner EL, Martin RM, Donovan JL. "Let's get the best quality research we can": Public awareness and acceptance of consent to use existing data in health research: a systematic review and qualitative study. BMC Med Res Methodol 2013;13:72 [FREE Full text] [doi: 10.1186/1471-2288-13-72] [Medline: 23734773]
15. Riordan F, Papoutsi C, Reed J, Marston C, Bell D, Majeed A. Patient and public attitudes towards informed consent models and levels of awareness of Electronic Health Records in the UK. Int J Med Inform 2015 Apr;84(4):237-247 [FREE Full text] [doi: 10.1016/j.ijmedinf.2015.01.008] [Medline: 25649841]
16. Taylor MJ, Taylor N. Health research access to personal confidential data in England and Wales: Assessing any gap in public attitude between preferable and acceptable models of consent. Life Sci Soc Policy 2014 Dec;10:15 [FREE Full text] [doi: 10.1186/s40504-014-0015-6] [Medline: 26085451]
17. The Collection, Linking and Use of Data in Biomedical Research and Health Care: Ethical Issues. London, UK: Nuffield Council on Bioethics; 2015. URL: http://nuffieldbioethics.org/wp-content/uploads/Biological_and_health_data_web.pdf [accessed 2016-03-06] [WebCite Cache ID 6foiOFViz]
18. The Independent Information Governance Oversight Panel's Report to the care.data Programme Board on the care.data Pathfinder Stage. Leeds, UK: National Data Guardian; 2014 Dec 18. URL: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/389219/IIGOP_care.data.pdf [accessed 2016-03-04] [WebCite Cache ID 6fl0MEbST]
19. Williams H, Spencer K, Sanders C, Lund D, Whitley EA, Kaye J, et al. Dynamic consent: A possible solution to improve patient confidence and trust in how electronic patient records are used in medical research. JMIR Med Inform 2015;3(1):e3 [FREE Full text] [doi: 10.2196/medinform.3525] [Medline: 25586934]
20. Kaye J, Whitley EA, Lund D, Morrison M, Teare H, Melham K. Dynamic consent: A patient interface for twenty-first century research networks. Eur J Hum Genet 2015 Feb;23(2):141-146 [FREE Full text] [doi: 10.1038/ejhg.2014.71] [Medline: 24801761]
21. American Thoracic Society, European Respiratory Society. American Thoracic Society/European Respiratory Society International Multidisciplinary Consensus Classification of the Idiopathic Interstitial Pneumonias. This joint statement of the American Thoracic Society (ATS), and the European Respiratory Society (ERS) was adopted by the ATS board of directors, June 2001 and by the ERS Executive Committee, June 2001. Am J Respir Crit Care Med 2002 Jan 15;165(2):277-304. [doi: 10.1164/ajrccm.165.2.ats01] [Medline: 11790668]
22. Kjeken I, Ziegler C, Skrolsvik J, Bagge J, Smedslund G, Tøvik A, et al. How to develop patient-centered research: Some perspectives based on surveys among people with rheumatic diseases in Scandinavia. Phys Ther 2010 Mar;90(3):450-460 [FREE Full text] [doi: 10.2522/ptj.20080381] [Medline: 20110341]
23. Patton MQ. Qualitative Evaluation & Research Methods. 3rd edition. Thousand Oaks, CA: Sage Publications; 2002.
24. YouTube: Health eResearch Centre. Manchester, UK: University of Manchester; 2014 Oct 22. Health records and research: How your data is managed URL: https://www.youtube.com/watch?v=duFb4IfuOpQ [accessed 2016-02-11] [WebCite Cache ID 6fDerNYST]
25. Strauss AL. Qualitative Analysis for Social Scientists. Cambridge, UK: Cambridge University Press; 1987.
26. Teare HJ, Morrison M, Whitley EA, Kaye J. Towards 'Engagement 2.0': Insights from a study of dynamic consent with biobank participants. Digit Health 2015 Sep 28;1:1-13. [doi: 10.1177/2055207615605644]

27. Big Data: Public Views on the Collection, Sharing and Use of Personal Data by Government and Companies. Didcot, UK: Sciencewise; 2014 Apr. URL: http://www.sciencewise-erc.org.uk/cms/assets/Uploads/SocialIntelligenceBigData.pdf [accessed 2016-03-04] [WebCite Cache ID 6fl11s0er]
28. Buckley BS, Murphy AW, MacFarlane AE. Public attitudes to the use in research of personal health information from general practitioners' records: A survey of the Irish general public. J Med Ethics 2011 Jan;37(1):50-55. [doi: 10.1136/jme.2010.037903] [Medline: 21071570]
29. Dixon-Woods M, Tarrant C. Why do people cooperate with medical research? Findings from three studies. Soc Sci Med 2009 Jun;68(12):2215-2222. [doi: 10.1016/j.socscimed.2009.03.034] [Medline: 19394741]
30. Promoting a ‘Research Active’ Nation. London, UK: National Institute for Health Research; 2014 Apr. URL: http://www.nihr.ac.uk/documents/get-involved/Promoting%20A%20Research%20Active%20Nation_NIHR%20Strategic%20Plan_May%202014.pdf [accessed 2016-03-04] [WebCite Cache ID 6fl1CK261]
31. Sanders C, Rogers A, Bowen R, Bower P, Hirani S, Cartwright M, et al. Exploring barriers to participation and adoption of telehealth and telecare within the Whole System Demonstrator trial: A qualitative study. BMC Health Serv Res 2012;12:220 [FREE Full text] [doi: 10.1186/1472-6963-12-220] [Medline: 22834978]
32. Greenhalgh T, Hinder S, Stramer K, Bratan T, Russell J. Adoption, non-adoption, and abandonment of a personal electronic health record: Case study of HealthSpace. BMJ 2010;341:c5814 [FREE Full text] [Medline: 21081595]
33. Willison DJ, Steeves V, Charles C, Schwartz L, Ranford J, Agarwal G, et al. Consent for use of personal information for health research: Do people with potentially stigmatizing health conditions and the general public differ in their opinions? BMC Med Ethics 2009;10:10 [FREE Full text] [doi: 10.1186/1472-6939-10-10] [Medline: 19630941]
34. Dougados M, Soubrier M, Antunez A, Balint P, Balsa A, Buch MH, et al. Prevalence of comorbidities in rheumatoid arthritis and evaluation of their monitoring: Results of an international, cross-sectional study (COMORA). Ann Rheum Dis 2014 Jan;73(1):62-68 [FREE Full text] [doi: 10.1136/annrheumdis-2013-204223] [Medline: 24095940]

Abbreviations

EnCoRe: Ensuring Consent and Revocation
EPR: electronic patient record
GP: general practitioner
IT: information technology
MRC: Medical Research Council
NHS: National Health Service
NIHR: National Institute for Health Research
PPI: patient and public involvement

Edited by G Eysenbach; submitted 18.08.15; peer-reviewed by T Poder, H Williams, C Papoutsi; comments to author 01.10.15; revised version received 12.11.15; accepted 14.11.15; published 15.04.16.

Please cite as:
Spencer K, Sanders C, Whitley EA, Lund D, Kaye J, Dixon WG
Patient Perspectives on Sharing Anonymized Personal Health Data Using a Digital System for Dynamic Consent and Research Feedback: A Qualitative Study
J Med Internet Res 2016;18(4):e66
URL: http://www.jmir.org/2016/4/e66/
doi:10.2196/jmir.5011
PMID:27083521

Original Paper

Opportunities for Web-based Drug Repositioning: Searching for Potential Antihypertensive Agents with Hypotension Adverse Events

Kejian Wang[1], PhD; Mei Wan[2], MS; Rui-Sheng Wang[3], PhD; Zuquan Weng[2], PhD

[1]CoMed Technology & Consulting Co., Ltd., Hong Kong, China
[2]College of Biological Science and Engineering, Fuzhou University, Fuzhou, China
[3]Japan National Institute of Occupational Safety and Health, Kawasaki, Japan

Corresponding Author:
Zuquan Weng, PhD
College of Biological Science and Engineering
Fuzhou University
N Ring Rd
Fuzhou
China
Phone: 1 5017667350
Fax: 1 5017667350
Email: wengzq@fzu.edu.cn

Abstract

Background: Drug repositioning refers to the process of developing new indications for existing drugs. As a phenotypic indicator of drug response in humans, clinical side effects may provide straightforward signals and unique opportunities for drug repositioning.

Objective: We aimed to identify drugs frequently associated with hypotension adverse reactions (ie, the opposite condition of hypertension), which could be potential candidates as antihypertensive agents.

Methods: We systematically searched the electronic records of the US Food and Drug Administration (FDA) Adverse Event Reporting System (FAERS) through the openFDA platform to assess the association between hypotension incidence and antihypertensive therapeutic effect regarding a list of 683 drugs.

Results: Statistical analysis of FAERS data demonstrated that those drugs frequently co-occurring with hypotension events were more likely to have antihypertensive activity. Ranked by the statistical significance of frequent hypotension reporting, the well-known antihypertensive drugs were effectively distinguished from others (with an area under the receiver operating characteristic curve > 0.80 and a normalized discounted cumulative gain of 0.77). In addition, we found a series of antihypertensive agents (particularly drugs originally developed for treating nervous system diseases) among the drugs with top significant reporting, suggesting the good potential of Web-based and data-driven drug repositioning.

Conclusions: We found several candidate agents among the hypotension-related drugs on our list that may be redirected for lowering blood pressure. More important, we showed that a pharmacovigilance system could alternatively be used to identify antihypertensive agents and sustainably create opportunities for drug repositioning.

(J Med Internet Res 2016;18(4):e76) doi:10.2196/jmir.4541

KEYWORDS

Web-based drug repositioning; FDA Adverse Event Reporting System; FAERS; openFDA; big data; antihypertensive drugs; hypotension

Introduction

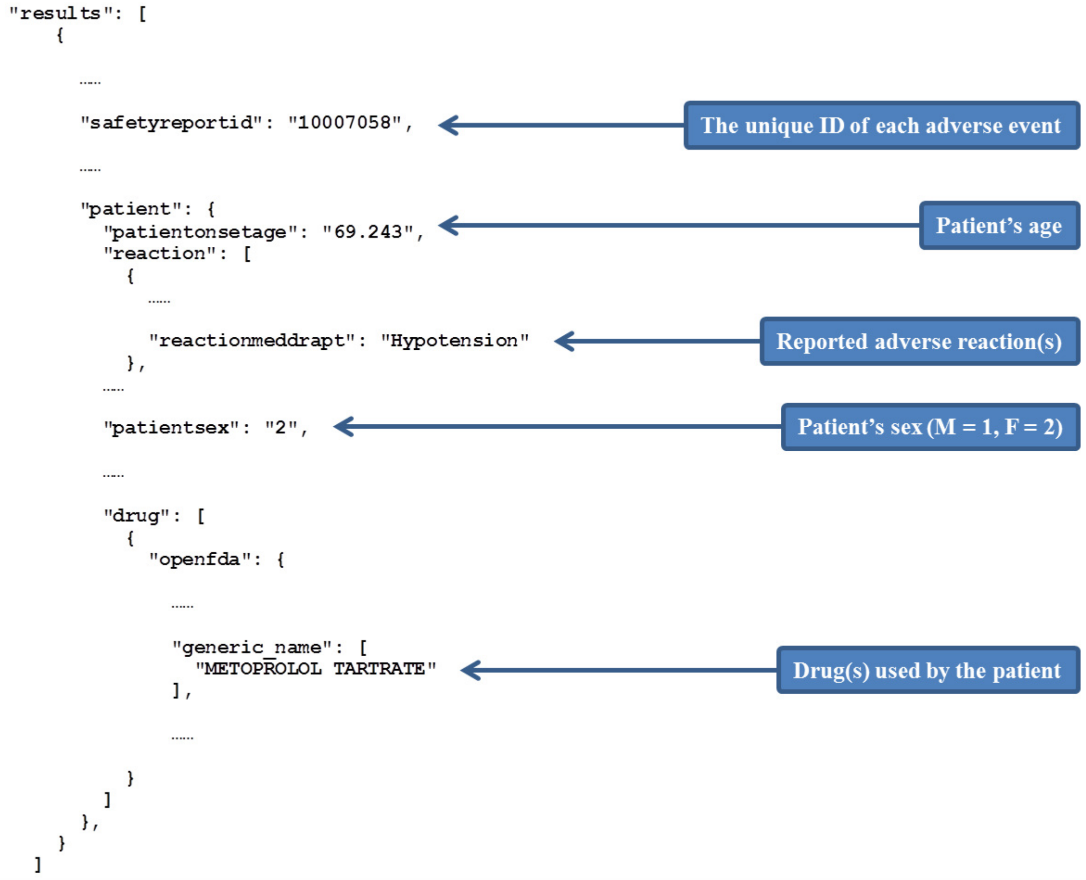
Drug repositioning, also referred to as drug repurposing, is the process of developing new indications for existing drugs [1]. The financial advantage of drug repositioning over traditional drug development is that much of the cost and time spent in the early stage can be bypassed. In addition, the risk of failure caused by adverse reactions can be better controlled, since the toxicity of the repurposed drugs has already been tested [2]. For these reasons, both the pharmaceutical industry and academic communities are paying particular attention to this field. An increasing number of *in silico* [3,4] and *in vitro* [5-7] approaches have been developed to efficiently scan the existing pharmacopoeia for new usage. Other than these strategies primarily focused on preclinical information, side effects data are increasingly used for rational drug repositioning [8-10], due to the direct reflection of the clinical reality of actual patients [11]. The collection of side effects information based on clinical trials is conventionally a time-consuming and labor-intensive process. Making things more difficult, most raw data are not freely available to the public. However, the big data concept and Internet-related technologies are making it easier to access and analyze side effects records.

Hypertension, characterized by aberrantly high arterial blood pressure, is a chronic medical condition affecting almost one billion people worldwide [12]. Despite the availability of several blood pressure-decreasing drugs, hypertension is not effectively controlled in more than half of patients receiving antihypertensive treatment [13]. Among the multiple reasons contributing to this unsatisfactory clinical outcome, an undeniable reality is that most pharmaceutical companies, considering the potential costs and profits, have abandoned antihypertensive drug development [14]. As drug pipelines dry up, new drug development is expensive and time consuming, thus diminishing future profitability [15]. Worse still, the high failure rate (estimated to be >90%) due to toxicity and other reasons has made new drug development a highly risky investment [16]. Therefore, alternative strategies are urgently needed to drive the productivity and cost-effectiveness of antihypertensive drug development. Assuming that hypotension, with the symptom of abnormally low blood pressure, can be regarded as the opposite condition of hypertension, then potential antihypertensive agents may be discovered among drugs that induce hypotension as a side effect.

Established by the US Food and Drug Administration (FDA), the FDA Adverse Event Reporting System (FAERS) [17] is one of the most comprehensive sources of pharmacovigilance big data worldwide. With adverse drug events spontaneously submitted by consumers and indirectly reported by health care professionals, FAERS supports not only FDA's safety surveillance on all approved drugs, but also the research of scientists and clinicians. Each report provides a variety of clinical information, particularly the drug(s) used by the patient and the adverse reaction(s) that the patient experienced. FAERS' raw data used to be released without sufficient formatting, thus making a high throughput analysis very difficult. Fortunately, the FDA's Office of Informatics and Technology Innovation launched a new initiative, OpenFDA, in 2013 [18,19], whose primary goal was to facilitate public access to high-value FDA data (including FAERS) by providing user-friendly and open-source application programming interfaces.

By using the OpenFDA platform and relevant statistical methods [20], we aimed to examine the co-occurrence of specific drugs and hypotension as a side effect, that is, the drug and a hypotension reaction co-occurred in the same adverse event report (Figure 1). Instead of serendipitous searching, we examined the correlation between hypotension incidence and antihypertensive property for a total of 683 unique drugs, so as to identify those that frequently co-occurred with hypotension adverse events. As an exploratory effort in systematically analyzing drug adverse events, we hoped that this study would provide a unique insight into Web-based and data-driven drug repositioning.

Figure 1. Co-occurrence of drug and adverse reaction. An adverse event with unique identifier (ID) can be queried in the US Food and Drug Administration (FDA) Adverse Event Reporting System (FAERS), which returns results as JavaScript Object Notation by default. In this illustrated adverse event (with partial content displayed), the use of metoprolol and the incidence of hypotension reaction are reported simultaneously, which is defined as co-occurrence.



```
"results": [
    {

      ......

      "safetyreportid": "10007058",

      ......

      "patient": {
        "patientonsetage": "69.243",
        "reaction": [
          {
            ......

            "reactionmeddrapt": "Hypotension"
          },
        ......

        "patientsex": "2",

        ......

        "drug": [
          {
            "openfda": {

              ......

              "generic_name": [
                "METOPROLOL TARTRATE"
              ],

              ......

          }
        ]
      },
    }
  ]
```

Methods

FAERS Data Query

We retrieved the raw FAERS reports (from January 2004 to September 2013) from the public adverse events dataset of OpenFDA in December 2014, according to the official tutorial on query tools [9]. We primarily investigated the 1000 most commonly reported generic names. By merging multiple generic names corresponding to the same drug product (eg, “abacavir” and “abacavir sulfate”) and excluding combinations of multiple active ingredients (eg, “oxycodone and acetaminophen”), we identified a total of 683 unique drugs. For each drug, we searched for hypotension-related events using the terms “hypotension,” “blood pressure decreased,” and “orthostatic hypotension.”

Statistical Analysis

For each drug of interest, we constructed a 2×2 contingency table to serve as the framework for analysis of all FAERS reports. For the reports with adverse event co-occurrence with the drug, we defined the numbers of events including hypotension as n11 and excluding hypotension as n10. For the reports not related to the drug of interest, we defined the numbers of events with hypotension as n01 and without hypotension as n00. Then, we calculated the hypotension reporting odds ratio (ROR) as (n11 × n00) / (n10 × n01). We determined the statistical significance of the ROR by Fisher exact test.

Normalized Discounted Cumulative Gain

Normalized discounted cumulative gain (NDCG) measures the relevance of a document based on its position in the result list. The gain is accumulated from the top of the result list to the bottom, with the gain of each result discounted at lower ranks. We ranked each of the 683 drugs by the significance of its ROR and judged it on a relevance score, with 1 meaning approved antihypertensives and 0 meaning other, irrelevant drugs. We calculated the crude discounted cumulative gain (DCG) by equation (a) in Figure 2.

To normalize the DCG value, we produced an ideal ordering for the 683 drugs (the ranking with maximum possible DCG) to calculate the ideal discounted cumulative gain. Then the NDCG, ranging from 0 to 1, was calculated by equation (b) in Figure 2. The closer the NDGC value is to 1, the better the performance. In addition, we calculated the NDCG for the top 20 drugs (NDCG@20) by equation (c) in Figure 2. Similarly, the closer the NDGC@20 value is to 1, the better the performance in terms of identifying the top 20 drugs.

Figure 2. Equations for calculating crude discounted cumulative gain (DCG), normalized discounted cumulative gain (NDCG), and NDCG for the top 20 drugs (NDCG@20). DCG@20: DCG of the top 20 drugs; IDCG: ideal discounted cumulative gain; IDCG@20: IDCG of the top 20 drugs; rel_1: relevance score for approved antihypertensive drugs; rel_i: relevance score.

(a)

$$\mathrm{DCG} = rel_1 + \sum_{i=2}^{683} \frac{rel_i}{\log_2 i}$$

(b)

$$\mathrm{NDCG} = \frac{DCG}{IDCG}$$

(c)

$$\mathrm{DCG@20} = rel_1 + \sum_{i=2}^{20} \frac{rel_i}{\log_2 i}$$

$$\mathrm{NDCG@20} = \frac{DCG@20}{IDCG@20}$$

Results

Co-occurrence of Well-Known Antihypertensive Drugs and Hypotension Reports

To justify the basic assumption that hypotension events may suggest antihypertensive activity, we expected the approved antihypertensive drugs to be reported more often for hypotension. To achieve broad drug coverage, we examined 683 study drugs with a single active ingredient for hypotension adverse events (Multimedia Appendix 1). As was done previously in a series of FAERS-based studies [21,22], we calculated the significance level of the ROR [23] to assess the association between drug use and hypotension incidence, according to the numbers of hypotension and nonhypotension reports that co-occurred with the drug of interest (see Methods).

We ranked all of the 683 study drugs by the significance level (in terms of unadjusted *P* value), which pinpointed well-known FDA-approved antihypertensives according to drug indication information that we retrieved from DrugBank [24]. We found that the approved antihypertensives were effectively distinguished from other study drugs. The approved antihypertensives were highly represented among the top significant drugs (ie, with the lowest *P* values), and the area under the receiver operating characteristic (ROC) curve was >0.80 (Figure 3). When the percentage of nonantihypertensive drugs (ie, false-positive rate) was <10%, the proportion of approved antihypertensives was nearly 40% and the partial area under the ROC curve was 0.026. In addition, we calculated the NDCG [25-29] to measure the quality of the ranking of approved antihypertensives (see Methods). The overall NDCG and NDCG@20 of our model were 0.77 and 0.59, respectively, showing a remarkable performance. Among the approved antihypertensives, 86% (48/56) achieved an unadjusted $P<10^{-10}$. On the other hand, the percentage for other drugs was only 39.4% (247/627). Such a dramatic difference (Figure 4) (odds ratio 9.21, Fisher exact test $P=8.96\times10^{-12}$) indicated that hypotension reporting in FAERS could serve as a phenotypic indicator and effectively reflect the well-known antihypertensive activity.

Figure 3. Correlation between antihypertensive activity and hypotension reporting in the US Food and Drug Administration (FDA) Adverse Event Reporting System (FAERS), showing the classification between approved antihypertensives and other drugs.

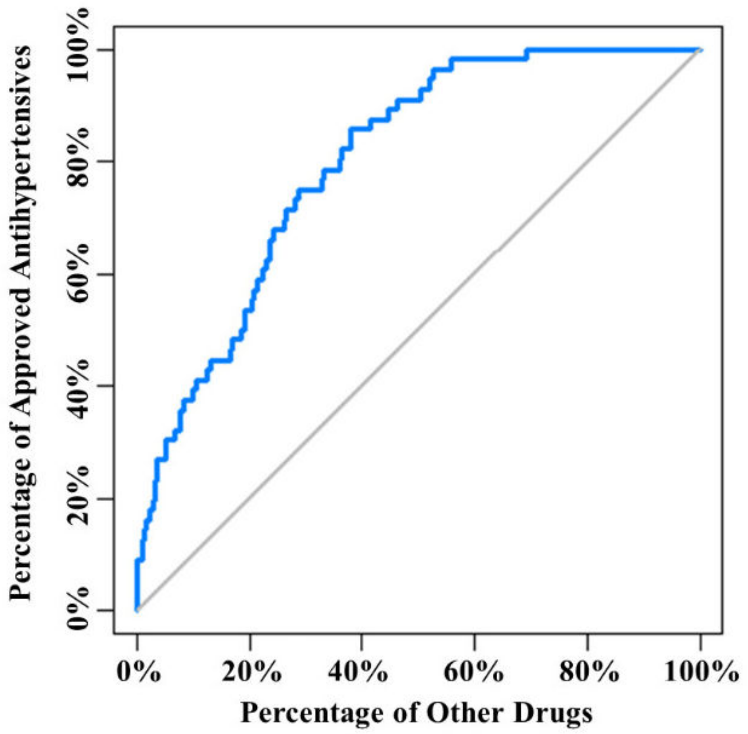
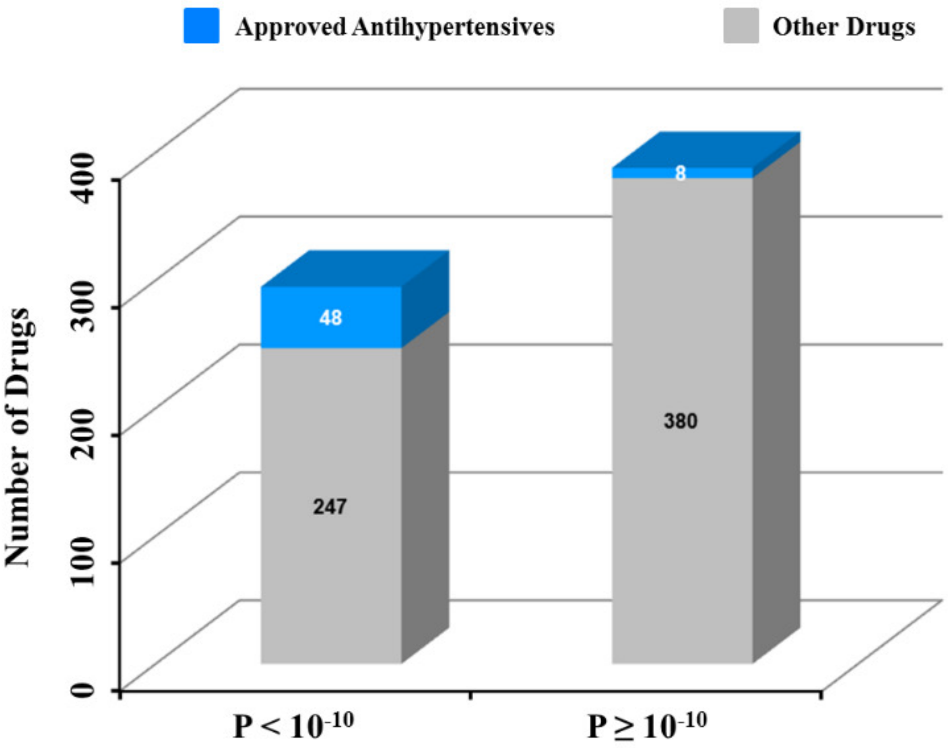


Figure 4. Correlation between antihypertensive activity and hypotension reporting in the US Food and Drug Administration (FDA) Adverse Event Reporting System (FAERS), showing the number of drugs with significant hypotension reporting with regard to the cut-off of $P<10^{-10}$.



Frequent Hypotension Reporting as an Indicator of Potential Antihypertensive Agents

Since the correlation has been proved between the reporting of hypotension as a side effect and the antihypertensive indication, we speculated that the drugs still not approved for treatment of hypertension but frequently co-occurring with hypotension adverse events might lead to hidden opportunities for drug repositioning. To test this hypothesis, we preliminarily surveyed all of the top 20 drugs as ranked by the significance level (Table 1).

Among the top 11 drugs with a *P* value close to zero, 5 are approved antihypertensive drugs: metoprolol, spironolactone, furosemide, lisinopril, and carvedilol. Surprisingly, most of the other top-ranked drugs were also reported for hypotensive effects, particularly in humans. Aspirin, in addition to its importance as a widely used analgesic and anti-inflammatory drug, has been repeatedly reported to improve blood pressure control in hypertensive patients [30-32]. Digoxin, originally indicated for arrhythmias and heart diseases, was also found in various double-blind studies to significantly decrease diastolic blood pressure during overnight sleep [33,34]. As an essential mineral in the body, potassium has been clinically proven to lower blood pressure in humans [35-37]. And according to the package inserts (ie, the warnings and precautions subsection), as approved by the FDA, the anesthetic drugs morphine [38] and propofol [39] have a warning for inducing hypotension in clinical trials. This leads to the conclusion that FAERS can be efficiently screened for additional molecules, besides the approved antihypertensive drugs, with potential antihypertensive properties of the additional molecules supported by independent clinical evidence.

Table 1. The top 20 drugs that most frequently co-occurred with hypotension adverse events.

Drug	ATC[a] code	*P* value	ROR[b]	Adverse events co-occurring with drug (n)		Adverse events not co-occurring with the drug (n)	
				Hypotension	Not hypotension	Hypotension	Not hypotension
Metoprolol	C07AB02[c]	0	2.99	3287	62,998	64,291	3,683,647
Spironolactone	C03DA01[c]	0	3.71	1581	24,021	65,997	3,722,624
Furosemide	C03CA01[c]	0	3.49	5173	86,868	62,405	3,659,777
Lisinopril	C09AA03[c]	0	2.69	3465	73,779	64,113	3,672,866
Carvedilol	C07AG02[c]	0	4.11	1955	26,954	65,623	3,719,691
Propofol	N01AX10	0	7.81	1098	7903	66,480	3,738,742
Digoxin	C01AA05	0	4.44	2259	28,945	65,319	3,717,700
Potassium	N/A	0	3.51	1533	24,635	66,045	3,722,010
Morphine	N02AA01	0	3.02	1683	31,375	65,895	3,715,270
Warfarin	B01AA03	0	2.52	3076	69,491	64,502	3,677,154
Aspirin	A01AD05 B01AC06 N02BA01	0	2.46	6107	145,276	61,471	3,601,369
Amlodipine	C08CA01[c]	3.95E–315	2.29	2854	70,799	64,724	3,675,846
Isosorbide	C01DA14[c]	5.12E–319	4.16	1139	15,390	66,439	3,731,255
Clopidogrel	B01AC04	9.83E–306	2.45	2395	55,283	65,183	3,691,362
Acetaminophen	N02BE01	8.37E–302	2.51	2266	51,149	65,312	3,695,496
Atenolol	C07AB03[c]	4.26E–298	2.52	2223	49,993	65,355	3,696,652
Midazolam	N05CD08	1.30E–296	7.13	641	5023	66,937	3,741,622
Ondansetron	A04AA01	9.23E–295	3.51	1293	20,700	66,285	3,725,945
Lorazepam	N05BA06	3.13E–284	2.62	1965	42,386	65,613	3,704,259
Ramipril	C09AA05[c]	2.90E–275	2.97	1539	29,213	66,039	3,717,432

[a] ATC: Anatomical Therapeutic Chemical Classification System.

[b] ROR: reporting odds ratio.

[c] Approved antihypertensive drug.

In addition, we observed a possible correlation between the Anatomical Therapeutic Chemical (ATC) Classification System code and antihypertensive activity in the top 20 agents. As Table 1 shows, except for drugs designed for cardiac therapy (ie, indicated by the ATC first-level code C), other candidate agents were originally developed for treating the alimentary tract and metabolism (ie, ATC first-level code A), blood and blood-forming organs (ie, ATC first-level code B), or the nervous system (ie, ATC first-level code N). In particular, drugs acting on the nervous system were well represented in the top-ranked drugs, as evidence by a relatively high NDCG@20 value of 0.35 (Table 2). This observation was consistent with the commonly known interplay between blood pressure and the nervous system [40], suggesting that the hypotensive effects of nervous system agents can be detected effectively with patient-centric pharmacovigilance data. Therefore, along the direction of this study, particular attention may be paid to nervous system agents in further searches for novel antihypertensives.

Table 2. Noncardiac drug classes highly represented in the top 20 drugs.

ATC[a] code (first level)	Drug class	Ranks in top 20	NDCG@20[b]
N	Nervous system	6, 9, 11, 15, 17, 19	0.35
B	Blood and blood-forming organs	10, 11, 14	0.17
A	Alimentary tract and metabolism	11, 18	0.12

[a] ATC: Anatomical Therapeutic Chemical Classification System.

[b] NDCG@20: normalized discounted cumulative gain for the top 20 drugs.

Discussion

In recent years, conventional target-based drug development has been facing various challenges, such as increasingly higher costs and lower productivity. In consequence, both the pharmaceutical industry and the academic community are earnestly seeking new means of drug innovation. In this process, many realize that, along with improvements in clinical informatics, the role of patients can be changed from a minimally informed recipient to a potential participant in drug development [41,42]. By providing raw download access to structured datasets, OpenFDA has created a researcher-friendly portal for quick and easy querying. With patient-centric data being a highly cost-effective and valuable resource, we can make this systematic effort and corroborate the evidence linking hypotension as a side effect to an antihypertensive therapeutic effect.

For hypertension and many other medical conditions, rationally finding unknown therapeutic agents has always been a valuable but difficult task. Even though various methods are established to accelerate the process of selecting repositioning candidates, there remains at least one major barrier between theoretical therapeutic effects and real benefits. Due to the complexity of the human body, many repositioning approaches that are focused on omics and preclinical data may not always be consistent with patients' clinical outcomes [43]. Regarding this issue, the unique advantage of clinical informatics analysis using FAERS should be appreciable. Basically, hypotension adverse events are observations on humans, as opposed to molecular entities, cell cultures, or animal models. The antihypertensive signals suggested by FAERS tend to be more interpretable and straightforward, presenting relatively fewer challenges to bench-to-bed translation.

Pharmacovigilance data are not the only source of side effect information, since package inserts (also known as drug labels) determined by regulatory agencies have also been studied for drug repositioning [10,44,45]. However, in several ways, our study suggests that the patient-based FAERS data play an irreplaceable role in side effect analysis. As the basis of drug labeling, clinical trials are mostly conducted among a relatively small number of people and over a limited period of time. In contrast, postmarketing drug surveillance is a long-term effort that involves the general population. Therefore, the incidence of various side effects can be more comprehensively monitored through the patients' self-reports. In addition, drug labels usually focus on the existence of certain side effects, while the prevalence is mentioned less often. However, the frequency of hypotension is an essential parameter for estimating the size of the applicable population, which is directly related to the potential profitability of the repositioned drug. In this regard, the FAERS data enabled us to calculate the significance level of the hypotension ROR for all of the study drugs, efficiently screening for the promising candidates.

Despite the richness of the information, some limitations related to the voluntary nature of FAERS merit additional attention. Since the details about the patients' medical history and the context of adverse events may not be sufficiently addressed in many spontaneously submitted reports, the co-occurrence of a specific drug and a side effect may not be directly interpreted as a cause-and-effect relationship [46]. To address this issue, the patient-expert relationship needs to be enhanced following the concept of participatory medicine [47,48]. For instance, an Internet-based community can be established to connect patients and experts. The first-hand information reported by patients can then be promptly scrutinized and curated by clinical experts, thus improving the reliability of the raw data. Beyond being a source of information, patients can therefore play a more proactive role in clinical informatics research.

Given the lack of a novel candidate molecule, as well as a new mechanism of action, in antihypertensive drug development for a long time now, our work may have both realistic and long-term implications. First, our results provide a collection of candidate agents that may decrease blood pressure. Since most of the top frequently reported drugs for hypotension events are either well-known antihypertensive drugs or were proved later to show antihypertensive activity, there are grounds for believing that novel antihypertensive agents can be found among other study drugs that we identified (see Multimedia Appendix 1). We therefore suggest that the highly significant signals of certain drugs should be followed up by experimental or clinical investigations. Second, and more important, this study has shown a sustainable way of detecting potential antihypertensive agents. As numerous adverse events are consistently submitted to FAERS and periodically released through the OpenFDA platform, not only the marketed drugs addressed in this study, but also new drugs approved in the future may be highlighted for subsequent hypotension adverse events. Third, we believe the rationale of the opposite condition, linking the hypotension side effect and hypertension treatment, can be expanded to other indications with elaborate study designs. As long as FAERS continues to be updated, more opportunities for drug repositioning may be persistently discovered. Eventually, the newfound agents and the underlying mechanism of action would promote the discovery and development of various therapies.

Acknowledgments

Supported by the Drug Repositioning Programme of CoMed Technology & Consulting Co., Limited, Hong Kong; and Start-up Funds of Minjiang Scholar (510170), Fujian, China.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Supplementary data regarding all 683 study drugs.

[XLSX File (Microsoft Excel File), 69KB - jmir_v18i4e76_app1.xlsx]

References

1. Ashburn TT, Thor KB. Drug repositioning: identifying and developing new uses for existing drugs. Nat Rev Drug Discov 2004 Aug;3(8):673-683. [doi: 10.1038/nrd1468] [Medline: 15286734]
2. Tobinick EL. The value of drug repositioning in the current pharmaceutical market. Drug News Perspect 2009 Mar;22(2):119-125. [doi: 10.1358/dnp.2009.22.2.1303818] [Medline: 19330170]
3. Yang L, Chen J, Shi L, Hudock MP, Wang K, He L. Identifying unexpected therapeutic targets via chemical-protein interactome. PLoS One 2010;5(3):e9568 [FREE Full text] [doi: 10.1371/journal.pone.0009568] [Medline: 20221449]
4. Luo H, Chen J, Shi L, Mikailov M, Zhu H, Wang K, et al. DRAR-CPI: a server for identifying drug repositioning potential and adverse drug reactions via the chemical-protein interactome. Nucleic Acids Res 2011 Jul;39(Web Server issue):W492-W498 [FREE Full text] [doi: 10.1093/nar/gkr299] [Medline: 21558322]
5. Kunkel SD, Suneja M, Ebert SM, Bongers KS, Fox DK, Malmberg SE, et al. mRNA expression signatures of human skeletal muscle atrophy identify a natural compound that increases muscle mass. Cell Metab 2011 Jun 8;13(6):627-638 [FREE Full text] [doi: 10.1016/j.cmet.2011.03.020] [Medline: 21641545]
6. Wang K, Sun J, Zhou S, Wan C, Qin S, Li C, et al. Prediction of drug-target interactions for drug repositioning only based on genomic expression similarity. PLoS Comput Biol 2013;9(11):e1003315 [FREE Full text] [doi: 10.1371/journal.pcbi.1003315] [Medline: 24244130]
7. Sirota M, Dudley JT, Kim J, Chiang AP, Morgan AA, Sweet-Cordero A, et al. Discovery and preclinical validation of drug indications using compendia of public gene expression data. Sci Transl Med 2011 Aug 17;3(96):96ra77 [FREE Full text] [doi: 10.1126/scitranslmed.3001318] [Medline: 21849665]
8. Campillos M, Kuhn M, Gavin A, Jensen LJ, Bork P. Drug target identification using side-effect similarity. Science 2008 Jul 11;321(5886):263-266 [FREE Full text] [doi: 10.1126/science.1158140] [Medline: 18621671]
9. Tatonetti NP, Denny JC, Murphy SN, Fernald GH, Krishnan G, Castro V, et al. Detecting drug interactions from adverse-event reports: interaction between paroxetine and pravastatin increases blood glucose levels. Clin Pharmacol Ther 2011 Jul;90(1):133-142 [FREE Full text] [doi: 10.1038/clpt.2011.83] [Medline: 21613990]
10. Yang L, Agarwal P. Systematic drug repositioning based on clinical side-effects. PLoS One 2011;6(12):e28025 [FREE Full text] [doi: 10.1371/journal.pone.0028025] [Medline: 22205936]
11. Wang K, Weng Z, Sun L, Sun J, Zhou S, He L. Systematic drug safety evaluation based on public genomic expression (Connectivity Map) data: myocardial and infectious adverse reactions as application cases. Biochem Biophys Res Commun 2015 Feb 13;457(3):249-255. [doi: 10.1016/j.bbrc.2014.12.096] [Medline: 25576362]
12. Kearney PM, Whelton M, Reynolds K, Muntner P, Whelton PK, He J. Global burden of hypertension: analysis of worldwide data. Lancet 2005;365(9455):217-223. [doi: 10.1016/S0140-6736(05)17741-1] [Medline: 15652604]

13. Hajjar I, Kotchen TA. Trends in prevalence, awareness, treatment, and control of hypertension in the United States, 1988-2000. JAMA 2003 Jul 9;290(2):199-206. [doi: 10.1001/jama.290.2.199] [Medline: 12851274]
14. Feig PU, Roy S, Cody RJ. Antihypertensive drug development: current challenges and future opportunities. J Am Soc Hypertens 2010;4(4):163-173. [doi: 10.1016/j.jash.2010.04.003] [Medline: 20537610]
15. Morgan S, Grootendorst P, Lexchin J, Cunningham C, Greyson D. The cost of drug development: a systematic review. Health Policy 2011 Apr;100(1):4-17. [doi: 10.1016/j.healthpol.2010.12.002] [Medline: 21256615]
16. DiMasi JA, Hansen RW, Grabowski HG, Lasagna L. Cost of innovation in the pharmaceutical industry. J Health Econ 1991 Jul;10(2):107-142. [Medline: 10113009]
17. Young RC, Friedman MA, Schilsky RL, Sigal EV. Drug safety and drug efficacy: two sides of the same coin. Clin Cancer Res 2007 May 1;13(9):2533-2534 [FREE Full text] [doi: 10.1158/1078-0432.CCR-07-0608] [Medline: 17473179]
18. Kuehn BM. FDA opens adverse event data. JAMA 2014 Jul 23;312(4):328. [doi: 10.1001/jama.2014.9120]
19. Hampton T. Data mining approach shows promise in detecting unexpected drug interactions. JAMA 2011 Jul 13;306(2):144. [doi: 10.1001/jama.2011.918] [Medline: 21750287]
20. Sakaeda T, Tamon A, Kadoyama K, Okuno Y. Data mining of the public version of the FDA Adverse Event Reporting System. Int J Med Sci 2013;10(7):796-803 [FREE Full text] [doi: 10.7150/ijms.6048] [Medline: 23794943]
21. Piccinni C, Motola D, Marchesini G, Poluzzi E. Assessing the association of pioglitazone use and bladder cancer through drug adverse event reporting. Diabetes Care 2011 Jun;34(6):1369-1371 [FREE Full text] [doi: 10.2337/dc10-2412] [Medline: 21515844]
22. Tamura T, Sakaeda T, Kadoyama K, Okuno Y. Omeprazole- and esomeprazole-associated hypomagnesaemia: data mining of the public version of the FDA Adverse Event Reporting System. Int J Med Sci 2012;9(5):322-326 [FREE Full text] [doi: 10.7150/ijms.4397] [Medline: 22745572]
23. Rothman KJ, Lanes S, Sacks ST. The reporting odds ratio and its advantages over the proportional reporting ratio. Pharmacoepidemiol Drug Saf 2004 Aug;13(8):519-523. [doi: 10.1002/pds.1001] [Medline: 15317031]
24. Law V, Knox C, Djoumbou Y, Jewison T, Guo AC, Liu Y, et al. DrugBank 4.0: shedding new light on drug metabolism. Nucleic Acids Res 2014 Jan;42(Database issue):D1091-D1097 [FREE Full text] [doi: 10.1093/nar/gkt1068] [Medline: 24203711]
25. Jarvelin K, Kekalainen J. Cumulated gain-based evaluation of IR techniques. ACM Trans Inf Syst 2002;20(4):422-446.
26. Zhang W, Ji L, Chen Y, Tang K, Wang H, Zhu R, et al. When drug discovery meets web search: Learning to Rank for ligand-based virtual screening. J Cheminform 2015;7:5 [FREE Full text] [doi: 10.1186/s13321-015-0052-z] [Medline: 25705262]
27. Ye H, Liu Q, Wei J. Construction of drug network based on side effects and its application for drug repositioning. PLoS One 2014;9(2):e87864 [FREE Full text] [doi: 10.1371/journal.pone.0087864] [Medline: 24505324]
28. Claveau V, Hamon T, Le MS, Grabar N. Health consumer-oriented information retrieval. Stud Health Technol Inform 2015;210:80-84. [Medline: 25991106]
29. Ullah MZ, Aono M, Seddiqui MH. Estimating a ranked list of human genetic diseases by associating phenotype-gene with gene-disease bipartite graphs. ACM Trans Intell Syst Technol 2015;6(4):1-21.
30. Hermida RC, Ayala DE, Calvo C, López JE. Aspirin administered at bedtime, but not on awakening, has an effect on ambulatory blood pressure in hypertensive patients. J Am Coll Cardiol 2005 Sep 20;46(6):975-983 [FREE Full text] [doi: 10.1016/j.jacc.2004.08.071] [Medline: 16168278]
31. Hermida RC, Ayala DE, Calvo C, López JE, Mojón A, Rodríguez M, et al. Differing administration time-dependent effects of aspirin on blood pressure in dipper and non-dipper hypertensives. Hypertension 2005 Oct;46(4):1060-1068 [FREE Full text] [doi: 10.1161/01.HYP.0000172623.36098.4e] [Medline: 16087788]
32. Magen E, Viskoper JR, Mishal J, Priluk R, London D, Yosefy C. Effects of low-dose aspirin on blood pressure and endothelial function of treated hypertensive hypercholesterolaemic subjects. J Hum Hypertens 2005 Sep;19(9):667-673. [doi: 10.1038/sj.jhh.1001910] [Medline: 16034448]
33. Grossmann M, Jamieson MJ, Kirch W. Effects of digoxin and digitoxin on circadian blood pressure profile in healthy volunteers. Eur J Clin Invest 1998 Sep;28(9):701-706. [Medline: 9767368]
34. Kirch C, Grossmann M, Fischer S, Neumeister V, Kirch W. Effect of digoxin on circadian blood pressure values in patients with congestive heart failure. Eur J Clin Invest 2000 Apr;30(4):285-289. [Medline: 10759875]
35. Whelton PK, He J, Cutler JA, Brancati FL, Appel LJ, Follmann D, et al. Effects of oral potassium on blood pressure. Meta-analysis of randomized controlled clinical trials. JAMA 1997 May 28;277(20):1624-1632. [Medline: 9168293]
36. Haddy FJ, Vanhoutte PM, Feletou M. Role of potassium in regulating blood flow and blood pressure. Am J Physiol Regul Integr Comp Physiol 2006 Mar;290(3):R546-R552 [FREE Full text] [doi: 10.1152/ajpregu.00491.2005] [Medline: 16467502]
37. Houston MC. The importance of potassium in managing hypertension. Curr Hypertens Rep 2011 Aug;13(4):309-317. [doi: 10.1007/s11906-011-0197-8] [Medline: 21403995]
38. Qian Y, Ye X, Du W, Ren J, Sun Y, Wang H, et al. A computerized system for detecting signals due to drug-drug interactions in spontaneous reporting systems. Br J Clin Pharmacol 2010 Jan;69(1):67-73 [FREE Full text] [doi: 10.1111/j.1365-2125.2009.03557.x] [Medline: 20078614]

39. Dallenbach MF, Bovier PA, Desmeules J. Detecting drug interactions using personal digital assistants in an out-patient clinic. QJM 2007 Nov;100(11):691-697 [FREE Full text] [doi: 10.1093/qjmed/hcm088] [Medline: 17932098]
40. Parati G, Esler M. The human sympathetic nervous system: its relevance in hypertension and heart failure. Eur Heart J 2012 May;33(9):1058-1066 [FREE Full text] [doi: 10.1093/eurheartj/ehs041] [Medline: 22507981]
41. Powell J, Buchan I. Electronic health records should support clinical research. J Med Internet Res 2005;7(1):e4 [FREE Full text] [doi: 10.2196/jmir.7.1.e4] [Medline: 15829476]
42. Xie B, Wang M, Feldman R, Zhou L. Internet use frequency and patient-centered care: measuring patient preferences for participation using the health information wants questionnaire. J Med Internet Res 2013;15(7):e132 [FREE Full text] [doi: 10.2196/jmir.2615] [Medline: 23816979]
43. Buchan NS, Rajpal DK, Webster Y, Alatorre C, Gudivada RC, Zheng C, et al. The role of translational bioinformatics in drug discovery. Drug Discov Today 2011 May;16(9-10):426-434. [doi: 10.1016/j.drudis.2011.03.002] [Medline: 21402166]
44. Kuhn M, Campillos M, Letunic I, Jensen LJ, Bork P. A side effect resource to capture phenotypic effects of drugs. Mol Syst Biol 2010;6:343 [FREE Full text] [doi: 10.1038/msb.2009.98] [Medline: 20087340]
45. Bisgin H, Liu Z, Kelly R, Fang H, Xu X, Tong W. Investigating drug repositioning opportunities in FDA drug labels through topic modeling. BMC Bioinformatics 2012;13 Suppl 15:S6 [FREE Full text] [doi: 10.1186/1471-2105-13-S15-S6] [Medline: 23046522]
46. Ahmad SR. Adverse drug event monitoring at the Food and Drug Administration. J Gen Intern Med 2003 Jan;18(1):57-60 [FREE Full text] [Medline: 12534765]
47. Gerber BS, Eiser AR. The patient physician relationship in the Internet age: future prospects and the research agenda. J Med Internet Res 2001;3(2):E15 [FREE Full text] [doi: 10.2196/jmir.3.2.e15] [Medline: 11720957]
48. Hart A, Henwood F, Wyatt S. The role of the Internet in patient-practitioner relationships: findings from a qualitative research study. J Med Internet Res 2004 Sep 30;6(3):e36 [FREE Full text] [doi: 10.2196/jmir.6.3.e36] [Medline: 15471762]

Abbreviations

ATC: Anatomical Therapeutic Chemical Classification System
DCG: crude discounted cumulative gain
FAERS: FDA Adverse Event Reporting System
FDA: US Food and Drug Administration
NDCG: normalized discounted cumulative gain
NDCG@20: normalized discounted cumulative gain for the top 20 drugs
ROC: receiver operating characteristic
ROR: reporting odds ratio

Edited by G Eysenbach; submitted 16.04.15; peer-reviewed by H Ye; comments to author 12.11.15; revised version received 02.12.15; accepted 04.01.16; published 01.04.16.

Please cite as:
Wang K, Wan M, Wang RS, Weng Z
Opportunities for Web-based Drug Repositioning: Searching for Potential Antihypertensive Agents with Hypotension Adverse Events
J Med Internet Res 2016;18(4):e76
URL: http://www.jmir.org/2016/4/e76/
doi:10.2196/jmir.4541
PMID:27036325

Original Paper

Answers to Health Questions: Internet Search Results Versus Online Health Community Responses

Shaheen Kanthawala[1], MA; Amber Vermeesch[2], RN, FNP-C, MSN, PhD; Barbara Given[3], RN, FAAN, PhD; Jina Huh[4], PhD

[1]Department of Media and Information, Michigan State University, East Lansing, MI, United States

[2]School of Nursing, University of Portland, Portland, OR, United States

[3]College of Nursing, Michigan State University, East Lansing, MI, United States

[4]Department of Medicine, University of California San Diego, La Jolla, CA, United States

Corresponding Author:
Shaheen Kanthawala, MA
Department of Media and Information
Michigan State University
404 Wilson Road, Room 517
East Lansing, MI, 48824
United States
Phone: 1 517 281 8044
Fax: 1 517 355 1292
Email: kanthawa@msu.edu

Abstract

Background: About 6 million people search for health information on the Internet each day in the United States. Both patients and caregivers search for information about prescribed courses of treatments, unanswered questions after a visit to their providers, or diet and exercise regimens. Past literature has indicated potential challenges around quality in health information available on the Internet. However, diverse information exists on the Internet—ranging from government-initiated webpages to personal blog pages. Yet we do not fully understand the strengths and weaknesses of different types of information available on the Internet.

Objective: The objective of this research was to investigate the strengths and challenges of various types of health information available online and to suggest what information sources best fit various question types.

Methods: We collected questions posted to and the responses they received from an online diabetes community and classified them according to Rothwell’s classification of question types (fact, policy, or value questions). We selected 60 questions (20 each of fact, policy, and value) and the replies the questions received from the community. We then searched for responses to the same questions using a search engine and recorded the

Results: Community responses answered more questions than did search results overall. Search results were most effective in answering value questions and least effective in answering policy questions. Community responses answered questions across question types at an equivalent rate, but most answered policy questions and the least answered fact questions. Value questions were most answered by community responses, but some of these answers provided by the community were incorrect. Fact question search results were the most clinically valid.

Conclusions: The Internet is a prevalent source of health information for people. The information quality people encounter online can have a large impact on them. We present what kinds of questions people ask online and the advantages and disadvantages of various information sources in getting answers to those questions. This study contributes to addressing people’s online health information needs.

(J Med Internet Res 2016;18(4):e95) doi:10.2196/jmir.5369

KEYWORDS

health communication; online health communities; question types classification; self-management; health-related Internet behavior use; risk of misinformation; Internet; diabetes

Introduction

On average, about 6 million individuals in the United States search for health information on the Internet per day [1]. This number is greater than the 2.27 million physician office visits per day [1].

People look for health information on the Internet as patients or as caregivers [2]. They look for information such as a newly prescribed course of treatment, unanswered questions after visiting providers, or information about a change in diet or exercise habits [3]. Patients consult the Internet over their providers' suggestions, or challenge their providers' suggestions based on information these patients find on the Internet [4]. Most people who find the information they were looking for believe it is of good quality and trustworthy [5]. This could prove to be potentially problematic for situations such as when health information seekers with low health literacy skills are unable to evaluate the information they find [6]. Past research has noted the need for consumer education on Internet navigation and suggested the incorporation of decision aids into health information websites [7].

The Internet provides users access to a myriad of health-related sources, such as government and professional organization websites, medical journals, mailing lists, articles, and online support groups [5]. The Internet can put a person in touch with others with similar conditions. Especially in cases where the patient has a chronic illness, where a large part of the disease management occurs at home, social media environments provide a primary resource for people to get help from peer patients [8]. According to a Pew Research Center survey [9], 23% of patients online who have chronic illnesses have used the Internet to find others with similar health conditions. These patients found getting emotional support and quick remedies from fellow patients helpful [10]. Thus, the Internet provides a variety of resources that are advantageous for one situation over another.

In this study, we compared how 2 frequently used sources of health information on the Internet—answers from other patients (eg, replies in online health communities) and Internet search results (eg, querying search engines such as Google.com)—have unique advantages and disadvantages. We discuss implications for how online health information seekers can be further supported to receive high-quality information. Depending on the types of questions health information seekers have, search engines could provide more helpful resources than would online health communities, or vice versa. Identifying the most appropriate Internet health information resource is challenging for patients. We addressed this challenge by investigating what informative sources are most appropriate for the types of questions information seekers have.

Background

In 2013, the Pew Research Center found that 8 out of 10 health-related inquiries started with a search engine [2]. About 77% of online health information seekers stated they used websites such as Google, Bing, or Yahoo. Of the survey respondents, 13% visited specialized websites such as WebMD, 2% visited more general websites such as Wikipedia, and 1% began their online health information seeking with a social network through websites such as Facebook [2].

In 1996, the National Library of Medicine had reported that 7 million of their annual Internet searches were health related [11]. In 2003, Google reported that 6.75 million of their daily search logs were health related [11]. In 2012, Google accounted for around two-thirds of US Internet searching, and this share is increasing [12]. These statistics show the exponential growth of Internet health information seeking.

The Internet provides a resource for patients who have similar health conditions to connect with one another. Connecting with other patients provides them with an outlet to share similar experiences and receive informational and emotional support [3,10,13,14]. A group of researchers studied what patients posted in online health communities to understand patients' information needs from such communities [3]. Patients needed expertise coming from clinicians as well as expertise coming from fellow patients. Some patients also posted “desperate calls for help,” such as what to do when their blood sugar monitors detected extremely high blood sugar [3,15]. As such, questions about their health concerns vary greatly. One might ask factual information about a medical condition, others' opinions about a certain recipe, or how to maneuver through a side effect.

Researchers approached taxonomy of questions in several ways based on their study context and the purpose in classifying questions [16-19]. One taxonomy of questions classified them based on the individual's domain knowledge level. For instance, a novel situation is where the person is unfamiliar with the domain. Thus, the person lacks prior knowledge of how to approach the problem. Misindexed knowledge refers to when the individual has prior knowledge but this information has not been correctly classified under the cues for a particular schema. Incorrect or incompletely understood knowledge is when previous experience and knowledge may have been incomplete or incorrect [16].

Based on tutoring transcripts, Nielsen et al [18] developed a taxonomy of questions built on whether the question was asking for a description, explanation, comparison, or preference. They then used these question types to automatically generate questions for educational assessment [18]. Ely and colleague's [20] evidence taxonomy helped to identify clinical questions requiring answers with evidence and whether the question was specific to an individual patient. Tutos and Mollá [19] applied this evidence taxonomy to identify clinical questions in a search engines context.

For questions asked in a social medium, Efron and Winget [17] developed question classification in the context of microblogging (eg, Twitter.com). They organized questions into 9 categories that address the purpose of the question being asked more than the type of question. For instance, Efron and Winget described some questions asked in microblogging as rhetorical in nature, where the questions invited action or coordinated action among the participants of a particular microblogging thread.

For our purpose of identifying the types of questions asked by patients and caregivers online, we needed a classification schema that could encompass the variety of patients' and caregivers' information-seeking needs as discussed in prior work [3]. Patients' and caregivers' questions posed online can be unstructured and incomplete [3]. The question taxonomies discussed so far were limited for our purpose for two reasons: the taxonomies covered only a subgroup of question types seen in patients' and caregivers' questions online; or they assumed that the questions were structured and well formulated.

Rothwell's classification of questions [21], primarily designed to understand questions asked in small groups, most appropriately addressed these specialized needs of classifying patients' and caregivers' questions online. Rothwell argued that questions could be phrased as *fact*, *value*, or *policy* questions. A question of fact asks whether something is true and to what extent. This question can be answered with the help of objective evidence. A question of value asks for an evaluation of the desirability of an object, idea, event, or person. Such a question cannot be answered with objective evidence, since these answers are subjective views of the responder. A question of policy asks whether a specific course of action should be undertaken to solve a problem [21].

To extract patients' questions online, we used patients' posts from online health communities. WebMD, one of the top 1000 websites worldwide, reported in 2012 [22] that they had 111.8 million unique monthly visitors out of the estimated 117.8 million unique monthly visitors to all general health-related sites [23]. This website was one of the most popular health discussion boards for patients available online with about 19.5 million visits as of December 2012 [24]. The website consists of multiple health communities where people ask questions and get responses from the community members.

Study Objective

Questions posed to these communities provide insights into the types of questions patients have about their health issues. Because of the diverse content of patients' questions, what constitutes an optimal source and content to answer those patients' questions can vary greatly. We used patients' questions posted on the WebMD diabetes community to understand how those questions can benefit from 2 main sources of health information on the Internet: a search engine versus responses from peers in online health communities. Our research questions were (1) What health information do search engines provide versus online health communities? (2) How clinically accurate is information in search engines versus that in online health communities? (3) What is the most appropriate source of health information for different question types?

Below, we describe how we operationalized these research questions.

Methods

Collection of Questions

We collected patients' questions and community responses from the WebMD online diabetes community. We chose the diabetes community over other communities because of the balanced amount of questions across various question types a diabetes context allows [3]. We wrote a script to download publicly available WebMD online diabetes community posts to a local MySQL database (version 5.6, Oracle) with a Sequel Pro interface (open source software under MIT license). Our institutional review board decided that our study did not require their regulation because the data were equivalent to public observation.

Our dataset contained 71,177 community posts between 2007 and 2014. These consisted of 9576 thread-initiating posts and 61,592 replies to those posts. The thread-initiating posts contained patients' questions, emotional support-seeking messages, or information dissemination [3]. From our prior work, we learned that thread-initiating posts with shorter lengths included more patient questions than longer posts, which tended to be rapport building. We filtered the data down to 1555 thread-initiating posts with fewer than 200 characters, that is, short posts. Next, to examine the most recent questions posted by patients, we organized the posts by posting date. We coded down the list from the most recent to older posts coding for (1) whether the post was a question and, if so, (2) which type of question it was based on our codebook.

We iteratively modified Rothwell's classification of questions to develop the following codebook:

- Questions of fact: These questions ask whether something is true and to what extent, requesting objective, factual information (eg, "What are the normal ranges for blood sugar?"; "What could it mean if you have a sweet taste in your mouth?").
- Questions of policy: These questions ask whether a specific course of action should be undertaken to solve a problem (eg, "I just got diagnosed with type 2 diabetes. What should I do next?"; "After overnight fasting I experienced a sugar spike. What should I do to bring it back down?").
- Questions of value: These questions ask for an evaluation of an idea, object, or event of a person (eg, "Has anyone experienced tingly fingers as a side effect to diabetes?"; "Can someone describe their experience with foot surgery and healing associated with this kind of surgery?").

We then selected the most recent 100 thread-initiating posts. One coder (henceforth referred to as the nonclinical coder) started to code the 100 posts for whether the post was a question, starting with the most recent. If it was a question, this coder coded the question type. The coder continued the process until we had 20 questions under each question type. As a result, we had 60 patient questions in total to investigate. Half of the questions from the total of 60 questions were randomly selected and given to another coder to measure interrater reliability between the 2 coders.

The average interrater reliability between the coders was 0.79. We then convened to resolve the disagreement. We dropped the disagreed item and recoded until we found another fact question to have an equal number of questions for each question type, which resulted in 20 fact questions, 20 value questions, and 20 policy questions.

Collection of Search Engine Results

To determine what information a search engine would retrieve, we queried each of the resulting 60 questions on Google.com.

We chose Google.com because it is the most used search engine as of 2015, according to websites that measure search engine traffic (ie, alexa.com and amazon.com) based on combined measure of page views and unique site users [25]. We used the full sentence of each question as the query to mimic how an online health information seeker asks questions through a search engine [26]. Additionally, we chose to keep the questions as they were posed to keep the sentiment of the question intact, that is, whether it was a fact, a value, or a policy question.

User searches comprise both keyword-based searches and sentence-based searches. Older search engines, such as AltaVista or AOL, were based on the model of keyword-based searches. With advancing Internet use, search engines recognized supporting sentence-based searches, especially in the context of users asking for answers to their questions. Consequently, Microsoft developed a patent on parsing searches of Frequently Asked Questions pages [27]. Hence, while keyword searches might be more prevalent for general search engine use, for the specific user needs context we are addressing—asking questions on the Internet—a sentence-based search model is more appropriate. Thus, we used the questions posted in online communities to address the types of search queries that users would post in the context of getting answers.

Research has shown that people explored only the first few results through a perfunctory examination of the search results [28]. One study used eye tracking to determine how the ranking of a link in search results affected the amount of attention it received. Their results indicated that people spent almost equal time looking at the first and second links—the viewing attention span drastically dropped from the third link onward. However, substantially fewer participants clicked the second link than the first [29]. To mimic people's practices around reading Internet search engine results, we limited our data collection to the first 3 search results for each question. We excluded search results that linked back to the WebMD online health community, because this information source would be a duplicate of our other set of data on the community responses. Searching for information through a search engine directs users to WebMD and, therefore, direct question searching on the Internet could lead individuals to information and discussions in online health communities, such as this one.

Because each question had 3 online search results, each category resulted in 60 search results. However, some searches yielded fewer than 3 results, while others resulted only in the original WebMD question and its responses. This explains the lower than 60 search results (57 results) in the fact-type questions below.

Collection of Community Responses

For each question, we returned to its original post in the WebMD community and collected the responses to these posts. The number of responses to these posts varied from 0 to 30 each. We collected all of these responses, including the responses to other responses. We did not restrict the total number of community responses per question for the analysis based on the finding that online health community users attempt to read all replies to the question unless the user deems the replies to be unrelated to the topic.

Analysis of Search Results and Community Responses

For each search result and community response, the nonclinical coder answered the following questions: (1) How complete was the answer? (2) What kind of information did it or did it not provide? Then, we organized these observations as advantages and disadvantages for each search result and community response. The second nonclinical coder followed the same analysis process for one-quarter of the search results and community answers. The 2 coders examined each other's analyses and discussed disagreements. The resulting discussion informed the first nonclinical coder's analysis for the rest of the search results and community answers. The third coder—a family nurse practitioner—answered the following questions for all community answers and search results: (1) Did the information answer the questions? (2) How clinically relevant is the information? and (3) How clinically valid is the information? The third coder was also given a space for adding open-ended, qualitative comments. This clinical coder will be referred to as clinician A from here onward.

We defined *clinically relevant* as clinical information pertinent to the question asked. Even if the information did not answer the question, if it was relevant to the question, we considered the information as clinically relevant. We defined *clinically valid* as the accuracy of the information provided to the posed question, based on the knowledge of a clinician. Even if the information did not answer the question, the information was clinically valid if it was accurate. Through discussion with clinician A, we concluded that the factors influencing this coder's decision on whether a piece of information was clinically valid included the accuracy of the information on the website, the context from which the patient was posting the question, the safety (ie, level of potential harm to the patient) of the information presented, the health literacy level of the website, and whether the website was advertising a product. The website's mode of delivery was also considered, along with whether websites required additional clicks to follow links or download videos or PDFs.

We also had a fourth coder (henceforth referred to as clinician B)—a nursing faculty member with credentials as a registered nurse and a Fellow of the American Academy of Nursing. We provided this fourth coder a random sample of 15 questions from the total 60 questions to assess agreement between the clinicians. Both clinicians followed the same method of coding search results and community findings. Once the coding was completed, we discussed disagreements and common themes among the results.

Results

In this section, we walk through the content of the collected questions for each question type. We then describe our coding results. Multimedia Appendix 1 presents sample questions for each question type, alongside search results and community responses.

Question Content Under Each Question Type

In this section, we describe what kinds of questions our dataset included under each question type. *Fact questions* asked about factual information regarding diabetes medications and their effects (7/20 questions), fluctuating blood sugar level issues (7/20), diet and exercise and their effects on diabetes (4/20), blood pressure levels (1/20), and diabetes types (eg, brittle diabetes) (1/20). *Value questions* asked about people's experiences with medication or medicinal devices (12/20), food products or diet supplements for diabetes (3/20), diabetes-related symptoms (3/20), and other illnesses such as a stomach virus or a suspicious mammogram (2/20). *Policy questions* asked what course of action should be taken regarding medication and the side effects caused by diabetes and its medication (7/20), blood sugar levels (6/20), diabetes treatments (5/20), weight loss (1/20), and diet (1/20).

Analysis of Search Results and Community Responses for the 3 Question Types

Fact Questions: Most Clinically Valid Search Results

The search results for fact questions provided a variety of information sources, including video, question-and-answer websites containing health care providers' answers to patients' questions, overviews of factual information about the requested topic, examples for information seekers to follow, and weblinks to other potential resources, including well-reputed sources such as the American Diabetes Association website. Some even led to online tests to help users determine an answer to their question (such as a prediabetes test). Multimedia Appendix 1 provides a complete example of a question along with all of its search results.

Among these search results, clinician A identified 19/57 results to be correctly answering the questions. The rest were coded as incorrectly answering the questions due to the following reasons: the information was tangentially relevant, it did not directly address the question, the answer was incorrectly phrased, it appeared on an unreliable tabloid or advertisement webpage, or the information was outdated or old. The clinician also commented that accessibility to information was a potential challenge in information provided by the search results, such as in the case of large videos that take a long time to download and highly technical resources that add complications to understanding the material.

Clinician A also found 28/57 search results to be clinically relevant and 37/57 to be clinically valid. For instance, the question "Can Insulin alter the efficacy of Coumadin therapy?" led to a search result about what Coumadin is. This page provided an overview of the generic form of the drug (warfarin) and its brand name drugs, including Coumadin. This page was coded by clinician A as clinically relevant but not clinically valid. This was because the information on the page was correct, according to the clinician, but did not answer the question.

Our analysis showed that these search results provided a varying quality of information in terms of how much the information answered the questions: some partially addressed the question, and others provided a complete and comprehensive answer, while some gave an overview about the question's topic but did not fully address it. For instance, a poster's question was "Is Byetta a non insulin medication and can it be taken with Metformin? [sic]." To this question, the top search result was a website containing information about oral and noninsulin injectable medications for diabetes. This website contained information on Byetta being a noninsulin medication but it lacked information on whether it could be consumed with metformin, which was part of what the question was asking.

Some search results did not answer the questions at all. These results included information irrelevant to the questions with large amounts of text on the page, which would be overwhelming for lay users. The search results coded as not answering questions by both the nonclinical coder and clinician A presented extraneous information such as a class study guide, commercials for products, or a response to the question but not in a diabetes context. Some search results also did not include sources or citations from which the information was derived, thereby making the validity of the information questionable.

Community responses for fact questions contained personal experiences, anecdotes, and assessment of information provided by other responders. The community responses also included reassurances and compliments to the poster, psychosocially supporting them. For instance, to the question "I am new to diabetes. I have noticed that my blood sug[a]r goes to 180–200 when I exercise. The more strenuous the exercise the higher the blood sug[a]r. Does anyone know why this happens?", a poster responded: "Exercise is a form of stress on the body. Whenever you have stress, your liver secretes sugar. Even though it goes up during exercise, you may notice it dropping low after you are done. This does happen to me too, and I make sure to drink plenty of water while I work out." This conversation describes how a community response answers a poster's question while also providing personal experience with the problem and a way to resolve it based on the responder's own method. This was the most critical difference between the search results and community responses. Another prevalent answer to posters' questions was responders asking the posters to discuss their questions with their health care providers in case it was something that needed medical intervention. The responders also assessed the accuracy of other people's responses. Responders also denied answering questions due to the liability of the community.

Clinician A determined that 35/66 community responses answered fact questions. Similar to fact question search results, those community responses coded as not answering questions included general encouragement for the poster only (for instance, "it is great that you are being proactive about your health now"), responses to conversations with other individuals involved in the post, or just responses that were clinically incorrect. Additionally, 37/66 of the community responses were both clinically relevant and clinically valid.

Fact question community responses not only advised a poster to visit a health professional, but also provided suggestions about what should be discussed during this visit, referred posters to external resources, gave compliments and reassurances about the difficult time the poster was going through, and even alerted a poster to incorrect or dangerous information provided by other responders.

Overall, community responses presented more information related to answering the question content than did the search results.

Finally, we calculated the interrater reliability of clinicians A and B using Cohen kappa score. The overall kappa for fact questions was .46, indicating a weak level of agreement between the 2 clinicians [30]. A more detailed discussion explaining this lower interrater reliability is at the end of the Results section.

Policy Questions: Most Answered Through Community Responses, Least Answered Through Search Results

Search results of policy questions included diabetes management pages, blogs, stepwise instructions on how to solve a problem, and patient forums. However, not all answers were available within an article reached through a search result—often, the information was present (fully or tangentially) in the discussion or comments section of the page. Because these were how-to questions, the search results also led to different social media platforms such as Facebook, where other similar questions fully or partially answered the poster's question. For instance, the question "hi i just checked my blood sugar and its 490 how can i get it down my vision is blurry? [sic]" leads to a Facebook type 1 diabetes page post discussing eyesight fluctuation for a 25-year-old with newly diagnosed type 1 diabetes. This post is not an exact answer to the poster's query, but it partially explores the poster's questions. Furthermore, the comments the Facebook post received could help the poster get his or her answer. Some results did not answer the questions asked.

Our coding results indicated that a very low number of policy questions were answered through search results. Among the search results of policy questions, clinician A identified that 8/60 results correctly answered the questions. This coder also found that 8/60 search results were clinically relevant but 24/60 were clinically valid.

A search engine does not directly answer "what or how should I do [something]?" questions. For instance, a person asking about how to lower blood sugar first thing in the morning, due to high early-morning numbers, is led to a page describing the "dawn phenomenon"—a condition many diabetics experience wherein their morning sugar levels are higher than usual [31]. This helps patients with the information, but would not necessarily answer their question of what to do to remedy it.

Other search results included discussion pages related to the question posed, a stepwise to-do response to the question (such as ways to lose baby weight), external resources, advice on what the poster can do next (for instance, the next steps of having diabetes diagnosed), and access to social media pages displaying similar situations to the poster's.

Community responses to policy questions provided personal experiences with similar problems and gave posters insight on how to deal with their problems. These responses also prompted posters to think about other questions related to the situation, advised them to visit their health care provider, provided additional resources, reassured posters that they were not alone, and made them aware of potential dangers. For instance, to the question "is anyone on that can tell me how to lower my blood sugar. What can I eat right now to lower it. I had to many carb's and it's 202 usually it's 90 to 101 [sic]", a response provided stated "Hi, The only way to get your BG lowered from a spike is to exercise and drink a lot of water. There are no foods to bring your glucose level down. At least this is my meager experience and if someone else has a better idea I hope they share with you. Good luck and watch those carbs [sic]."

Clinician A observed that 30/49 community responses answered the questions posed, 30/49 responses were clinically relevant, and 31/49 responses were clinically valid. This imbalance between clinical relevance and clinical validity occurred because 1 poster posed a question about having a blood sugar level of 490 and asked how to bring it down so as to get rid of blurry vision. A responder suggested this person should wash his or her hands and try again and, if the sugar reading stayed as high, should call 911 or head to the emergency room immediately. This post was marked clinically valid because all the information it provided was clinically accurate; however, it was not marked as clinically relevant, because the information was not relevant to the actual question asked and did not help answer it (ie, how to get rid of the blurry vision).

Other community responses to policy questions stressed the dangers of a situation a poster may be in; provided potential solutions to help solve a problem, including suggesting home remedies and advising the poster to visit a health professional; offered personal stories, anecdotes, and opinions; and helped detect emergencies from the situation presented by the poster.

Finally, we calculated the interrater reliability of clinicians A and B using Cohen kappa score. The overall kappa for policy questions was .52, indicating weak agreement between the clinicians [30].

Value Questions: Community Responses With Mixed Quality Assessment by Clinicians

Search results for value questions included personal experiences and, therefore, a lot of the search results only partially answered the question or provided an overview of the subject. Search results included products and their reviews, patients' stories through blog posts, and discussion forums. One result helped with alleviating nervousness, while another showed how others also had similar symptoms. A portion of the questions asked about people's experiences with medicine or health issues (as elaborated in detail above); therefore, the search results led to online stores and reviews for these products, such as Amazon, where other customers provided reviews for the health information seeker to review. Lastly, some questions were answered, but not necessarily in a diabetes context. For instance, a poster's question was "I had my first [mammogram] last week and it came back suspicious. Had to go for more [picture] and an ultrasound and now for a biopsy. I am way past scared to death. Can anyone help me [sic]?" the first search result to this question led to a webpage discussion about mammograms in patients. This information provided encouragement to the poster as requested.

Among the search results of value questions, clinician A identified that 23/60 results correctly answered the questions. This clinician also found that 23/60 search results were clinically relevant and 29/60 were clinically valid. These results indicate that, while some questions were answered, still others were clinically relevant and clinically valid, but did not answer the question.

Other search results for value questions included a comprehensive overview of the subject in question; discussed side effects from credible sources; had comment sections on some webpages discussing the subject; helped differentiate between myths and facts related to the question; led to question-and-answer pages that help answer the question posed; resulted in product reviews for questions about specific products, thereby informing the poster about the quality and effectiveness of the product; included personal experience stories and encouragement, which reduced posters' nervousness; and included some results that were also backed up with statistical evidence.

Community responses, on the other hand, included details about things the question poser should be cautious about, provided side effects of medication, tips, and suggestions, conducted online searches, and found answers for the poster.

People are enthusiastic about providing their opinion. The downside of this is that there is no way to verify the answers provided. For instance, a poster posed the question "Have any of you had the Bayer Contour meter just readout "HIGH"? No numeric reading just "HIGH". I suspect that is a very bad sign [sic]." A response to this was, "Most meters read up to about 500 or 600. Anything higher than that and it simply greets you and says "HI". Did you test again after washing your hands? If you did call your doc and see someone immediately. This is not good [sic]."

For community responses, clinician A noted that 63/104 responses provided answered the questions posed. Additionally, 51/104 were clinically relevant and 47/104 were clinically valid.

The types of community responses received by a poster for value questions helped them be wary about new trends; informed them about medicinal side effects and provided insights about how this information was obtained; alerted people to any potential dangers; provided personal experiences, opinions and anecdotes, and tips and solutions to help resolve issues; advised the questioner to visit a health professional; corrected or clarified misinformation; redirected the poster to more resources and information; and reassured and encouraged the poster about his or her current situation.

Finally, we calculated the interrater reliability of clinicians A and B using Cohen kappa score. The overall kappa for fact questions was .46, indicating a weak agreement between the 2 clinicians [30].

Clinician Interrater Reliability

As can be noted, the Cohen kappa score between both clinicians was not very high, signifying weak interrater reliability. Some of the difference in coding between them can be explained through the following reasons. First, we found many of the search results on webpages containing large amounts of text, thereby making the process of locating the response on the page a difficult one. This sometimes led to 1 clinician coding the search result as answering the question posted, whereas the other did not. This observation depicts how different people, including clinicians, interpret different kinds of search results in terms of whether they answer the question asked. If both coders coded a question differently, the subsequent questions about clinical relevancy and clinical validity also tended to follow different paths. Sometimes, in a situation like this, the search result was coded as answering the question, but the validity or relevance of the answer was queried, therefore leading to different codes between the 2 clinicians. This variability in the results between the 2 clinicians speaks to the complexity of the problem, that is, the difficulty of defining and assessing the quality of information on the Internet.

Finally, the 2 coders interpreted the accuracy and relevance of some responses differently, therefore leading to lower kappa scores. This result points to the importance of the way information is shared and interpreted on the Internet and how better guidance and direction for gathering information is necessary. This study contributes to understanding the various factors clinicians consider and how these factors lead to their evaluations.

Table 1 summarizes the advantages and disadvantages of both sources. These are the characteristics of the overall findings of each category—every point does not apply to every finding. This is followed by an overview of clinician A's findings in Table 2.

Table 1. Summary of advantages and disadvantages of search results versus community responses to questions posted to an online diabetes community.

Type of question	Advantages	Disadvantages
Search results		
Fact		
	Provides some answers to the questions.	Does not always provide an answer to questions posed.
	Provides an overview of the subject.	If question is answered, it could be only a partial response.
	Multiple websites provide a wide range of information.	Can provide irrelevant responses.
	External links can route health information seekers to various resources.	Can answer out of the context it is posed in (ie, out of a diabetes context).
	Responses from reputed websites, such as the American Diabetes Association, can be assumed to be accurate.	Accuracy of responses is not always known.
	Provides a test to help posters determine their answer (eg, a prediabetes test).	Websites can contain large amount of content, thereby preventing easy location of response.
		Some results are commercials for products, leading to biased information.
		Does not answer unusual questions.
Policy		
	Gives access to the discussion and conversation pages related to the question posed.	Some results do not answer the question directly, leaving the question poser to make extrapolations.
	Provides step-by-step responses to the questions posed (eg, ways to lose baby weight).	Some results do not provide required answers, ie, are irrelevant.
	Provides external resources to relevant information.	Some results answer questions partially or tangentially.
	Provides the next steps for poster (eg, next steps of having a diabetes diagnosis).	Websites can contain a large amount of content, thereby preventing easy location of response.
	Provides access to social media results, such as Facebook, showcasing similar cases.	Some results are commercials, leading to biased information.
Value		
	Some responses answer questions precisely.	Some results do not provide required answers, ie, are irrelevant.
	Some results provide a good overview of the question topic.	Some results answer questions partially or tangentially.
	Some responses discuss side effects from credible sources.	Website can be very large and contain a lot of content, thereby preventing easy location of response.
	The comments sections of webpages help discuss the subject.	Some results are commercials, leading to biased information.
	Answers differentiate between myths and facts of the subject.	
	Question-and-answer pages help answer poster's question.	
	Some pages lead to product reviews that help answer the question.	
	User experiences and encouragement on different result pages help alleviate poster's nervousness.	
	Some results back up claims through statistical evidence.	
Community responses		
Fact		
	Provides personal experiences, opinions, and anecdotes.	Some questions do not get responses.
	Advises poster to visit a health professional.	Cannot check accuracy of responses.
	Provides examples and external resources.	Question may be deferred to a health professional, thereby delaying response.
	Provides compliments and reassurances for the difficult time the poster is going through.	Some questions are answered only partially.
	Alerts poster of potential dangers (including those from other people's responses).	Cannot answer due to liability of the forum.

Type of question	Advantages	Disadvantages
	Redirects to a person or resource with more information.	Does not always provide a complete or relevant response.
	Provides alternative options, external resources, and potential talking points to discuss with one’s health care professional.	Some responses are irrelevant or potentially dangerous.
Policy		
	Responses stress the dangers of the situation.	Some questions do not get responses.
	Provides tips or solutions to resolve issue.	No way to check accuracy of responses.
	Provides personal experiences, opinions, and anecdotes.	Some questions are answered only partially.
	Advises poster to visit a health care professional.	Some responses are not in line with other responders.
	Provides home remedies.	
	Detects emergency cases.	
Value		
	Some results help posters be wary of latest trends.	Some questions do not get responses or are irrelevant.
	Some results provide effects and side effects along with insights about how this information was found.	
	Alerts poster to potential dangers.	No way to check accuracy of responses.
	Provides personal experiences, opinions, and anecdotes.	Some questions are answered only partially.
	Provides tips or solutions to resolve issue.	
	Advises poster to visit a health care professional.	
	Responses help clear misinformation for poster.	
	Redirects to a person or resource with more information.	
	Provides reassurances and encouragement to poster.	

Table 2. Evaluation of all community answers and search results by Clinician A.

Type of question	Answers the question		Is clinically relevant		Is clinically valid	
	No.	%	No.	%	No.	%
Fact						
Search results (n=57)	19	33	28	49	37	65
Community responses (n=66)	35	53	37	56	37	56
Policy						
Search results (n=60)	8	13	8	13	24	40
Community responses (n=49)	30	61	30	61	31	63
Value						
Search results (n=60)	23	38	23	38	29	48
Community responses (n=104)	63	61	51	49	47	45

Table 2 breaks down the questions according to clinician A’s analyses. This table lists the proportion of positive responses to the selected questions in each category, that is, whether the results or responses answered the questions, were clinically relevant, and were clinically valid. For example, 19 of the 57 observed search results for fact-type questions answered the question posed. The remaining results either did not answer the question posed or contained unknown or not applicable information, such as webpages that would not open or community posts with no answers.

For example, the question “recently lost a kidney/ureter to cancer. sugar levels are moderately changed and i need to know the actual effects i can expect as my body adjusts to life with only one kidney [sic].” leads to a human anatomy textbook online, which clinician A deemed to have an “unknown” value.

The community responses, on the other hand, varied in number, since we included every answer to a question posed. The total numbers indicated above are the number of responses each category of community responses received. However, there were questions with no responses at all, such as the question “Help. I just gave myself my pre dinner shot in my leg. It formed a huge bump on my leg. Will the insulin still get in me or do I need to take my shot again???[sic]" had no responses at all and therefore clinician A marked it as "not applicable."

Discussion

Principle Findings

Our analyses showed that community responses answered more questions than did search results across the board. These community responses were more clinically relevant than their search result counterparts. However, clinical validity varied, with search results being more clinically valid than community responses for fact and value questions, but community responses being more clinically valid than search results for policy questions. These observations show that answers from both sources contain clinically accurate information, which does not necessarily answer questions.

Search results for fact questions showed that results were often clinically valid but not clinically relevant. Fact question search results had the highest validity among all question types. Even if search results were clinically relevant, only a small portion of those results completely addressed the questions in the query. Questions left unanswered can be attributed to imprecise question phrasing. Wording questions so that their meaning is clear and concise will lead to more relevant search results than will questions with meandering and unclear content. We need mechanisms to help patients formulate queries and questions online.

We learned from the search results of policy questions that nearly half of the search results to policy questions contained clinically valid information. However, most of this information was irrelevant to answering the questions and the information was incomplete. A reason for this finding could be the way policy questions are framed. These questions require information about how or what needs to be done in a particular situation. Factoring in the multiple variables in an individual's question would result in varied search results, none of which answer the question as required.

Analysis of value question search results showed that search results often did not answer the question posed, nor was the information clinically relevant. However, a greater number of results were clinically valid, showing that accurate information does not always lead to answers. This observation is in line with fact and policy question search results, where a larger number of results were clinically valid but did not answer the question posed.

We observed a similar pattern in the community responses to fact questions, where a greater number of questions were clinically valid and clinically relevant, but fewer questions were answered.

For community responses to policy questions, clinician A observed that a much higher number of questions were answered as compared with policy question search results. This shows how the specific nature of policy questions makes them a better fit for an audience who are familiar with the issues of the community, thereby providing posters with the information required.

Community responses to value questions were opposite to fact and policy community responses. Value questions were the only question type to which the community provided a greater number of answers, but a smaller fraction of the information provided was accurate. In comparison, fact and policy questions tended to get superfluous but clinically valid responses, leading to more unanswered questions. This finding is important because it portrays the risks of unchecked information being exchanged in online health communities. Value questions ask others for their personal experiences and evaluations, which is a warning to posters about the unmoderated nature of the information.

The combined analyses of the coders and the clinicians indicated that policy and value questions get more community responses than do fact questions. This observation could be because these 2 categories ask responders to provide their personal experiences with a situation, and knowing an answer to a question or not (as with fact questions) is not a criterion.

Community responses also make question posters feel better about their situation and remind them that they are not alone through reassurances and compliments. The responses warn question posters about potentially dangerous actions or incorrect advice they get from other responders.

We also observed that community responses consistently got a higher percentage of questions answered. as opposed to a search results. Therefore, while past research indicates that people go online and search for health-related information, these individuals are more likely to get their questions answered through online social support groups of people in similar situations to their own.

Recommendations

Based on the advantages and disadvantages of the search results and community responses (Table1), there are potential recommendations for stakeholders involved.

Patients must be vigilant about the information they find through search results by keeping track of the sources of websites and the validity of the information provided. Formation of the question asked in search results also plays a role in the kinds of responses it gets; therefore, posing a clear question while searching for results on the Internet is important. Community responses are provided by similar others, that is, other individuals facing health situations similar to the question poster's. While this indicates a familiarity with a question posed or a situation described, it is important to verify medically related information or steps to be taken with one's health care provider so as to prevent negative health consequences.

Guiding patients toward accurate information obtained through search results helps providers by not having to correct misconceptions patients build through information they gather via search results. It would be helpful to patients for providers to guide them in searching for information online and in determining whether information is trustworthy.

Online health communities play a critical role in providing social support to people going through similar health issues.

Self-management of health conditions plays an important role in management of chronic illnesses [32], such as diabetes. Introducing patients to such a resource and encouraging them to use it is helpful, for both the patient and the provider. Patients can build a network of resources additional to their providers—a support system that is available to them, preventing them from feeling isolated and being frightened of their situation. Providers can get additional time that can be allocated to patients in greater need or to themselves.

Researchers should compare different online communities to analyze information sharing for a variety of health conditions. Such health information can help technology developers create more efficient health communities, with more opportunities and resources for patients participating in them. One important finding was the sharing of incorrect or dangerous information by other participants in an online community. Developers could create a way to tag dangerous posts, based on feedback provided by other users, which would require moderators of the community to evaluate such posts and rectify the information provided as needed. Classification of questions based on our codebook may help researchers and developers in the future to tag questions needing professional evaluation. Answers to policy questions provide direction to individuals in need, and value questions give personal evaluation, both of which we have observed to have incorrect information. In future studies, researchers should also focus on the best way to formulate questions to gain the most accurate information through search results. Additionally, it is important to develop a way to help patients analyze whether the information they gather through Internet search results is accurate. Patients look for health information online in large numbers. Therefore, ensuring the accuracy of this information is crucial.

Informaticists should analyze the way information is shared in online health communities, especially through the relations between participants within these communities. Participants roles in these communities provide insights into how relationships form and how these relationships lead to an exchange of information. These insights include the kinds of information they share and receive, and whether it is clinical or social in nature. Additionally, informaticists should look into the questions posted, both those searched for on the Internet and those posted in online communities. Such an analysis would prove valuable to determine which questions would be ideally suited to be answer through a search engine versus those ideally posted in an online community (ie, which would provide the most accurate and complete information to a specific question).

Finally, the trouble faced by both of our clinicians in interpreting similar information differently is a commentary on how difficult it is to find and assess health information online. This observation is important so future websites can address this wide-ranging quality issue.

Limitations

As mentioned above, because the questions posted in the search engine came from a diabetes health community, not all of the questions mention their diabetes context. Community members assume this diabetes context. For example, an evaluation of a particular diet would not ask members to keep in mind that the diet was for a person with diabetes because it was posed in a diabetes community. Because we decided to search unaltered questions in the search engine, the answers we got could have been less efficient (answers not in a diabetes context) than if we had modified the questions to include this parameter. Future studies could make the context explicit to study the differences in answers it would produce.

Searching complete questions could misguide a search engine. Future research can focus on alternatives such as carrying out keyword searches with pertinent information from the question as opposed to using the question as an exact-phrase search.

Additionally, more than 3 search results can be included in the search result answers. While most people do not venture beyond the first page of search engine results, and even there they focus on simply the first few search results [28,29], adding additional search links will give a comprehensive insight into the kind of responses available.

Finally, Google’s page ranking method ranks high-quality websites higher than other websites with a lower level of authority in the related topic (eg, the total number of incoming links from government institutions such as the US Centers for Disease Control and Prevention). As a result, our findings on the high accuracy of online health information in the first 3 search results are biased toward what Google already ranked as having high authority in the topic. However, most searchers check the first 2 search results, thereby correcting this imbalance [29].

Conclusion

We evaluated the responses people get to health information they seek online from 2 different avenues: search engine results and online community responses. Our findings indicate how question types matter for determining information quality and sources. Health care practitioners, informatics researchers, and policy makers should consider the strengths and weaknesses of each information source based on the types of questions information seekers have. Our study contributes to improving online health information quality, making self-management of health more efficient and lowering costs for medical professionals and patients.

Acknowledgments

This work has been partially funded by NIH grant 1 K01 LM011980-01.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Examples of questions of each category and their community responses.

[PDF File (Adobe PDF File), 21KB - jmir_v18i4e95_app1.pdf]

References

1. Fox S, Rainie L. Vital decisions: how Internet users decide what information to trust when they or their loved ones are sick. Washington, DC: Pew Internet & American Life Project; 2002 May 22. URL: http://www.pewinternet.org/files/old-media//Files/Reports/2002/PIP_Vital_Decisions_May2002.pdf [accessed 2015-11-24] [WebCite Cache ID 6dHz6H2kj]
2. Fox S, Duggan M. Health online 2013. Washinton, DC: Pew Internet & American Life Project; 2013 Jan 15. URL: http://www.pewinternet.org/~/media/Files/Reports/PIP_HealthOnline.pdf [accessed 2015-11-20] [WebCite Cache ID 6dAt8Zj54]
3. Huh J, McDonald DW, Hartzler A, Pratt W. Patient moderator interaction in online health communities. AMIA Annu Symp Proc 2013;2013:627-636 [FREE Full text] [Medline: 24551364]
4. Ahmad F, Hudak PL, Bercovitz K, Hollenberg E, Levinson W. Are physicians ready for patients with Internet-based health information? J Med Internet Res 2006;8(3):e22 [FREE Full text] [doi: 10.2196/jmir.8.3.e22] [Medline: 17032638]
5. Morahan-Martin JM. How internet users find, evaluate, and use online health information: a cross-cultural review. Cyberpsychol Behav 2004 Oct;7(5):497-510. [Medline: 15667044]
6. Diviani N, van den Putte B, Giani S, van Weert JC. Low health literacy and evaluation of online health information: a systematic review of the literature. J Med Internet Res 2015;17(5):e112 [FREE Full text] [doi: 10.2196/jmir.4018] [Medline: 25953147]
7. Perez SL, Paterniti DA, Wilson M, Bell RA, Chan MS, Villareal CC, et al. Characterizing the processes for navigating Internet health information using real-time observations: a mixed-methods approach. J Med Internet Res 2015;17(7):e173 [FREE Full text] [doi: 10.2196/jmir.3945] [Medline: 26194787]
8. Moorhead SA, Hazlett DE, Harrison L, Carroll JK, Irwin A, Hoving C. A new dimension of health care: systematic review of the uses, benefits, and limitations of social media for health communication. J Med Internet Res 2013;15(4):e85 [FREE Full text] [doi: 10.2196/jmir.1933] [Medline: 23615206]
9. Fox S. Peer-to-peer healthcare. Washington, DC: Pew Internet & American Life Project; 2011 Feb 28. URL: http://www.pewinternet.org/files/old-media//Files/Reports/2011/Pew_P2PHealthcare_2011.pdf [accessed 2016-04-20] [WebCite Cache ID 6gumaIyUR]
10. Huh J, Yetisgen-Yildiz M, Pratt W. Text classification for assisting moderators in online health communities. J Biomed Inform 2013 Dec;46(6):998-1005 [FREE Full text] [doi: 10.1016/j.jbi.2013.08.011] [Medline: 24025513]
11. Eysenbach G, Kohler C. What is the prevalence of health-related searches on the World Wide Web? Qualitative and quantitative analysis of search engine queries on the internet. AMIA Annu Symp Proc 2003:225-229 [FREE Full text] [Medline: 14728167]
12. Wang L, Wang J, Wang M, Li Y, Liang Y, Xu D. Using Internet search engines to obtain medical information: a comparative study. J Med Internet Res 2012;14(3):e74 [FREE Full text] [doi: 10.2196/jmir.1943] [Medline: 22672889]
13. Huh J, Ackerman MS. Collaborative help in chronic disease management: supporting individualized problems. CSCW 2012;2012:853-862 [FREE Full text] [doi: 10.1145/2145204.2145331] [Medline: 25360442]
14. Huh J. Clinical questions in online health communities: the case of "See your doctor" threads. CSCW 2015;2015:1488-1499 [FREE Full text] [doi: 10.1145/2675133.2675259] [Medline: 26146665]
15. Hartzler A, Pratt W. Managing the personal side of health: how patient expertise differs from the expertise of clinicians. J Med Internet Res 2011;13(3):e62 [FREE Full text] [doi: 10.2196/jmir.1728] [Medline: 21846635]
16. Ram A. A theory of questions and question asking. J Learning Sci 1991;1(3-4):273-318. [doi: 10.1207/s15327809jls0103&4_2]
17. Efron M, Winget M. Questions are content: a taxonomy of questions in a microblogging environment. Proc Am Soc Info Sci Technol 2011 Feb 03;47(1):1-10. [doi: 10.1002/meet.14504701208]
18. Nielsen RD, Buckingham J, Knoll G, Marsh B, Palen L. Proceedings of the Workshop on the Question Generation Shared Task and Evaluation Challenge. 2008 Sep. A Taxonomy of Questions for Question Generation URL: http://www.rodneynielsen.com/papers/nielsenr_QG08_A_taxonomy_of_questions.pdf [accessed 2015-11-21] [WebCite Cache ID 6dCCpt8q6]
19. Tutos A, Mollá D. A study on the use of search engines for answering clinical questions. 2010 Presented at: Fourth Australasian Workshop on Health Informatics and Knowledge Management; January 2010; Brisbane, Australia p. 61-68 URL: http://crpit.com/confpapers/CRPITV108Tutos.pdf [WebCite Cache ID 6gxfdYnuh]
20. Ely JW, Osheroff JA, Gorman PN, Ebell MH, Chambliss ML, Pifer EA, et al. A taxonomy of generic clinical questions: classification study. BMJ 2000 Aug 12;321(7258):429-432 [FREE Full text] [Medline: 10938054]
21. Rothwell JD. In Mixed Company: Communicating in Small Groups. Boston, MA: Wadsworth; 2010.
22. WebMD Announces Fourth Quarter and Year End Financial Results. New York, NY: PR Newswire; 2015 Feb 24. URL: http://www.prnewswire.com/news-releases/webmd-announces-fourth-quarter-and-year-end-financial-results-300040633.html [accessed 2016-04-20] [WebCite Cache ID 6gun22V0t]

23. Lauckner C, Hsieh G. The presentation of health-related search results and its impact on negative emotional outcomes. 2015 Presented at: SIGCHI Conference on Human Factors in Computing Systems; April 27-May 2, 2013; Paris, France p. 333-342 URL: http://faculty.washington.edu/garyhs/docs/lauckner-chi2013-health.pdf [WebCite Cache ID 6gxfremJq] [doi: 10.1145/2470654.2470702]
24. Lefebvre RC, Bornkessel AS. Digital social networks and health. Circulation 2013 Apr 30;127(17):1829-1836 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.112.000897] [Medline: 23630086]
25. Site Overview: Google.com. San Francisco, CA: Alexa Internet, Inc; 2015. URL: http://www.alexa.com/siteinfo/https://www.google.com [accessed 2015-11-21] [WebCite Cache ID 6dCDHFI8D]
26. Ozmutlu S, Ozmutlu HC, Spink A. Are people asking questions of general Web search engines? Online Inf Rev 2003 Dec;27(6):396-406. [doi: 10.1108/14684520310510037]
27. Wang HF, Lee KF, Yang Q. Search Engine with Natural Language-Based Robust Parsing of User Query and Relevance Feedback Learning. US patent 0040243568 A1. 2004 Dec 02. URL: http://www.google.com/patents/US20040243568 [accessed 2016-04-18] [WebCite Cache ID 6gs1bCJNc]
28. Eysenbach G, Köhler C. How do consumers search for and appraise health information on the world wide web? Qualitative study using focus groups, usability tests, and in-depth interviews. BMJ 2002 Mar 9;324(7337):573-577 [FREE Full text] [Medline: 11884321]
29. Granka LA, Joachims T, Gay G. Eye-tracking analysis of user behavior in WWW search. 2004 Jul Presented at: 27th annual international ACM SIGIR conference on Research and development in information retrieval; July 2004; Sheffield, UK p. 478-479. [doi: 10.1145/1008992.1009079]
30. McHugh ML. Interrater reliability: the kappa statistic. Biochem Med (Zagreb) 2012;22(3):276-282 [FREE Full text] [Medline: 23092060]
31. Monnier L, Colette C, Dejager S, Owens D. The dawn phenomenon in type 2 diabetes: how to assess it in clinical practice? Diabetes Metab 2015 Apr;41(2):132-137. [doi: 10.1016/j.diabet.2014.10.002] [Medline: 25457475]
32. Lorig KR, Sobel DS, Stewart AL, Brown BW, Bandura A, Ritter P, et al. Evidence suggesting that a chronic disease self-management program can improve health status while reducing hospitalization: a randomized trial. Med Care 1999 Jan;37(1):5-14. [Medline: 10413387]

Edited by G Eysenbach; submitted 30.11.15; peer-reviewed by N Diviani, D Xu, A Irwin; comments to author 03.01.16; revised version received 19.02.16; accepted 18.03.16; published 28.04.16.

Please cite as:
Kanthawala S, Vermeesch A, Given B, Huh J
Answers to Health Questions: Internet Search Results Versus Online Health Community Responses
J Med Internet Res 2016;18(4):e95
URL: http://www.jmir.org/2016/4/e95/
doi:10.2196/jmir.5369
PMID:27125622

Original Paper

A Web-Based Computer-Tailored Alcohol Prevention Program for Adolescents: Cost-Effectiveness and Intersectoral Costs and Benefits

Ruben MWA Drost[1*], MSc; Aggie TG Paulus[1*], PhD; Astrid F Jander[2*], PhD; Liesbeth Mercken[2*], PhD; Hein de Vries[2*], PhD; Dirk Ruwaard[1*], MD, PhD; Silvia MAA Evers[1,3*], LLM, PhD

[1]Department of Health Services Research, School for Public Health and Primary Care (CAPHRI), Faculty of Health, Medicine and Life Sciences, Maastricht University, Maastricht, Netherlands

[2]Department of Health Promotion, School for Public Health and Primary Care (CAPHRI), Faculty of Health, Medicine and Life Sciences, Maastricht University, Maastricht, Netherlands

[3]Trimbos, Netherlands Institute of Mental Health and Addiction, Utrecht, Netherlands

[*]all authors contributed equally

Corresponding Author:
Ruben MWA Drost, MSc
Department of Health Services Research
School for Public Health and Primary Care (CAPHRI)
Faculty of Health, Medicine and Life Sciences, Maastricht University
Duboisdomein 30
Maastricht, 6229 GT
Netherlands
Phone: 31 433881729
Fax: 31 433884162
Email: r.drost@maastrichtuniversity.nl

Abstract

Background: Preventing excessive alcohol use among adolescents is important not only to foster individual and public health, but also to reduce alcohol-related costs inside and outside the health care sector. Computer tailoring can be both effective and cost-effective for working with many lifestyle behaviors, yet the available information on the cost-effectiveness of computer tailoring for reducing alcohol use by adolescents is limited as is information on the costs and benefits pertaining to sectors outside the health care sector, also known as intersectoral costs and benefits (ICBs).

Objective: The aim was to assess the cost-effectiveness of a Web-based computer-tailored intervention for reducing alcohol use and binge drinking by adolescents from a health care perspective (excluding ICBs) and from a societal perspective (including ICBs).

Methods: Data used were from the Alcoholic Alert study, a cluster randomized controlled trial with randomization at the level of schools into two conditions. Participants either played a game with tailored feedback on alcohol awareness after the baseline assessment (intervention condition) or received care as usual (CAU), meaning that they had the opportunity to play the game subsequent to the final measurement (waiting list control condition). Data were recorded at baseline (T0=January/February 2014) and after 4 months (T1=May/June 2014) and were used to calculate incremental cost-effectiveness ratios (ICERs), both from a health care perspective and a societal perspective. Stochastic uncertainty in the data was dealt with by using nonparametric bootstraps (5000 simulated replications). Additional sensitivity analyses were conducted based on excluding cost outliers. Subgroup cost-effectiveness analyses were conducted based on several background variables, including gender, age, educational level, religion, and ethnicity.

Results: From both the health care perspective and the societal perspective for both outcome measures, the intervention was more costly and more effective in comparison with CAU. ICERs differed for both perspectives, namely €40 and €79 from the health care perspective to €62 and €144 for the societal perspective per incremental reduction of one glass of alcohol per week and one binge drinking occasion per 30 days, respectively. Subgroup analyses showed, from both perspectives and for both outcome measures, that the intervention was cost-effective for older adolescents (aged 17-19 years) and those at a lower educational level and, from a health care perspective, the male and nonreligious adolescent subgroups.

Conclusions: Computer-tailored feedback could be a cost-effective way to target alcohol use and binge drinking among adolescents. Including ICBs in the economic evaluation had an impact on the cost-effectiveness results of the analysis. It could be worthwhile to aim the intervention specifically at specific subgroups.

Trial Registration: Nederlands Trial Register: NTR4048; http://www.trialregister.nl/trialreg/admin/rctview.asp?TC=4048 (Archived by Webcite at http://www.webcitation.org/6c7omN8wG)

(J Med Internet Res 2016;18(4):e93) doi:10.2196/jmir.5223

KEYWORDS

adolescents; alcohol use; cluster randomized controlled trial; game; computer tailoring; education; criminal justice; costs and cost analysis; economic evaluation; intersectoral costs and benefits

Introduction

Excessive alcohol use and alcohol use disorders are major causes of death and disability worldwide [1-4]. In 2012, approximately 3.3 million deaths, or 5.9% of all global deaths, were attributable to alcohol use [4]. In addition, alcohol use led to an estimated total of 139 million disability-adjusted life years, representing 5.1% of the global burden of disease and injury in that year [4]. For all age groups, outliers on the proportion of alcohol-related deaths can be seen in the World Health Organization European Region, varying from 10% for the population aged 80 years and older to 25% for the 20 to 39 age group. It is particularly striking that in this region 10% of adolescent deaths (those aged 15 to 19 years) were attributable to alcohol.

Apart from the impact of alcohol use on morbidity and mortality, the harmful use of alcohol may also lead to significant societal costs [3-7]. For example, in the European Union alone, alcohol-attributable costs were estimated at €125 billion in 2003 [4]. These encompass health care services, such as hospitalizations, home health care, and ambulatory care, but also costs outside the health care sector, such as costs resulting from productivity losses and costs in the criminal justice system. Examples of the latter include vehicle crashes, increased crime, and arrests. Studies have shown that youthful drinkers are at greater risk of being victimized and perpetrating youth violence, low educational attainment, and low college expectations [8,9], putting a financial burden on the criminal justice system and educational sector.

Preventing excessive alcohol use in the whole population and in the young population in particular is thus important, not only to improve the health of individuals and of the whole population, but also to reduce alcohol-related costs inside and outside the health care sector. Computer tailoring could be a means to achieve these goals. Computer tailoring is a behavioral intervention that can be effective in changing health behaviors in general, including the use of alcohol [10-15]. Within a computer-tailored intervention, the content is adapted to individual characteristics of respondents [16]. Often a questionnaire is used as a screening instrument, assessing behavior, relevant sociodemographics, and motivational factors [17,18]. Respondent answers are collected into a data file and automatically matched with tailored feedback messages [16,19]. An advantage of tailored information is that it is perceived as more relevant than nontailored information [20,21]. Moreover, through the Internet, these programs are accessible by a growing percentage of the world population—42.3% had access to Internet in 2014 in comparison to 5.9% in 2000 [22]—and can be accessed wherever and whenever it suits the respondents. For Europe and its younger population (aged 16-24 years), the current Internet penetration is even higher: 70.5% [22] and 91.0% [23], respectively. Taking advantage of this high accessibility by providing Internet-based behavior interventions might significantly limit the need for and burden on health professionals and could, in turn, lead to less administrative costs [19]. Subsequently, the resulting savings of time and resources contribute to a more efficient health care system.

Although the effectiveness of multiple Internet-based and computer-tailored interventions have been studied [24,25], limited information exists on their cost-effectiveness [26]. To date, studies which have included cost-effectiveness analyses (CEAs) focused on adults aged 18 to 65 years [21,27] and 18 to 69 years [3], whereas the cost-effectiveness of such interventions for adolescents younger than 18 years has not yet been studied. Available economic analyses of interventions aimed at reducing alcohol use among adolescents are those of interventions which are not Internet-based and not computer-tailored [28,29]. Furthermore, as acknowledged by Smit et al [3], these studies show a limitation in terms of disregarding many of the costs and benefits in sectors outside the health care sector, also known as intersectoral costs and benefits (ICBs) [30]. Excluding ICBs, such as costs and benefits in the educational and criminal justice sector, may significantly affect the results of the CEAs of interventions [30]. Moreover, including and reporting ICBs within economic evaluations could support decision making regarding the large-scale implementation of such programs. Therefore, the aim of this study is to answer the question of whether a Web-based computer-tailored intervention for adolescents for reducing the use of alcohol is cost-effective from both a health care and societal perspective, and to assess the impact of including ICBs on the outcomes of the analysis.

Methods

Design

Data used was from the Alcoholic Alert study, which was designed as a cluster randomized controlled trial (RCT) with randomization at the level of schools into two conditions [31]. Participants either played a game on alcohol awareness after the baseline assessment (intervention condition) or received care as usual (CAU), meaning they had the opportunity to play the game subsequent to the final measurement (waiting list control condition). Providing this opportunity was due to ethical considerations. Data were recorded at baseline (T0=January/February 2014) and after 4 months (T1=May/June 2014). These were used to conduct two sets of comparative CEAs; one was performed from a health care perspective (including health care costs, excluding ICBs) and one from a societal perspective (including both health care costs and ICBs). Because of a rapid change in Dutch government policy on the minimum legal drinking age (ie, 18 years vs 17 years as of January 1, 2014), the abovementioned time frame and start date differed from the original design [31]. The Alcoholic Alert study was approved by the Medical Ethics Committee of Atrium Orbis Zuyd (METC number: 12-N-104) and was registered in the Dutch Trial Register (NTR 4048).

Randomization

Randomization was performed at school level to prevent contamination between participants. Randomization was conducted by drawing lots. After randomization, 21 schools were assigned to the intervention condition and 23 schools were assigned to the control condition.

Sample

The study population consisted of Dutch adolescents (aged 15-19 years) attending school [31]. Participants included students at schools of higher secondary education, lower secondary education, and lower vocational training. To have enough power for the evaluation, a participant target was made based on the following criteria: 10% reduction in binge drinking occasions (ie, for girls, at least four glasses or, for boys, five glasses of alcohol-containing drinks in one occasion) [32] during the preceding 30 days between the intervention and control group, with an intraclass correlation of .02, a power of 80%, and a significance level of .05. Furthermore, taking into account the drop in power due to an expected dropout of 50% at follow-up, it was estimated that at least 34 schools should be included at T0 [31].

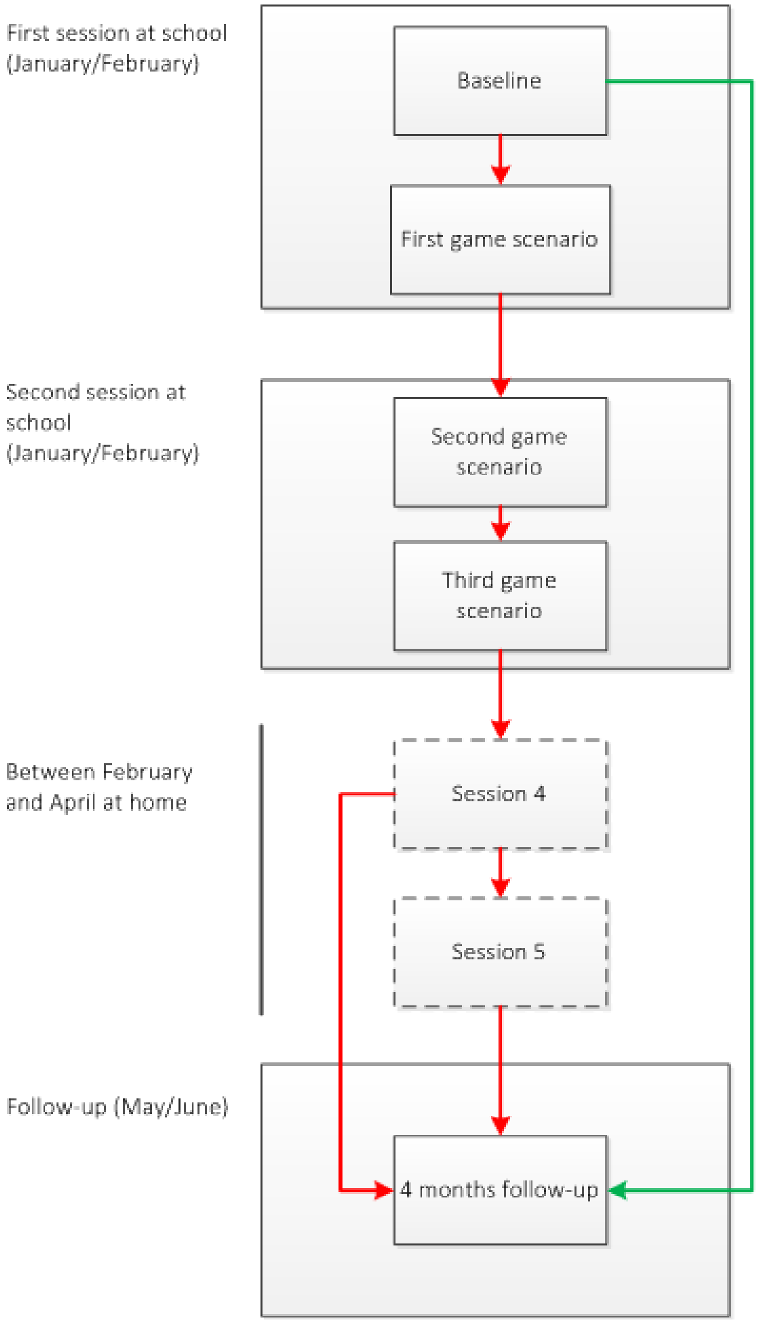
To reach the required number of 34, schools were recruited via several media; schools first received flyers with information about the Alcoholic Alert study, after which they were contacted via telephone and email. If schools enrolled in the study, students of the schools were eligible to participate. However, they could do so only if they provided informed consent by clicking a checkbox, which preceded the Web-based questionnaire at T0 [31].

Intervention

Adolescents in the intervention condition participated in a Web-based computer-tailored alcohol reduction program called Alcoholic Alert [31]. After completing a Web-based questionnaire on the Alcoholic Alert website at T0, the participants entered a game called “Watskeburt” (Dutch slang for “What Happened?!”). In the game, the participant played a character whose goal it was to find out what happened after a night of heavy drinking. Participants received in-game questions concerning alcohol-related sociocognitive factors, including attitude, social influences, self-efficacy expectations, and action plans toward alcohol drinking. These questions were based on the I-Change model, which is an integrated model explaining motivational and behavioral change [33]. Based on their answers, they received computer-tailored feedback on these determinants. They played in three game scenarios within three sessions. A week after playing the third game scenario, the participants were asked to revisit the intervention website to answer several questions. In this fourth session, they were asked about their drinking behavior during the preceding week and then they received computer-tailored feedback on their alcohol use compared to Dutch drinking guidelines. Subsequently, the participants were asked whether they had an event (eg, party, wedding) in the upcoming 30 days then they were challenged to drink less than usual and were asked for the maximum amount they wanted to drink. An email, with a reminder of accepting the challenge, was sent to them a day before the event. Two days after the event, during a fifth session, they were asked to visit the intervention website and fill in their alcohol use. If the participant failed the challenge, they received computer-tailored feedback with tailored advice and had the opportunity to take on a new challenge. If the participant met the challenge, he or she received congratulations and the intervention was over (Figure 1).

Participants receiving CAU also filled in the Web-based questionnaire at T0 and T1, but they did not have access to the game and did not receive computer-tailored feedback until after the final measurement. Further information on the intervention can be found elsewhere [31].

Figure 1. Flowchart of the intervention (based on Jander et al [31]). Red line: routing intervention condition; green line: routing control line; dashed boxes: intervention parts that had to be done at home.



Measurements and Outcomes

The measurements at T0 and T1 were performed at school after participants received instructions from their teachers [31]. The Web-based questionnaires used for the measurements included items related to alcohol drinking behavior, use of services within the health care sector, and ICBs. In addition, several background variables were measured at T0, including gender, age, educational level (higher secondary level, lower secondary level, and lower vocational), religion (Catholic, Protestant, Muslim, other religion, no religion), and ethnicity (Dutch, Antillean, Belgian, German, Surinamese, Moroccan, Turkish, other) [31].

In this study, the outcome measures were weekly alcohol use and the number of binge drinking occasions in the preceding 30 days. Weekly alcohol use was assessed by asking participants on which days in the past week they had been drinking and, if they did, how many glasses of alcohol they had on these days. Based on this information, the total amount of glasses of alcohol was calculated [17]. From this, as for binge drinking occasions, the weekly alcohol use at T1 was subtracted from the alcohol use at T0. This led to positive scores in case of a reduction in weekly alcohol use or number of binge drinking occasions, and negative scores in case weekly alcohol use or the number of binge drinking occasions increased.

Resource Use and Costing

The following costs related to the Alcoholic Alert intervention were identified as important and measured: (1) intervention costs, (2) health care costs (ie, costs for services inside the health care sector), (3) intersectoral costs (ie, costs for services outside the health care sector), and (4) costs of substance use (eg, use of hard drugs). Because alcohol use was the outcome measure in this study, the costs of alcohol use were not included in the CEA. Doing so would have led to double counting [34].

Costs (in Euros) were measured irrespective of who bore them and were indexed for the reference year 2014 using price indexes from Statistics Netherlands [35]. Calculations for specific costs of service utilization and substance use can be obtained from the first author (RD).

Intervention Costs

Intervention costs were divided into costs made during the development of the intervention and costs for running the intervention. Costs incurred during the development included game development costs (€20,328) and automatic tailoring software license and development costs (€8367.15). These costs and costs for other personnel involved in the development and application of the intervention for this study, such as costs incurred contacting schools, recruiting participants, and analyzing data are sunk costs [36]. Therefore, these are not included in the CEA. This is further justified by the fact that the intervention itself is Web-based and universal for adolescents, meaning the intervention has a wide reach and the development costs per participant drop to a minimum when it is used widely and structurally.

Costs for running the intervention include hosting costs for the website, tailored feedback software, and participants' time investments. As for the sunk costs, the website hosting costs (€300 per year) per participant drop to a minimum if used widely and are not included in the CEA. Tailored feedback software costs were €7 per participant per week. In case the participant met his or her challenge in the fifth session, he or she would receive tailored feedback over a period of 3 to 6 weeks. The mean intervention duration was 4 weeks, so the tailored feedback costs were €28. The five sessions took up 1.5 hours at school (€8.30 per hour) and 1 hour of free time (€12.50). The total of 2.5 hours was valued at €25 per participant. In sum, the total intervention costs per participant were an estimated €53.

Health Care Costs

Health care costs were calculated by multiplying volumes of health services by related cost prices. Health services measured included contacts with the general practitioner, emergency care, hospital stays, ambulance rides, and mental health services. Cost prices were drawn from the Dutch manual for costing in economic evaluations [37].

Intersectoral Costs

The ICB-related costs were calculated by multiplying volumes of services and time investments outside the health care sector with related cost prices. The ICBs were classified in sectors according to a classification scheme for ICBs by Drost et al [30]. The sectors in this scheme included education, labor and social security, household and leisure, and criminal justice system. The services and time investments measured included school absenteeism and contacts with an attendance officer (education), work absenteeism (labor and social security), failing to perform household and other activities, contacts with youth and family center and family care (household and leisure), and contacts with (youth) police services, court proceedings, and child (health) protection services (criminal justice system). Cost prices were drawn from a Dutch manual for intersectoral costs and benefits of (preventive) interventions [38]. Cost prices not mentioned in the manual were extracted from the Institute for Medical Technology Assessment (iMTA) questionnaire on intensive youth care [39]. For ICBs that required valuation of time, such as failing to perform household activities, some additional information was drawn from the Dutch report "The Netherlands in a Day" (free translation) [40].

Costs of Substance Use

In addition to alcohol use, use of other substances were measured as well. These included packs of cigarettes, use of soft drugs, and use of hard drugs. Cost prices were found on the website of the Jellinek Clinic, which is a renowned Dutch institution specializing in preventing and treating alcohol and substance abuse [41].

Data Preparation

The basis of the analysis was the dataset used for the Alcoholic Alert effect study [24]. However, because cost measures were used for conducting the economic evaluation, some additional data cleaning was required to create a dataset that was suitable for conducting the CEA.

First, because the digital questionnaires contained open-ended questions, participants had the opportunity to fill in unrealistic answers. It was a small subsample that systematically filled in these unrealistic answers, but to improve the validity of the results, these respondents were excluded from analysis. To clear the data of these respondents, limits were set for each variable. Participants breaching these limits by providing unrealistic answers were excluded from analysis. For example, because the 4-month recall period between T0 and T1 amounted to 120 days, any respondent claiming to have stayed more than 120 days in a hospital was excluded. In general, to reduce the chance of wrongfully excluding participants, limits were set high, but within the range of credibility. As for all steps during the process of data preparation and analysis, the limits were discussed in author meetings and were agreed on by all authors. A list of these limits can be obtained in Multimedia Appendix 1.

Second, the dataset was cleared of respondents who at baseline did not answer a single question related to costs. Based on these two steps, the sample at T0, and accordingly at T1, was smaller and different in composition in the CEA than the sample used for the effect study [24].

Finally, to assess the school-based part of the variance in this cluster RCT, intracluster correlation coefficients (ICCs) were calculated for both the weekly alcohol use and binge drinking occasions outcome measures. The ICCs were calculated based on the following formula: ICC or $\rho = s_b^2/(s_b^2 + s_w^2)$, where s_b^2= the variance between clusters and s_w^2= the variance within clusters [42]. Input for both outcome effect sizes was generated using SPSS version 20 by running linear regression mixed models. These analyses and corresponding calculations resulted in ρ=.06 for binge drinking occasions and ρ=.01 for weekly alcohol use, which shows that the within-cluster variances for both effect sizes were much greater than the between-cluster variances [42]. Based on these results, no re-estimations of effects were required.

Analysis

Descriptive Statistics

Descriptive statistics were used to describe the characteristics of the sample at T0 and at T1. Differences between the intervention and control conditions were assessed in SPSS version 20 using independent samples *t* tests for continuous variables and chi-square tests for discrete variables. The same software was used to conduct stepwise linear regression analyses to assess the dependence of the outcome measures on these variables.

Cost-Effectiveness Analysis

The base scenario of this study included CEAs from the two perspectives mentioned earlier. We calculated costs of services utilization in three steps: (1) assessment of the services and time consumed in the 4-month period between T0 and T1, (2) calculation of the associated costs in Euros, and (3) calculation of the incremental cost-effectiveness ratio (ICER) using the formula $(C_i–C_c)/(E_i–E_c)$. Here *C* represents the average total costs per participant during the 4-month period between T0 and T1 and *E* represents the mean difference in the number of glasses of alcohol or binge drinking occasions at T1 in comparison with the number measured at T0 in the intervention (C_i and E_i) and in the control (C_c and E_c) condition.

Stochastic uncertainty in the data was dealt with using nonparametric bootstraps. By using the bootstrapping technique, means and confidence intervals were calculated and 5000 ICERs were simulated, which were plotted in cost-effectiveness planes. These planes provided a visual representation of the probability of the intervention being cost-effective in comparison with the control condition by showing the distribution of ICERs across four quadrants: (1) more effective and more costly in the northeast quadrant (NE), (2) more effective and less costly in the southeast quadrant (SE), (3) less effective and less costly in the southwest quadrant (SW), and (4) less effective and more costly in the northwest quadrant (NW) [43].

An ICER in the SE and NW is negative, indicating that the intervention is dominant over (SE) or inferior to (NW) the control condition. An ICER in the SW or NE is positive, indicating that from a cost-effectiveness viewpoint the intervention is more favorable than the control condition only when the ICER is lower than the willingness to pay (WTP) per unit effect. Because no threshold (ie, maximum WTP) was available for the weekly alcohol use outcome measure, a cost-effectiveness acceptability curve (CEAC) was created for both perspectives. The CEAC showed the likelihood of the intervention being favorable over CAU for several hypothetical thresholds.

Sensitivity and Subgroup Analyses

Apart from the analyses to deal with stochastic uncertainty, several other analyses were conducted. First, to assess the impact of cost outliers, ICERs were calculated based on data in which cost outliers were excluded. Based on the output of descriptive statistics, it was decided to exclude participants when total costs were greater than €5000. Based on this approach, in the analyses that were conducted from the health care perspective, one participant was excluded. In the analyses that were conducted from the societal perspective, four participants were excluded. Second, to assess the effect of the uptake, the costs of cigarette use, and of the use of soft and hard drugs in the analyses conducted from the societal perspective, additional analyses were conducted without these costs.

Finally, given the heterogenic composition of the study sample, several subgroup CEAs were conducted based on the background variables measured at T0. These included analyses based on dichotomized background variables, including gender (male, female), age (15-16, ≥17), educational level (low, high), religion (religious, not religious), and ethnicity (Dutch, non-Dutch). Again, for all these analyses, stochastic uncertainty was dealt with using nonparametric bootstraps.

Results

Dropout and Sample Characteristics

Figure 2 shows a flowchart with the number of participating schools and adolescents at T0 and T1. In total, 44 schools were randomized into the control condition or intervention condition. Of the schools randomized to the control condition, five withdrew their participation before T0 (one secondary lower education, one lower vocational training, two secondary higher education, one secondary education mixed). In addition, three schools in the control condition (all secondary higher education) and two schools in the intervention condition (one lower vocational education, one higher secondary education) did not start the baseline assessment and did not respond to the emails and phone calls [24].

In total, 2649 adolescents from 34 schools participated in the baseline questionnaire. Of these, and different from the effect study [24], an additional 91 participants (3.4%) were excluded from analysis based on providing unrealistic answers to the cost questions. Another 65 participants were excluded from analysis because they did not answer the cost questions at T0. This resulted in 2493 adolescents who were included in the baseline analysis. Of these, 1538 were in the intervention condition and 955 were in the control condition (Figure 2). The *t* tests and chi-square tests conducted on the baseline sample showed that the adolescents in the two conditions significantly differed on multiple characteristics. Adolescents in the intervention condition were significantly younger, more often female, had a higher educational level, more often indicated being religious, were less likely to be a drinker, were less often a binge drinker, had less binge drinking occasions, and had a lower weekly alcohol use than adolescents in the control condition (Table 1).

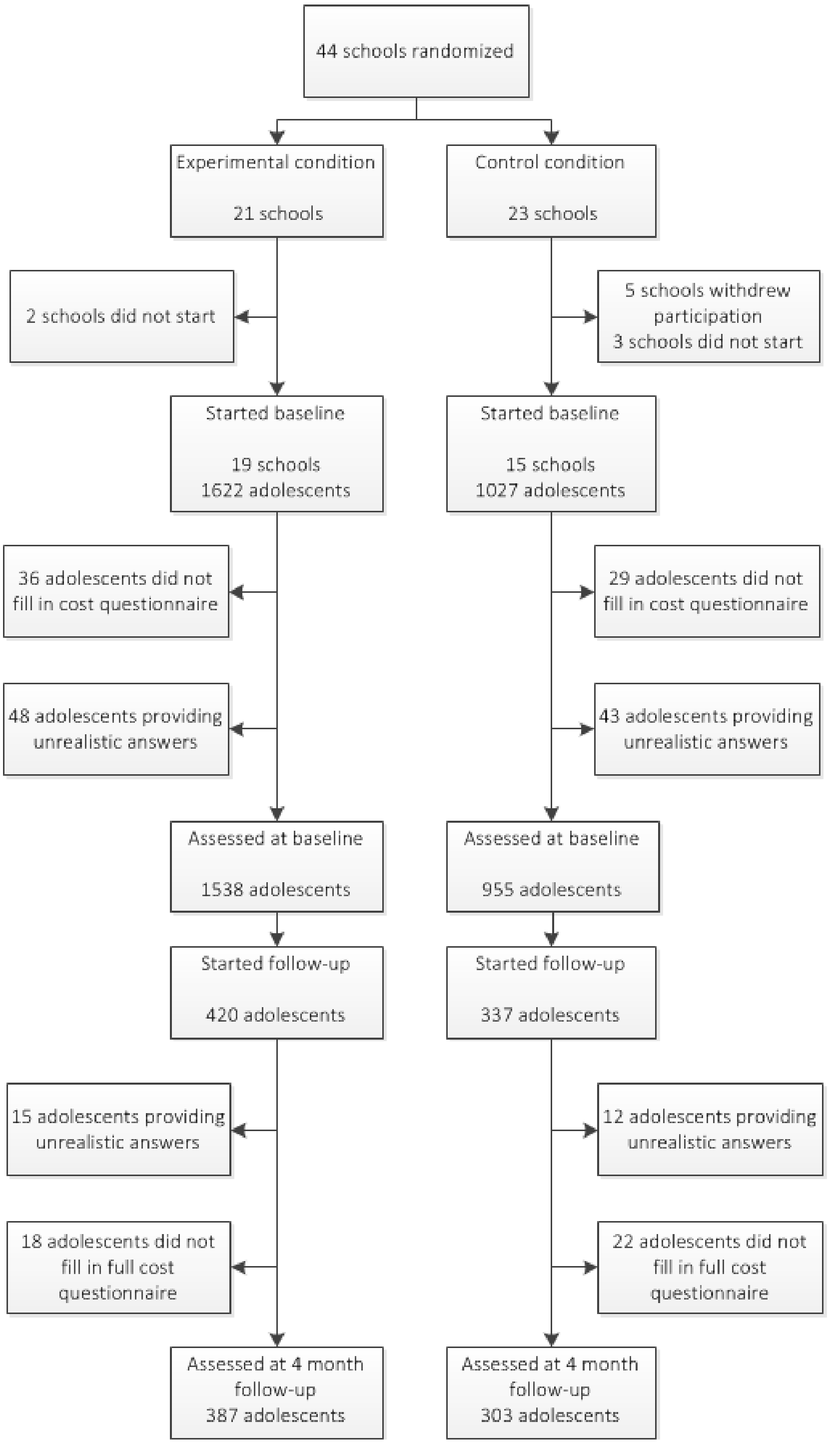
Of the 2493 adolescents, 757 participated in the cost questionnaire at T1 (response rate 30.36%). Of these 757, another 27 participants (3.6%) were excluded from analysis based on having provided unrealistic answers to the cost questions. An additional 40 were excluded from analysis because they did not answer the cost questions at T1, resulting in 690 participants to be analyzed at baseline. Of these, 387 were in the intervention condition and 303 were in the control condition. Here, adolescents in the intervention condition were more often female, had a higher educational level, more often indicated being religious, and had a lower weekly alcohol use than participants in the control condition did (Table 1).

For the T1 sample, results of the linear regression analyses show that weekly alcohol use was dependent on gender, age and educational level (R^2=.146). For binge drinking occasions, a significant proportion of the variance could be explained by age and educational level (R^2=.136).

Table 1. Baseline characteristics and differences at T0 and at T1.

Variable	Total (N=2493)	Intervention (n=1538)	Control (n=955)	Baseline difference (T0)			Follow-up difference (T1)		
				t_{2492}	χ^2_1	P	t_{954}	χ^2_1	P
Age (15-19 years), mean (SD)	16.3 (1.2)	16.0 (1.1)	16.7 (1.2)	15.08		<.001	1.78		.08
Gender, n (%)					45.6	<.001		16.5	<.001
Male	1295 (51.95)	717 (46.62)	578 (60.5)						
Female	1198 (48.05)	821 (53.38)	377 (39.5)						
Educational level, n (%)					73.4	<.001		6.7	.006
High	1483 (59.49)	1017 (66.12)	466 (48.8)						
Low	1010 (40.51)	521 (33.88)	489 (51.2)						
Religion, n (%)					10.0	.002		15.4	<.001
No religion	1465 (58.76)	866 (56.31)	599 (62.7)						
Religion	1028 (41.24)	672 (43.69)	356 (37.3)						
Catholic	593 (23.79)	397 (25.81)	196 (20.5)						
Protestant	174 (6.98)	130 (8.45)	44 (4.6)						
Muslim	150 (6.02)	75 (4.88)	75 (7.9)						
Other	111 (4.45)	70 (4.55)	41 (4.3)						
Ethnicity, n (%)					1.2	.27		0.0	.51
Dutch	2225 (89.25)	1381 (89.79)	844 (88.4)						
Non-Dutch	268 (10.75)	157 (10.21)	111 (11.6)						
Antillean	5 (0.20)	3 (0.20)	2 (0.2)						
Belgian	10 (0.40)	5 (0.33)	5 (0.5)						
German	13 (0.52)	10 (0.65)	3 (0.3)						
Surinamese	26 (1.04)	18 (1.17)	8 (0.8)						
Moroccan	35 (1.40)	15 (0.98)	20 (2.1)						
Turkish	48 (1.93)	21 (1.37)	27 (2.8)						
Other	131 (5.25)	85 (5.53)	46 (4.8)						
Alcohol use, n (%)									
Never drinkers	664 (26.63)	459 (29.84)	205 (21.5)		21.2	<.001		0.7	.39
Binge drinkers	1271 (50.98)	724 (47.07)	547 (57.3)		24.5	<.001		3.6	.06
Binge drinking occasions, mean (SD)	2.2 (4.0)	2.0 (4.2)	2.4 (3.8)	2.51		.01	1.18		.24
Weekly alcohol use, mean (SD)	3.9 (8.8)	3.2 (8.1)	5.0 (9.6)	4.96		<.001	2.17		.03

Figure 2. Flowchart describing the dropout of participants.



Costs

Table 2 shows that the total health care costs per adolescent were lower in the intervention group (€85.65) than in the control group (€124.49). This difference can largely be explained by the difference in costs for reported hospital stays. However, costs of ICBs (€162.68) and substance use (€36.30) were higher in the intervention group than in the control group (€112.61 and €24.64, respectively). The difference in mean costs for court proceedings is noticeable and explains much of the difference in ICB-related costs. Large differences in specific costs were caused by outliers in either the intervention or the CAU arm. As indicated previously, sensitivity analyses were conducted without these outliers. When the intervention costs were included both from the health care and societal perspective, costs were higher in the intervention group. The *z* score for each cost category was positive and higher than the reference value of 1.96 [44], indicating that costs were skewed and tailed to the right (Table 2).

Table 2. Mean and bootstrapped median costs per adolescent measured at T1 covering a 4-month period between T0 and T1 (€, 2014) and *z* scores per cost category.

Type of costs	Intervention group (€)		Control group (€)		Skewness (z score)[b]
	Mean (SD)	Median (σ^2)[a]	Mean (SD)	Median (σ^2)[a]	
Intervention costs	53.00 (0.00)	53.00 (0.00)			
Health care costs					
General practitioner	25.42 (3.39)	25.22 (11.49)	23.84 (4.36)	23.50 (11.49)	73.44
Emergency care	10.16 (3.14)	10.06 (9.86)	8.08 (3.74)	7.65 (13.99)	97.98
Hospital stays	27.22 (12.38)	26.88 (135.26)	67.90 (58.94)	66.98 (3473.92)	258.45
Ambulance rides	7.34 (2.95)	6.74 (8.70)	9.37 (6.18)	8.64 (38.19)	174.83
Mental health care	16.07 (6.37)	15.59 (40.58)	14.16 (5.72)	13.78 (32.71)	115.83
Total health care costs	85.65 (20.62)	84.89 (425.18)	124.49 (67.91)	120.38 (4611.76)	229.86
Total health care perspective	139.16 (20.77)	138.04 (431.39)	127.45 (68.64)	122.12 (4711.45)	229.09
Educational sector costs					
School absenteeism	51.88 (7.13)	51.44 (50.84)	66.77 (11.03)	65.74 (121.66)	76.31
Attendance officer	0.03 (0.03)	0.04 (0.00)	2.53 (2.16)	2.40 (4.67)	272.90
Total educational sector costs	51.52 (7.16)	51.17 (51.27)	69.30 (11.35)	68.37 (128.82)	74.60
Labor and social security costs					
Work absenteeism	9.67 (3.49)	9.20 (12.18)	6.03 (1.41)	6.01 (1.99)	140.85
Total labor and social security costs	9.63 (3.39)	9.30 (11.49)	6.09 (1.40)	6.01 (1.96)	140.85
Household and leisure costs					
Failure to perform household activities	8.58 (3.01)	8.31 (9.06)	10.92 (3.39)	10.74 (11.49)	98.06
Failure to perform other activities	25.77 (13.20)	24.67 (174.24)	8.17 (1.60)	8.02 (2.56)	182.19
Youth and family center	0.00 (0.00)	0.00 (0.00)	0.33 (0.34)	0.33 (0.12)	282.45
Family care	1.26 (1.27)	1.49 (1.61)	0.18 (0.18)	0.19 (0.03)	270.73
Total household and leisure costs	35.38 (14.80)	33.55 (219.04)	19.23 (4.21)	18.81 (17.72)	163.43
Criminal justice system costs					
Police services	4.68 (2.61)	4.45 (6.81)	6.67 (2.50)	6.47 (6.25)	126.44
Youth police services	1.58 (1.72)	1.91 (2.96)	0.28 (0.20)	0.28 (0.04)	277.89
Court proceedings	55.10 (34.45)	50.92 (1186.80)	9.95 (7.39)	10.15 (54.61)	197.40
Child protection services	0.29 (0.31)	0.36 (0.10)	0.68 (0.72)	0.72 (0.52)	227.11
Child health protection services	3.13 (3.73)	4.06 (13.91)	0.46 (0.45)	0.48 (0.20)	277.12
Total criminal justice system costs	66.32 (38.43)	61.51 (1476.86)	18.77 (9.79)	17.40 (95.84)	182.55
Total intersectoral costs	162.68 (41.85)	158.61 (1751.42)	112.61 (18.83)	111.26 (354.57)	124.84
Substance use costs					
Cigarettes	30.68 (6.25)	30.46 (39.06)	18.82 (3.94)	18.82 (15.52)	48.86
Soft drugs	4.30 (1.95)	4.00 (3.80)	5.44 (3.01)	5.12 (9.06)	153.72
Hard drugs	1.15 (1.08)	1.32 (1.17)	0.20 (0.11)	0.19 (0.01)	275.97
Total costs of substance use	36.30 (7.58)	36.09 (57.46)	24.64 (5.85)	24.30 (34.22)	61.42
Total societal perspective	336.45 (53.31)	331.41 (2841.96)	263.52 (70.70)	255.75 (4998.49)	135.60

[a] The presented median cost is the 50th percentile of 1000 bootstrap replications.

[b] The *z* score for each cost category is positive and higher than the reference value of 1.96 [44] indicating that costs were skewed and tailed to the right.

Incremental Costs

Table 3 shows costs per condition. Note that these means slightly differ from the costs presented in Table 2, for costs in Table 2 are bootstrapped means, whereas costs in Table 3 are means drawn from raw data. For both perspectives, costs were higher in the intervention condition. The incremental costs (ie, the difference in mean costs per adolescent between the intervention and control condition) varied per perspective, namely €13.76 from the health care perspective and €74.03 from the societal perspective.

Table 3. Summary statistics for the base case sensitivity cost-effectiveness bootstrap analyses.

Perspective[a] and condition	Costs (€)[b]	Effect[c]	ICER[d]	NE	NW (inferior)	SW	SW (dominant)
Base case analyses							
Weekly alcohol use							
Health care							
Control (n=303)	125.32	–1.51					
Intervention (n=387)	139.08	–0.78	40	55%	10%	6%	30%
Societal							
Control (n=303)	262.68	–1.51					
Intervention (n=387)	336.71	–0.78	62	60%	14%	3%	23%
Binge drinking occasions							
Health care							
Control (n=303)	125.32	–0.33					
Intervention (n=387)	139.08	0.16	79	60%	4%	2%	34%
Societal							
Control (n=303)	262.68	–0.33					
Intervention (n=387)	336.71	0.16	144	69%	5%	1%	25%
Sensitivity analyses excluding outliers							
Weekly alcohol use							
Health care							
Control (n=302)	59.47	–1.55					
Intervention (n=387)	139.08	–0.78	72	82%	17%	0%	1%
Societal							
Control (n=302)	193.85	–1.55					
Intervention (n=384)	269.19	–0.66	67	80%	12%	1%	7%
Binge drinking occasions							
Health care							
Control (n=302)	59.47	–0.33					
Intervention (n=387)	139.08	0.16	140	93%	6%	0%	1%
Societal							
Control (n=302)	193.85	–0.33					
Intervention (n=384)	269.19	0.21	124	87%	4%	0%	9%

[a] Bootstrap analyses were conducted from two perspectives: the health care perspective and the societal perspective.

[b] Mean costs per adolescent at 2014 prices.

[c] Reduction in per week alcohol use or binge drinking occasions between T0 and T1, with negative values indicating an increase at T1 compared to T0.

[d] The presented ICER is the 50th percentile of 5000 bootstrap replications of the ICER.

Incremental Effects

Table 3 shows the effects per condition. In comparison with the control condition, the intervention was incrementally effective in reducing the weekly use of alcohol and number of binge drinking occasions. At T1, adolescents in the control condition drank a mean 1.51 glasses of alcohol per week more than at T0. In the intervention condition, there was an increase of 0.78

glasses, resulting in a mean incremental effect of 0.73 glasses per week. Furthermore, in the control condition, there was an increase of 0.33 binge drinking occasions. In the intervention condition, there was a decrease of 0.16, resulting in a mean incremental effect of 0.49 binge drinking occasions per 30 days. For both outcome measures, this did not change with perspective; a change of perspective within the base case scenario stipulated only a change in costs.

Incremental Cost-Effectiveness Ratios

From both perspectives, the mean costs were higher for the intervention condition in comparison with the control condition. Since the intervention was more effective than CAU on both outcome measures, this resulted in positive ICERs (Table 3). However, ICERs differed for both perspectives, namely €40 and €79 from the health care perspective, and €62 and €144 from the societal perspective per incremental reduction of one glass of alcohol per week and binge drinking occasion per 30 days, respectively.

The cost-effectiveness planes (Figures 3 and 4, left side) show differences in distributions of the 5000 simulated ICERs across the four quadrants between the CEAs carried out from the two perspectives. Corresponding with the median ICERs presented in Table 3, the majority of simulated ICERs for all base case analyses are located in the NE quadrant. However, the distribution of the simulated ICERs among the quadrants differs between the perspectives. Notable is the shift of the cloud of ICERs toward the SE quadrant in the analysis carried out from the health care perspective (ie, 30% for weekly alcohol use and 34% for binge drinking occasions) in comparison to the analyses carried out from the societal perspective (ie, 23% and 25%, respectively).

The preceding percentages equal the probabilities of the intervention being cost-effective at a WTP max of €0 in the CEACs (Figures 3 and 4, right side). These results show that for low WTP thresholds the probability of the intervention being cost-effective over the control intervention is higher from a health care perspective than it is from the societal perspective. For all base case analyses, the vast majority of simulated incremental effects were in the NE quadrant; therefore, these probabilities increase to approximately 80% when the WTP max increases. The probabilities of the intervention being cost-effective do not differ much between the two perspectives for WTP thresholds greater than €500 (Figures 3 and 4, right side).

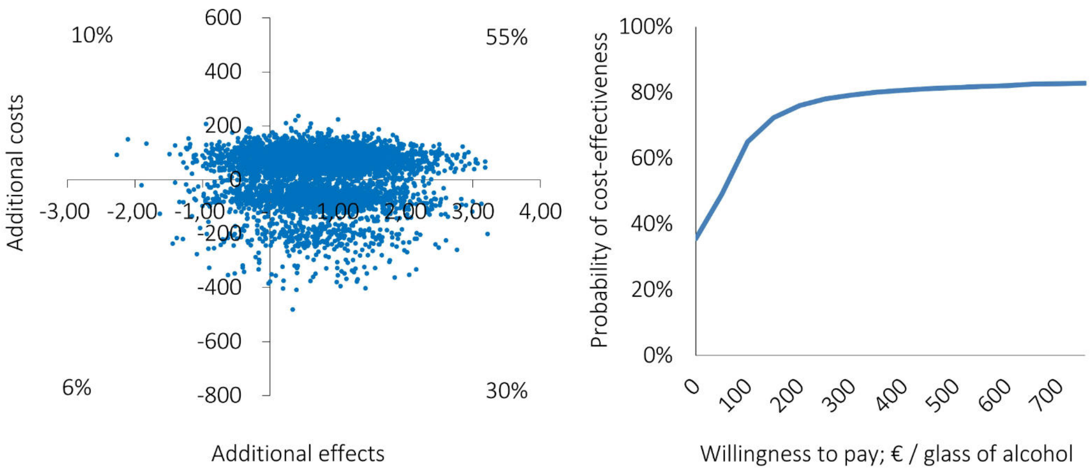
Sensitivity and Subgroup Analyses

The results of the sensitivity analyses (ie, excluding cost outliers) attest to the robustness of the base case analyses (Table 3). From the societal perspective, ICERs were close to similar. From the health care perspective, ICERs increased and were higher than those of the societal perspective. However, as for the base case analyses, the probability of the intervention being cost-effective remained dependent on the WTP max. The results of the analyses conducted from a societal perspective minus the costs of drugs and cigarette use were similar to the results of the analyses conducted in which these costs were included (Figures 3 and 4).

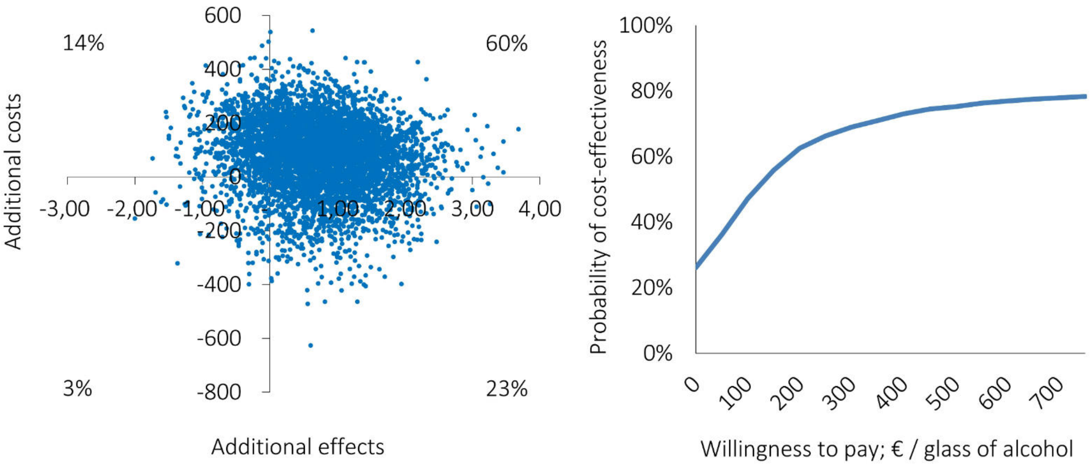
Subgroup analyses showed, from both the health care and the societal perspective, and for both outcome measures, that the intervention was cost-effective for the older adolescents and those at a lower educational level (Tables 4 and 5). From a health care perspective, it was found to be cost-effective for the male and nonreligious adolescent subgroups as well. The intervention was not cost-effective for those with a non-Dutch ethnicity or for female adolescents for the weekly alcohol use outcome measure. For all other subgroups, ICERs were positive, meaning the intervention was cost-effective depending on the WTP max. The corresponding cost-effectiveness planes and CEACs of all subgroup analyses can be obtained in Multimedia Appendices 2 and 3.

Figure 3. Cost-effectiveness planes (left side) and corresponding CEACs (right side) of the economic evaluations based on the weekly alcohol use outcome measure, which were conducted from the health care perspective (upper), societal perspective including drugs and cigarette use (middle), and societal perspective excluding drugs and cigarette use (lower).

Health care perspective



Societal perspective (including drugs and cigarette use)



Societal perspective (excluding drugs and cigarette use)

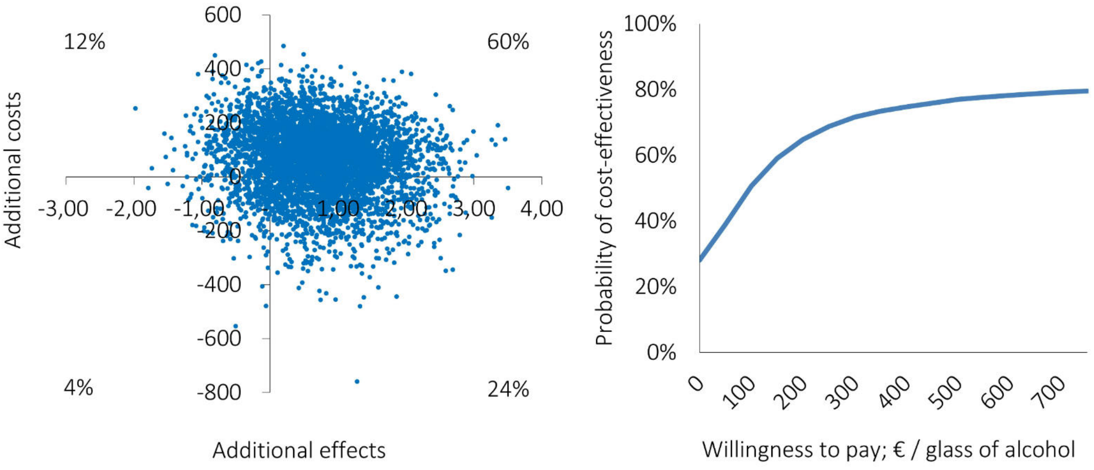


Figure 4. Cost-effectiveness planes (left side) and corresponding CEACs (right side) of the economic evaluations based on the binge drinking occasions outcome measure, which were conducted from the health care perspective (upper), societal perspective including drugs and cigarette use (middle), and societal perspective excluding drugs and cigarette use (lower).

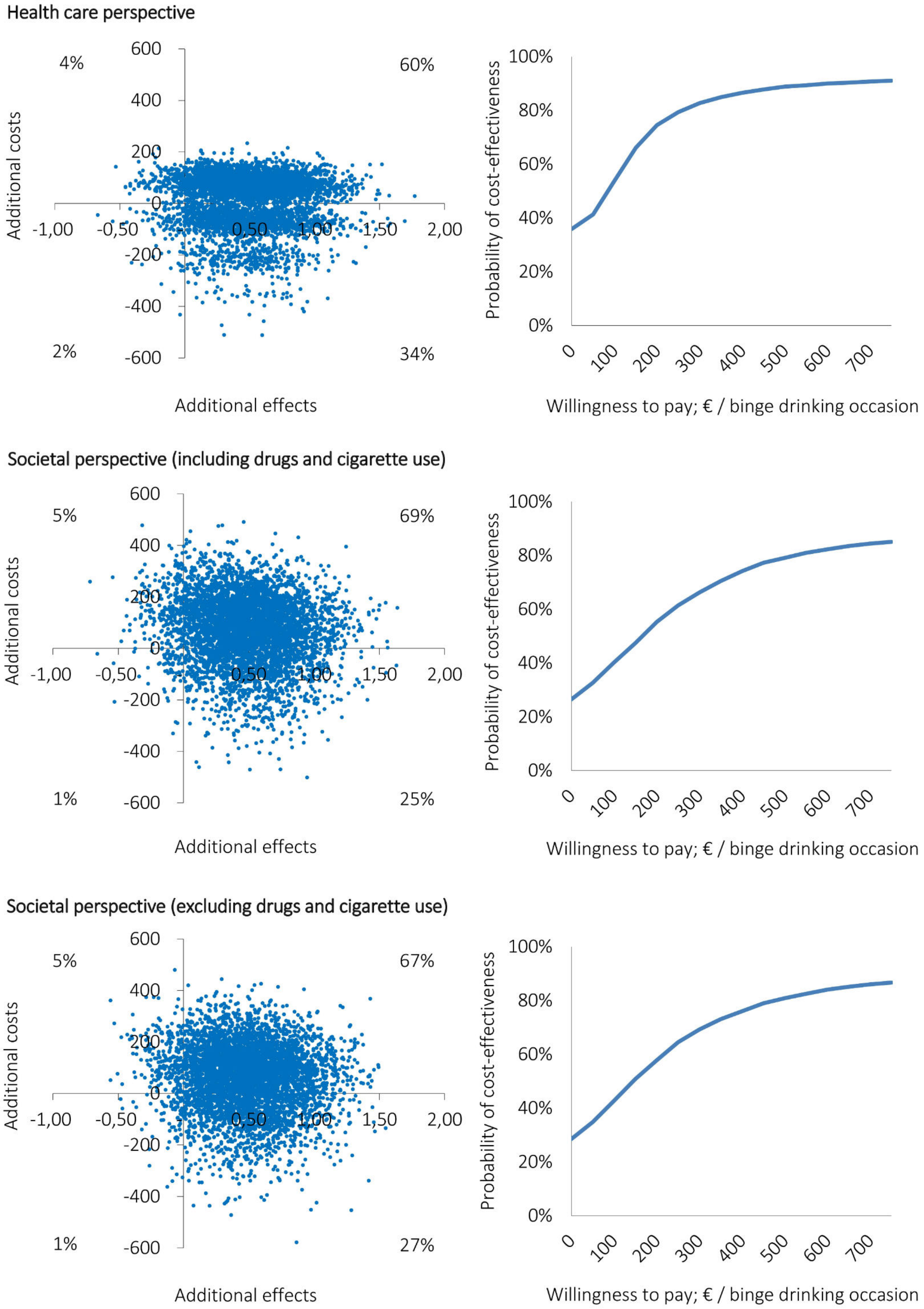


Table 4. Summary statistics for the subgroup sensitivity cost-effectiveness bootstrap analyses based on the weekly alcohol use outcome measure.

Perspective[a] and condition	Costs (€)[b]	Effect[c]	ICER[d]	NE	NW (inferior)	SW	SE (dominant)
Gender subgroups							
Male							
Health care							
Control (n=162)	193.96	–2.19					
Intervention (n=147)	164.53	–0.80	Dominant	43%	2%	3%	52%
Societal							
Control (n=162)	337.21	–2.19					
Intervention (n=147)	352.43	–0.80	21	56%	3%	2%	39%
Female							
Health care							
Control (n=141)	46.46	–0.73					
Intervention (n=240)	123.48	–0.77	Inferior	48%	52%	0%	0%
Societal							
Control (n=141)	177.04	–0.73					
Intervention (n=240)	327.09	–0.77	Inferior	46%	52%	1%	1%
Age subgroups							
Younger adolescents (15-16 years)							
Health care							
Control (n=200)	50.05	–1.24					
Intervention (n=281)	145.79	–0.73	108	80%	20%	0%	0%
Societal							
Control (n=200)	177.07	–1.24					
Intervention (n=281)	334.90	–0.73	149	77%	20%	0%	3%
Older adolescents (≥17 years)							
Health care							
Control (n=103)	271.48	–2.03					
Intervention (n=106)	121.27	–0.92	Dominant	28%	5%	11%	56%
Societal							
Control (n=103)	428.90	–2.03					
Intervention (n=106)	341.53	–0.92	Dominant	32%	7%	10%	51%
Educational level subgroups							
Low							
Health care							
Control (n=98)	263.02	–2.20					
Intervention (n=91)	117.35	–0.60	Dominant	31%	4%	7%	58%
Societal							
Control (n=98)	435.78	–2.20					
Intervention (n=91)	282.38	–0.60	Dominant	26%	4%	8%	63%
High							
Health care							
Control (n=205)	59.50	–1.18					

Perspective[a] and condition	Costs (€)[b]	Effect[c]	ICER[d]	NE	NW (inferior)	SW	SE (dominant)
Intervention (n=296)	145.76	–0.83	102	73%	26%	0%	1%
Societal							
Control (n=205)	179.93	–1.18					
Intervention (n=296)	353.42	–0.83	172	70%	28%	0%	1%
Religion subgroups							
Religion							
Health care							
Control (n=97)	46.91	–2.05					
Intervention (n=181)	148.11	–0.71	66	92%	8%	0%	0%
Societal							
Control (n=97)	155.80	–2.05					
Intervention (n=181)	336.35	–0.71	110	90%	9%	0%	1%
No religion							
Health care							
Control (n=206)	162.24	–1.25					
Intervention (n=206)	131.13	–0.84	Dominant	37%	14%	13%	37%
Societal							
Control (n=206)	313.00	–1.25					
Intervention (n=206)	337.04	–0.84	5	42%	19%	8%	31%
Ethnicity subgroups							
Dutch							
Health care							
Control (n=278)	128.83	–1.65					
Intervention (n=356)	136.90	–0.84	36	53%	8%	5%	34%
Societal							
Control (n=278)	262.96	–1.65					
Intervention (n=356)	334.67	–0.84	57	60%	11%	3%	26%
Other							
Health care							
Control (n=25)	86.35	0.08					
Intervention (n=31)	164.09	–0.10	Inferior	32%	51%	4%	13%
Societal							
Control (n=25)	259.53	0.08					
Intervention (n=31)	360.20	–0.10	Inferior	23%	49%	7%	21%

[a] Bootstrap analyses were conducted from two perspectives: the health care perspective and the societal perspective.

[b] Costs per adolescent at 2014 prices.

[c] Reduction in per week alcohol use between T0 and T1, with negative values indicating an increase at T1 compared to T0.

[d] The presented ICER is the 50th percentile of 5000 bootstrap replications of the ICER.When an ICER is negative, then it is labeled as being either “dominant” (suggesting that the intervention is both more effective and less costly than CAU) or “inferior” (suggesting that the intervention is both less effective and more costly than CAU).

Table 5. Summary statistics for the subgroup sensitivity cost-effectiveness bootstrap analyses based on the binge drinking occasions outcome measure.

Perspective[a] and condition		Costs (€)[b]	Effect[c]	ICER[d]	NE	NW (inferior)	SW	SE (dominant)
Gender subgroups								
Male								
Health care								
	Control (n=162)	193.96	–0.57					
	Intervention (n=147)	164.53	0.08	Dominant	44%	1%	1%	54%
Societal								
	Control (n=162)	337.21	–0.57					
	Intervention (n=147)	352.43	0.08	46	57%	1%	1%	42%
Female								
Health care								
	Control (n=141)	46.46	–0.04					
	Intervention (n=240)	123.48	0.21	179	81%	19%	0%	0%
Societal								
	Control (n=141)	177.04	–0.04					
	Intervention (n=240)	327.09	0.21	291	78%	19%	0%	2%
Age subgroups								
Younger adolescents (15-16 years)								
Health care								
	Control (n=200)	50.05	–0.26					
	Intervention (n=281)	145.79	0.13	276	71%	29%	0%	0%
Societal								
	Control (n=200)	177.07	–0.26					
	Intervention (n=281)	334.90	0.13	343	68%	29%	1%	2%
Older adolescents (≥17 years)								
Health care								
	Control (n=103)	271.48	–0.45					
	Intervention (n=106)	121.27	0.94	Dominant	31%	0%	0%	68%
Societal								
	Control (n=103)	428.90	–0.45					
	Intervention (n=106)	341.53	0.94	Dominant	39%	0%	0%	60%
Educational level subgroups								
Low								
Health care								
	Control (n=98)	263.02	–0.53					
	Intervention (n=91)	117.35	1.04	Dominant	35%	0%	0%	64%
Societal								
	Control (n=98)	435.78	–0.53					
	Intervention (n=91)	282.38	1.04	Dominant	30%	0%	0%	70%
High								
Health care								
	Control (n=205)	59.50	–0.23					

Perspective[a] and condition		Costs (€)[b]	Effect[c]	ICER[d]	NE	NW (inferior)	SW	SE (dominant)
	Intervention (n=296)	145.76	–0.10	231	69%	30%	0%	1%
Societal								
	Control (n=205)	179.93	–0.23					
	Intervention (n=296)	353.42	–0.10	435	69%	29%	0%	1%
Religion subgroups								
Religion								
Health care								
	Control (n=97)	46.91	–0.34					
	Intervention (n=181)	148.11	0.32	148	96%	4%	0%	0%
Societal								
	Control (n=97)	155.80	–0.34					
	Intervention (n=181)	336.35	0.32	256	95%	4%	0%	1%
No religion								
Health care								
	Control (n=206)	162.24		–0.32				
	Intervention (n=206)	131.13	0.02	Dominant	43%	7%	6%	45%
Societal								
	Control (n=206)	313.00	–0.32					
	Intervention (n=206)	337.04	0.02	47	50%	8%	4%	37%
Ethnicity subgroups								
Dutch								
Health care								
	Control (n=278)	128.83	–0.36					
	Intervention (n=356)	136.90	0.18	71	59%	3%	2%	37%
Societal								
	Control (n=278)	262.96	–0.36					
	Intervention (n=356)	334.67	0.18	139	69%	4%	1%	26%
Other								
Health care								
	Control (n=25)	86.35	0.00					
	Intervention (n=31)	164.09	0.03	Inferior	36%	48%	3%	13%
Societal								
	Control (n=25)	259.53	0.00					
	Intervention (n=31)	360.20	0.03	Inferior	24%	48%	4%	24%

[a] Bootstrap analyses were conducted from two perspectives: the health care perspective and the societal perspective.

[b] Costs per adolescent at 2014 prices.

[c] Reduction in per week alcohol use between T0 and T1, with negative values indicating an increase at T1 compared to T0.

[d] The presented ICER is the 50th percentile of 5000 bootstrap replications of the ICER.When an ICER is negative, then it is labeled as being either "dominant" (suggesting that the intervention is both more effective and less costly than CAU) or "inferior" (suggesting that the intervention is both less effective and more costly than CAU).

Discussion

Principal Results

To the best of our knowledge, this was the first cost-effectiveness analysis of a Web-based intervention conducted from both the health care and societal perspective that also incorporated the possible impact of ICBs on the cost-effectiveness results. From both a health care and a societal perspective, our study shows the intervention was incrementally more effective and more costly than CAU. This counts for both the analyses in which the weekly alcohol use outcome measure was incorporated and the analyses based on the binge drinking occasions outcome measure.

Although the intervention was incrementally effective in targeting weekly alcohol use, there was an increase in the number of glasses of alcohol between T0 and T1 in both arms of the trial. This can be explained based on the estimation that approximately one-third of the study sample had his or her birthday during this 4-month period between T0 and T1, of which some reached the legal drinking age of 18 years. Furthermore, all adolescents in the sample aged 4 months, which increased the chance of them starting to drink or drink more. This is also true for the younger Dutch adolescents because many start drinking before the legal drinking age [45]. However, contrary to weekly alcohol use, the number of binge drinking occasions did not increase in the intervention arm; a small decrease of a mean 0.16 binge drinking occasions was noticed compared to an increase of a mean 0.33 in the CAU arm. Therefore, relative to the overall alcohol intake and compared to CAU, it can tentatively be concluded that adolescents in the intervention arm became less irresponsible about drinking.

Our research also shows that the inclusion of ICBs in the economic evaluation impacted the cost-effectiveness results of the analysis, especially for certain subgroups. From a health care perspective, the intervention is cost-effective for the male, lower education, older adolescent, and nonreligious subgroups. However, from a societal perspective (which includes ICBs), the intervention is clearly cost-effective only for the lower education and older adolescent subgroups.

The inferiority of the intervention for certain subgroups could, among other reasons, partly be explained based on the finding that the baseline consumption for these subgroups was relatively low compared to that of their counterparts. For example, the baseline mean weekly alcohol use in the female subgroup was 1.49 glasses (SD 3.57) compared to weekly mean 4.18 glasses (SD 8.42) in the male subgroup. In so far as the following can be concluded based on an analysis of the smallest subgroup (n=56), this also goes for the non-Dutch subgroup (mean 2.16, SD 5.00 glasses and mean 1.21, SD 2.44 binge drinking occasions) versus the Dutch subgroup (mean 2.74, SD 6.48 glasses and mean 1.64, SD 2.77 binge drinking occasions). In these subgroups, there was less effect to be gained. Consequently, this could be related to the possibility of these adolescents not identifying themselves as being part of the target group of, and being affected by, the intervention.

Strengths and Limitations

Some of the strengths of this study are its relatively large sample size and its randomized design. The cluster randomization at the school level minimized the risk of contamination between the study conditions. Furthermore, the large heterogenic study sample was a good representation of the Dutch adolescent school-going population and allowed for subgroup analyses on multiple background variables.

The use of a societal perspective along with a health care perspective was a major strength of this study. The societal perspective is argued to be dominant over other perspectives [46-48]. This is because of, but not restricted to, health economics' foundations in welfare economics, which means that an economic evaluation should include the impact of an intervention on the whole society [46]. However, not only the choice of perspective, but also the way this was implemented in the study design, can be considered a major strength. Because the study population consisted of school-going adolescents, limiting the societal perspective to including merely labor productivity costs would not have properly reflected this impact. For this study, the results show that labor productivity costs make up just a small part of the total costs of ICBs (Table 2). By including costs within the educational sector and criminal justice system, we managed to provide a better reflection of the economic impact of this intervention on society.

Apart from these strengths, the findings of this study need to be placed in the context of the study's limitations. First, both at T0 and T1, the composition of adolescents in the intervention condition was significantly different from that of the control condition for various characteristics, including gender, educational level, and religion. This might have been caused by (1) the cluster RCT design instead of randomization at the individual level and/or (2) the large dropout before the baseline assessment within the control condition in comparison with the intervention condition. Although the results of the regression analysis showed a relationship between some of the background variables and the outcome measures, uncertainty around the ICERs that were calculated in the base case analyses was dealt with through various strategies. The sensitivity analyses attest to the robustness of the findings. Furthermore, the heterogeneity of the sample was addressed extensively by calculating ICERs and conducting bootstrap analyses for all subgroups based on all background variables.

Second, the follow-up period of this CEA (ie, 4 months) might be regarded as short. Costs and (health) benefits that fall beyond these 4 months were not assessed. Future studies, including additional follow-up measures and cost-effectiveness modeling, could be interesting. Other studies have shown that the cumulative cost savings in the life span of health promotion interventions for adolescents could be high [49].

Third, the authors decided not to further modify the original dataset by imputation and restricted the analysis to complete cases. A missing completely at random analysis (MCAR) in SPSS version 20 based on the N=2493 sample that started at baseline showed 71% to 72% per cost variable at T1. Furthermore, the missing values were not at random (*P*<.001). Although this was expected considering the large dropout between T0 and T1, the same goes for the n=757 sample that started follow-up (*P*<.001). Given the nonnormality of cost variables, the nonrandomness of missing values, and the large dropout as is common in many Web-based interventions [19,50-55], it was concluded that additional imputation would have manipulated the original dataset too much. This counts for both the basic imputation methods, such as expectation maximization and last observation carried forward, as well as for the more advanced methods, such as Markov chain Monte Carlo technique with predictive mean matching [56]. As for imputation strategies, the chosen strategy might have led to biased results. Nevertheless, the alternative of replacing more than 70% of the values, which would have been needed in this study, would have increased the risk of a type II error [51]. Imputation would have resulted in an increased chance of underestimating the intervention's effectiveness and an unrealistic representation of costs.

Fourth, measurements were based on self-reports, which could have led to an underestimation of service use, alcohol use, and use of other goods in comparison with daily diaries [24,57]. As for any measurement based on recalling services or goods used, this is because of forgetting [17]. However, in this study, recall periods were kept short. For example, respondents were asked for their alcohol use in the previous week and not in a typical week. Furthermore, the recall period for the cost measurement questions was only 4 months. In addition, because the groups were randomized, this underestimation is likely to be equally distributed among the intervention and control group. Therefore, it is unlikely that the ICERs were affected.

Finally, within the setting of this study, respondents were free to fill in the answers themselves during the measurements. As mentioned earlier, limits needed to be set to exclude respondents who provided unrealistic answers. The choice made by the authors to exclude whole cases might have affected the outcomes of the analysis. However, this choice is justifiable based on the finding that the vast majority of unrealistic answers were not even close to the limits set by the authors and were far higher than the credible range. For example, there were seven respondents who claimed to have spent more than 200 nights in a hospital bed in the previous 4 months, of which three said to have spent more than 200,000. The 27 adolescents at T1 who filled in unrealistic answers (only 3.6% of the n=757 sample that started follow-up) had a mean 2.9 unrealistic answers. Based on the data, it was clear that the vast majority of these 27 adolescents did this deliberately and systematically. Although the outcome of the analyses might have been affected by the limits that were set by the authors, these limits were carefully considered, discussed, and decided a priori to the analyses. This was done to minimize the chance of biased results. Furthermore, the impact of cost outliers on the outcomes of the base case analyses has been covered in sensitivity analyses in which cost outliers were excluded.

Recommendations

Computer-tailored feedback can be a cost-effective way to target alcohol use and binge drinking among adolescents. In the Netherlands, despite the Dutch government's change of policy to reduce the minimum legal drinking age, 33.4% of Dutch adolescents were drinkers in 2014 [45]. This is because in practice drinking rules are set not only by Dutch law, but also by parents or caregivers and alcoholic beverages can easily be obtained via family and friends. Therefore, effective and cost-effective interventions targeting adolescent drinking behavior are still very much needed. The high dissemination capabilities of the Alcoholic Alert intervention, combined with its solid basis in the I-Change model and low intervention costs could make it an interesting investment for reducing alcohol use among adolescents.

Because the cost-effectiveness for the whole sample is dependent on the WTP max per effect, it is difficult to make strong recommendations on whether the intervention should be implemented from an economic point of view. Contrary to the generic outcome measure quality-adjusted life year (QALY) [58,59], and as is common for the majority of specific outcome measures, no guidelines are available that provide a reference cost-effectiveness threshold for reducing the consumption of alcohol. However, the CEACs provide decision supportive information because these provide cost-effectiveness probabilities for a wide range of hypothetical thresholds for all analyses. These also show that, from both the health care and the societal perspective, the intervention is cost-effective for older adolescents and for those at a lower educational level, regardless of which threshold is set. From a health economic viewpoint, it is recommended that these specific groups be targeted. When adopting a health care perspective, the same goes for the male and nonreligious adolescent subgroups.

In general, policy makers should be aware of the impact of the perspective chosen for the analysis on its outcomes. Omitting ICBs could negatively affect the reliability and informative value of analyses that are conducted from a societal perspective. Therefore, it is recommended that researchers should carefully make a priori considerations on the costs to be included because leaving out important costs could lead to biased results [60]. Finally, as in this study, high attrition rates could affect the outcomes of CEAs. High attrition rates are common in eHealth interventions [61,62]. It is recommended that more research should be conducted on adherence to eHealth interventions and that these interventions be implemented in practice, thus increasing their effectiveness, cost-effectiveness, and impact on public health.

Acknowledgments

This study was funded by ZonMW, the Netherlands Organisation for Health Research and Development (grant numbers: 200120006 and 200400010). Intervention development, data collection and analysis took place at Maastricht University. The game was developed in collaboration with Goal 043. The authors want to thank Claire Schoren, Kim Eijmael, and Mathieu van Adrichem for their contributions in developing the intervention.

Authors' Contributions

AJ, LM, and HV were in charge of the intervention design and data collection. RD, AP, DR, and SE provided advice on data collection and were in charge of conducting the economic analysis. All authors were involved in all stages of the development, analysis, and writing of the manuscript.

Conflicts of Interest

Hein de Vries is scientific director of Vision2Health, a company that licenses evidence-based innovative computer-tailored health communication tools. No other authors reported any conflicts of interest.

Multimedia Appendix 1

Additional File 1 List of limits of realistic answers to the questions in the cost questionnaires at T0 and T1.

[PDF File (Adobe PDF File), 35KB - jmir_v18i4e93_app1.pdf]

Multimedia Appendix 2

Additional File 2 Cost-effectiveness planes (left side) and corresponding CEACs (right side) of the economic evaluations based on the weekly alcohol use outcome measure, which were conducted from the health care perspective (upper), societal perspective including drugs and cigarette use (middle) and societal perspective excluding drugs and cigarette use (below) for the subgroups: male/female, younger/older, lower education/higher education, religious/non-religious, and Dutch/non-Dutch.

[PDF File (Adobe PDF File), 8MB - jmir_v18i4e93_app2.pdf]

Multimedia Appendix 3

Additional File 3 Cost-effectiveness planes (left side) and corresponding CEACs (right side) of the economic evaluations based on binge drinking occasions outcome measure, which were conducted from the health care perspective (upper), societal perspective including drugs and cigarette use (middle) and societal perspective excluding drugs and cigarette use (below) for the subgroups: male/female, younger/older, lower education/higher education, religious/non-religious, and Dutch/non-Dutch.

[PDF File (Adobe PDF File), 8MB - jmir_v18i4e93_app3.pdf]

References

1. Rehm J, Taylor B, Room R. Global burden of disease from alcohol, illicit drugs and tobacco. Drug Alcohol Rev 2006 Nov;25(6):503-513. [doi: 10.1080/09595230600944453] [Medline: 17132570]
2. Ezzati M, Lopez A, Rodgers A, Murray C. World Health Organization. 2004. Comparative quantification of health risks: global and regional burden of diseases attributable to selected major risks URL: http://www.who.int/healthinfo/global_burden_disease/cra/en/ [accessed 2016-04-05] [WebCite Cache ID 6gXdYQLrz]
3. Smit F, Lokkerbol J, Riper H, Majo MC, Boon B, Blankers M. Modeling the cost-effectiveness of health care systems for alcohol use disorders: how implementation of eHealth interventions improves cost-effectiveness. J Med Internet Res 2011;13(3):e56 [FREE Full text] [doi: 10.2196/jmir.1694] [Medline: 21840836]
4. World Health Organization. 2014. Global status report on alcohol and health URL: http://www.who.int/substance_abuse/publications/global_alcohol_report/en/ [accessed 2016-04-05] [WebCite Cache ID 6gXdf8yc4]
5. Mohapatra S, Patra J, Popova S, Duhig A, Rehm J. Social cost of heavy drinking and alcohol dependence in high-income countries. Int J Public Health 2010 Jun;55(3):149-157. [doi: 10.1007/s00038-009-0108-9] [Medline: 20024666]
6. Rehm J, Mathers C, Popova S, Thavorncharoensap M, Teerawattananon Y, Patra J. Global burden of disease and injury and economic cost attributable to alcohol use and alcohol-use disorders. Lancet 2009 Jun 27;373(9682):2223-2233. [doi: 10.1016/S0140-6736(09)60746-7] [Medline: 19560604]
7. Smit F, Cuijpers P, Oostenbrink J, Batelaan N, de GR, Beekman A. Costs of nine common mental disorders: implications for curative and preventive psychiatry. J Ment Health Policy Econ 2006 Dec;9(4):193-200. [Medline: 17200596]
8. World Health Organization. 2002. World report on violence and health URL: http://www.who.int/violence_injury_prevention/violence/world_report/en/ [accessed 2016-04-05] [WebCite Cache ID 6gXdisqW6]
9. Youth Violence and Alcohol. Geneva: World Health Organization; 2006. URL: http://www.who.int/violence_injury_prevention/violence/world_report/factsheets/fs_youth.pdf [accessed 2016-04-05] [WebCite Cache ID 6gXdmk4u8]
10. Chiauzzi E, Green TC, Lord S, Thum C, Goldstein M. My student body: a high-risk drinking prevention web site for college students. J Am Coll Health 2005;53(6):263-274 [FREE Full text] [doi: 10.3200/JACH.53.6.263-274] [Medline: 15900990]
11. Kypri K, Saunders JB, Williams SM, McGee RO, Langley JD, Cashell-Smith ML, et al. Web-based screening and brief intervention for hazardous drinking: a double-blind randomized controlled trial. Addiction 2004 Nov;99(11):1410-1417. [doi: 10.1111/j.1360-0443.2004.00847.x] [Medline: 15500594]

12. Lustria ML, Cortese J, Noar SM, Glueckauf RL. Computer-tailored health interventions delivered over the Web: review and analysis of key components. Patient Educ Couns 2009 Feb;74(2):156-173. [doi: 10.1016/j.pec.2008.08.023] [Medline: 18947966]
13. Matano RA, Koopman C, Wanat SF, Winzelberg AJ, Whitsell SD, Westrup D, et al. A pilot study of an interactive web site in the workplace for reducing alcohol consumption. J Subst Abuse Treat 2007 Jan;32(1):71-80. [doi: 10.1016/j.jsat.2006.05.020] [Medline: 17175400]
14. Riper H, Kramer J, Smit F, Conijn B, Schippers G, Cuijpers P. Web-based self-help for problem drinkers: a pragmatic randomized trial. Addiction 2008 Feb;103(2):218-227. [doi: 10.1111/j.1360-0443.2007.02063.x] [Medline: 18199300]
15. van der Wulp NY, Hoving C, Eijmael K, Candel MJ, van Dalen W, De Vries H. Reducing alcohol use during pregnancy via health counseling by midwives and internet-based computer-tailored feedback: a cluster randomized trial. J Med Internet Res 2014;16(12):e274 [FREE Full text] [doi: 10.2196/jmir.3493] [Medline: 25486675]
16. de Vries H, Brug J. Computer-tailored interventions motivating people to adopt health promoting behaviours: introduction to a new approach. Patient Educ Couns 1999 Feb;36(2):99-105. [Medline: 10223015]
17. Lemmens P, Tan ES, Knibbe RA. Measuring quantity and frequency of drinking in a general population survey: a comparison of five indices. J Stud Alcohol 1992 Sep;53(5):476-486. [Medline: 1405641]
18. van Keulen HM, Mesters I, Brug J, Ausems M, Campbell M, Resnicow K, et al. Vitalum study design: RCT evaluating the efficacy of tailored print communication and telephone motivational interviewing on multiple health behaviors. BMC Public Health 2008;8:216 [FREE Full text] [doi: 10.1186/1471-2458-8-216] [Medline: 18565222]
19. Smit ES, Evers SM, de Vries H, Hoving C. Cost-effectiveness and cost-utility of Internet-based computer tailoring for smoking cessation. J Med Internet Res 2013;15(3):e57 [FREE Full text] [doi: 10.2196/jmir.2059] [Medline: 23491820]
20. Kreuter MW, Wray RJ. Tailored and targeted health communication: strategies for enhancing information relevance. Am J Health Behav 2003 Dec;27 Suppl 3:S227-S232. [Medline: 14672383]
21. Schulz DN, Smit ES, Stanczyk NE, Kremers SP, de Vries H, Evers SM. Economic evaluation of a web-based tailored lifestyle intervention for adults: findings regarding cost-effectiveness and cost-utility from a randomized controlled trial. J Med Internet Res 2014;16(3):e91 [FREE Full text] [doi: 10.2196/jmir.3159] [Medline: 24650860]
22. Internet World Stats. 2014. Internet users in the world URL: http://www.internetworldstats.com/stats.htm [accessed 2016-12-09] [WebCite Cache ID 6c7pfXiGN]
23. Eurostat. 2015. Internet use statistics - individuals URL: http://ec.europa.eu/eurostat/statistics-explained/index.php/Internet_use_statistics_-_individuals [accessed 2015-10-08] [WebCite Cache ID 6c838lG2c]
24. Jander A, Crutzen R, Mercken L, Candel M, de Vries H. Effects of a web-based computer-tailored game to reduce binge drinking among dutch adolescents: a cluster randomized controlled trial. J Med Internet Res 2016;18(2):e29 [FREE Full text] [doi: 10.2196/jmir.4708] [Medline: 26842694]
25. Voogt CV, Kleinjan M, Poelen EA, Lemmers LA, Engels RC. The effectiveness of a web-based brief alcohol intervention in reducing heavy drinking among adolescents aged 15-20 years with a low educational background: a two-arm parallel group cluster randomized controlled trial. BMC Public Health 2013;13:694 [FREE Full text] [doi: 10.1186/1471-2458-13-694] [Medline: 23895403]
26. Suijkerbuijk A, Van Gils P, De Wit G. Bilthoven: National Institute for Public Health and the Environment. 2014. De kosteneffectiviteit van interventies gericht op verslaving aan alcohol en middelen URL: http://www.rivm.nl/dsresource?objectid=rivmp:251950&type=org&disposition=inline&ns_nc=1
27. Blankers M, Nabitz U, Smit F, Koeter MW, Schippers GM. Economic evaluation of internet-based interventions for harmful alcohol use alongside a pragmatic randomized controlled trial. J Med Internet Res 2012;14(5):e134 [FREE Full text] [doi: 10.2196/jmir.2052] [Medline: 23103771]
28. Ingels JB, Corso PS, Kogan SM, Brody GH. Cost-effectiveness of the strong African American families-teen program: 1-year follow-up. Drug Alcohol Depend 2013 Dec 1;133(2):556-561 [FREE Full text] [doi: 10.1016/j.drugalcdep.2013.07.036] [Medline: 23998376]
29. Nelson JP. Binge drinking and alcohol prices: a systematic review of age-related results from econometric studies, natural experiments and field studies. Health Econ Rev 2015;5:6 [FREE Full text] [doi: 10.1186/s13561-014-0040-4] [Medline: 25853004]
30. Drost RM, Paulus AT, Ruwaard D, Evers SM. Inter-sectoral costs and benefits of mental health prevention: towards a new classification scheme. J Ment Health Policy Econ 2013 Dec;16(4):179-186. [Medline: 24526586]
31. Jander A, Crutzen R, Mercken L, de Vries H. A Web-based computer-tailored game to reduce binge drinking among 16 to 18 year old Dutch adolescents: development and study protocol. BMC Public Health 2014;14:1054 [FREE Full text] [doi: 10.1186/1471-2458-14-1054] [Medline: 25301695]
32. Wechsler H, Dowdall GW, Davenport A, Rimm EB. A gender-specific measure of binge drinking among college students. Am J Public Health 1995 Jul;85(7):982-985. [Medline: 7604925]
33. de Vries H, Mudde A, Leijs I, Charlton A, Vartiainen E, Buijs G, et al. The European Smoking Prevention Framework Approach (EFSA): an example of integral prevention. Health Educ Res 2003 Oct;18(5):611-626 [FREE Full text] [Medline: 14572020]

34. Johannesson M. Avoiding double-counting in pharmacoeconomic studies. Pharmacoeconomics 1997 May;11(5):385-388. [Medline: 10168027]
35. Centraal Bureau voor de Statistiek. 2014. Statline URL: http://statline.cbs.nl/statweb/ [accessed 2016-05-06] [WebCite Cache ID 6c7pYgjVK]
36. Arkes H, Blumer C. The psychology of sunk cost. Organ Behav Hum Dec 1985 Feb;35(1):124-140. [doi: 10.1016/0749-5978(85)90049-4]
37. Hakkaart-van RL, Tan S, Bouwmans C. Methoden en referentieprijzen voor economische evaluaties in de gezondheidszorg. Diemen: College voor Zorgverzekeringen; 2011. Handleiding voor kostenonderzoek URL: http://www.zorginstituutnederland.nl/binaries/content/documents/zinl-www/documenten/publicaties/overige-publicaties/1007-handleiding-voor-kostenonderzoek/1007-handleiding-voor-kostenonderzoek/Handleiding+voor+kostenonderzoek.pdf [accessed 2015-10-09] [WebCite Cache ID 6c93WU4sJ]
38. Drost RM, Paulus AT, Ruwaard D, Evers SM. Maastricht University. 2014. Handleiding intersectorale kosten en baten van (preventieve) interventies: Classificatie, identificatie en kostprijzen URL: http://www.maastrichtuniversity.nl/web/file?uuid=140bf3cc-a0e8-41ea-be66-545663a93f62&owner=5b078ea8-4508-4eee-bb18-80a308782b25
39. Bouwmans CA, Schawo SJ, Jansen DE, Vermeulen KM, Reijneveld SA, Hakkaart-van Roijen L. iMTA Questionnaire Intensive Youth Care. 2012. URL: https://www.bmg.eur.nl/fileadmin/ASSETS/bmg/Onderzoek/Onderzoeksrapporten___Working_Papers/2012.06_-_Handleiding_Vragenlijst_Intensieve_Jeugdzorg.pdf [accessed 2016-11-20] [WebCite Cache ID 6c7qepCdu]
40. Cloïn M, Kamphuis C, Schols M, Tiessen-Raaphorst A, Verbeek D. The Hague: Sociaal Cultureel Planbureau. 2011. Nederland in een dag: Tijdsbesteding in Nederland vergeleken met die in vijftien andere Europese landen URL: http://www.scp.nl/Publicaties/Alle_publicaties/Publicaties_2011/Nederland_in_een_dag [accessed 2016-04-05] [WebCite Cache ID 6gXeqax6C]
41. Jellinek. Hoeveel kosten drugs? URL: https://www.jellinek.nl/vraag-antwoord/hoeveel-kosten-drugs/ [accessed 2015-10-08] [WebCite Cache ID 6c7pQEam8]
42. Killip S, Mahfoud Z, Pearce K. What is an intracluster correlation coefficient? Crucial concepts for primary care researchers. Ann Fam Med 2004;2(3):204-208 [FREE Full text] [Medline: 15209195]
43. Drummond M, O'Brien B, Stoddart G, Torrance G. Methods for the Economic Evaluation of Health Care Programmes. New York: Oxford University Press; 2005.
44. Cramer D, Howitt D. The Sage Dictionary of Statistics: A Practical Resource for Students in the Social Sciences. London: SAGE; 2004.
45. Centraal Bureau voor de Statistiek. 2015. Leefstijl en (preventief) gezondheidsonderzoek; persoonskenmerken URL: http://statline.cbs.nl/Statweb/publication/?DM=SLNL&PA=83021NED&D1=11-16&D2=a&D3=0&D4=l&VW=T [accessed 2016-08-20] [WebCite Cache ID 6c7pHUss6]
46. Byford S, Raftery J. Perspectives in economic evaluation. BMJ 1998 May 16;316(7143):1529-1530 [FREE Full text] [Medline: 9582152]
47. Jönsson B. Ten arguments for a societal perspective in the economic evaluation of medical innovations. Eur J Health Econ 2009 Oct;10(4):357-359. [doi: 10.1007/s10198-009-0173-2] [Medline: 19618224]
48. Knies S, Severens JL, Ament AJ, Evers SM. The transferability of valuing lost productivity across jurisdictions. differences between national pharmacoeconomic guidelines. Value Health 2010 Aug;13(5):519-527. [doi: 10.1111/j.1524-4733.2010.00699.x] [Medline: 20712601]
49. Ahmad S. Closing the youth access gap: the projected health benefits and cost savings of a national policy to raise the legal smoking age to 21 in the United States. Health Policy 2005 Dec;75(1):74-84. [doi: 10.1016/j.healthpol.2005.02.004] [Medline: 16298230]
50. Blankers M, Koeter MW, Schippers GM. Missing data approaches in eHealth research: simulation study and a tutorial for nonmathematically inclined researchers. J Med Internet Res 2010;12(5):e54 [FREE Full text] [doi: 10.2196/jmir.1448] [Medline: 21169167]
51. Eysenbach G. The law of attrition. J Med Internet Res 2005;7(1):e11 [FREE Full text] [doi: 10.2196/jmir.7.1.e11] [Medline: 15829473]
52. McKay HG, Danaher BG, Seeley JR, Lichtenstein E, Gau JM. Comparing two web-based smoking cessation programs: randomized controlled trial. J Med Internet Res 2008;10(5):e40 [FREE Full text] [doi: 10.2196/jmir.993] [Medline: 19017582]
53. Shahab L, McEwen A. Online support for smoking cessation: a systematic review of the literature. Addiction 2009 Nov;104(11):1792-1804. [doi: 10.1111/j.1360-0443.2009.02710.x] [Medline: 19832783]
54. Webb TL. Commentary on Shahab & McEwen (2009): Understanding and preventing attrition in online smoking cessation interventions: a self-regulatory perspective. Addiction 2009 Nov;104(11):1805-1806. [doi: 10.1111/j.1360-0443.2009.02751.x] [Medline: 19832784]
55. West R, Hajek P, Stead L, Stapleton J. Outcome criteria in smoking cessation trials: proposal for a common standard. Addiction 2005 Mar;100(3):299-303. [doi: 10.1111/j.1360-0443.2004.00995.x] [Medline: 15733243]
56. Acock A. Working with missing values. J Marriage Fam 2005 Nov;67(4):1012-1028. [doi: 10.1111/j.1741-3737.2005.00191.x]

57. Sobell LC, Cellucci T, Nirenberg TD, Sobell MB. Do quantity-frequency data underestimate drinking-related health risks? Am J Public Health 1982 Aug;72(8):823-828. [Medline: 7091478]
58. Devlin N, Parkin D. Does NICE have a cost-effectiveness threshold and what other factors influence its decisions? A binary choice analysis. Health Econ 2004 May;13(5):437-452. [doi: 10.1002/hec.864] [Medline: 15127424]
59. Neumann PJ, Cohen JT, Weinstein MC. Updating cost-effectiveness--the curious resilience of the $50,000-per-QALY threshold. N Engl J Med 2014 Aug 28;371(9):796-797. [doi: 10.1056/NEJMp1405158] [Medline: 25162885]
60. Evers SM, Hiligsmann M, Adarkwah CC. Risk of bias in trial-based economic evaluations: identification of sources and bias-reducing strategies. Psychol Health 2015 Jan;30(1):52-71. [doi: 10.1080/08870446.2014.953532] [Medline: 25141273]
61. de Vries H, Logister M, Krekels G, Klaasse F, Servranckx V, van Osch L. Internet based computer tailored feedback on sunscreen use. J Med Internet Res 2012;14(2):e48 [FREE Full text] [doi: 10.2196/jmir.1902] [Medline: 22547528]
62. Kohl LF, Crutzen R, de Vries NK. Online prevention aimed at lifestyle behaviors: a systematic review of reviews. J Med Internet Res 2013;15(7):e146 [FREE Full text] [doi: 10.2196/jmir.2665] [Medline: 23859884]

Abbreviations

CAU: care as usual
CEA: cost-effectiveness analysis
CEAC: cost-effectiveness acceptability curve
ICB: intersectoral costs and benefit
ICC: intracluster correlation coefficient
ICER: incremental cost-effectiveness ratio
MCAR: missing completely at random
NE: northeast quadrant
NW: northwest quadrant
RCT: randomized controlled trial
SE: southeast quadrant
SW: southwest quadrant
WTP max: maximum willingness to pay

Edited by G Eysenbach; submitted 09.10.15; peer-reviewed by F Vannieuwenborg, U Frick; comments to author 12.11.15; revised version received 23.12.15; accepted 07.01.16; published 21.04.16.

Please cite as:
Drost RMWA, Paulus ATG, Jander AF, Mercken L, de Vries H, Ruwaard D, Evers SMAA
A Web-Based Computer-Tailored Alcohol Prevention Program for Adolescents: Cost-Effectiveness and Intersectoral Costs and Benefits
J Med Internet Res 2016;18(4):e93
URL: http://www.jmir.org/2016/4/e93/
doi:10.2196/jmir.5223
PMID:27103154

Original Paper

A Patient-Centered Framework for Evaluating Digital Maturity of Health Services: A Systematic Review

Kelsey Flott[1], BA, MSc; Ryan Callahan[1], BA; Ara Darzi[1], PC, KBE, FRS, FMedSci, HonFREng; Erik Mayer[1], BSc, MBBS, MRCS, PhD, FRCS (Urol)

Centre for Health Policy, Institute of Global Health Innovation, Imperial College London, London, United Kingdom

Corresponding Author:
Kelsey Flott, BA, MSc
Centre for Health Policy
Institute of Global Health Innovation
Imperial College London
St Mary's Campus
Praed Street
London, W2 1NY
United Kingdom
Phone: 44 7909248515
Fax: 44 2075941484
Email: k.flott14@imperial.ac.uk

Abstract

Background: Digital maturity is the extent to which digital technologies are used as enablers to deliver a high-quality health service. Extensive literature exists about how to assess the components of digital maturity, but it has not been used to design a comprehensive framework for evaluation. Consequently, the measurement systems that do exist are limited to evaluating digital programs within one service or care setting, meaning that digital maturity evaluation is not accounting for the needs of patients across their care pathways.

Objective: The objective of our study was to identify the best methods and metrics for evaluating digital maturity and to create a novel, evidence-based tool for evaluating digital maturity across patient care pathways.

Methods: We systematically reviewed the literature to find the best methods and metrics for evaluating digital maturity. We searched the PubMed database for all papers relevant to digital maturity evaluation. Papers were selected if they provided insight into how to appraise digital systems within the health service and if they indicated the factors that constitute or facilitate digital maturity. Papers were analyzed to identify methodology for evaluating digital maturity and indicators of digitally mature systems. We then used the resulting information about methodology to design an evaluation framework. Following that, the indicators of digital maturity were extracted and grouped into increasing levels of maturity and operationalized as metrics within the evaluation framework.

Results: We identified 28 papers as relevant to evaluating digital maturity, from which we derived 5 themes. The first theme concerned general evaluation methodology for constructing the framework (7 papers). The following 4 themes were the increasing levels of digital maturity: resources and ability (6 papers), usage (7 papers), interoperability (3 papers), and impact (5 papers). The framework includes metrics for each of these levels at each stage of the typical patient care pathway.

Conclusions: The framework uses a patient-centric model that departs from traditional service-specific measurements and allows for novel insights into how digital programs benefit patients across the health system.

Trial Registration: N/A

(J Med Internet Res 2016;18(4):e75) doi:10.2196/jmir.5047

KEYWORDS

digital maturity; evaluation; health information exchange; patient-centered care

Introduction

Digital technologies are transforming health services by providing new mechanisms for accessing personal medical records, submitting incident reports, and communicating across care settings. In England, the government has placed the role of these technologies high on the agenda by setting a 3-year target for a fully paperless National Health Service (NHS). The *NHS Five Year Forward View*, which sets the strategic direction for the health service in England, focuses heavily on improving NHS digital technology with the aim of integrating all electronic medical records (EMRs) [1]. Furthermore, the continuing ambition of integrating health and social care relies on connected information technology (IT) systems. Amid this political climate and the influx of digital technologies, the potential for improvement is vast and certainly not limited to service operations; there is also scope to significantly improve the patient experience. There is an opportunity to depart from traditional service arrangements, overcome geographical boundaries, and even reconfigure services around the patients and their needs by harnessing digital technologies. However, these are not automatic byproducts of augmented IT programs. Like any care intervention, digital technologies need to be rigorously evaluated and monitored to ensure they operate in the way they are intended and cultivate a better experience across patient pathways. To conduct constructive appraisals, an evaluation framework is needed to make sure each factor that influences a digital system’s success is captured.

Digital maturity—the extent to which digital technologies are used as enablers to deliver a high-quality health service—is an emerging concept across developed health care systems, and there is no established measurement that accounts for all of its intricacies [2]. The adoption of digital solutions for EMRs across care services in Canada (facilitated by the Canada Health Infoway) provides existing examples of digital systems that have achieved a high level of maturity. However, the discourse surrounding digital maturity is dominated by ideas for its potential to improve services in the future. For instance, digital maturity is extensively cited as an aspirational goal necessary to join IT systems across care settings in order to effectively integrate health and social care services [3].

Example frameworks for evaluation from the health sector often focus on the operational benefits that individual services receive from digital technologies [4-6]. However, digital maturity is not about the success of one technological system and the benefits to one service’s particular stakeholders; rather, it is the advancement of the entire health service. To improve patient experience across care pathways, digital maturity must be measured in a way that addresses all the intersections it has across care settings and must be conceptualized in a way that dissolves the entrenched, service-specific standard for how digital systems are assessed.

The aim of our study was to propose a novel framework for evaluating digital maturity based on these principles. We summarized the existing evidence about how best to evaluate digital maturity and its component parts, and the merits of current digital maturity evaluations. Moreover, we synthesized evidence around what should be reflected in evaluations and developed a new framework capable of measuring digital maturity across the patient pathway and presenting it as a patient-centric, sectorwide achievement.

Theory of Digital Maturity

The concept of digital maturity originated in the field of public service improvement. As more government services became IT enabled, they did so in siloes, meaning a single user (the citizen) would have to attempt to access information or obtain services from several different departments. Although the term digital maturity was not used at the time, the framework proposed by Layne and Lee in 2001 had 4 stages of integration to bring about citizen-centric e-government [7]. Their framework demonstrates that at its most mature level e-government represents the advancements and interworking of an entire field for the citizen; in health care this would be translated into the advancement of an entire field respectful of the way patients experience care.

Digital maturity builds on existing evidence about digital literacy. In healthcare this can primarily be captured in the idea of eHealth literacy or the ability of people to use information and communications technologies to improve or enable health care [8]. For health systems to support advancement of digital maturity, staff must be digitally literate and help patients improve their digital literacy. It is important to note, however, that while digital literacy can facilitate digital maturity, digital maturity should also be responsive to the whole patient population and account for their needs regardless of their digital literacy.

Gottschalk [9] and colleagues’ research moved beyond the concept of usability and linked the concept of maturity to greater interoperability. Their research posited a trajectory for how the content of interoperability tends to evolve—or mature—over time from the simplest level to more advanced, more integrated levels organized around the citizen. Because of the shift to organize services around the citizen, organizations would need to be increasingly interoperable, not just in terms of technical issues, but also in realizing benefits and setting goals. This is also mirrored in the health service as it becomes more integrated and individual services are required to communicate effectively across the patient pathway.

According to this research, digital maturity encompasses not only the resources and ability to use a system, but also how interoperable it is with other systems and ultimately its impact on the public. To understand how these aspects of digital maturity can be measured in the health service, a substantial body of literature provides guidance on how to build an evaluation framework for an information system.

Methods

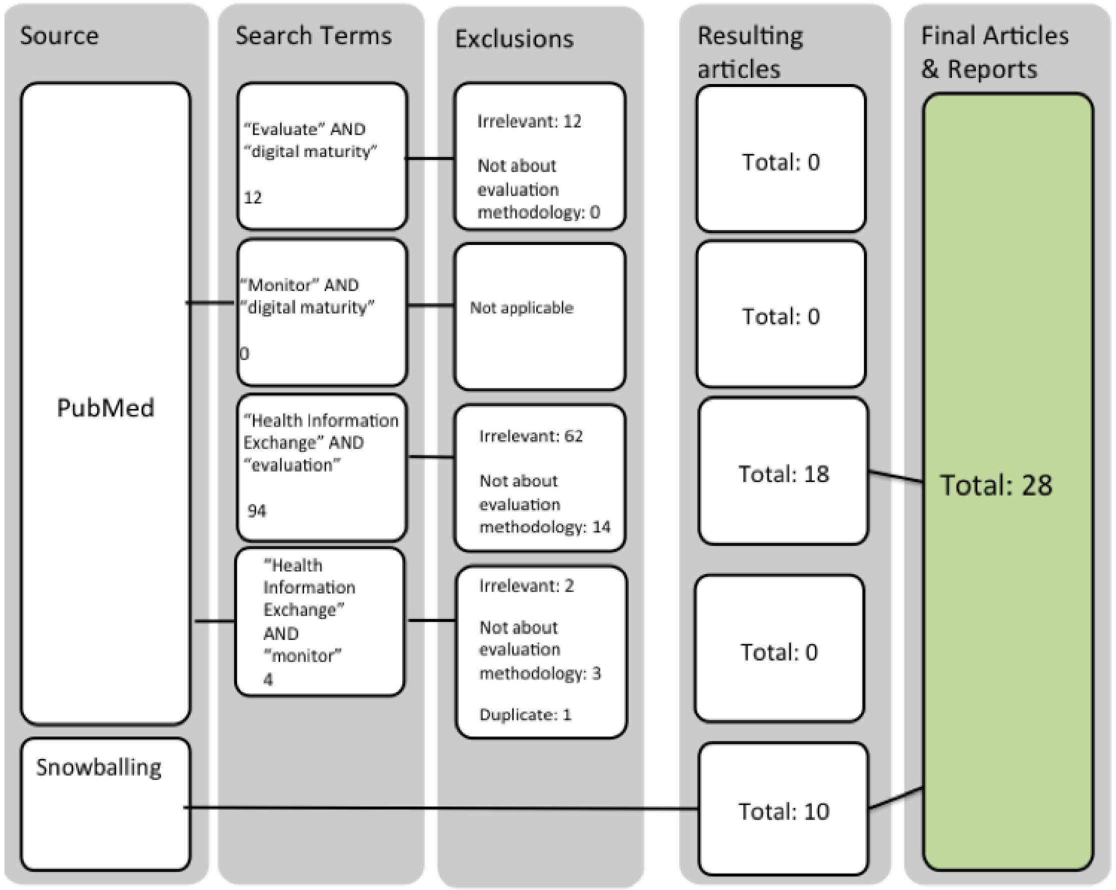
Search Strategy

We searched the PubMed database for literature relevant to constructing a digital maturity evaluation framework and reports about the most prominent international examples of existing evaluation frameworks. The search strategy for locating sources included broad terms such as “evaluate” AND “digital maturity,” as well as more specific terms such as “health information exchange” AND “evaluation.” These specific terms were essential because, in the health service, digital maturity is usually misrepresented and not thought of as a universal advancement, but rather is defined by the individual digital systems or programs that aim to support maturity. Health information exchanges (HIEs), while still only a part of digital maturity, share similar goals with overall maturity in that they aim to mobilize health care information electronically across systems. Therefore, HIE was a useful proxy term for digital maturity in this literature search. We filtered results for relevance to health care and medicine. Finally, we intentionally made the search strategy broad to return papers about all evaluations of digital maturity whether they were specific to one care setting or across care pathways. We searched the PubMed database for literature published between 1995 and 2015 using the following Boolean search strings: (1) “evaluate” AND “digital maturity”, (2) “monitor” AND “digital maturity”, (3) “health information exchange” AND “evaluation”, (4) “health information exchange” AND “monitor.”

Review Strategy

As the flow chart in Figure 1 shows, the search strategy returned over 110 papers, but most were either irrelevant to the goals of the review or did not provide information that contributed to a better understanding of digital maturity evaluation. Inclusion criteria specified that papers must be pertinent to the health service and maintain a focus on the evaluation and some component of digital maturity, such as HIEs. We applied exclusion criteria if papers solely concerned the experience or outcomes associated with digitally mature systems rather than evaluation. Papers were also excluded if they only reported the results of an evaluation and contained no methodological insight. The search returned 18 peer reviewed papers, from which snowballing techniques yielded 7 more peer reviewed papers and 3 relevant reports, totaling 28 included sources.

Figure 1. Literature search and review strategy flow chart.



Analysis and Framework Development

We analyzed the papers to draw out their contribution to digital system evaluation. This included identifying how to evaluate systems, what to measure, and at what point. It also included a specific examination of how the patient perspective was accounted for within the evaluation. For instance, this included exploring whether the evaluation methodology discussed how to make sure a system was evaluated across the patient pathway or if indicators of success were relevant to patients.

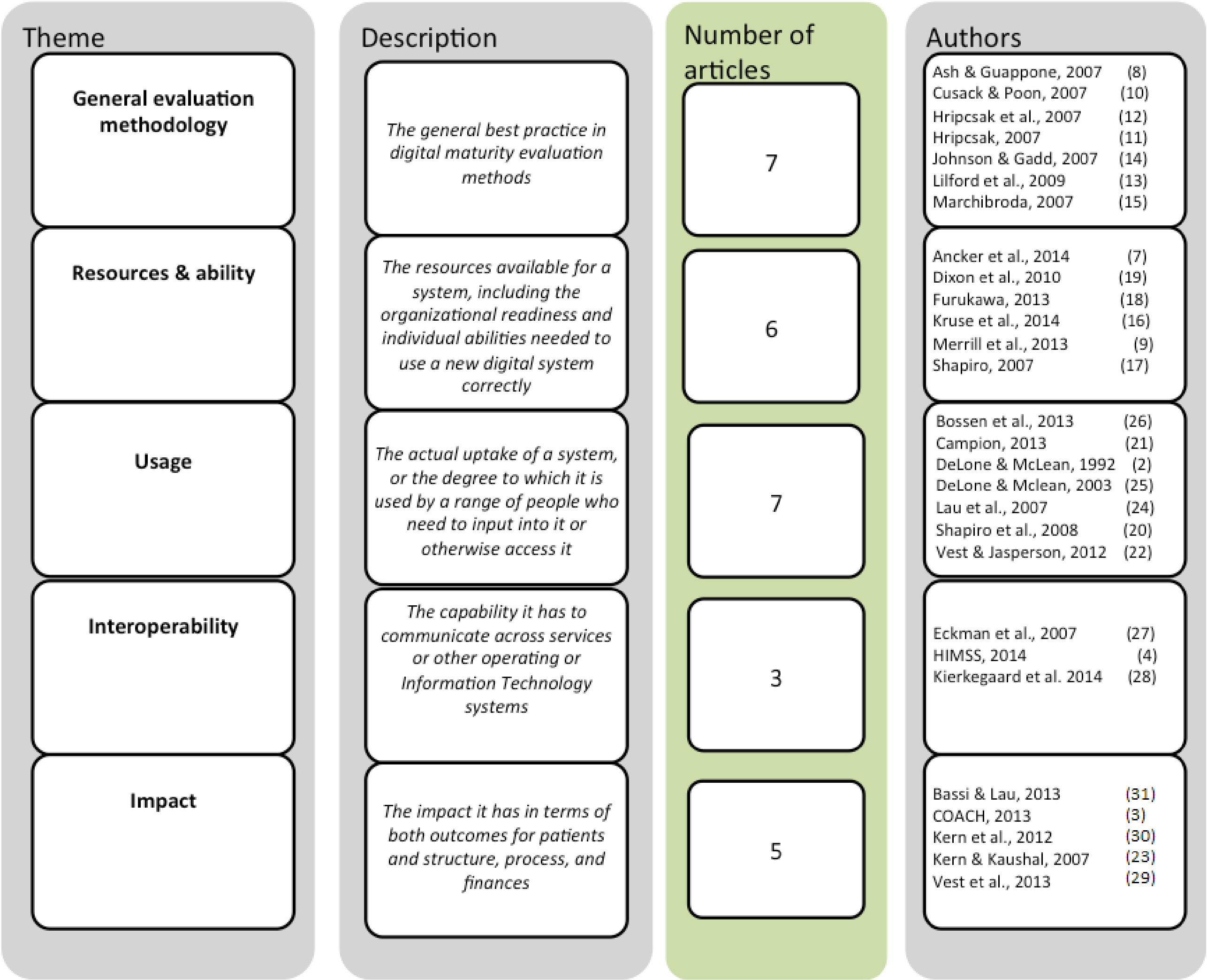
We then used the results from the literature search and review to construct a framework. During the analysis we drew out 5 themes (these themes and associated papers are discussed in the Results section). We translated these themes into the levels of analysis for the evaluation framework and used the indicators within the themes as individual scoring points within each level of analysis (see Framework Development subsection in the Discussion). The resulting framework is situated within a patient-centered paradigm, meaning it measures each of the levels at each major point along the care pathway. This is a departure from existing measurements of digital systems that are rooted in the service-specific goals and are therefore limited in their scope to provoke improvements across the whole care pathway.

Results

We analyzed findings from the 28 papers identified in the literature search and grouped them into 5 themes to generate the skeleton of a new digital maturity framework: general evaluation methodology, resources and ability, usage, interoperability, and impact. These are detailed in Figure 2.

The literature review confirmed the importance of measuring digital maturity in a way that accounts for its multidimensional nature. The 28 papers reviewed fell into 5 distinct themes, similar to those identified in the background about the concept of digital maturity. The 5 themes—general evaluation methodology, resources and ability, usage, interoperability, and impact—received varying levels of attention in the literature, with the final 2 being the most limited. We discuss these themes, and their associated indicators identified from the literature, in detail to provide context to the development of the evaluation framework. Across all of these themes most papers did not refer to the patient centricity of their approach but rather assumed a service-centric approach, meaning the evaluation considered only one or two settings of care. Although some indicated the importance of certain indicators to patients, any discussion about how the evaluation accounted for success across entire care pathways was distinctly absent.

Figure 2. Themes from literature review.



Theme	Description	Number of articles	Authors
General evaluation methodology	*The general best practice in digital maturity evaluation methods*	7	Ash & Guappone, 2007 (8) Cusack & Poon, 2007 (10) Hripcsak et al., 2007 (12) Hripcsak, 2007 (11) Johnson & Gadd, 2007 (14) Lilford et al., 2009 (13) Marchibroda, 2007 (15)
Resources & ability	*The resources available for a system, including the organizational readiness and individual abilities needed to use a new digital system correctly*	6	Ancker et al., 2014 (7) Dixon et al., 2010 (19) Furukawa, 2013 (18) Kruse et al., 2014 (16) Merrill et al., 2013 (9) Shapiro, 2007 (17)
Usage	*The actual uptake of a system, or the degree to which it is used by a range of people who need to input into it or otherwise access it*	7	Bossen et al., 2013 (26) Campion, 2013 (21) DeLone & McLean, 1992 (2) DeLone & Mclean, 2003 (25) Lau et al., 2007 (24) Shapiro et al., 2008 (20) Vest & Jasperson, 2012 (22)
Interoperability	*The capability it has to communicate across services or other operating or Information Technology systems*	3	Eckman et al., 2007 (27) HIMSS, 2014 (4) Kierkegaard et al. 2014 (28)
Impact	*The impact it has in terms of both outcomes for patients and structure, process, and finances*	5	Bassi & Lau, 2013 (31) COACH, 2013 (3) Kern et al., 2012 (30) Kern & Kaushal, 2007 (23) Vest et al., 2013 (29)

General Evaluation Methodology

Many papers emphasized the multifaceted nature of digital maturity specifically for HIEs and the deliberate way evaluations must account for this. Similar to the depictions of digital maturity above, not only do HIEs have many different components, but they also take different forms and have different capacities as they mature [10-13]. Barriers to development, issues around leadership, and commitment are paramount to early-stage evaluations, while stakeholder motivations for engaging with the system are more important as the system advances [11]. As Hripcsak [14] noted,

> *An HIE project undergoes a series of steps, from early conception to mature maintenance. Reviewing the steps can uncover possible unintended and unexpected effects.*

Following a United Hospital Fund (New York, USA) meeting in 2006 to review the best practice for evaluating HIE programs, a series of research projects made inroads into a common evaluation approach and established that their evolutionary nature must be central to appraisals [14]. In other words, using a rigid scoring system or a single metric to evaluate such dynamic and fluid systems is untenable; rather, evaluations must use a comprehensive framework approach [15]. Although a framework should aim to be as objective as possible, considerable evidence suggests that mixed methods add invaluable richness to the data and improve its ability to drive improvement policy [11,16]. Given that assessing how much digital maturity has improved health care quality is very complex, an evaluation framework needs to incorporate nuanced, qualitative feedback from staff and patients [11]. One influential proposal for evaluating digital systems comprehensively and incorporating mixed methods was the “smallball” approach, which disaggregates the components of HIEs and measures them individually across different points in time [17]. Ultimately, digital maturity and the details of operational strategy ought to be closely linked. Just as there can be no single metric for strategy evaluation, findings from the literature explain the futility in narrowing an evaluation framework. A narrow framework would compromise its ability to measure the entirety of digital maturity across the patient pathway in delivering on strategic health system priorities.

In addition to these evaluation strategies, the literature also reveals that HIEs and digital maturity have a wide range of stakeholders to whom each of the components matter differently. Comprehensive evaluation is not just the inclusion of a variety of metrics, but also separate assessments of those metrics as they pertain differently to patients, providers, and policy makers. A framework must span care settings and produce measures from all service levels, as well as the regional and national levels [18].

Resources and Ability

The studies described above compel the development and use of a multifaceted evaluation framework, and a related body of literature suggests what metrics the framework should incorporate. This research indicates the importance of including readiness measures as a way to assess the extent to which HIEs evolve within an environment conducive to their success. This includes measures of organizations’ existing technology, cultural norms, and leadership to provide context to system functioning and user uptake [10].

More specifically, studies demonstrate the value of gathering information about institutional resources and existing programs that, if insufficient or hostile to the entrance of an HIE, hinder the success of digital systems [19,20]. This refers to finances, staff capacity, experience and willingness, and the existing protocols for information exchange [21]. For instance, inconsistent goals, project rework, and underdeveloped resources are chronic barriers to HIE success, which should be accounted for in the evaluation [12]. Further research explains the importance of quantifying implementation effort when the system is younger, and advancing into usage and then cost metrics only as the system matures [22].

Usage Measures

Usage can be evaluated in a variety of ways: it can be defined as the volume of information transmitted, the duration and specific activity of users, or simply as the number of login sessions [23,24]. Campion and colleagues [24] used login sessions to explore 3 communities in New York State and the differences in usage among the various stakeholders. They explained that one important measure is whether patient summary data are displayed by default on logging into the system, as this was one of the most significant predictors of patient usage [24].

One of the largest studies about measuring HIE usage demonstrates that usage measurements can be more robust and meaningful when differentiated by activity during use [25]. Vest and Jasperson [25] stratified usage into 5 classifications: minimal usage, repetitive searching, clinical information, mixed information, and demographic information. These types of usage varied by the user’s role. For instance, minimal usage was highest among physicians and clinical information was highest among nurses. This is critical to include in an evaluation framework, as it quantifies how the system is being used and to whom changes would be most impactful [25]. These different types of usage, and their associations with specific roles, help target improvement strategy and evaluate where return on investment can be maximized.

Usability needs to be at the forefront of design, as engendering use is the crux of system development and, without it, indicators at later stages of maturation might be irrelevant [26]. Some evaluation strategies already exist to measure this area. For instance, the information benefits evaluation framework, a nonindustry-specific evaluation model used to evaluate HIEs in Canada, addresses usage and usage type extensively [4,27]. However, while it is appropriate for measuring usage, this framework is not necessarily suitable for measuring digital maturity as a sectorwide advancement for patients, as it neglects organizational, cultural, policy, and other external factors [27]. DeLone and McLean, authors of the parent system, more recently conducted a 10-year review of the system and suggested that a renovated system should include service quality as a new dimension of information system success, which is intended to enhance its applicability throughout the health service [28,29].

Interoperability

Moving beyond usage, studies have suggested measures for evaluating digital systems’ ability to communicate across settings, as this is central to range and depth of their impact in an integrated, patient-centric health service. In addition to syntactic interoperability, this also includes semantic interoperability. Semantic interoperability, or the harmonization of clinical terminology across care providers, settings, and systems, is particularly important for developing a workable exchange of information [30].

The interoperability of systems for all stakeholders is vital to the systems’ effectiveness in terms of achieving an apparatus for patient centricity; however, research suggests that HIEs are often faulty, with poor ability to communicate across service settings and care sectors [31]. Additionally, evaluating to what extent HIEs are connected across geographic regions is crucial to appraising digital maturity in its entirety, rather than a series of one-off exchanges of information [32].

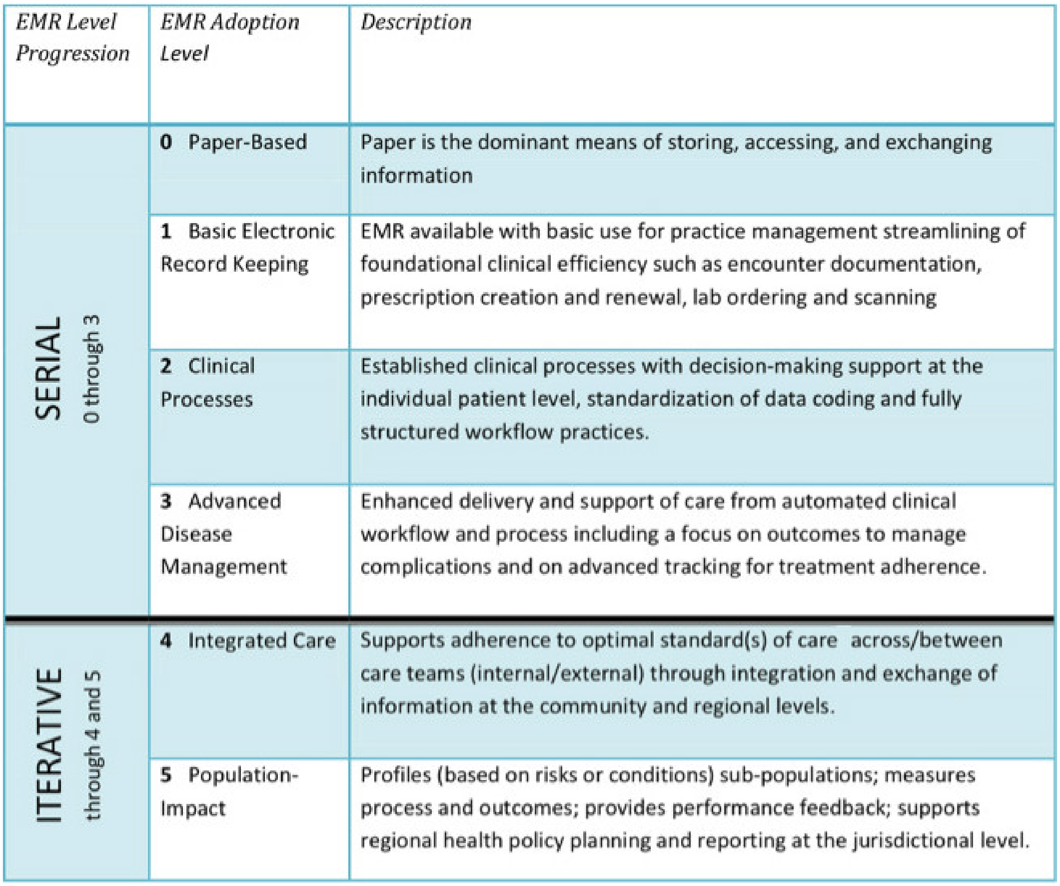
A health care-specific model for evaluating digital systems is the continuity of care maturity model, which addresses the “convergence of interoperability, information exchange, care coordination, patient engagement and analytics” [6]. It builds on the EMR adoption model, a framework that helps services benchmark their success in using EMRs. This model does move toward a whole systems approach, as it demonstrates the importance of measuring each point along the digital maturity continuum, from establishing electronic systems, to making sure they are interoperable, to evaluating their effectiveness. However, at its core it is designed to be used by individual services to improve their digital functioning, and consequently it has no mechanism for detecting holes in maturity in other services or care settings that might affect overall maturity of the system. This undermines its ability to capture digital maturity holistically and generate improvements that will be relevant across the patient pathway.

Impact

Moving toward impact, studies suggest that it is necessary to depart from service-specific measures and assess impact across the health care landscape [26]. This can be done by determining how the information structure or other digital program offers public utility [33]. Impact in terms of cost can also be evaluated across stakeholders. Cost metrics focus on the functionalities enabled by HIEs that save money, including HIEs’ ability to generate alerts when there are orders placed for expensive medications or redundant laboratory orders [34]. A systematic review of HIE cost evaluations demonstrates that cost savings were associated with HIE use to a small degree, but there needed to be better, more standardized ways of measuring and reporting cost evaluations [35].

Canada’s Health Informatics Association (COACH), as Figure 3 [5] shows, reviewed 4 impact evaluations of EMR systems. They proposed an EMR-specific evaluation framework that progresses from serial to iterative stages. Although this evaluation model exemplifies measuring a system across a variety of metrics over a period of maturity, it also does not necessarily capture digital maturity as a multiservice or whole systems concept. As a result, it cannot always indicate whether the digital system under investigation is compromised by lags in other care settings.

Figure 3. Canadian electronic medical records (EMR) adoption and maturity model. Reproduced with permission from Canada's Health Informatics Association (COACH) [3].



EMR Level Progression	*EMR Adoption Level*	*Description*
SERIAL 0 through 3	**0** Paper-Based	Paper is the dominant means of storing, accessing, and exchanging information
	1 Basic Electronic Record Keeping	EMR available with basic use for practice management streamlining of foundational clinical efficiency such as encounter documentation, prescription creation and renewal, lab ordering and scanning
	2 Clinical Processes	Established clinical processes with decision-making support at the individual patient level, standardization of data coding and fully structured workflow practices.
	3 Advanced Disease Management	Enhanced delivery and support of care from automated clinical workflow and process including a focus on outcomes to manage complications and on advanced tracking for treatment adherence.
ITERATIVE through 4 and 5	**4** Integrated Care	Supports adherence to optimal standard(s) of care across/between care teams (internal/external) through integration and exchange of information at the community and regional levels.
	5 Population-Impact	Profiles (based on risks or conditions) sub-populations; measures process and outcomes; provides performance feedback; supports regional health policy planning and reporting at the jurisdictional level.

Discussion

Limitations

The remit of PubMed (the database we used to search for papers) includes publications on health care, life sciences, and biomedicine; while it is unlikely that PubMed excluded any relevant sources, we could have augmented the results of the literature review through searching in other databases. More important, relevant sources may exist in harder-to-identify gray literature. Furthermore, although we broadened the search terms to be as inclusive as possible, there is still a bias toward digital maturity for information exchange rather encompassing digital maturity application in all aspects of service delivery

Analysis

Despite the wealth of information in the literature about how best to evaluate the components of digital maturity, these have not been synthesized into a digital maturity evaluation framework. The first theme, the general evaluation methodology, demonstrates a strong emphasis in the literature on digital maturity being a multidimensional concept. Furthermore, it is clear from the other 4 themes—resources and ability, usage measures, interoperability, and impact—that these must all be captured to evaluate digital maturity holistically. Equipped with the general evaluation methodology and indicators of the 4 levels of maturity, it is necessary to establish at what level they should be measured. The literature also exposes the range of stakeholders who benefit from digital maturity and the growing need in the health service to account for each of the 4 dimensions across different points along the care pathway.

However, our analysis found no evidence in the existing literature of attempts to evaluate digital maturity across the entire care pathway despite the fact that this is necessary to account for all indicators of success. The tendency of existing evaluations, as evidenced by the information benefits framework, the continuity of care maturity model, and the COACH models, is that they place the goals of a specific service at the center of the evaluation, or otherwise fail to capture information about maturity across the whole system. When a framework like this is applied at the service level, or to assess the maturity of a single program within a single service, it risks overlooking the external issues that are most fundamental to the success of a system and overall digital maturity. For instance, these models do not capture where there are resource shortages in one setting or poor interoperability in another. This is critical, as evidence suggests substantial disparities in the quality and use of digital technologies across care settings. A recent survey (Centre for Health Policy, Imperial College London, unpublished data, 2015) demonstrated that IT in secondary care lags the IT systems in primary care, and that secondary care doctors do not have access to all necessary records from primary care. This is not only because primary care records are not interoperable, but also because the digital resources within secondary care are not as advanced.

The evaluation of digital maturity should be able to map the advances of digital systems such that they can facilitate integrated care, better coordination, and improved patient experience across a whole pathway. Measures produced from existing frameworks cannot indicate improvements that would have a sectorwide benefit and be most meaningful to patients who experience care as a pathway, not an individual service. The framework proposed below is designed to measure digital maturity as a driver of integration and improved patient experience. It works from a paradigm with patients at the epicenter, surrounded by their community, primary care, and secondary care to mirror the patient pathway and measure digital maturity across it.

Framework Development

As opposed to current approaches, this paradigm enables a new way of thinking about the 4 primary areas of digital maturity across a patient pathway. Thinking about the service landscape like this, with the goals and needs of the patient at the center, it becomes possible to envisage new digital solutions that cross the boundaries of traditional service arrangements. Our framework for evaluating systems across all care settings promotes more holistic quality improvement and surpasses frameworks that assess the impact for patients in only one portion of their pathway. We used the 5 themes we identified in the literature review to build a comprehensive framework; we used the general evaluation methodology theme to inform the approach to framework construction, while we translated the other 4 themes into 4 levels of maturity to measure across the patient pathway.

The framework (Multimedia Appendix 1) presents open-ended questions for an evaluator to answer. The open-ended quality is deliberate, as it is meant to provide a standard set of questions but allow for nuances to be captured. The scope for qualitative appraisal is important to measure the evolving progress of digital maturity; however, it is also useful to apply a metric for success. The framework operates in a user-friendly way, in that each question can receive 1 point for a positive answer, meaning each column and row will have a total score. Evaluators can pinpoint areas for improvement by looking at individual negative scores and can identify broader areas that need work through the column and row scores. The overall score out of 58 also provides a comprehensive score for longitudinal benchmarking.

For example, it is possible to analyze mHealth in this framework. mHealth uses phones or tablets to collect patient data including vital signs and general personal health monitoring. mHealth technology can then be used to relay these data to clinicians and community health workers in near real time. Data can also be shared between health settings. If applied well, mHealth technologies have a distinct potential for digital maturity. Such technologies could be evaluated in the home setting to see whether they work for the target population, are easy to use for all patients in the target population, whether they work with the devices patients already have and know how to use, and whether they make a difference without an overwhelming cost or burden. The maturity of the mHealth technologies could then be evaluated at the community level for data transferred to community health workers and the primary and secondary care levels to understand its maturity across the care pathway.

While this framework could be used to evaluate the success of a digital system at 1 service level by isolating 1 row, the overall maturity score is dependent on a sectorwide patient understanding. It highlights where gaps in maturity exist, which presents an opportunity to address the specific shortcomings. This allows subsequent improvement work to be thoroughly patient centric, as it will be intended to support digital maturity across their care pathway. If there is a gap in any box, a solution can be designed that will have a ripple effect and bolster the whole maturity score; this precipitates digital maturity along a patient pathway, so that success in one area is not stymied by failure in another.

Next Steps

Understanding the utility of this framework will require a trial period of applying it in different health care contexts and comparing scores on certain digital systems. This will help identify whether any areas are missing from the framework. Furthermore, it will be important to gather feedback from evaluators to make sure that the framework is user friendly and well received. Application of the framework has the potential to introduce a standard approach to evaluating digital maturity and mobilize internal benchmarking of digital systems' maturity.

Conclusion

The idea that digital maturity is a sectorwide advancement centered on a principal group has been established in other sectors as evidenced by the citizen-centric method of evaluation in e-government. However, in the health sector the progression and success of digital systems has been measured primarily within the confines of individual services' or care settings' performance. This does not capture the entirety of digital maturity but, more important, it does not indicate whether a new digital system is capable of helping patients at all points along their pathway.

In order for advancements in digital technologies to proliferate patient benefit in terms of care coordination and enhanced information, digital maturity needs to be conceptualized as a sectorwide, patient-centric measure. Using the literature available on theories behind how to measure the parts of digital maturity, best practice on how to gather indicators of its component parts, and examples of existing evaluation frameworks, our study proposes a contemporary framework that captures 4 key domains of digital maturity across the patient pathway, to pinpoint how digital maturity can be most meaningfully improved.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Digital Maturity Evaluation Framework.

[PDF File (Adobe PDF File), 39KB - jmir_v18i4e75_app1.pdf]

References

1. NHS England. Five Year Forward View Internet. London; 2014 Oct. URL: https://www.england.nhs.uk/wp-content/uploads/2014/10/5yfv-web.pdf [accessed 2016-03-14] [WebCite Cache ID 6g001blqz]
2. National Information Board. Personalised Health and Care 2020: Using Data and Technology to Transform Outcomes for Patients and Citizens: A Framework for Action. 2014. URL: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/384650/NIB_Report.pdf [accessed 2016-03-14] [WebCite Cache ID 6g02aLh8o]
3. Ham C, Curry N. Integrated Care: What Is It? Does It Work? What Does It Mean for the NHS?. London, UK: The King's Fund; 2011.
4. DeLone WH, McLean ER. Information systems success: the quest for the dependent variable. Inf Syst Res 1992;3:60-95 [FREE Full text] [WebCite Cache ID http://herbsleb.org/SCALEpapers/delone-information-1992.pdf]
5. COACH Canada's Health Informatics Association. Canadian EMR Adoption and Maturity Model: A Multi-Jurisdiction Collaborative and Common EMR Adoption and Maturity Model. 2013 Feb. URL: http://www.coachorg.com/en/resourcecentre/EMR-Adoption-White-Paper.asp [accessed 2016-03-14] [WebCite Cache ID 6g00JhZl7]
6. HIMSS. Continuity of Care Maturity Model Going beyond EMRAM 2014. 2014. URL: http://www.himssanalytics.org/provider-solutions [accessed 2016-03-14] [WebCite Cache ID 6g00MJTi8]
7. Layne K, Lee J. Developing fully functional E-government: a four stage model. Government Inf Q 2001 Jun;18(2):122-136. [doi: 10.1016/S0740-624X(01)00066-1]
8. Neter E, Brainin E. eHealth literacy: extending the digital divide to the realm of health information. J Med Internet Res 2012;14(1):e19 [FREE Full text] [doi: 10.2196/jmir.1619] [Medline: 22357448]
9. Gottschalk P. Maturity levels for interoperability in digital government. Government Inf Q 2009 Jan;26(1):75-81. [doi: 10.1016/j.giq.2008.03.003]
10. Ancker J, Miller M, Patel V, Kaushal R. Sociotechnical challenges to developing technologies for patient access to health information exchange data. J Am Med Inform Assoc 2014;21(4):664-670 [FREE Full text] [doi: 10.1136/amiajnl-2013-002073] [Medline: 24064443]

11. Ash JS, Guappone KP. Qualitative evaluation of health information exchange efforts. J Biomed Inform 2007 Dec;40(6 Suppl):S33-S39 [FREE Full text] [doi: 10.1016/j.jbi.2007.08.001] [Medline: 17904914]
12. Merrill J, Deegan M, Wilson R, Kaushal R, Fredericks K. A system dynamics evaluation model: implementation of health information exchange for public health reporting. J Am Med Inform Assoc 2013 Jun;20(e1):e131-e138 [FREE Full text] [doi: 10.1136/amiajnl-2012-001289] [Medline: 23292910]
13. Cusack C, Poon E. Health Information Technology Evaluation Toolkit (AHRQ Publication No. 08-0005-EF).: Agency for Healthcare Research and Quality; 2007 Oct. URL: https://www.healthit.gov/unintended-consequences/sites/default/files/pdf/ModuleIIpdf1.5.pdf [accessed 2016-03-14] [WebCite Cache ID 6g00oM3N6]
14. Hripcsak G. Developing common methods for evaluating health information exchange. J Biomed Inform 2007 Dec;40(6 Suppl):S1-S2 [FREE Full text] [doi: 10.1016/j.jbi.2007.08.005] [Medline: 17911045]
15. Hripcsak G, Kaushal R, Johnson KB, Ash JS, Bates DW, Block R, et al. The United Hospital Fund meeting on evaluating health information exchange. J Biomed Inform 2007 Dec;40(6 Suppl):S3-10 [FREE Full text] [doi: 10.1016/j.jbi.2007.08.002] [Medline: 17919986]
16. Lilford R, Foster J, Pringle M. Evaluating eHealth: how to make evaluation more methodologically robust. PLoS Med 2009 Nov;6(11):e1000186 [FREE Full text] [doi: 10.1371/journal.pmed.1000186] [Medline: 19956674]
17. Johnson KB, Gadd C. Playing smallball: approaches to evaluating pilot health information exchange systems. J Biomed Inform 2007 Dec;40(6 Suppl):S21-S26 [FREE Full text] [doi: 10.1016/j.jbi.2007.08.006] [Medline: 17931981]
18. Marchibroda JM. Health information exchange policy and evaluation. J Biomed Inform 2007 Dec;40(6 Suppl):S11-S16 [FREE Full text] [doi: 10.1016/j.jbi.2007.08.008] [Medline: 17981099]
19. Kruse C, Regier V, Rheinboldt K. Barriers over time to full implementation of health information exchange in the United States. JMIR Med Inform 2014;2(2):e26 [FREE Full text] [doi: 10.2196/medinform.3625] [Medline: 25600635]
20. Shapiro JS. Evaluating public health uses of health information exchange. J Biomed Inform 2007 Dec;40(6 Suppl):S46-S49 [FREE Full text] [doi: 10.1016/j.jbi.2007.08.003] [Medline: 17919985]
21. Furukawa M. Meaningful use: a roadmap for the advancement of health information exchange. Isr J Health Policy Res 2013;2(1):26 [FREE Full text] [doi: 10.1186/2045-4015-2-26] [Medline: 23880399]
22. Dixon B, Zafar A, Overhage J. A Framework for evaluating the costs, effort, and value of nationwide health information exchange. J Am Med Inform Assoc 2010;17(3):295-301 [FREE Full text] [doi: 10.1136/jamia.2009.000570] [Medline: 20442147]
23. Shapiro JS, Vaidya SR, Kuperman G. Preparing for the evaluation of health information exchange. AMIA Annu Symp Proc 2008:1128. [Medline: 18999179]
24. Campion T, Edwards A, Johnson S, Kaushal R. Health information exchange system usage patterns in three communities: practice sites, users, patients, and data. Int J Med Inform 2013 Sep;82(9):810-820. [doi: 10.1016/j.ijmedinf.2013.05.001] [Medline: 23743323]
25. Vest JR, Jasperson S. How are health professionals using health information exchange systems? Measuring usage for evaluation and system improvement. J Med Syst 2012 Oct;36(5):3195-3204 [FREE Full text] [doi: 10.1007/s10916-011-9810-2] [Medline: 22127521]
26. Kern LM, Kaushal R. Health information technology and health information exchange in New York State: new initiatives in implementation and evaluation. J Biomed Inform 2007 Dec;40(6 Suppl):S17-S20 [FREE Full text] [doi: 10.1016/j.jbi.2007.08.010] [Medline: 17945542]
27. Lau F, Hagens S, Muttitt S. A proposed benefits evaluation framework for health information systems in Canada. Healthc Q 2007;10(1):112-6, 118. [Medline: 17326376]
28. Delone W, McLean E. The DeLone and McLean model of information systems success: a ten-year update. J Manage Inf Syst 2003;19(4):9-30.
29. Bossen C, Jensen L, Udsen F. Evaluation of a comprehensive EHR based on the DeLone and McLean model for IS success: approach, results, and success factors. Int J Med Inform 2013 Oct;82(10):940-953. [doi: 10.1016/j.ijmedinf.2013.05.010] [Medline: 23827768]
30. Garde S, Knaup P, Hovenga E, Heard S. Towards semantic interoperability for electronic health records. Methods Inf Med 2007;46(3):332-343. [doi: 10.1160/ME5001] [Medline: 17492120]
31. Eckman B, Bennett C, Kaufman J, Tenner J. Varieties of interoperability in the transformation of the health-care information infrastructure. IBM Syst J 2007;46(1):19-41. [doi: 10.1147/sj.461.0019]
32. Kierkegaard P, Kaushal R, Vest JR. How could health information exchange better meet the needs of care practitioners? Appl Clin Inform 2014;5(4):861-877 [FREE Full text] [doi: 10.4338/ACI-2014-06-RA-0055] [Medline: 25589903]
33. Vest J, Campion T, Kaushal R. Challenges, alternatives, and paths to sustainability for health information exchange efforts. J Med Syst 2013 Dec;37(6):9987. [doi: 10.1007/s10916-013-9987-7] [Medline: 24141531]
34. Kern LM, Wilcox A, Shapiro J, Dhopeshwarkar RV, Kaushal R. Which components of health information technology will drive financial value? Am J Manag Care 2012 Aug;18(8):438-445 [FREE Full text] [Medline: 22928759]
35. Bassi J, Lau F. Measuring value for money: a scoping review on economic evaluation of health information systems. J Am Med Inform Assoc 2013;20(4):792-801 [FREE Full text] [doi: 10.1136/amiajnl-2012-001422] [Medline: 23416247]

Abbreviations

COACH: Canada’s Health Informatics Association
EMR: electronic medical record
HIE: health information exchange
IT: information technology
NHS: National Health Service

Edited by A Moorhead; submitted 18.08.15; peer-reviewed by AM Almeida, L Alpay; comments to author 09.10.15; revised version received 25.11.15; accepted 03.12.15; published 14.04.16.

Please cite as:
Flott K, Callahan R, Darzi A, Mayer E
A Patient-Centered Framework for Evaluating Digital Maturity of Health Services: A Systematic Review
J Med Internet Res 2016;18(4):e75
URL: http://www.jmir.org/2016/4/e75/
doi:10.2196/jmir.5047
PMID:27080852

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org